THE COMMUNICATIONS HANDBOOK

SECOND EDITION

Editor-in-Chief

Jerry D. Gibson

Southern Methodist University
Dallas, Texas

CRC PRESS

Boca Raton London New York Washington, D.C.

Library of Congress Cataloging-in-Publication Data

Catalog record is available from the Library of Congress

Direct all inquiries to CRC Press LLC, 2000 N.W. Corporate Blvd., Boca Raton, Florida 33431.

Visit the CRC Press Web site at www.crcpress.com

International Standard Book Number 0-8493-0967-0
Printed in the United States of America 1 2 3 4 5 6 7 8 9 0
Printed on acid-free paper

Preface

The handbook series published by CRC Press represents a truly unique approach to disseminating technical information. Starting with the first edition of *The Electrical Engineering Handbook*, edited by Richard Dorf and published in 1993, this series is dedicated to the idea that a reader should be able to pull one of these handbooks off the shelf and, at least 80% of the time, find what he or she needs to know about a subject area. As handbooks, these books are also different in that they are more than just a dry listing of facts and data, filled mostly with tables. In fact, a hallmark of these handbooks is that the articles or chapters are designed to be relatively short, written as tutorials or overviews, so that once the reader locates the broader topic, it is easy to find an answer to a specific question.

Of course, the authors are the key to achieving the overall goal of the handbook, and, having read all of the chapters personally, the results are impressive. The chapters are authoritative, to-the-point, and enjoyable to read. Answers to frequently asked questions, facts, and figures are available almost at a glance. Since the authors are experts in their field, it is understandable that the content is excellent. Additionally, the authors were encouraged to put some of their own interpretations and insights into the chapters, which greatly enhances the readability. However, I am most impressed by the ability of the authors to condense so much information into so few pages. These chapters are unlike any ever written—they are not research journal articles, they are not textbooks, they are not long tutorial review articles for magazines. They really are a new format.

In reading drafts of the chapters, I applied two tests. If the chapter covered a topic with which I was familiar, I checked to see if it contained what I thought were the essential facts and ideas. If the chapter was in an area of communications less familiar to me, I looked for definitions of terms I had heard of or for a discussion that informed me why this area is important and what is happening in the field today. I was amazed at what I learned.

Using *The Communications Handbook* is simple. Look up your topic of interest either in the Table of Contents or the Index. Go directly to the relevant chapter. As you are reading the chosen chapter, you may wish to refer to other chapters in the same main heading. If you need some background information in the general communications field, you need only consult Section I, Basic Principles. There is no need to read *The Communications Handbook* beginning-to-end, start-to-finish. Look up what you need right now, read it, and go back to the task at hand.

The pleasure of such a project as this is in working with the authors, and I am gratified to have had this opportunity and to be associated with each of them. The first edition of this handbook was well-received and served an important role in providing information on communications to technical and non-technical readers alike. I hope that this edition is found to be equally useful.

It is with great pleasure that I acknowledge the encouragement and patience of my editor at CRC Press, Nora Konopka, and the efforts of both Nora and Helena Redshaw in finally getting this edition in print.

Jerry D. Gibson
Editor-in-Chief

Editor-in-Chief

Jerry D. Gibson currently serves as chairman of the Department of Electrical Engineering at Southern Methodist University in Dallas, Texas. From 1987 to 1997, he held the J. W. Runyon, Jr. Professorship in the Department of Electrical Engineering at Texas A&M University. He also has held positions at General Dynamics—Fort Worth, the University of Notre Dame, and the University of Nebraska—Lincoln, and during the fall of 1991, Dr. Gibson was on sabbatical with the Information Systems Laboratory and the Telecommunications Program in the Department of Electrical Engineering at Stanford University.

He is co-author of the books *Digital Compression for Multimedia* (Morgan–Kaufmann, 1998) and *Introduction to Nonparametric Detection with Applications* (Academic Press, 1975 and IEEE Press, 1995) and author of the textbook *Principles of Digital and Analog Communications* (Prentice-Hall, 2nd ed., 1993). He is editor-in-chief of *The Mobile Communications Handbook* (CRC Press, 2nd ed., 1999) and *The Communications Handbook* (CRC Press, 1996) and editor of the book *Multimedia Communications: Directions and Innovations* (Academic Press, 2000).

Dr. Gibson was associate editor for speech processing for the *IEEE Transactions on Communications* from 1981 to 1985 and associate editor for communications for the *IEEE Transactions on Information Theory* from 1988 to 1991. He has served as a member of the Speech Technical Committee of the IEEE Signal Processing Society (1992–1995), the IEEE Information Theory Society Board of Governors (1990–1998), and the Editorial Board for the *Proceedings of the IEEE*. He was president of the IEEE Information Theory Society in 1996. Dr. Gibson served as technical program chair of the 1999 IEEE Wireless Communications and Networking Conference, technical program chair of the 1997 Asilomar Conference on Signals, Systems, and Computers, finance chair of the 1994 IEEE International Conference on Image Processing, and general co-chair of the 1993 IEEE International Symposium on Information Theory. Currently, he serves on the steering committee for the Wireless Communications and Networking Conference.

In 1990, Dr. Gibson received The Fredrick Emmons Terman Award from the American Society for Engineering Education, and, in 1992, was elected fellow of the IEEE "for contributions to the theory and practice of adaptive prediction and speech waveform coding." He was co-recipient of the 1993 IEEE Signal Processing Society Senior Paper Award for the field of speech processing.

His research interests include data, speech, image, and video compression, multimedia over networks, wireless communications, information theory, and digital signal processing.

Contributors

Joseph A. Bannister
The Aerospace Corporation
Marina del Rey, California

Melbourne Barton
Telcordia Technologies
Red Bank, New Jersey

Vijay K. Bhargava
University of Victoria
Victoria, British Columbia, Canada

Ezio Biglieri
Politecnico di Torino
Torino, Italy

Anders Bjarklev
Technical University of Denmark
Lyngby, Denmark

Daniel J. Blumenthal
University of California
Santa Barbara, California

Helmut Bölcskei
ETH Zurich
Zurich, Switzerland

Madhukar Budagavi
Texas Instruments
Dallas, Texas

James J. Caffery, Jr.
University of Cincinnati
Cincinnati, Ohio

Pierre Catala
Texas A&M University
College Station, Texas

Wai-Yip Chan
Illinois Institute of Technology
Chicago, Illinois

Li Fung Chang
Mobilink Telecom
Middletown, New Jersey

Biao Chen
University of Texas
Richardson, Texas

Matthew Cheng
Mobilink Telecom
Middletown, New Jersey

Giovanni Cherubini
IBM Research
Ruschlikon, Switzerland

Niloy Choudhury
Network Elements, Inc.
Beaverton, Oregon

Youn Ho Choung
TRW Electronics & Technology
Rolling Hills Estates, California

Leon W. Couch, II
University of Florida
Gainesville, Florida

Donald C. Cox
Stanford University
Stanford, California

Rene L. Cruz
University of California
La Jolla, California

Marc Delprat
Alcatel Mobile Network Division
Velizy, France

Paul Diament
Columbia University
New York, New York

Robert L. Douglas
Harding University
Searcy, Arkansas

Eric Dubois
University of Ottawa
Ottawa, Ontario, Canada

Niloy K. Dutta
University of Connecticut
Storrs, Connecticut

Bruce R. Elbert
Application Technology Strategy, Inc.
Thousand Oaks, California

Ahmed K. Elhakeem
Concordia University
Montreal, Quebec, Canada

Ivan J. Fair
University of Alberta
Edmonton, Alberta, Canada

Michael D. Floyd
Motorola Semiconductor Products
Austin, Texas

Lew E. Franks
University of Massachusetts
Amherst, Massachusetts

Susan A. R. Garrod
Purdue University
West Lafayette, Indiana

Costas N. Georghiades
Texas A&M University
College Station, Texas

Ira Gerson
AUVO Technologies, Inc.
Itasca, Illinois

Jerry D. Gibson
Southern Methodist University
Dallas, Texas

Paula R.C. Gomez
FEEC/UNICAMP
Campinas, Brazil

Steven D. Gray
Nokia Research Center
Espoo, Finland

David Haccoun
École Polytechnique de Montréal
Montreal, Quebec, Canada

Frederick Halsall
University of Wales, Swansea
East Sussex, England

Jeff Hamilton
General Instrument Corporation
Doylestown, Pennsylvania

Lajos Hanzo
University of Southampton
Southampton, England

Roger Haskin
IBM Almaden Research Center
San Jose, California

Tor Helleseth
University of Bergen
Bergen, Norway

Garth D. Hillman
Motorola Semiconductor Products
Austin, Texas

Michael L. Honig
Northwestern University
Evanston, Illinois

Hwei P. Hsu
Farleigh Dickinson University
Teaneck, New Jersey

Erwin C. Hudson
WildBlue Communications, Inc.
Denver, Colorado

Yeonging Hwang
Pinnacle EM Wave
Los Altos Hills, California

Louis J. Ippolito, Jr.
ITT Industries
Ashburn, Virginia

Bijan Jabbari
George Mason University
Fairfax, Virginia

Ravi K. Jain
Telcordia Technologies
Morristown, New Jersey

Varun Kapoor
Ericsson Wireless Communication
San Diego, California

Byron L. Kasper
Agere Systems
Irwindale, California

Mark Kolber
General Instrument Corporation
Horsham, Pennsylvania

Boneung Koo
Kyonggi University
Seoul, Korea

P. Vijay Kumar
University of Southern California
Los Angeles, California

Vinod Kumar
Alcatel Research & Innovation
Marcoussis, France

B. P. Lathi
California State University
Carmichael, California

Allen H. Levesque
Worcester Polytechnic Institute
Worcester, Massachusetts

Curt A. Levis
The Ohio State University
Columbus, Ohio

Chung-Sheng Li
IBM T.J. Watson Research Center
Hawthorne, New York

Yi-Bing Lin
National Chaio Tung University
Hsinchu, Taiwan

Joseph L. LoCicero
Illinois Institute of Technology
Chicago, Illinois

John Lodge
Communications Research Centre
Ottawa, Ontario, Canada

Mari W. Maeda
IPA from CNRI/DARPA/ITO
Arlington, Virginia

Nicholas Malcolm
Hewlett-Packard (Canada) Ltd.
Burnaby, British Columbia Canada

Masud Mansuripur
University of Arizona
Tucson, Arizona

Gerald A. Marin
University of Central Florida
Orlando, Florida

Nasir D. Memon
Polytechnic University
Brooklyn, New York

Paul Mermelstein
INRS-Telecommunications
Montreal, Quebec, Canada

Toshio Miki
NTT Mobile Communications
Network, Inc.
Tokyo, Japan

Laurence B. Milstein
University of California
La Jolla, California

Abdi R. Modarressi
BellSouth Science and
Technology
Atlanta, Georgia

Seshadri Mohan
Comverse Network Systems
Wakefield, Massachusetts

Michael Moher
Communications Research Centre
Ottawa, Ontario, Canada

Jaekyun Moon
University of Minnesota
Minneapolis, Minnesota

Madihally J. Narasimha
Stanford University
Stanford, California

A. Michael Noll
University of Southern California
Los Angeles, California

Peter Noll
Technische Universität
Berlin, Germany

Peter P. Nuspl
W. L. Pritchard & Co.
Bethesda, Maryland

Michael O'Flynn
San Jose State University
San Jose, California

Tero Ojanperä
Nokia Group
Espoo, Finland

Raif O. Onvural
Orologic
Camarillo, California

Geoffrey C. Orsak
Southern Methodist University
Dallas, Texas

Kaveh Pahlavan
Worcester Polytechnic Institute
Worcester, Massachusetts

Joseph C. Palais
Arizona State University
Tempe, Arizona

Bernd-Peter Paris
George Mason University
Fairfax, Virginia

Bhasker P. Patel
Illinois Institute of Technology
Chicago, Illinois

Achille Pattavina
Politecnico di Milano
Milan, Italy

Arogyaswami J. Paulraj
Stanford University
Stanford, California

Ken D. Pedrotti
University of California
Santa Cruz, California

Roman Pichna
Nokia
Espoo, Finland

Samuel Pierre
Université du Québec
Quebec, Canada

Alistair J. Price
Corvis Corporation
Columbia, Maryland

David R. Pritchard
Norlight Telecommunications
Skokie, Illinois

Wilbur Pritchard
W. L. Pritchard & Co.
Bethesda, Maryland

John G. Proakis
Northeastern University
Boston, Massachusetts

Bala Rajagopalan
Tellium, Inc.
Ocean Port, New Jersey

Ramesh R. Rao
University of California
La Jolla, California

A. L. Narasimha Reddy
Texas A&M University
College Station, Texas

Whitham D. Reeve
Reeve Engineers
Anchorage, Alaska

Daniel Reininger
Semandex Networks, Inc.
Princeton, New Jersey

Bixio Rimoldi
Swiss Federal Institute of Technology
Lausanne, Switzerland

Thomas G. Robertazzi
SUNY at Stony Brook
Stony Brook, New York

Martin S. Roden
California State University
Los Angeles, California

Izhak Rubin
University of California
Los Angeles, California

Khalid Sayood
University of Nebraska—Lincoln
Lincoln, Nebraska

Charles E. Schell
General Instrument Corporation
Churchville, Pennsylvania

Erchin Serpedin
Texas A & M University
College Station, Texas

A. Udaya Shankar
University of Maryland
College Park, Maryland

Marvin K. Simon
Jet Propulsion Laboratory
Pasadena, California

Suresh Singh
Portland State University
Portland, Oregon

Bernard Sklar
Communications Engineering Services
Tarzana, California

David R. Smith
George Washington University
Ashburn, Virginia

Jonathan M. Smith
University of Pennsylvania
Philadelphia, Pennsylvania

Richard G. Smith
Cenix, Inc.
Center Valley, Pennsylvania

Raymond Steele
Multiple Access Communications
Southampton, England

Gordon L. Stüber
Georgia Institute of Technology
Atlanta, Georgia

Raj Talluri
Texas Instruments
Dallas, Texas

Len Taupier
General Instrument Corporation
Hatboro, Pennsylvania

Jahangir A. Tehranil
Iranian Society of Consulting Engineering
Tehran, Iran

Rüdiger Urbanke
Swiss Federal Institute of Technology
Lausanne, Switzerland

Harrell J. Van Norman
Unisys Corporation
Miamisburg, Ohio

Qiang Wang
University of Victoria
Victoria, Canada

Richard H. Williams
University of New Mexico
Corrales, New Mexico

Maynard A. Wright
Acterna
San Diego, California

Michel Daoud Yacoub
FEEC/UNICAMP
Campinas, Brazil

Shinji Yamashita
The University of Tokyo
Tokyo, Japan

Lie-Liang Yang
University of Southampton
Southampton, England

William C. Young
Bell Communications Research
Middletown, New Jersey

Wei Zhao
Texas A&M University
College Station, Texas

Roger E. Ziemer
University of Colorado
Colorado Springs, Colorado

Contents

SECTION I Basic Principles

SECTION II Telephony

SECTION III Networks

SECTION IV Optical

SECTION V Satellite

SECTION VI Wireless

SECTION VII Source Compression

SECTION VIII Data Recording

I

Basic Principles

1

Complex Envelope Representations for Modulated Signals*

Leon W. Couch, II
University of Florida

1.1 Introduction

What is a general representation for bandpass digital and analog signals? How do we represent a **modulated signal**? How do we evaluate the spectrum and the power of these signals? These are some of the questions that are answered in this chapter.

A *baseband* waveform has a spectral magnitude that is nonzero for frequencies in the vicinity of the origin (i.e., $f = 0$) and negligible elsewhere. A *bandpass* waveform has a spectral magnitude that is nonzero for frequencies in some band concentrated about a frequency $f = \pm f_c$ (where $f_c \gg 0$), and the spectral magnitude is negligible elsewhere. f_c is called the *carrier frequency*. The value of f_c may be arbitrarily assigned for mathematical convenience in some problems. In others, namely, **modulation** problems, f_c is the frequency of an oscillatory signal in the transmitter circuit and is the assigned frequency of the transmitter, such as 850 kHz for an AM broadcasting station.

In communication problems, the information source signal is usually a baseband signal, for example, a transistor–transistor logic (TTL) waveform from a digital circuit or an audio (analog) signal from a microphone. The communication engineer has the job of building a system that will transfer the information from this source signal to the desired destination. As shown in Fig. 1.1, this usually requires the use of a bandpass signal, $s(t)$, which has a bandpass spectrum that is concentrated at $\pm f_c$, where f_c is selected so that $s(t)$ will propagate across the communication channel (either a wire or a wireless channel).

Modulation is the process of imparting the source information onto a bandpass signal with a carrier frequency f_c by the introduction of amplitude and/or phase perturbations. This bandpass signal is called the *modulated* signal $s(t)$, and the baseband source signal is called the *modulating* signal $m(t)$. Examples of

**Source:* Couch, L. W., II. 2001. *Digital and Analog Communication Systems,* 6th ed., Prentice Hall, Upper Saddle River, NJ.

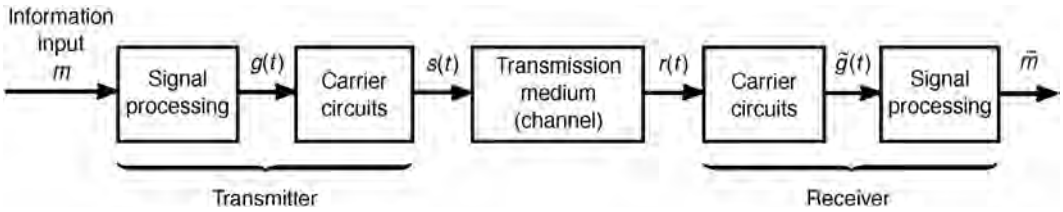


FIGURE 1.1 Bandpass communication system. *Source:* Couch, L.W., II. 2001. *Digital and Analog Communication Systems,* 6th ed., Prentice Hall, Upper Saddle River, NJ, p. 231. With permission.

exactly how modulation is accomplished are given later in this chapter. This definition indicates that modulation may be visualized as a mapping operation that maps the source information onto the bandpass signal $s(t)$ that will be transmitted over the channel.

As the modulated signal passes through the channel, noise corrupts it. The result is a bandpass signal-plus-noise waveform that is available at the receiver input, $r(t)$, as illustrated in Fig. 1.1. The receiver has the job of trying to recover the information that was sent from the source; $\tilde{m}$ denotes the corrupted version of m.

1.2 Complex Envelope Representation

All bandpass waveforms, whether they arise from a modulated signal, interfering signals, or noise, may be represented in a convenient form given by the following theorem. $v(t)$ will be used to denote the **bandpass waveform** canonically. That is, $v(t)$ can represent the signal when $s(t) \equiv v(t)$, the noise when $n(t) \equiv v(t)$, the filtered signal plus noise at the channel output when $r(t) \equiv v(t)$, or any other type of bandpass waveform.*

Theorem 1.1 *Any physical bandpass waveform can be represented by*

$$v(t) = Re\{g(t)e^{j\omega_c t}\} \tag{1.1a}$$

$Re\{\cdot\}$ denotes the real part of $\{\cdot\}$. $g(t)$ is called the complex envelope of $v(t)$, and f_c is the associated carrier frequency (hertz) where $\omega_c = 2\pi f_c$. Furthermore, two other equivalent representations are

$$v(t) = R(t)\cos[\omega_c t + \theta(t)] \tag{1.1b}$$

and

$$v(t) = x(t)\cos\omega_c t - y(t)\sin\omega_c t \tag{1.1c}$$

where

$$g(t) = x(t) + jy(t) = |g(t)|e^{j\angle g(t)} \equiv R(t)e^{j\theta(t)} \tag{1.2}$$

$$x(t) = Re\{g(t)\} \equiv R(t)\cos\theta(t) \tag{1.3a}$$

$$y(x) = Im\{g(t)\} \equiv R(t)\sin\theta(t) \tag{1.3b}$$

$$R(t) \triangleq |g(t)| \equiv \sqrt{x^2(t) + y^2(t)} \tag{1.4a}$$

$$\theta(t) \triangleq \angle g(t) = \tan^{-1}\left(\frac{y(t)}{x(t)}\right) \tag{1.4b}$$

*The symbol $\equiv$ denotes an equivalence and the symbol $\triangleq$ denotes a definition.

The waveforms $g(t)$, $x(t)$, $y(t)$, $R(t)$, and $\theta(t)$ are all **baseband waveforms**, and, except for $g(t)$, they are all real waveforms. $R(t)$ is a nonnegative real waveform. Equation (1.1) is a low-pass-to-bandpass transformation. The $e^{j\omega_c t}$ factor in Eq. (1.1a) shifts (i.e., translates) the spectrum of the baseband signal $g(t)$ from baseband up to the carrier frequency f_c. In communications terminology, the frequencies in the baseband signal $g(t)$ are said to be *heterodyned* up to f_c. The **complex envelope**, $g(t)$, is usually a complex function of time and it is the generalization of the phasor concept. That is, if $g(t)$ happens to be a complex constant, then $v(t)$ is a pure sine wave of frequency f_c and this complex constant is the phasor representing the sine wave. If $g(t)$ is not a constant, then $v(t)$ is not a pure sine wave because the amplitude and phase of $v(t)$ varies with time, caused by the variations of $g(t)$.

Representing the complex envelope in terms of two real functions in Cartesian coordinates, we have

$$g(x) \equiv x(t) + jy(t) \tag{1.5}$$

where $x(t) = \mathrm{Re}\{g(t)\}$ and $y(t) = \mathrm{Im}\{g(t)\}$. $x(t)$ is said to be the *in-phase modulation* associated with $v(t)$, and $y(t)$ is said to be the *quadrature modulation* associated with $v(t)$. Alternatively, the polar form of $g(t)$, represented by $R(t)$ and $\theta(t)$, is given by Eq. (1.2), where the identities between Cartesian and polar coordinates are given by Eqs. (1.3) and (1.4). $R(t)$ and $\theta(t)$ are real waveforms, and, in addition, $R(t)$ is always nonnegative. $R(t)$ is said to be the *amplitude modulation* (AM) on $v(t)$, and $\theta(t)$ is said to be the *phase modulation* (PM) on $v(t)$.

The usefulness of the complex envelope representation for bandpass waveforms cannot be overemphasized. In modern communication systems, the bandpass signal is often partitioned into two channels, one for $x(t)$ called the I (in-phase) channel and one for $y(t)$ called the Q (quadrature-phase) channel. In digital computer simulations of bandpass signals, the sampling rate used in the simulation can be minimized by working with the complex envelope, $g(t)$, instead of with the bandpass signal, $v(t)$, because $g(t)$ is the baseband equivalent of the bandpass signal [1].

1.3 Representation of Modulated Signals

Modulation is the process of encoding the source information $m(t)$ (modulating signal) into a bandpass signal $s(t)$ (modulated signal). Consequently, the modulated signal is just a special application of the bandpass representation. The *modulated signal* is given by

$$s(t) = \mathrm{Re}\{g(t)e^{j\omega_c t}\} \tag{1.6}$$

where $\omega_c = 2\pi f_c$. f_c is the carrier frequency. The complex envelope $g(t)$ is a function of the modulating signal $m(t)$. That is,

$$g(t) = g[m(t)] \tag{1.7}$$

Thus $g[\cdot]$ performs a mapping operation on $m(t)$. This was shown in Fig. 1.1.

Table 1.1 gives an overview of the *big picture* for the modulation problem. Examples of the mapping function $g[m]$ are given for amplitude modulation (AM), double-sideband suppressed carrier (DSB-SC), phase modulation (PM), frequency modulation (FM), single-sideband AM suppressed carrier (SSB-AM-SC), single-sideband PM (SSB-PM), single-sideband FM (SSB-FM), single-sideband envelope detectable (SSB-EV), single-sideband square-law detectable (SSB-SQ), and quadrature modulation (QM). For each $g[m]$, Table 1.1 also shows the corresponding $x(t)$ and $y(t)$ quadrature modulation components and the corresponding $R(t)$ and $\theta(t)$ amplitude and phase modulation components. Digitally modulated bandpass signals are obtained when $m(t)$ is a digital baseband signal, for example, the output of a transistor transistor logic (TTL) circuit.

Obviously, it is possible to use other $g[m]$ functions that are not listed in Table 1.1. The question is: are they useful? $g[m]$ functions are desired that are easy to implement and that will give desirable spectral

TABLE 1.1 Complex Envelope Functions for Various Types of Modulation[a]

Type of Modulation	Mapping Function $g(m)$	Corresponding Quadrature Modulation $x(t)$	$y(t)$
AM	$A_c[1+m(t)]$	$A_c[1+m(t)]$	0
DSB-SC	$A_c m(t)$	$A_c m(t)$	0
PM	$A_c e^{jD_p m(t)}$	$A_c \cos[D_p m(t)]$	$A_c \sin[D_p m(t)]$
FM	$A_c e^{jD_f \int_{-\infty}^{t} m(\sigma)d\sigma}$	$A_c \cos\left[D_f \int_{-\infty}^{t} m(\sigma)d\sigma\right]$	$A_c \sin\left[D_f \int_{-\infty}^{t} m(\sigma)d\sigma\right]$
SSB-AM-SC[b]	$A_c[m(t) \pm j\hat{m}(t)]$	$A_c m(t)$	$\pm A_c \hat{m}(t)$
SSB-PM[b]	$A_c e^{jD_p[m(t) \pm j\hat{m}(t)]}$	$A_c e^{\mp D_p \hat{m}(t)} \cos[D_p m(t)]$	$A_c e^{\mp D_p \hat{m}(t)} \sin[D_p m(t)]$
SSB-FM[b]	$A_c e^{jD_f \int_{-\infty}^{t} [m(\sigma) \pm j\hat{m}(\sigma)]d\sigma}$	$A_c e^{\mp D_f \int_{-\infty}^{t} \hat{m}(\sigma)d\sigma} \cos\left[D_f \int_{-\infty}^{t} m(\sigma)d\sigma\right]$	$A_c e^{\mp D_f \int_{-\infty}^{t} \hat{m}(\sigma)d\sigma} \sin\left[D_f \int_{-\infty}^{t} m(\sigma)d\sigma\right]$
SSB-EV[b]	$A_c e^{\{\ln[1+m(t)] \pm j\hat{\ln}\lvert 1+m(t)\rvert\}}$	$A_c[1+m(t)]\cos\{\hat{\ln}[1+m(t)]\}$	$\pm A_c[1+m(t)]\sin\{\hat{\ln}[1+m(t)]\}$
SSB-SQ[b]	$A_c e^{(1/2)\{\ln[1+m(t)] \pm j\hat{\ln}\lvert 1+m(t)\rvert\}}$	$A_c\sqrt{1+m(t)}\cos\left\{\frac{1}{2}\hat{\ln}[1+m(t)]\right\}$	$\pm A_c\sqrt{1+m(t)}\sin\left\{\frac{1}{2}\hat{\ln}[1+m(t)]\right\}$
QM	$A_c[m_1(t) + jm_2(t)]$	$A_c m_1(t)$	$A_c m_2(t)$

Type of Modulation	Corresponding Amplitude and Phase Modulation $R(t)$	$\theta(t)$	Linearity	Remarks
AM	$A_c\lvert 1+m(t)\rvert$	$\begin{cases} 0, & m(t) > -1 \\ 180°, & m(t) < -1 \end{cases}$	L[c]	$m(t) > -1$ required for envelope detection
DSB-SC	$A_c\lvert m(t)\rvert$	$\begin{cases} 0, & m(t) > 0 \\ 180°, & m(t) < 0 \end{cases}$	L	Coherent detection required
PM	A_c	$D_p m(t)$	NL	D_p is the phase deviation constant (rad/volt)
FM	A_c	$D_f \int_{-\infty}^{t} m(\sigma)d\sigma$	NL	D_f is the frequency deviation constant (rad/volt-sec)
SSB-AM-SC[b]	$A_c\sqrt{[m(t)]^2 + [\hat{m}(t)]^2}$	$\tan^{-1}[\pm\hat{m}(t)/m(t)]$	L	Coherent detection required
SSB-PM[b]	$A_c e^{\pm D_p \hat{m}(t)}$	$D_p m(t)$	NL	
SSB-FM[b]	$A_c e^{\pm D_f \int_{-\infty}^{t} \hat{m}(\sigma)d\sigma}$	$D_f \int_{-\infty}^{t} m(\sigma)d\sigma$	NL	
SSB-EV[b]	$A_c\lvert 1+m(t)\rvert$	$\pm\hat{\ln}[1+m(t)]$	NL	$m(t) > -1$ is required so that the $\ln(\cdot)$ will have a real value
SSB-SQ[b]	$A_c\sqrt{1+m(t)}$	$\pm\frac{1}{2}\hat{\ln}[1+m(t)]$	NL	$m(t) > -1$ is required so that the $\ln(\cdot)$ will have a real value
QM	$A_c\sqrt{m_1^2(t) + m_2^2(t)}$	$\tan^{-1}[m_2(t)/m_1(t)]$	L	Used in NTSC color television; requires coherent detection

Source: Couch, L.W., II. 2001. *Digital and Analog Communication Systems,* 5th ed., Prentice Hall, Upper Saddle River, NJ, p. 235–236. With permission.

[a] $A_c > 0$ is a constant that sets the power level of the signal as evaluated by use of Eq. (1.11); L, linear; NL, nonlinear; and $[\hat{\cdot}]$ is the Hilbert transform (a −90° phase-shifted version of $[\cdot]$). For example, $\hat{m}(t) = m(t) * \frac{1}{\pi t} = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{m(\lambda)}{t-\lambda}d\lambda$.

[b] Use upper signs for upper sideband signals and lower signals for lower sideband signals.

[c] In the strict sense, AM signals are not linear because the carrier term does not satisfy the linearity (superposition) condition.

properties. Furthermore, in the receiver, the inverse function $m[g]$ is required. The inverse should be single valued over the range used and should be easily implemented. The inverse mapping should suppress as much noise as possible so that $m(t)$ can be recovered with little corruption.

1.4 Generalized Transmitters and Receivers

A more detailed description of transmitters and receivers, as first shown in Fig. 1.1, will now be illustrated.

There are two canonical forms for the generalized transmitter, as indicated by Eqs. (1.1b) and (1.1c). Equation (1.1b) describes an AM-PM type circuit, as shown in Fig. 1.2. The baseband signal processing circuit generates $R(t)$ and $\theta(t)$ from $m(t)$. The R and θ are functions of the modulating signal $m(t)$, as given in Table 1.1, for the particular modulation type desired. The signal processing may be implemented either by using nonlinear analog circuits or a digital computer that incorporates the R and θ algorithms under software program control. In the implementation using a digital computer, one analog-to-digital converter (ADC) will be needed at the input of the baseband signal processor, and two digital-to-analog converters (DACs) will be needed at the output. The remainder of the AM-PM canonical form requires radio frequency (RF) circuits, as indicated in the figure.

Figure 1.3 illustrates the second canonical form for the generalized transmitter. This uses in-phase and quadrature-phase (IQ) processing. Similarly, the formulas relating $x(t)$ and $y(t)$ to $m(t)$ are shown in

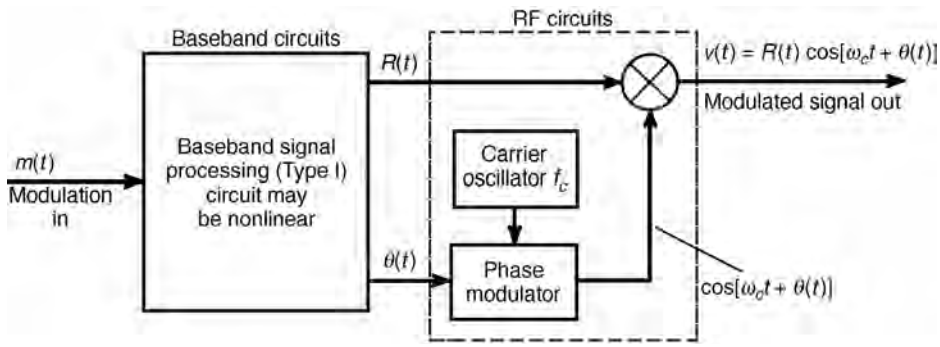


FIGURE 1.2 Generalized transmitter using the AM-PM generation technique. *Source:* Couch, L.W., II. 2001. *Digital and Analog Communication Systems,* 6th ed., Prentice Hall, Upper Saddle River, NJ, p. 282. With permission.

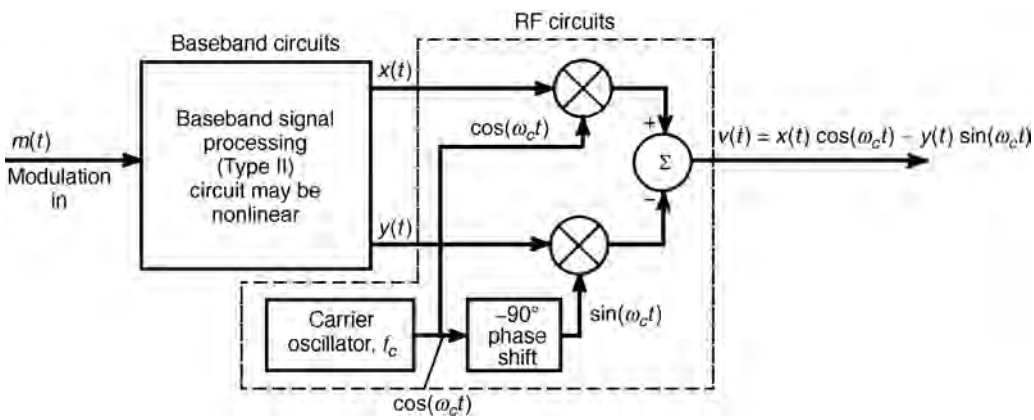


FIGURE 1.3 Generalized transmitter using the quadrature generation technique. *Source:* Couch, L.W., II. 2001. *Digital and Analog Communication Systems,* 6th ed., Prentice Hall, Upper Saddle River, NJ, p. 283. With permission.

Table 1.1, and the baseband signal processing may be implemented by using either analog hardware or digital hardware with software. The remainder of the canonical form uses RF circuits as indicated. Modern transmitters and receivers, such as those used in cellular telephones, often use IQ processing.

Analogous to the transmitter realizations, there are two canonical forms of receiver. Each one consists of RF carrier circuits followed by baseband signal processing, as illustrated in Fig. 1.1. Typically, the carrier circuits are of the superheterodyne-receiver type which consist of an RF amplifier, a down converter (mixer plus local oscillator) to some intermediate frequency (IF), an IF amplifier, and then detector circuits [1]. In the first canonical form of the receiver, the carrier circuits have amplitude and phase detectors that output $\tilde{R}(t)$ and $\tilde{\theta}(t)$, respectively. This pair, $\tilde{R}(t)$ and $\tilde{\theta}(t)$, describe the polar form of the received complex envelope, $\tilde{g}(t)$. $\tilde{R}(t)$ and $\tilde{\theta}(t)$ are then fed into the signal processor, which uses the inverse functions of Table 1.1 to generate the recovered modulation, $\tilde{m}(t)$. The second canonical form of the receiver uses quadrature product detectors in the carrier circuits to produce the Cartesian form (IQ processing) of the received complex envelope, $\tilde{x}(t)$ and $\tilde{y}(t)$. $\tilde{x}(t)$ and $\tilde{y}(t)$ are then inputted to the signal processor, which generates $\tilde{m}(t)$ at its output.

Once again, it is stressed that any type of signal modulation (see Table 1.1) may be generated (transmitted) or detected (received) by using either of these two canonical forms. Both of these forms conveniently separate baseband processing from RF processing. Digital signal processing (DSP) techniques are especially useful to realize the baseband processing portion. Furthermore, if DSP circuits are used, any desired modulation type can be realized by selecting the appropriate software algorithm. This is the basis for software radios [1].

1.5 Spectrum and Power of Bandpass Signals

The spectrum of the bandpass signal is the translation of the spectrum of its complex envelope. Taking the **Fourier transform** of Eq. (1.1a), the spectrum of the bandpass waveform is [1]

$$V(f) = \frac{1}{2}[G(f-f_c) + G^*(-f-f_c)] \tag{1.8}$$

where $G(f)$ is the Fourier transform of $g(t)$,

$$G(f) = \int_{-\infty}^{\infty} g(t)e^{-j2\pi ft}\,dt$$

and the asterisk superscript denotes the complex conjugate operation. The *power spectra density* (PSD) of the bandpass waveform is [1]

$$\mathcal{P}_v(f) = \frac{1}{4}[\mathcal{P}_g(f-f_c) + \mathcal{P}_g(-f-f_c)] \tag{1.9}$$

where $\mathcal{P}_g(f)$ is the PSD of $g(t)$.

The average power dissipated in a resistive load is V_{rms}^2/R_L or $I_{\text{rms}}^2 R_L$, where V_{rms} is the rms value of the voltage waveform across the load and I_{rms} is the rms value of the current through the load. For bandpass waveforms, Eq. (1.1) may represent either the voltage or the current. Furthermore, the rms values of $v(t)$ and $g(t)$ are related by [1]

$$v_{\text{rms}}^2 = \langle v^2(t)\rangle = \frac{1}{2}\langle |g(t)|^2\rangle = \frac{1}{2}g_{\text{rms}}^2 \tag{1.10}$$

where $\langle\cdot\rangle$ denotes the time average and is given by

$$\left\langle \left[\quad\right] \right\rangle = \lim_{t\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}\left[\quad\right]dt$$

Thus, if $v(t)$ of Eq. (1.1) represents the bandpass voltage waveform across a resistive load, the average power dissipated in the load is

$$P_L = \frac{v_{\text{rms}}^2}{R_L} = \frac{\langle v^2(t)\rangle}{R_L} = \frac{\langle |g(t)|^2\rangle}{2R_L} = \frac{g_{\text{rms}}^2}{2R_L} \tag{1.11}$$

where g_{rms} is the rms value of the complex envelope, and R_L is the resistance of the load.

1.6 Amplitude Modulation

Amplitude modulation (AM) will now be examined in more detail. From Table 1.1, the complex envelope of an AM signal is

$$g(t) = A_c[1 + m(t)] \tag{1.12}$$

so that the spectrum of the complex envelope is

$$G(f) = A_c\delta(f) + A_c M(f) \tag{1.13}$$

Using Eq. (1.6), we obtain the AM signal waveform

$$s(t) = A_c[1 + m(t)]\cos\omega_c t \tag{1.14}$$

and, using Eq. (1.8), the AM spectrum

$$S(f) = \frac{1}{2}A_c[\delta(f-f_c) + M(f-f_c) + \delta(f+f_c) + M(f+f_c)] \tag{1.15}$$

where $\delta(f) = \delta(-f)$ and, because $m(t)$ is real, $M^*(f) = M(-f)$. Suppose that the magnitude spectrum of the modulation happens to be a triangular function, as shown in Fig. 1.4(a). This spectrum might arise from an analog audio source where the bass frequencies are emphasized. The resulting AM spectrum, using Eq. (1.15), is shown in Fig. 1.4(b). Note that because $G(f-f_c)$ and $G^*(-f-f_c)$ do not overlap, the magnitude spectrum is

$$|S(f)| = \begin{cases} \frac{1}{2}A_c\delta(f-f_c) + \frac{1}{2}A_c|M(f-f_c)|, & f>0 \\ \frac{1}{2}A_c\delta(f+f)_c + \frac{1}{2}A_c|M(-f-f)_c|, & f<0 \end{cases} \tag{1.16}$$

The 1 in

$$g(t) = A_c[1 + m(t)]$$

causes delta functions to occur in the spectrum at $f = \pm f_c$, where f_c is the assigned carrier frequency. Also, from Fig. 1.4 and Eq. (1.16), it is realized that the bandwidth of the AM signal is 2B. That is, the bandwidth of the AM signal is twice the bandwidth of the baseband modulating signal.

The average power dissipated into a resistive load is found by using Eq. (1.11).

$$P_L = \frac{A_c^2}{2R_L}\langle |1 + m(t)|^2\rangle = \frac{A_c^2}{2R_L}[1 + 2\langle m(t)\rangle + \langle m^2(t)\rangle]$$

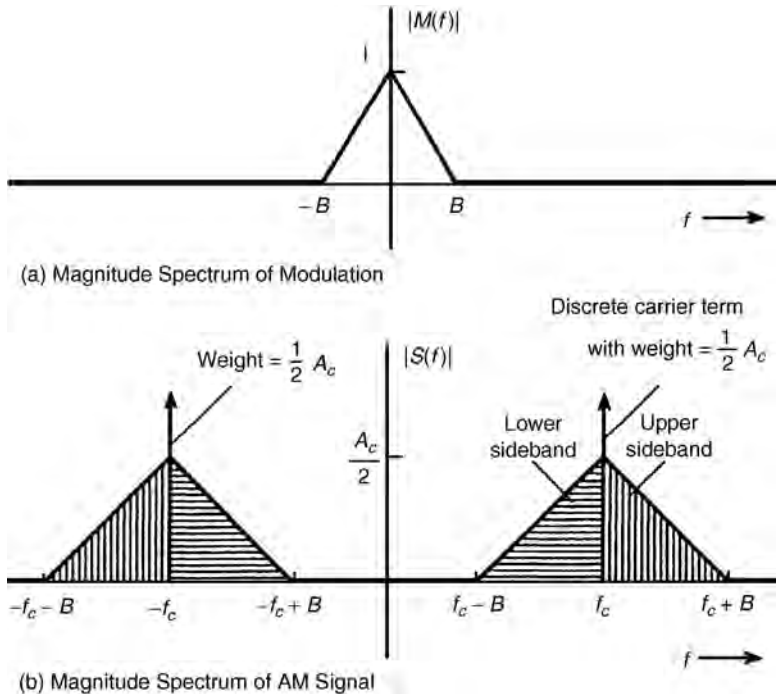


FIGURE 1.4 Spectrum of an AM signal. *Source:* Couch, L.W., II. 2001. *Digital and Analog Communication Systems,* 6th ed., Prentice Hall, Upper Saddle River, NJ, p. 239. With permission.

If we assume that the *dc* value of the modulation is zero, $\langle m(t)\rangle = 0$, then the average power dissipated into the load is

$$P_L = \frac{A_c^2}{2R_L}\langle 1 + m_{\text{rms}}^2\rangle \tag{1.17}$$

where m_{rms} is the rms value of the modulation, $m(t)$. Thus, the average power of an AM signal changes if the rms value of the modulating signal changes. For example, if $m(t)$ is a sine wave test tone with a peak value of 1.0 for 100% modulation,

$$m_{\text{rms}} = 1/\sqrt{2}$$

Assume that $A_c = 1000$ volts and $R_L = 50$ ohms, which are typical values used in AM broadcasting. Then, the average power dissipated into the 50 Ω load for this AM signal is

$$P_L = \frac{(1000)^2}{2(50)}\left[1 + \frac{1}{2}\right] = 15{,}000 \text{ watts} \tag{1.18}$$

The Federal Communications Commission (FCC) rated carrier power is obtained when $m(t) = 0$. In this case, Eq. (1.17) becomes $P_L = (1000)^2/100 = 10{,}000$ watts, and the FCC would rate this as a 10,000 watt AM station. The sideband power for 100% sine wave modulation is 5000 watts.

Now let the modulation on the AM signal be a binary digital signal such that $m(t) = \pm 1$ where +1 is used for a binary one and −1 is used for a binary 0. Referring to Eq. (1.14), this AM signal becomes an *on–off keyed* (OOK) digital signal where the signal is on when a binary one is transmitted and off when a binary zero is transmitted. For $A_c = 1000$ and $R_L = 50$ Ω, the average power dissipated would be 20,000 watts since $m_{\text{rms}} = 1$ for $m(t) = \pm 1$.

1.7 Phase and Frequency Modulation

Phase modulation (PM) and *frequency modulation* (FM) are special cases of angle-modulated signalling. In angle-modulated signalling, the complex envelope is

$$g(t) = A_c e^{j\theta(t)} \tag{1.19}$$

Using Eq. (1.6), the resulting *angle-modulated* signal is

$$s(t) = A_c \cos[\omega_c + \theta(t)] \tag{1.20}$$

For PM, the phase is directly proportional to the modulating signal:

$$\theta(t) = D_p m(t) \tag{1.21}$$

where the proportionality constant D_p is the phase sensitivity of the phase modulator, having units of radians per volt [assuming that $m(t)$ is a voltage waveform]. For FM, the phase is proportional to the integral of $m(t)$:

$$\theta(t) = D_f \int_{-\infty}^{t} m(\sigma)\, d\sigma \tag{1.22}$$

where the frequency deviation constant D_f has units of radians/volt-second. These concepts are summarized by the PM and FM entries in Table 1.1.

By comparing the last two equations, it is seen that if we have a PM signal modulated by $m_p(t)$, there is *also* FM on the signal corresponding to a *different* modulating waveshape that is given by

$$m_f(t) = \frac{D_p}{D_f}\left[\frac{dm_p(t)}{dt}\right] \tag{1.23}$$

where the subscripts f and p denote frequency and phase, respectively. Similarly, if we have an FM signal modulated by $m_f(t)$, the corresponding phase modulation on this signal is

$$m_p(t) = \frac{D_f}{D_p} \int_{-\infty}^{t} m_f(\sigma)\, d\sigma \tag{1.24}$$

By using Eq. (1.24), a PM circuit may be used to synthesize an FM circuit by inserting an integrator in cascade with the phase modulator input.

Other properties of PM and FM are that the **real envelope**, $R(t) = |g(t)| = A_c$, is a constant, as seen from Eq. (1.19). Also, $g(t)$ is a *nonlinear* function of the modulation. However, from Eqs. (1.21) and (1.22), $\theta(t)$ is a linear function of the modulation $m(t)$. Using Eq. (1.11), the average power dissipated by a PM or FM signal is the constant

$$P_L = \frac{A_c^2}{2R_L} \tag{1.25}$$

That is, the average power of a PM or FM signal does not depend on the modulating waveform $m(t)$.

The *instantaneous frequency deviation* for an FM signal from its carrier frequency is given by the derivative of its phase $\theta(t)$. Taking the derivative of Eq. (1.22), the *peak frequency deviation* is

$$\Delta F = \frac{1}{2\pi} D_f M_p \text{ Hz} \tag{1.26}$$

where $M_p = \max[m(t)]$ is the peak value of the modulation waveform and the derivative has been divided by 2π to convert from radians/sec to Hz units.

For FM and PM signals, Carson's rule estimates the transmission bandwidth containing approximately 98% of the total power. This FM or PM signal bandwidth is

$$B_T = 2(\beta + 1)B \tag{1.27}$$

where B is bandwidth (highest frequency) of the modulation. The modulation index β, is $\beta = \Delta F/B$ for FM and $\beta = \max[D_p m(t)] = D_p M_p$ for PM.

The AMPS (Advanced Mobile Phone System) analog cellular phones use FM signalling. A peak deviation of 12 kHz is specified with a modulation bandwidth of 3 kHz. From Eq. (1.27), this gives a bandwidth of 30 kHz for the AMPS signal and allows a channel spacing of 30 kHz to be used. To accommodate more users, narrow-band AMPS (NAMPS) with a 5 kHz peak deviation is used in some areas. This allows 10 kHz channel spacing if the carrier frequencies are carefully selected to minimize interference to used adjacent channels. A maximum FM signal power of 3 watts is allowed for the AMPS phones. However, hand-held AMPS phones usually produce no more than 600 mW, which is equivalent to 5.5 volts rms across the 50 Ω antenna terminals.

The GSM (Group Special Mobile) digital cellular phones use FM with *minimum frequency-shift-keying* (MSK) where the peak frequency deviation is selected to produce orthogonal waveforms for binary one and binary zero data. (Digital phones use a speech codec to convert the analog voice source to a digital data source for transmission over the system.) Orthogonality occurs when $\Delta F = 1/4R$ where R is the bit rate (bits/sec)[1]. Actually, GSM uses Gaussian shaped MSK (GMSK). That is, the digital data waveform (with rectangular binary one and binary zero pulses) is first filtered by a low-pass filter having a Gaussian shaped frequency response (to attenuate the higher frequencies). This Gaussian filtered data waveform is then fed into the frequency modulator to generate the GMSK signal. This produces a digitally modulated FM signal with a relatively small bandwidth.

Other digital cellular standards use QPSK signalling, as discussed in the next section.

1.8 QPSK, π/4 QPSK, QAM, and OOK Signalling

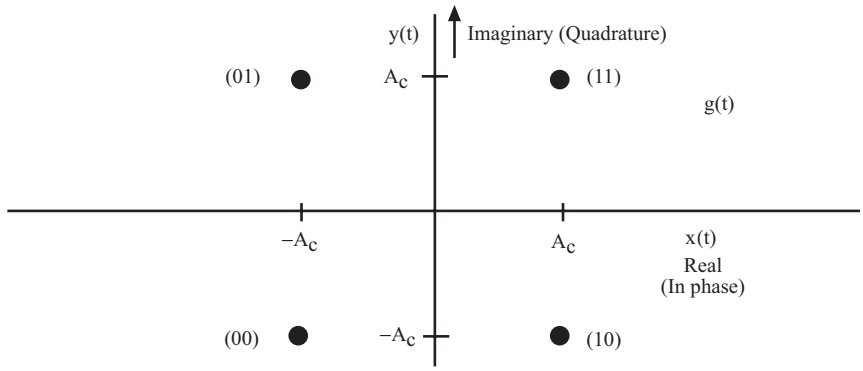
Quadrature phase-shift-keying (QPSK) is a special case of quadrature modulation, as shown in Table 1.1, where $m_1(t) = \pm 1$ and $m_2(t) = \pm 1$ are two binary bit streams. The complex envelope for QPSK is

$$g(t) = x(t) + jy(t) = A_c[m_1(t) + jm_2(t)]$$

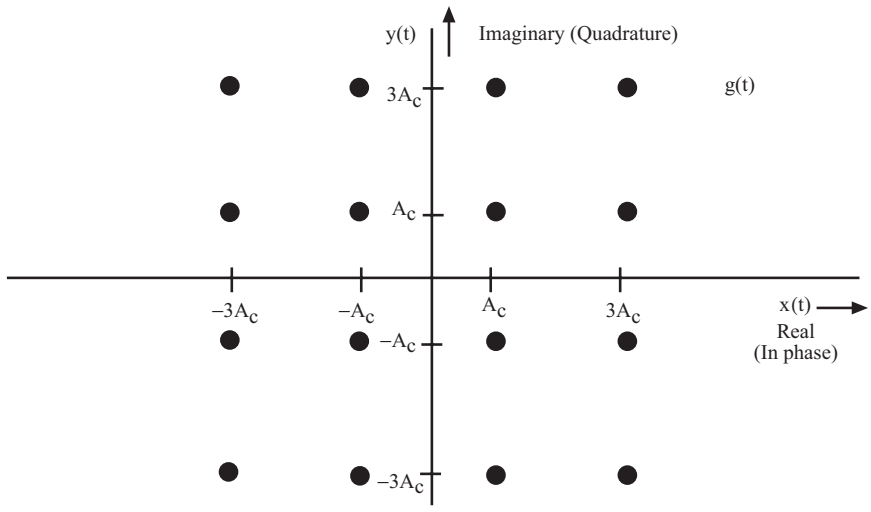
where $x(t) = \pm A_c$ and $y(t) = \pm A_c$. The permitted values for the complex envelope are illustrated by the QPSK **signal constellation** shown in Fig. 1.5a. The *signal constellation* is a plot of the permitted values for the complex envelope $g(t)$. QPSK may be generated by using the quadrature generation technique of Fig. 1.3, where the baseband signal processor is a serial-to-parallel converter that reads in two bits of data at a time from the serial binary input stream, $m(t)$, and outputs the first of the two bits to $x(t)$ and the second bit to $y(t)$. If the two input bits are both binary ones, (11), then $m_1(t) = +A_c$ and $m_2(t) = +A_c$. This is represented by the top right-hand dot for $g(t)$ in the signal constellation for QPSK signalling in Fig. 1.5a. Likewise, the three other possible two-bit words, (10), (01), and (00), are also shown. The QPSK signal is also equivalent to a four-phase phase-shift-keyed signal (4PSK) since all the points in the signal constellation fall on a circle where the permitted phases are $\theta(t) = 45°$, 135°, 225°, and 315°. There is no amplitude modulation on the QPSK signal since the distances from the origin to all the signal points on the signal constellation are equal.

For QPSK, the spectrum of $g(t)$ is of the sin x/x type since $x(t)$ and $y(t)$ consist of rectangular data pulses of value $\pm A_c$. Moreover, it can be shown that for equally likely independent binary one and binary zero data, the power spectral density of $g(t)$ for digitally modulated signals with M point signal constellations is [1]

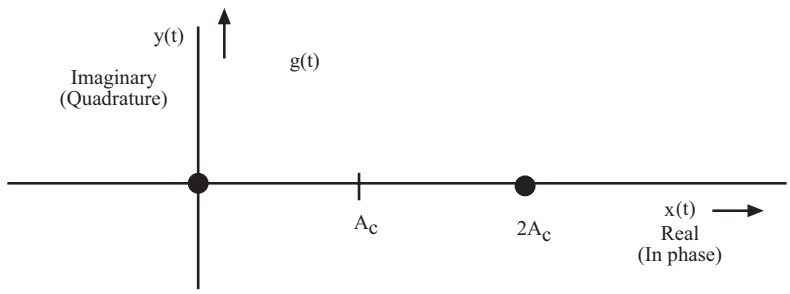
$$\mathcal{P}_g(f) = K\left(\frac{\sin \pi f \ell T_b}{\pi f \ell T_b}\right)^2 \tag{1.28}$$



(a) QPSK Signal Constellation



(b) 16 QAM Signal Constellation



(c) OOK Signal Constellation

FIGURE 1.5 Signal constellations (permitted values of the complex envelope).

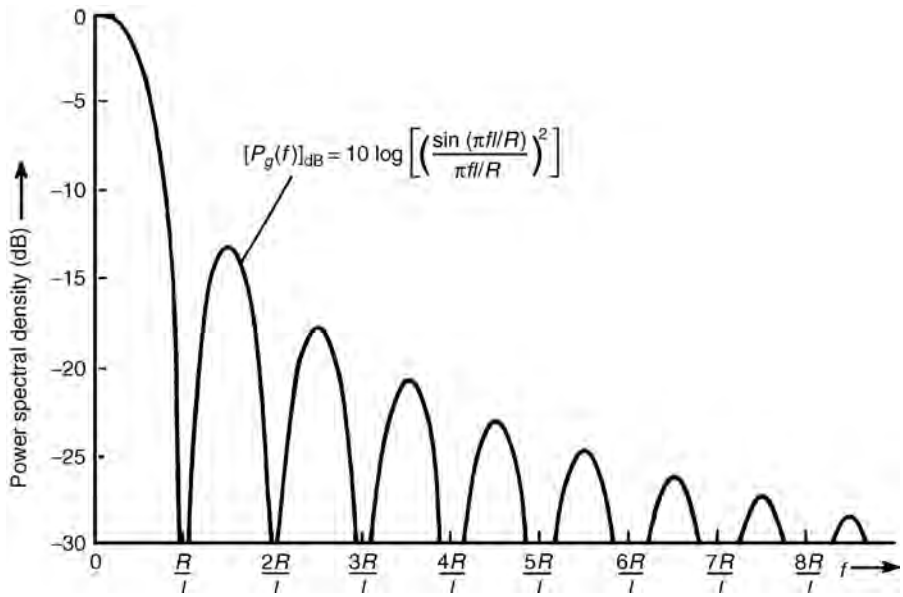


FIGURE 1.6 PSD for the complex envelope of MPSK and QAM where $M = 2^\ell$ and R is bit rate (positive frequencies shown). *Source:* Couch, L.W., II. 2001. *Digital and Analog Communication Systems,* 6th ed., Prentice Hall, Upper Saddle River, NJ, p. 358. With permission.

where K is a constant, $R = 1/T_b$ is the data rate (bits/sec) of $m(t)$, and $M = 2^\ell$. M is the number of points in the signal constellation. For QPSK, $M = 4$ and $\ell = 2$. This PSD for the complex envelope, $\mathcal{P}_g(f)$, is plotted in Fig. 1.6. The PSD for the QPSK signal ($\ell = 2$) is given by translating $\mathcal{P}_g(f)$ up to the carrier frequency as indicated by Eq. (1.9).

Referring to Fig. 1.6 or using Eq. (1.28), the first-null bandwidth of $g(t)$ is R/ℓ Hz. Consequently, the null-to-null bandwidth of the modulated RF signal is

$$B_{\text{null}} = \frac{2R}{\ell} \text{ Hz} \tag{1.29}$$

For example, if the data rate of the baseband information source is 9600 bits/sec, then the null-to-null bandwidth of the QPSK signal would be 9.6 kHz since $\ell = 2$.

Referring to Fig. 1.6, it is seen that the sidelobes of the spectrum are relatively large, so, in practice, the sidelobes of the spectrum are filtered off to prevent interference to the adjacent channels. This filtering rounds off the edges of the rectangular data pulses and this causes some amplitude modulation on the QPSK signal. That is, the points in the signal constellation for the filtered QPSK signal would be fuzzy since the transition from one constellation point to another point is not instantaneous because the filtered data pulses are not rectangular. QPSK is the modulation used for digital cellular phones with the IS-95 Code Division Multiple Access (CDMA) standard.

$\pi/4$ QPSK is generated by alternating symbols between a signal constellation shown by Fig. 1.5a and a signal constellation (of Fig. 1.5a) rotated by 45°. $\pi/4$ QPSK has the same spectrum as QPSK. However, the peak-to-average power ratio of filtered $\pi/4$ QPSK is much lower than that for filtered QPSK because there are no 180° phase transitions (from symbol to symbol) in the $\pi/4$ QPSK signal. $\pi/4$ QPSK is the modulation technique used for IS-136 TDMA digital cellular phones.

Equation (1.28) and Fig. 1.6 also represent the spectrum for *quadrature modulation amplitude modulation* (QAM) signalling. QAM signalling allows more than two values for $x(t)$ and $y(t)$. For example,

QAM where $M = 16$ has 16 points in the signal constellation with 4 values for $x(t)$ and 4 values for $y(t)$ such as, for example, $x(t) = +A_c, -A_c, +3A_c, -3A_c$ and $y(t) = +A_c, -A_c, +3A_c, -3A_c$. This is shown in Fig. 1.5b. Each point in the $M = 16$ QAM signal constellation would represent a unique four-bit data word, as compared with the $M = 4$ QPSK signal constellation shown in Fig. 1.5a, where each point represents a unique two-bit data word. For an $R = 9600$ bits/sec information source data rate, an $M = 16$ QAM signal would have a null-to-null bandwidth of 4.8 kHz since $\ell = 4$.

For OOK signalling, as described at the end of Section 1.6, the signal constellation would consist of $M = 2$ points along the x axis where $x = 0, 2A_c$ and $y = 0$. This is illustrated in Fig. 1.5c. For an $R = 9600$ bit/sec information source data rate, an OOK signal would have a null-to-null bandwidth of 19.2 kHz since $\ell = 1$.

Defining Terms

Bandpass waveform: The spectrum of the waveform is nonzero for frequencies in some band concentrated about a frequency $f_c \gg 0$; f_c is called the carrier frequency.

Baseband waveform: The spectrum of the waveform is nonzero for frequencies near $f = 0$.

Complex envelope: The function $g(t)$ of a bandpass waveform $v(t)$ where the bandpass waveform is described by

$$v(t) = Re\{g(t)e^{j\omega_c t}\}$$

Fourier transform: If $w(t)$ is a waveform, then the Fourier transform of $w(t)$ is

$$W(f) = \Im[w(t)] = \int_{-\infty}^{\infty} w(t)e^{-j2\pi ft}\,dt$$

where f has units of hertz.

Modulated signal: The bandpass signal

$$s(t) = Re\{g(t)e^{j\omega_c t}\}$$

where fluctuations of $g(t)$ are caused by the information source such as audio, video, or data.

Modulation: The information source, $m(t)$, that causes fluctuations in a bandpass signal.

Real envelope: The function $R(t) = |g(t)|$ of a bandpass waveform $v(t)$ where the bandpass waveform is described by

$$v(t) = Re\{g(t)e^{j\omega_c t}\}$$

Signal constellation: The permitted values of the complex envelope for a digital modulating source.

Reference

1. Couch, L.W., II, *Digital and Analog Communication Systems,* 6th ed., Prentice Hall, Upper Saddle River, NJ, 2001.

Further Information

1. Bedrosian, E., The analytic signal representation of modulated waveforms. *Proc. IRE,* vol. 50, October, 2071–2076, 1962.
2. Couch, L.W., II, *Modern Communication Systems: Principles and Applications,* Macmillan Publishing, New York (now Prentice Hall, Upper Saddle River, NJ), 1995.

3. Dugundji, J., Envelopes and pre-envelopes of real waveforms. *IRE Trans. Information Theory,* vol. IT-4, March, 53–57, 1958.
4. Voelcker, H.B., Toward the unified theory of modulation—Part I: Phase-envelope relationships. *Proc. IRE,* vol. 54, March, 340–353, 1966.
5. Voelcker, H.B., Toward the unified theory of modulation—Part II: Zero manipulation. *Proc. IRE,* vol. 54, May, 735–755, 1966.
6. Ziemer, R.E. and Tranter, W.H., *Principles of Communications,* 4th ed., John Wiley & Sons, New York, 1995.

2
Sampling

Hwei P. Hsu
Fairleigh Dickinson University

2.1 Introduction

To transmit analog message signals, such as speech signals or video signals, by digital means, the signal has to be converted into digital form. This process is known as analog-to-digital conversion. The sampling process is the first process performed in this conversion, and it converts a continuous-time signal into a discrete-time signal or a sequence of numbers. Digital transmission of analog signals is possible by virtue of the sampling theorem, and the sampling operation is performed in accordance with the sampling theorem.

In this chapter, using the Fourier transform technique, we present this remarkable sampling theorem and discuss the operation of sampling and practical aspects of sampling.

2.2 Instantaneous Sampling

Suppose we sample an arbitrary analog signal $m(t)$ shown in Fig. 2.1(a) instantaneously at a uniform rate, once every T_s seconds. As a result of this sampling process, we obtain an infinite sequence of samples $\{m(nT_s)\}$, where n takes on all possible integers. This form of sampling is called *instantaneous sampling*. We refer to T_s as the **sampling interval** and to its reciprocal $1/T_s = f_s$ as the **sampling rate**. Sampling rate (samples per second) is often cited in terms of sampling frequency expressed in hertz.

Ideal Sampled Signal

Let $m_s(t)$ be obtained by multiplication of $m(t)$ by the unit impulse train $\delta_T(t)$ with period T_s [Fig. 2.1(c)], that is,

$$m_s(t) = m(t)\delta_{T_s}(t) = m(t)\sum_{n=-\infty}^{\infty}\delta(t-nT_s)$$
$$= \sum_{n=-\infty}^{\infty} m(t)\delta(t-nT_s) = \sum_{n=-\infty}^{\infty} m(nT_s)\delta(t-nT_s) \tag{2.1}$$

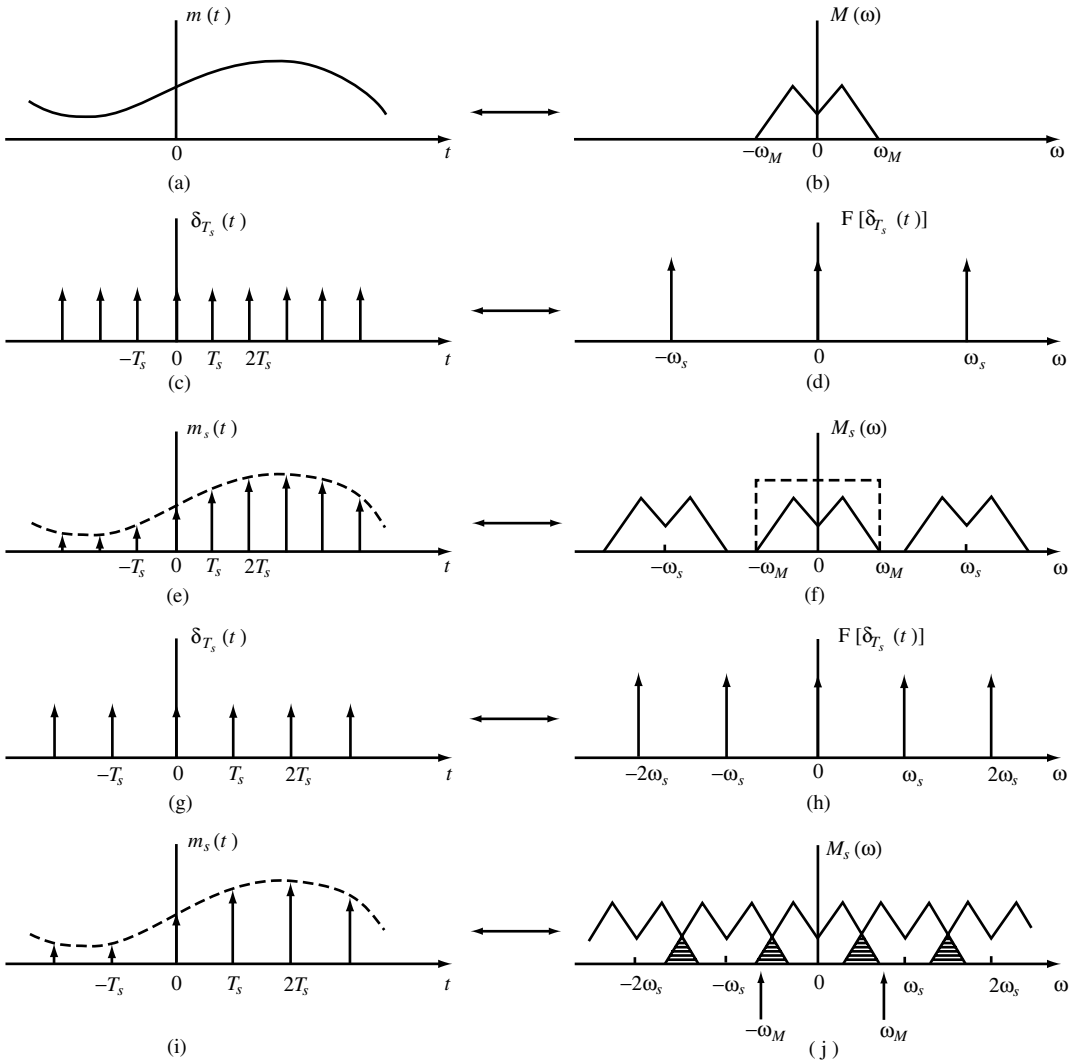


FIGURE 2.1 Illustration of instantaneous sampling and sampling theorem.

where we used the property of the δ function, $m(t)\delta(t-t_0) = m(t_0)\delta(t-t_0)$. The signal $m_s(t)$ [Fig. 2.1(e)] is referred to as the **ideal sampled signal**.

Band-Limited Signals

A real-valued signal $m(t)$ is called a **band-limited signal** if its Fourier transform $M(\omega)$ satisfies the condition

$$M(\omega) = 0 \quad \text{for } |\omega| > \omega_M \tag{2.2}$$

where $\omega_M = 2\pi f_M$ [Fig. 2.1(b)]. A band-limited signal specified by Eq. (2.2) is often referred to as a *low-pass signal*.

2.3 Sampling Theorem

The sampling theorem states that a band-limited signal $m(t)$ specified by Eq. (2.2) can be uniquely determined from its values $m(nT_s)$ sampled at uniform interval T_s if $T_s \le \pi/\omega_M = 1/(2f_M)$. In fact, when $T_s = \pi/\omega_M$, $m(t)$ is given by

$$m(t) = \sum_{n=-\infty}^{\infty} m(nT_s)\frac{\sin \omega_M(t - nT_s)}{\omega_M(t - nT_s)} \tag{2.3}$$

which is known as the **Nyquist–Shannon interpolation formula** and is also sometimes called the *cardinal series*. The sampling interval $T_s = 1/(2f_M)$ is called the *Nyquist interval* and the minimum rate $f_s = 1/T_s = 2f_M$ is known as the **Nyquist rate**.

Illustration of the instantaneous sampling process and the sampling theorem is shown in Fig. 2.1. The Fourier transform of the unit impulse train is given by [Fig. 2.1(d)]

$$\mathcal{F}\{\delta_{T_s}(t)\} = \omega_s \sum_{n=-\infty}^{\infty} \delta(\omega - n\omega_s) \qquad \omega_s = 2\pi/T_s \tag{2.4}$$

Then, by the convolution property of the Fourier transform, the Fourier transform $M_s(\omega)$ of the ideal sampled signal $m_s(t)$ is given by

$$\begin{aligned} M_s(\omega) &= \frac{1}{2\pi}\left[M(\omega) * \omega_s \sum_{n=-\infty}^{\infty} \delta(\omega - n\omega_s)\right] \\ &= \frac{1}{T_s}\sum_{n=-\infty}^{\infty} M(\omega - n\omega_s) \end{aligned} \tag{2.5}$$

where $*$ denotes convolution and we used the convolution property of the δ-function $M(\omega) * \delta(\omega - \omega_0) = M(\omega - \omega_0)$. Thus, the sampling has produced images of $M(\omega)$ along the frequency axis. Note that $M_s(\omega)$ will repeat periodically without overlap as long as $\omega_s \ge 2\omega_M$ or $f_s \ge 2f_M$ [Fig. 2.1(f)]. It is clear from Fig. 2.1(f) that we can recover $M(\omega)$ and, hence, $m(t)$ by passing the sampled signal $m_s(t)$ through an ideal low-pass filter having frequency response

$$H(\omega) = \begin{cases} T_s, & |\omega| \le \omega_M \\ 0, & \text{otherwise} \end{cases} \tag{2.6}$$

where $\omega_M = \pi/T_s$. Then

$$M(\omega) = M_s(\omega)H(\omega) \tag{2.7}$$

Taking the inverse Fourier transform of Eq. (2.6), we obtain the impulse response $h(t)$ of the ideal low-pass filter as

$$h(t) = \frac{\sin \omega_M t}{\omega_M t} \tag{2.8}$$

Taking the inverse Fourier transform of Eq. (2.7), we obtain

$$\begin{aligned} m(t) &= m_s(t) * h(t) \\ &= \sum_{n=-\infty}^{\infty} m(nT_s)\delta(t - nT_s) * \frac{\sin \omega_M t}{\omega_M t} \\ &= \sum_{n=-\infty}^{\infty} m(nT_s)\frac{\sin \omega_M(t - nT_s)}{\omega_M(t - nT_s)} \end{aligned} \tag{2.9}$$

which is Eq. (2.3).

The situation shown in Fig. 2.1(j) corresponds to the case where $f_s < 2f_M$. In this case, there is an overlap between $M(\omega)$ and $M(\omega - \omega_M)$. This overlap of the spectra is known as *aliasing* or *foldover*. When this aliasing occurs, the signal is distorted and it is impossible to recover the original signal $m(t)$ from the sampled signal. To avoid aliasing, in practice, the signal is sampled at a rate slightly higher than the Nyquist rate. If $f_s > 2f_M$, then, as shown in Fig. 2.1(f), there is a gap between the upper limit ω_M of $M(\omega)$ and the lower limit $\omega_s - \omega_M$ of $M(\omega - \omega_s)$. This range from ω_M to $\omega_s - \omega_M$ is called a *guard band*. As an example, speech transmitted via telephone is generally limited to $f_M = 3.3$ kHz (by passing the sampled signal through a low-pass filter). The Nyquist rate is, thus, 6.6 kHz. For digital transmission, the speech is normally sampled at the rate $f_s = 8$ kHz. The guard band is then $f_s - 2f_M = 1.4$ kHz. The use of a sampling rate higher than the Nyquist rate also has the desirable effect of making it somewhat easier to design the low-pass reconstruction filter so as to recover the original signal from the sampled signal.

2.4 Sampling of Sinusoidal Signals

A special case is the sampling of a sinusoidal signal having the frequency f_M. In this case, we require that $f_s > 2f_M$ rather than $f_s \geq 2f_M$. To see that this condition is necessary, let $f_s = 2f_M$. Now, if an initial sample is taken at the instant the sinusoidal signal is zero, then all successive samples will also be zero. This situation is avoided by requiring $f_s > 2f_M$.

2.5 Sampling of Bandpass Signals

A real-valued signal $m(t)$ is called a **bandpass signal** if its Fourier transform $M(\omega)$ satisfies the condition

$$M(\omega) = 0 \quad \text{except for} \quad \begin{cases} \omega_1 < \omega < \omega_2 \\ -\omega_2 < \omega < -\omega_1 \end{cases} \tag{2.10}$$

where $\omega_1 = 2\pi f_1$ and $\omega_2 = 2\pi f_2$ [Fig. 2.2(a)].

The sampling theorem for a band-limited signal has shown that a sampling rate of $2f_2$ or greater is adequate for a low-pass signal having the highest frequency f_2. Therefore, treating $m(t)$ specified by Eq. (2.10) as a special case of such a low-pass signal, we conclude that a sampling rate of $2f_2$ is adequate for the sampling of the bandpass signal $m(t)$. But it is not necessary to sample this fast. The minimum allowable sampling rate depends on f_1, f_2, and the bandwidth $f_B = f_2 - f_1$.

Let us consider the direct sampling of the bandpass signal specified by Eq. (2.10). The spectrum of the sampled signal is periodic with the period $\omega_s = 2\pi f_s$, where f_s is the sampling frequency, as in Eq. (2.4). Shown in Fig. 2.2(b) are the two right shifted spectra of the negative side spectrum $M_-(\omega)$. If the recovering of the bandpass signal is achieved by passing the sampled signal through an ideal bandpass

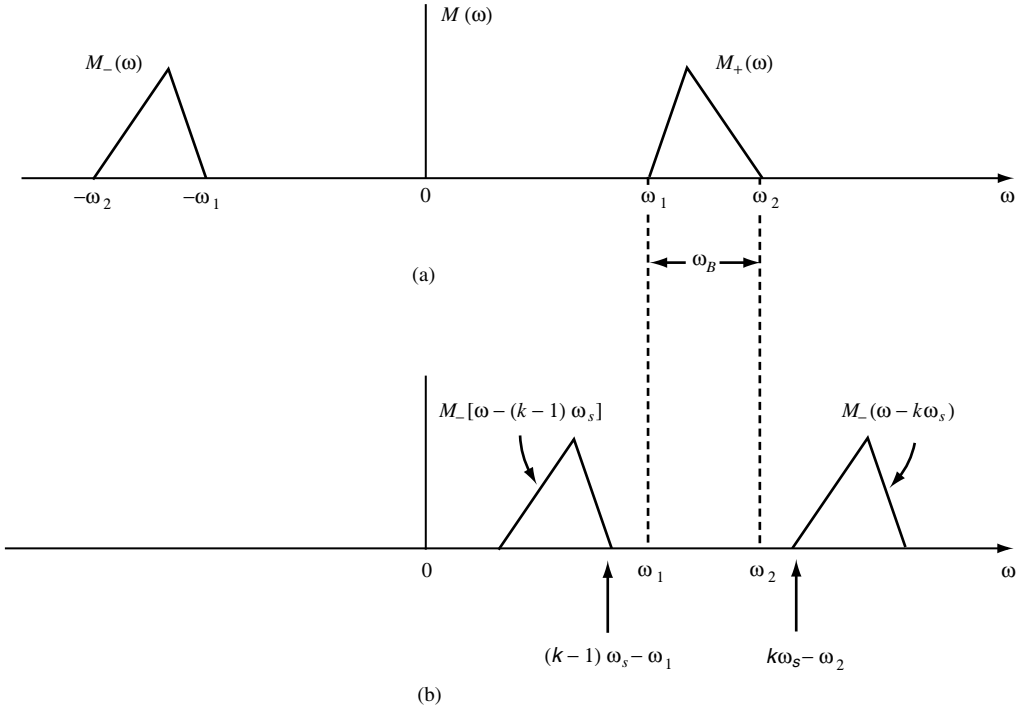


FIGURE 2.2 (a) Spectrum of a bandpass signal; (b) Shifted spectra of $M_-(\omega)$.

filter covering the frequency bands $(-\omega_2, -\omega_1)$ and (ω_1, ω_2), it is necessary that there be no aliasing problem. From Fig. 2.2(b), it is clear that to avoid overlap it is necessary that

$$\omega_s \geq 2(\omega_2 - \omega_1) \tag{2.11}$$

$$(k-1)\omega_s - \omega_1 \leq \omega_1 \tag{2.12}$$

and

$$k\omega_s - \omega_2 \geq \omega_2 \tag{2.13}$$

where $\omega_1 = 2\pi f_1$, $\omega_2 = 2\pi f_2$, and k is an integer ($k = 1, 2, \ldots$). Since $f_1 = f_2 - f_B$, these constraints can be expressed as

$$1 \leq k \leq \frac{f_2}{f_B} \leq \frac{k}{2}\frac{f_s}{f_B} \tag{2.14}$$

and

$$\frac{k-1}{2}\frac{f_s}{f_B} \leq \frac{f_2}{f_B} - 1 \tag{2.15}$$

A graphical description of Eqs. (2.14) and (2.15) is illustrated in Fig. 2.3. The unshaded regions represent where the constraints are satisfied, whereas the shaded regions represent the regions where the

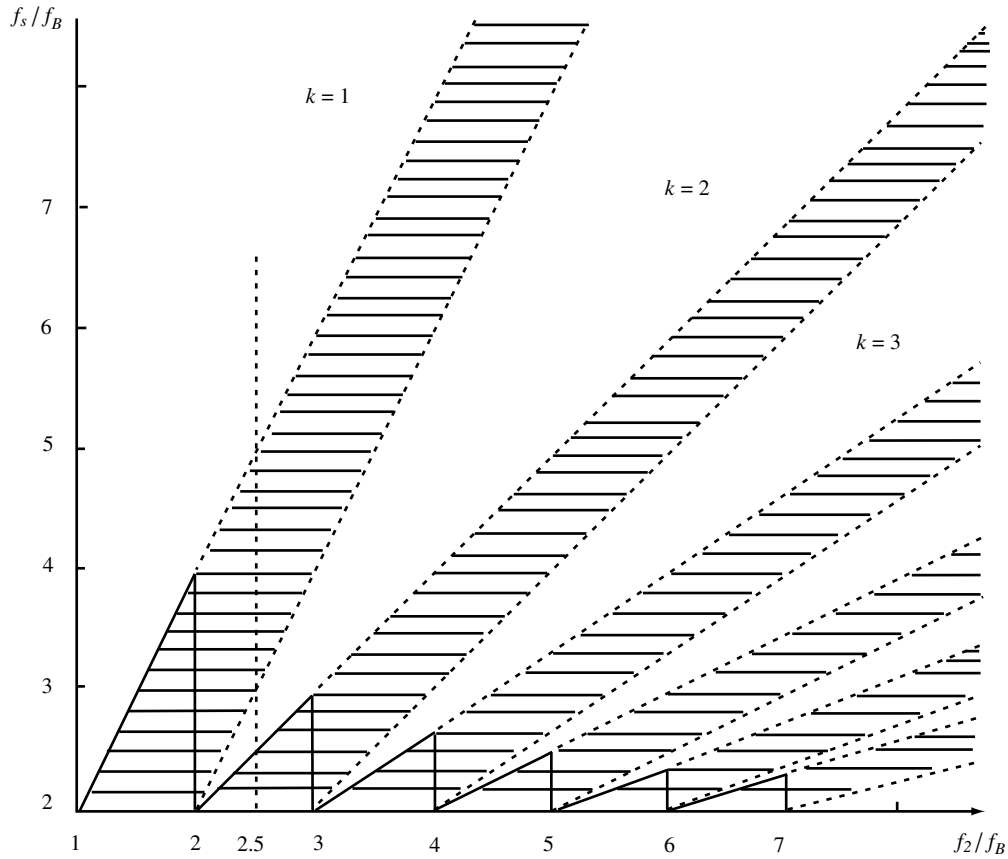


FIGURE 2.3 Minimum and permissible sampling rates for a bandpass signal.

constraints are not satisfied and overlap will occur. The solid line in Fig. 2.3 shows the locus of the minimum sampling rate. The minimum sampling rate is given by

$$\min\{f_s\} = \frac{2f_2}{m} \tag{2.16}$$

where m is the largest integer not exceeding f_2/f_B. Note that if the ratio f_2/f_B is an integer, then the minimum sampling rate is $2f_B$. As an example, consider a bandpass signal with $f_1 = 1.5$ kHz and $f_2 = 2.5$ kHz. Here, $f_B = f_2 - f_1 = 1$ kHz and $f_2/f_B = 2.5$. Then, from Eq. (2.16) and Fig. 2.3, we see that the minimum sampling rate is $2f_2/2 = f_2 = 2.5$ kHz, and allowable ranges of sampling rate are $2.5 \text{ kHz} \le f_s \le 3$ kHz and $f_s \ge 5$ kHz $(= 2f_2)$.

2.6 Practical Sampling

In practice, the sampling of an analog signal is performed by means of high-speed switching circuits, and the sampling process takes the form of *natural sampling* or **flat-top sampling**.

Natural Sampling

Natural sampling of a band-limited signal $m(t)$ is shown in Fig. 2.4. The sampled signal $m_{ns}(t)$ can be expressed as

$$m_{ns}(t) = m(t)x_p(t) \tag{2.17}$$

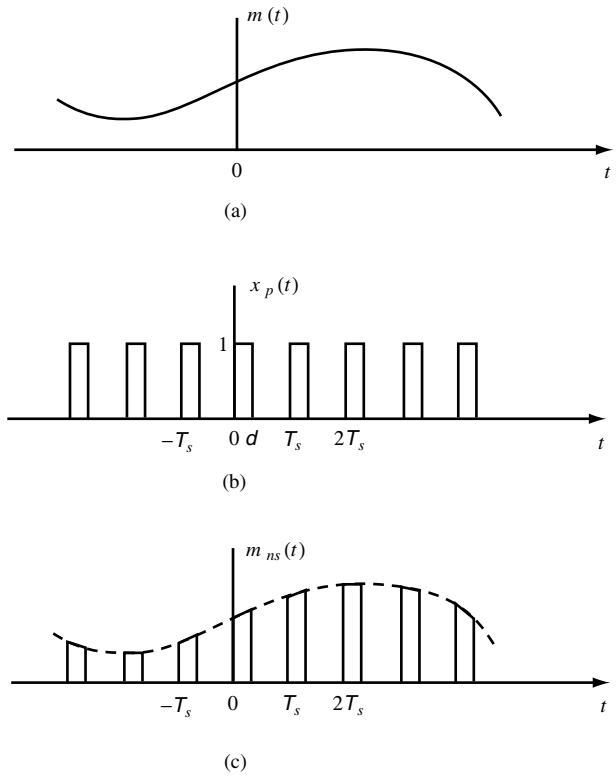


FIGURE 2.4 Natural sampling.

where $x_p(t)$ is the periodic train of rectangular pulses with fundamental period T_s, and each rectangular pulse in $x_p(t)$ has duration d and unit amplitude [Fig. 2.4(b)]. Observe that the sampled signal $m_{\text{ns}}(t)$ consists of a sequence of pulses of varying amplitude whose tops follow the waveform of the signal $m(t)$ [Fig. 2.4(c)].

The Fourier transform of $x_p(t)$ is

$$X_p(\omega) = \sum_{n=-\infty}^{\infty} c_n \delta(\omega - n\omega_s) \qquad \omega_s = 2\pi/T_s \tag{2.18}$$

where

$$c_n = \frac{d}{T_s} \frac{\sin(n\omega_s d/2)}{n\omega_s d/2} e^{-jn\omega_s d/2} \tag{2.19}$$

Then the Fourier transform of $m_{\text{ns}}(t)$ is given by

$$M_{\text{ns}}(\omega) = M(\omega) * X_p(\omega) = \sum_{n=-\infty}^{\infty} c_n M(\omega - n\omega_s) \tag{2.20}$$

from which we see that the effect of the natural sampling is to multiply the nth shifted spectrum $M(\omega - n\omega_s)$ by a constant c_n. Thus, the original signal $m(t)$ can be reconstructed from $m_{\text{ns}}(t)$ with no distortion by passing $m_{\text{ns}}(t)$ through an ideal low-pass filter if the sampling rate f_s is equal to or greater than the Nyquist rate $2f_M$.

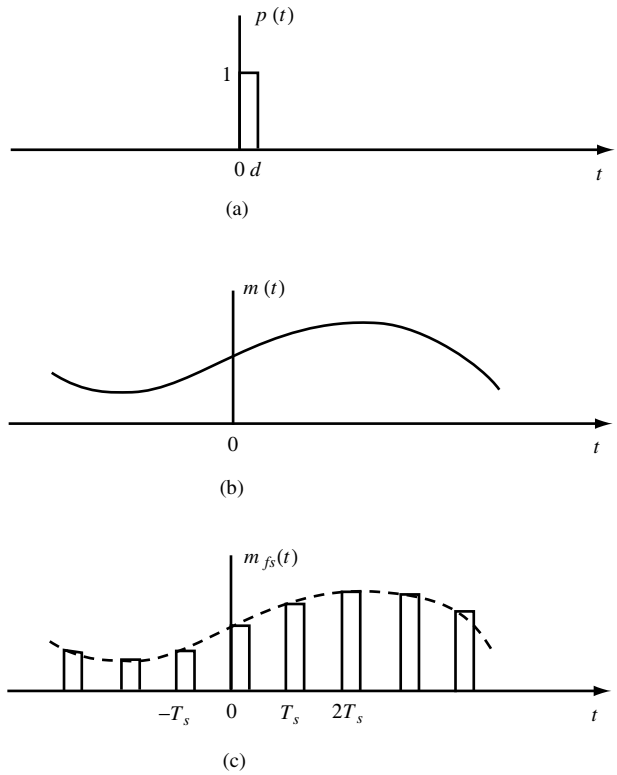


FIGURE 2.5 Flat-top sampling.

Flat-Top Sampling

The sampled waveform, produced by practical sampling devices that are the sample and hold types, has the form [Fig. 2.5(c)]

$$m_{\text{fs}}(t) = \sum_{n=-\infty}^{\infty} m(nT_s)p(t-nT_s) \tag{2.21}$$

where $p(t)$ is a rectangular pulse of duration d with unit amplitude [Fig. 2.5(a)]. This type of sampling is known as **flat-top sampling**. Using the ideal sampled signal $m_s(t)$ of Eq. (2.1), $m_{\text{fs}}(t)$ can be expressed as

$$m_{\text{fs}}(t) = p(t) * \left[\sum_{n=-\infty}^{\infty} m(nT_s)\delta(t-nT_s)\right] = p(t) * m_s(t) \tag{2.22}$$

Using the convolution property of the Fourier transform and Eq. (2.4), the Fourier transform of $m_{\text{fs}}(t)$ is given by

$$M_{\text{fs}}(\omega) = P(\omega)M_s(\omega) = \frac{1}{T_s}\sum_{n=-\infty}^{\infty} P(\omega)M(\omega-n\omega_s) \tag{2.23}$$

where

$$P(\omega) = d\frac{\sin(\omega d/2)}{\omega d/2}e^{-j\omega d/2} \tag{2.24}$$

From Eq. (2.23) we see that by using flat-top sampling we have introduced amplitude distortion and time delay, and the primary effect is an attenuation of high-frequency components. This effect is known as the *aperture effect*. The aperture effect can be compensated by an equalizing filter with a frequency response $H_{eq}(\omega) = 1/P(\omega)$. If the pulse duration d is chosen such that $d \ll T_s$, however, then $P(\omega)$ is essentially constant over the baseband and no equalization may be needed.

2.7 Sampling Theorem in the Frequency Domain

The sampling theorem expressed in Eq. (2.4) is the time-domain sampling theorem. There is a dual to this time-domain sampling theorem, i.e., the sampling theorem in the frequency domain.

Time-limited signals: A continuous-time signal $m(t)$ is called **time limited** if

$$m(t) = 0 \quad \text{for } |t| > |T_0| \tag{2.25}$$

Frequency-domain sampling theorem: The frequency-domain sampling theorem states that the Fourier transform $M(\omega)$ of a time-limited signal $m(t)$ specified by Eq. (2.25) can be uniquely determined from its values $M(n\omega_s)$ sampled at a uniform rate ω_s if $\omega_s \le \pi/T_0$. In fact, when $\omega_s = \pi/T_0$, then $M(\omega)$ is given by

$$M(\omega) = \sum_{n=-\infty}^{\infty} M(n\omega_s)\frac{\sin T_0(\omega - n\omega_s)}{T_0(\omega - n\omega_s)} \tag{2.26}$$

2.8 Summary and Discussion

The sampling theorem is the fundamental principle of digital communications. We state the sampling theorem in two parts.

Theorem 2.1. *If the signal contains no frequency higher than f_M Hz, it is completely described by specifying its samples taken at instants of time spaced $1/2f_M$ s.*

Theorem 2.2. *The signal can be completely recovered from its samples taken at the rate of $2f_M$ samples per second or higher.*

The preceding sampling theorem assumes that the signal is strictly band limited. It is known that if a signal is band limited, it cannot be time limited, and vice versa. In many practical applications, the signal to be sampled is time limited and, consequently, it cannot be strictly band limited. Nevertheless, we know that the frequency components of physically occurring signals attenuate rapidly beyond some defined bandwidth, and for practical purposes we consider these signals band limited. This approximation of real signals by band limited ones introduces no significant error in the application of the sampling theorem. When such a signal is sampled, we band limit the signal by filtering before sampling and sample at a rate slightly higher than the nominal Nyquist rate.

Defining Terms

Band-limited signal: A signal whose frequency content (Fourier transform) is equal to zero above some specified frequency.

Bandpass signal: A signal whose frequency content (Fourier transform) is nonzero only in a band of frequencies not including the origin.

Flat-top sampling: Sampling with finite width pulses that maintain a constant value for a time period less than or equal to the sampling interval. The constant value is the amplitude of the signal at the desired sampling instant.

Ideal sampled signal: A signal sampled using an ideal impulse train.

Nyquist rate: The minimum allowable sampling rate of $2f_M$ samples per second, to reconstruct a signal band limited to f_M hertz.

Nyquist-Shannon interpolation formula: The infinite series representing a time domain waveform in terms of its ideal samples taken at uniform intervals.

Sampling interval: The time between samples in uniform sampling.

Sampling rate: The number of samples taken per second (expressed in Hertz and equal to the reciprocal of the sampling interval).

Time-limited: A signal that is zero outside of some specified time interval.

References

1. Brown, J.L. Jr., First order sampling of bandpass signals—A new approach. *IEEE Trans. Information Theory,* IT-26(5), 613–615, 1980.
2. Byrne, C.L. and Fitzgerald, R.M., Time-limited sampling theorem for band-limited signals, *IEEE Trans. Information Theory,* IT-28(5), 807–809, 1982.
3. Hsu, H.P., *Applied Fourier Analysis,* Harcourt Brace Jovanovich, San Diego, CA, 1984.
4. Hsu, H.P., *Analog and Digital Communications,* McGraw-Hill, New York, 1993.
5. Hulthén, R., Restoring causal signals by analytical continuation: A generalized sampling theorem for causal signals. *IEEE Trans. Acoustics, Speech, and Signal Processing,* ASSP-31(5), 1294–1298, 1983.
6. Jerri, A.J., The Shannon sampling theorem—Its various extensions and applications: A tutorial review, *Proc. IEEE.* 65(11), 1565–1596, 1977.

Further Information

For a tutorial review of the sampling theorem, historical notes, and earlier references, see Jerri [6].

3

Pulse Code Modulation*

Leon W. Couch, II
University of Florida

3.1 Introduction

Pulse code modulation (PCM) is analog-to-digital conversion of a special type where the information contained in the instantaneous samples of an analog signal is represented by digital words in a serial bit stream.

If we assume that each of the digital words has n binary digits, there are $M = 2^n$ unique code words that are possible, each code word corresponding to a certain amplitude level. Each sample value from the analog signal, however, can be any one of an infinite number of levels, so that the digital word that represents the amplitude closest to the actual sampled value is used. This is called **quantizing**. That is, instead of using the exact sample value of the analog waveform, the sample is replaced by the closest allowed value, where there are M allowed values, and each allowed value corresponds to one of the code words.

PCM is very popular because of the many advantages it offers. Some of these advantages are as follows.

- Relatively inexpensive digital circuitry may be used extensively in the system.
- PCM signals derived from all types of analog sources (audio, video, etc.) may be time-division multiplexed with data signals (e.g., from digital computers) and transmitted over a common high-speed digital communication system.
- In long-distance digital telephone systems requiring repeaters, a *clean* PCM waveform can be regenerated at the output of each repeater, where the input consists of a noisy PCM waveform. The noise at the input, however, may cause bit errors in the regenerated PCM output signal.
- The noise performance of a digital system can be superior to that of an analog system. In addition, the probability of error for the system output can be reduced even further by the use of appropriate coding techniques.

These advantages usually outweigh the main disadvantage of PCM: a much wider bandwidth than that of the corresponding analog signal.

*Source: Couch, L.W., II. 2001. *Digital and Analog Communication Systems,* 6th ed., Prentice Hall, Upper Saddle River, NJ. With permission.

3.2 Generation of PCM

The PCM signal is generated by carrying out three basic operations: sampling, quantizing, and encoding (see Fig. 3.1). The sampling operation generates an instantaneously-sampled flat-top **pulse-amplitude modulated** (PAM) signal.

The quantizing operation is illustrated in Fig. 3.2 for the $M = 8$ level case. This quantizer is said to be *uniform* since all of the steps are of equal size. Since we are approximating the analog sample values by using a finite number of levels ($M = 8$ in this illustration), *error* is introduced into the recovered output analog signal because of the quantizing effect. The error waveform is illustrated in Fig. 3.2c. The quantizing error consists of the difference between the analog signal at the sampler input and the output of the quantizer. Note that the peak value of the error (±1) is one-half of the quantizer step size [2]. If we sample at the Nyquist rate ($2B$, where B is the absolute bandwidth, in hertz, of the input analog signal) or faster and there is negligible channel noise, there will still be noise, called *quantizing noise,* on the recovered analog waveform due to this error. The quantizing noise can also be thought of as a round-off error. The quantizer output is a *quantized* (i.e., only M possible amplitude values) PAM signal.

The PCM signal is obtained from the quantized PAM signal by encoding each quantized sample value into a digital word. It is up to the system designer to specify the exact code word that will represent a particular quantized level. If a Gray code of Table 3.1 is used, the resulting PCM signal is shown in Fig. 3.2d where the PCM word for each quantized sample is strobed out of the encoder by the next clock pulse. The Gray code was chosen because it has only 1-b change for each step change in the quantized level. Consequently, single errors in the received PCM code word will cause minimum errors in the recovered analog level, provided that the sign bit is not in error.

TABLE 3.1 Three-Bit Gray Code for $M = 8$ Levels

Quantized Sample Voltage	Gray Code Word (PCM Output)	
+7	110	
+5	111	
+3	101	
+1	100	
		Mirror image except for sign bit
−1	000	
−3	001	
−5	011	
−7	010	

Source: Couch, L.W., II. 2001. *Digital and Analog Communication Systems,* 6th ed., Prentice Hall, Upper Saddle River, NJ, p. 141. With permission.

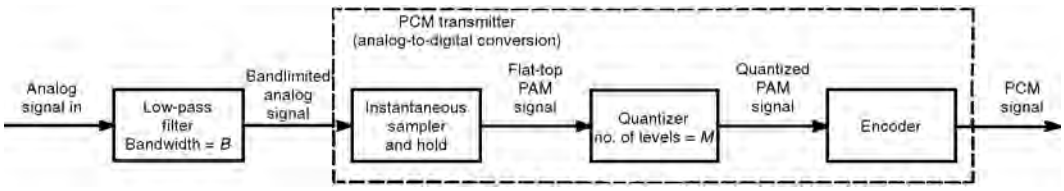


FIGURE 3.1 A PCM transmitter. *Source:* Couch, L.W., II. 2001. *Digital and Analog Communication Systems,* 6th ed., Prentice Hall, Upper Saddle River, NJ, p. 139. With permission.

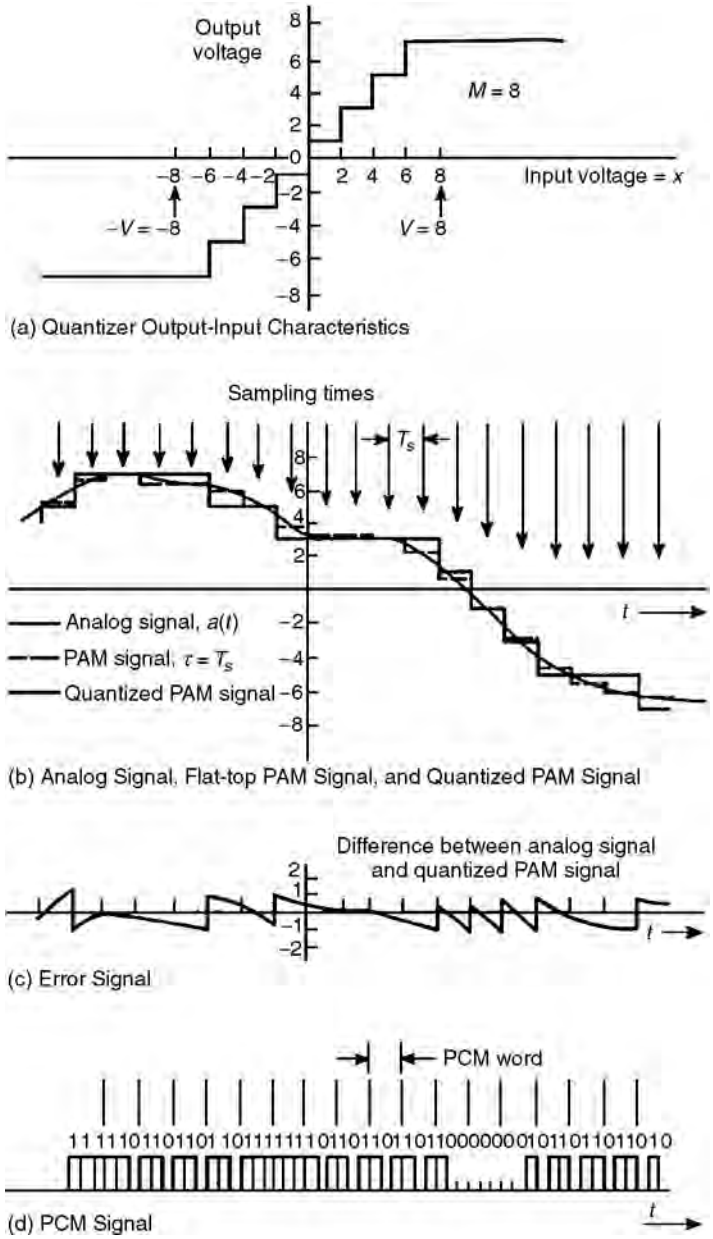


FIGURE 3.2 Illustration of waveforms in a PCM system. *Source:* Couch, L.W., II. 2001. *Digital and Analog Communication Systems,* 6th ed., Prentice Hall, Upper Saddle River, NJ, p. 140. With permission.

Here we have described PCM systems that represent the quantized analog sample values by *binary* code words. Of course, it is possible to represent the quantized analog samples by digital words using other than base 2. That is, for base q, the number of quantized levels allowed is $M = q^n$, where n is the number of q base digits in the code word. We will not pursue this topic since binary ($q = 2$) digital circuits are most commonly used.

3.3 Percent Quantizing Noise

The quantizer at the PCM encoder produces an error signal at the PCM decoder output, as illustrated in Fig. 3.2c. The peak value of this error signal may be expressed as a percentage of the maximum possible analog signal amplitude. Referring to Fig. 3.2c, a peak error of 1 V occurs for a maximum analog signal amplitude of $M = 8$ V, as shown in Fig. 3.1c. Thus, in general,

$$\frac{2P}{100} = \frac{1}{M} = \frac{1}{2^n}$$

or

$$2^n = \frac{50}{P} \tag{3.1}$$

where P is the peak percentage error for a PCM system that uses n bit code words. The design value of n needed in order to have less than P percent error is obtained by taking the base 2 logarithm of both sides of Eq. (3.1), where it is realized that $\log_2(x) = [\log_{10}(x)]/\log_{10}(2) = 3.32 \log_{10}(x)$. That is,

$$n \geq 3.32 \log_{10}\left(\frac{50}{P}\right) \tag{3.2}$$

where n is the number of bits needed in the PCM word in order to obtain less than P percent error in the recovered analog signal (i.e., decoded PCM signal).

3.4 Practical PCM Circuits

Three techniques are used to implement the analog-to-digital converter (ADC) encoding operation. These are the *counting* or *ramp, serial* or *successive approximation*, and *parallel* or *flash* encoders.

In the counting encoder, at the same time that the sample is taken, a ramp generator is energized and a binary counter is started. The output of the ramp generator is continuously compared to the sample value; when the value of the ramp becomes equal to the sample value, the binary value of the counter is read. This count is taken to be the PCM word. The binary counter and the ramp generator are then reset to zero and are ready to be reenergized at the next sampling time. This technique requires only a few components, but the speed of this type of ADC is usually limited by the speed of the counter. The Maxim ICL7126 CMOS ADC integrated circuit uses this technique.

The serial encoder compares the value of the sample with trial quantized values. Successive trials depend on whether the past comparator outputs are positive or negative. The trial values are chosen first in large steps and then in small steps so that the process will converge rapidly. The trial voltages are generated by a series of voltage dividers that are configured by (on-off) switches. These switches are controlled by digital logic. After the process converges, the value of the switch settings is read out as the PCM word. This technique requires more precision components (for the voltage dividers) than the ramp technique. The speed of the feedback ADC technique is determined by the speed of the switches. The National Semiconductor ADC0804 eight-bit ADC uses this technique.

The parallel encoder uses a set of parallel comparators with reference levels that are the permitted quantized values. The sample value is fed into all of the parallel comparators simultaneously. The high or low level of the comparator outputs determines the binary PCM word with the aid of some digital logic. This is a fast ADC technique but requires more hardware than the other two methods. The Harris CA3318 eight-bit ADC integrated circuit is an example of the technique.

All of the integrated circuits listed as examples have parallel digital outputs that correspond to the digital word that represents the analog sample value. For generation of PCM, the parallel output (digital word) needs to be converted to serial form for transmission over a two-wire channel. This is accomplished by using a parallel-to-serial converter integrated circuit, which is also known as a **serial-input-output** (SIO) chip. The SIO chip includes a shift register that is set to contain the parallel data (usually, from 8 or 16 input lines). Then the data are shifted out of the last stage of the shift register bit by bit onto a single output line to produce the serial format. Furthermore, the SIO chips are usually full duplex; that is, they have two sets of shift registers, one that functions for data flowing in each direction. One shift register converts parallel input data to serial output data for transmission over the channel, and, simultaneously, the other shift register converts received serial data from another input to parallel data that are available at another output. Three types of SIO chips are available: the *universal asynchronous receiver/transmitter* (UART), the *universal synchronous receiver/transmitter* (USRT), and the *universal synchronous/asynchronous receiver transmitter* (USART). The UART transmits and receives asynchronous serial data, the USRT transmits and receives synchronous serial data, and the USART combines both a UART and a USRT on one chip.

At the receiving end, the PCM signal is decoded back into an analog signal by using a digital-to-analog converter (DAC) chip. If the DAC chip has a parallel data input, the received serial PCM data are first converted to a parallel form using a SIO chip as described in the preceding paragraph. The parallel data are then converted to an approximation of the analog sample value by the DAC chip. This conversion is usually accomplished by using the parallel digital word to set the configuration of electronic switches on a resistive current (or voltage) divider network so that the analog output is produced. This is called a *multiplying* DAC since the analog output voltage is directly proportional to the divider reference voltage multiplied by the value of the digital word. The Motorola MC1408 and the National Semiconductor DAC0808 eight-bit DAC chips are examples of this technique. The DAC chip outputs samples of the quantized analog signal that approximate the analog sample values. This may be smoothed by a low-pass reconstruction filter to produce the analog output.

The Communications Handbook [6, *pp* 107–117] *and The Electrical Engineering Handbook* [5, *pp* 771–782] give more details on ADC, DAC, and PCM circuits.

3.5 Bandwidth of PCM

A good question to ask is: what is the spectrum of a PCM signal? For the case of PAM signalling, the spectrum of the PAM signal could be obtained as a function of the spectrum of the input analog signal because the PAM signal is a linear function of the analog signal. This is not the case for PCM. As shown in Figs. 3.1 and 3.2, the PCM signal is a nonlinear function of the input signal. Consequently, the spectrum of the PCM signal is not directly related to the spectrum of the input analog signal. It can be shown that the spectrum of the (serial) binary PCM signal depends on the bit rate, the correlation of the PCM data, and on the PCM waveform pulse shape (usually rectangular) used to describe the bits [2, 3]. From Fig. 3.2, the bit rate is

$$R = nf_s \tag{3.3}$$

where n is the number of bits in the PCM word ($M = 2^n$) and f_s is the sampling rate. For no aliasing, we require $f_s \geq 2B$ where B is the bandwidth of the analog signal (that is to be converted to the PCM signal). The dimensionality theorem [2, 3] shows that the bandwidth of the PCM waveform is bounded by

$$B_{\text{PCM}} \geq \frac{1}{2}R = \frac{1}{2}nf_s \tag{3.4}$$

where equality is obtained if a $(\sin x)/x$ type of pulse shape is used to generate the PCM waveform. The exact spectrum for the PCM waveform will depend on the pulse shape that is used as well as on the type

TABLE 3.2 Performance of a PCM System with Uniform Quantizing and No Channel Noise

Number of Quantizer Levels Used, M	Length of the PCM Word, n (bits)	Bandwidth of PCM Signal (First Null Bandwidth)[a]	Recovered Analog Signal Power-to-Quantizing Noise Power Ratios (dB) $(S/N)_{out}$
2	1	2B	6.0
4	2	4B	12.0
8	3	6B	18.1
16	4	8B	24.1
32	5	10B	30.1
64	6	12B	36.1
128	7	14B	42.1
256	8	16B	48.2
512	9	18B	54.2
1,024	10	20B	60.2
2,048	11	22B	66.2
4,096	12	24B	72.2
8,192	13	26B	78.3
16,384	14	28B	84.3
32,768	15	30B	90.3
65,536	16	32B	96.3

[a] B is the absolute bandwidth of the input analog signal.

Source: Couch, L.W., II. 2001. *Digital and Analog Communication Systems,* 6th ed., Prentice Hall, Upper Saddle River, NJ, p. 144. With permission.

of line encoding. For example, if one uses a rectangular pulse shape with polar nonreturn to zero (NRZ) line coding, the first null bandwidth is simply

$$B_{\text{PCM}} = R = nf_s \text{ Hz} \tag{3.5}$$

Table 3.2 presents a tabulation of this result for the case of the minimum sampling rate, $f_s = 2B$. Note that Eq. (3.4) demonstrates that the bandwidth of the PCM signal has a lower bound given by

$$B_{\text{PCM}} \geq nB \tag{3.6}$$

where $f_s > 2B$ and B is the bandwidth of the corresponding analog signal. Thus, for reasonable values of n, the bandwidth of the PCM signal will be significantly larger than the bandwidth of the corresponding analog signal that it represents. For the example shown in Fig. 3.2 where $n = 3$, the PCM signal bandwidth will be at least three times wider than that of the corresponding analog signal. Furthermore, if the bandwidth of the PCM signal is reduced by improper filtering or by passing the PCM signal through a system that has a poor frequency response, the filtered pulses will be elongated (stretched in width) so that pulses corresponding to any one bit will smear into adjacent bit slots. If this condition becomes too serious, it will cause errors in the detected bits. This pulse smearing effect is called **intersymbol interference (ISI)**.

3.6 Effects of Noise

The analog signal that is recovered at the PCM system output is corrupted by noise. Two main effects produce this noise or distortion: (1) quantizing noise that is caused by the M-step quantizer at the PCM transmitter and (2) bit errors in the recovered PCM signal. The bit errors are caused by *channel noise* as well as improper channel filtering, which causes ISI. In addition, if the input analog signal is not strictly

band limited, there will be some aliasing noise on the recovered analog signal [12]. Under certain assumptions, it can be shown that the recovered analog *average* signal power to the average noise power [2] is

$$\left(\frac{S}{N}\right)_{\text{out}} = \frac{M^2}{1 + 4(M^2 - 1)P_e} \tag{3.7}$$

where M is the number of uniformly spaced quantizer levels used in the PCM transmitter and P_e is the probability of bit error in the recovered binary PCM signal at the receiver DAC before it is converted back into an analog signal. Most practical systems are designed so that P_e is negligible. Consequently, if we assume that there are no bit errors due to channel noise (i.e., $P_e = 0$), the *S/N* due only to quantizing errors is

$$\left(\frac{S}{N}\right)_{\text{out}} = M^2 \tag{3.8}$$

Numerical values for these *S/N* ratios are given in Table 3.2.

To realize these *S/N* ratios, one critical assumption is that the peak-to-peak level of the analog waveform at the input to the PCM encoder is set to the design level of the quantizer. For example, referring to Fig. 3.2, this corresponds to the input traversing the range $-V$ to $+V$ volts where $V = 8$ V is the design level of the quantizer. Equation (3.7) was derived for waveforms with equally likely values, such as a triangle waveshape, that have a peak-to-peak value of $2V$ and an rms value of $V/\sqrt{3}$, where V is the design peak level of the quantizer.

From a practical viewpoint, the quantizing noise at the output of the PCM decoder can be categorized into four types depending on the operating conditions. The four types are overload noise, random noise, granular noise, and hunting noise. As discussed earlier, the level of the analog waveform at the input of the PCM encoder needs to be set so that its peak level does not exceed the design peak of V volts. If the peak input does exceed V, the recovered analog waveform at the output of the PCM system will have flat tops near the peak values. This produces *overload noise*. The flat tops are easily seen on an oscilloscope, and the recovered analog waveform sounds are distorted since the flat topping produces unwanted harmonic components.

The second type of noise, *random noise,* is produced by the random quantization errors in the PCM system under normal operating conditions when the input level is properly set. This type of condition is assumed in Eq. (3.8). Random noise has a white hissing sound. If the input level is not sufficiently large, the *S/N* will deteriorate from that given by Eq. (3.8); the quantizing noise will still remain more or less random.

If the input level is reduced further to a relatively small value with respect to the design level, the error values are not equally likely from sample to sample, and the noise has a harsh sound resembling gravel being poured into a barrel. This is called *granular noise*. This type of noise can be randomized (noise power decreased) by increasing the number of quantization levels and, consequently, increasing the PCM bit rate. Alternatively, granular noise can be reduced by using a nonuniform quantizer, such as the μ-law or A-law quantizers that are described in Section 3.7.

The fourth type of quantizing noise that may occur at the output of a PCM system is *hunting noise*. It can occur when the input analog waveform is nearly constant, including when there is no signal (i.e., zero level). For these conditions, the sample values at the quantizer output (see Fig. 3.2) can oscillate between two adjacent quantization levels, causing an undesired sinusoidal type tone of frequency $1/2f_s$ at the output of the PCM system. Hunting noise can be reduced by filtering out the tone or by designing the quantizer so that there is no vertical step at the constant value of the inputs, such as at 0-V input for the no signal case. For the no signal case, the hunting noise is also called *idle channel noise*. Idle channel noise can be reduced by using a horizontal step at the origin of the quantizer output–input characteristic instead of a vertical step as shown in Fig. 3.2.

Recalling that $M = 2^n$, we may express Eq. (3.8) in decibels by taking 10 $\log_{10}(\cdot)$ of both sides of the equation,

$$\left(\frac{S}{N}\right)_{\text{dB}} = 6.02n + \alpha \tag{3.9}$$

where n is the number of bits in the PCM word and $\alpha = 0$. This equation—called the 6-dB rule—points out the significant performance characteristic for PCM: an additional 6-dB improvement in S/N is obtained for each bit added to the PCM word. This is illustrated in Table 3.2. Equation (3.9) is valid for a wide variety of assumptions (such as various types of input waveshapes and quantification characteristics), although the value of α will depend on these assumptions [7]. Of course, it is assumed that there are no bit errors and that the input signal level is large enough to range over a significant number of quantizing levels.

One may use Table 3.2 to examine the design requirements in a proposed PCM system. For example, high fidelity enthusiasts are turning to digital audio recording techniques. Here, PCM signals are recorded instead of the analog audio signal to produce superb sound reproduction. For a dynamic range of 90 dB, it is seen that at least 15-b PCM words would be required. Furthermore, if the analog signal had a bandwidth of 20 kHz, the first null bandwidth for rectangular bit-shape PCM would be 2×20 kHz $\times$ 15 = 600 kHz. Consequently, video-type tape recorders are needed to record and reproduce high-quality digital audio signals. Although this type of recording technique might seem ridiculous at first, it is realized that expensive high-quality analog recording devices are hard pressed to reproduce a dynamic range of 70 dB. Thus, digital audio is one way to achieve improved performance. This is being proven in the marketplace with the popularity of the digital compact disk (CD). The CD uses a 16-b PCM word and a sampling rate of 44.1 kHz on each stereo channel [9, 10]. Reed–Solomon coding with interleaving is used to correct burst errors that occur as a result of scratches and fingerprints on the compact disk.

3.7 Nonuniform Quantizing: μ-Law and A-Law Companding

Voice analog signals are more likely to have amplitude values near zero than at the extreme peak values allowed. For example, when digitizing voice signals, if the peak value allowed is 1 V, weak passages may have voltage levels on the order of 0.1 V (20 dB down). For signals such as these with nonuniform amplitude distribution, the granular quantizing noise will be a serious problem if the step size is not reduced for amplitude values near zero and increased for extremely large values. This is called nonuniform quantizing since a variable step size is used. An example of a nonuniform quantizing characteristic is shown in Fig. 3.3.

The effect of nonuniform quantizing can be obtained by first passing the analog signal through a compression (nonlinear) amplifier and then into the PCM circuit that uses a uniform quantizer. In the U.S., a μ-law type of compression characteristic is used. It is defined [11] by

$$|w_2(t)| = \frac{\ln\left(1 + \mu|w_1(t)|\right)}{\ln(1+\mu)} \tag{3.10}$$

where the allowed peak values of $w_1(t)$ are ± 1 (i.e., $|w_1(t)| \le 1$) and μ is a positive constant that is a parameter. This compression characteristic is shown in Fig. 3.3(b) for several values of μ, and it is noted that $\mu \to 0$ corresponds to linear amplification (uniform quantization overall). In the United States, Canada, and Japan, the telephone companies use a $\mu = 255$ compression characteristic in their PCM systems [4].

In practice, the smooth nonlinear characteristics of Fig. 3.3b are approximated by piecewise linear chords as shown in Fig. 3.3d for the $\mu = 255$ characteristic [4]. Each chord is approximated by a uniform quantizer with 16 steps and an input step size that is set by the particular segment number. That is, 16

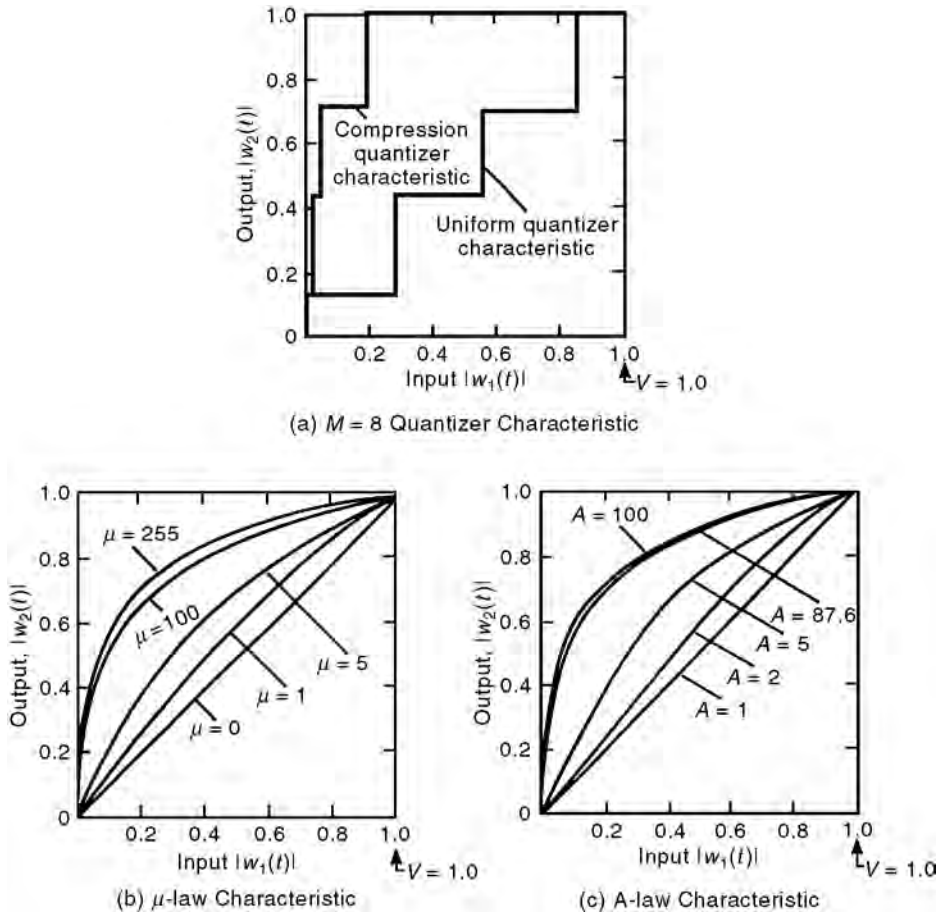


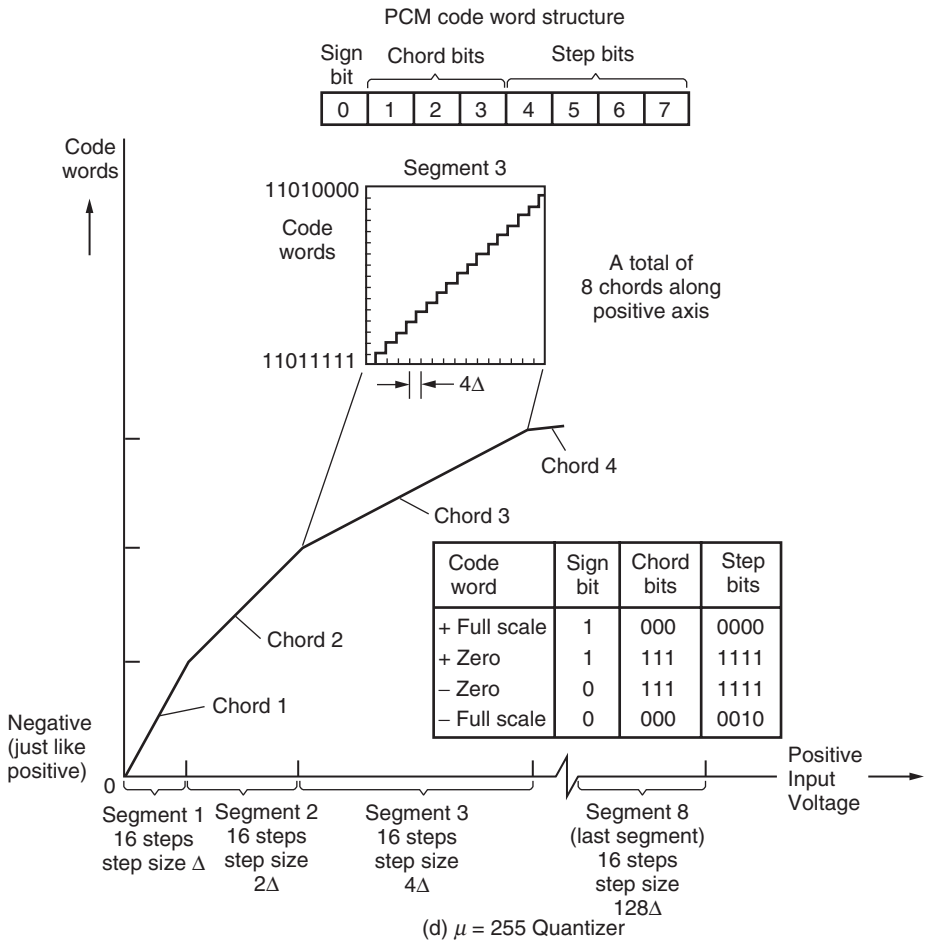
FIGURE 3.3 Compression characteristics (first quadrant shown). *Source:* Couch, L.W., II. 2001. *Digital and Analog Communication Systems,* 6th ed., Prentice Hall, Upper Saddle River, NJ, pp. 148–149. With permission.

steps (including a half-width step on each side of zero) of width Δ are used for Segment 1, 16 steps of width 2Δ are used for Segment 2, 16 steps of width 4Δ for Segment 3, etc. The value of Δ is chosen so that the full-scale value (last step of Segment 8) matches the peak value of the input analog signal. This segmenting technique is now accepted worldwide for the $\mu = 255$ characteristic. As shown in Fig. 3.3d, the eight-bit PCM code word consists of a sign bit that denotes a positive or negative input voltage, three chord bits that denote the segment number, and four step bits that denote the particular step within the segment. For more details, see the data sheet for the Motorola MC145500 series of PCM codecs, available on the World Wide Web from Motorola Semiconductor Products at the company's Web site, http://www.motorola.com.

Another compression law, used mainly in Europe, is the A-law characteristic. It is defined [1] by

$$|w_2(t)| = \begin{cases} \dfrac{A|w_1(t)|}{1+\ln A}, & 0 \le |w_1(t)| \le \dfrac{1}{A} \\ \dfrac{1+\ln(A|w_1(t)|)}{1+\ln A}, & \dfrac{1}{A} \le |w_1(t)| \le 1 \end{cases} \tag{3.11}$$

where $|w_1(t)| < 1$ and A is a positive constant. The A-law compression characteristic is shown in Fig. 3.3(c). The typical value for A is 87.6.



Code word	Sign bit	Chord bits	Step bits
+ Full scale	1	000	0000
+ Zero	1	111	1111
– Zero	0	111	1111
– Full scale	0	000	0010

FIGURE 3.3 (Continued)

When compression is used at the transmitter, *expansion* (i.e., decompression) must be used at the receiver output to restore signal levels to their correct relative values. The *expandor* characteristic is the inverse of the compression characteristic, and the combination of a compressor and an expandor is called a *compandor*.

Once again, it can be shown that the output *S/N* follows the 6-dB law [2]

$$\left(\frac{S}{N}\right)_{\text{dB}} = 6.02 + \alpha \tag{3.12}$$

where, for uniform quantizing,

$$\alpha = 4.77 - 20\log(V/x_{\text{rms}}) \tag{3.13}$$

and for sufficiently large input levels* for μ-law companding

$$\alpha \approx 4.77 - 20\log[\ln(1+\mu)] \tag{3.14}$$

*See Lathi, 1998 for a more complicated expression that is valid for any input level.

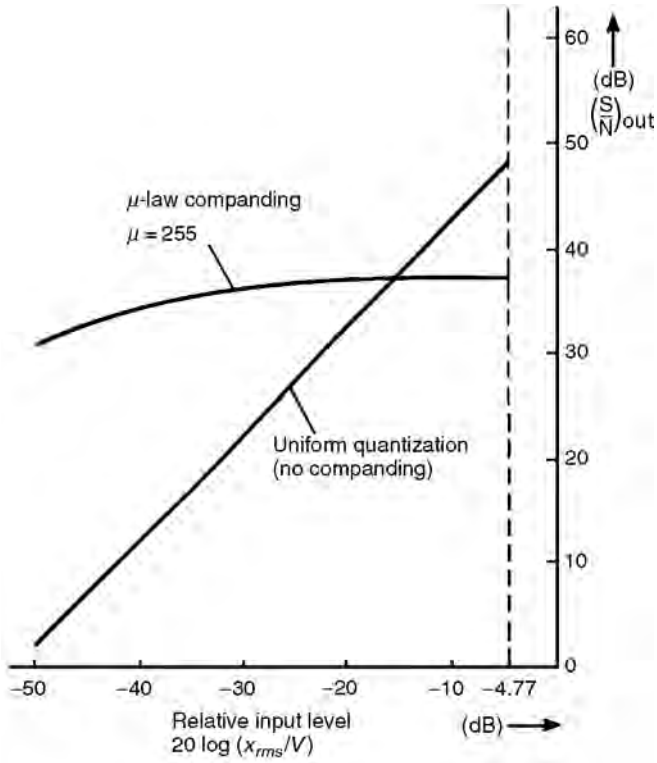


FIGURE 3.4 Output *S/N* of 8-b PCM systems with and without companding. *Source:* Couch, L.W., II. 2001. *Digital and Analog Communication Systems,* 6th ed., Prentice Hall, Upper Saddle River, NJ, p. 151. With permission.

and for *A*-law companding [7]

$$\alpha \approx 4.77 - 20\log[1 + \ln A] \tag{3.15}$$

n is the number of bits used in the PCM word, V is the peak design level of the quantizer, and x_{rms} is the rms value of the input analog signal. Notice that the output *S/N* is a function of the input level for the uniform quantizing (no companding) case but is relatively insensitive to input level for μ-law and *A*-law companding, as shown in Fig. 3.4. The ratio V/x_{rms} is called the *loading factor*. The input level is often set for a loading factor of 4 (12 dB) to ensure that the overload quantizing noise will be negligible. In practice, this gives $\alpha = -7.3$ for the case of uniform encoding as compared to $\alpha = 0$, which was obtained for the ideal conditions associated with Eq. (3.8).

3.8 Example: Design of a PCM System

Assume that an analog voice-frequency signal, which occupies a band from 300 to 3400 Hz, is to be transmitted over a binary PCM system. The minimum sampling frequency would be $2 \times 3.4 = 6.8$ kHz. In practice, the signal is oversampled, and in the U.S., a sampling frequency of 8 kHz is the standard used for voice-frequency signals in telephone communication systems. Assume that each sample value is represented by 8 b; then the bit rate of the PCM signal is

$$\begin{aligned} R &= (f_s \ \text{ samples}/s)(n \ \text{ b}/s) \\ &= (8k \ \text{ samples}/s)(8 \ \text{ b}/s) = 64 \text{ kb}/s \end{aligned} \tag{3.16}$$

Referring to the dimensionality theorem [Eq. (3.4)], we realize that the theoretically minimum absolute bandwidth of the PCM signal is

$$B_{\min} = \frac{1}{2}R = 32 \text{ kHz} \tag{3.17}$$

and this is realized if the PCM waveform consists of $(\sin x)/x$ pulse shapes. If rectangular pulse shaping is used, the absolute bandwidth is infinity, and the first null bandwidth [Eq. (3.5)] is

$$B_{\text{null}} = R = \frac{1}{T_b} = 64 \text{ kHz} \tag{3.18}$$

That is, we require a bandwidth of 64 kHz to transmit this digital voice PCM signal where the bandwidth of the original analog voice signal was, at most, 4 kHz. Using $n = 8$ in Eq. (3.1), the error on the recovered analog signal is ±0.2%. Using Eqs. (3.12) and (3.13) for the case of uniform quantizing with a loading factor, V/x_{rms}, of 10 (20 dB), we get, for uniform quantizing,

$$\left(\frac{S}{N}\right)_{\text{dB}} = 32.9 \text{ dB} \tag{3.19}$$

Using Eqs. (3.12) and (3.14) for the case of $\mu = 255$ companding, we get

$$\left(\frac{S}{N}\right) = 38.05 \text{ dB} \tag{3.20}$$

These results are illustrated in Fig. 3.4.

Defining Terms

Intersymbol interference: Filtering of a digital waveform so that a pulse corresponding to 1 b will smear (stretch in width) into adjacent bit slots.

Pulse amplitude modulation: An analog signal is represented by a train of pulses where the pulse amplitudes are proportional to the analog signal amplitude.

Pulse code modulation: A serial bit stream that consists of binary words which represent quantized sample values of an analog signal.

Quantizing: Replacing a sample value with the closest allowed value.

References

1. Cattermole, K.W., *Principles of Pulse-code Modulation,* American Elsevier, New York, 1969.
2. Couch, L.W., II, *Digital and Analog Communication Systems,* 6th ed., Prentice Hall, Upper Saddle River, NJ, 2001.
3. Couch, L.W., II, *Modern Communication Systems: Principles and Applications,* Macmillan Publishing, New York, 1995.
4. Dammann, C.L., McDaniel, L.D., and Maddox, C.L., D2 Channel Bank—Multiplexing and Coding. *B. S. T. J.,* 12(10), 1675–1700, 1972.
5. Dorf, R.C., *The Electrical Engineering Handbook,* CRC Press, Boca Raton, FL, 1993.
6. Gibson, J.D., *The Communications Handbook,* CRC Press, Boca Raton, FL, 1997.
7. Jayant, N.S. and Noll, P., *Digital Coding of Waveforms,* Prentice Hall, Englewood Cliffs, NJ, 1984.
8. Lathi, B.P., *Modern Digital and Analog Communication Systems,* 3rd ed., Oxford University Press, New York, NY, 1998.
9. Miyaoka, S., Digital Audio is Compact and Rugged. *IEEE Spectrum,* 21(3), 35–39, 1984.

10. Peek, J.B.H., Communication Aspects of the Compact Disk Digital Audio System. *IEEE Comm. Mag.,* 23(2), 7–15, 1985.
11. Smith, B., Instantaneous Companding of Quantized Signals. *B. S. T. J.,* 36(5), 653–709, 1957.
12. Spilker, J.J., *Digital Communications by Satellite,* Prentice Hall, Englewood Cliffs, NJ, 1977.

Further Information

Many practical design situations and applications of PCM transmission via twisted-pair T-1 telephone lines, fiber optic cable, microwave relay, and satellite systems are given in [2] and [3].

4

Probabilities and Random Variables

Michael O'Flynn
San Jose State University

4.1 Introduction

Probability theory is presented here for those who use it in communication theory and signal processing of random waveforms. The concept of **sampling a finite power waveform**, where a value of time t or n is uniformly chosen over a long time and $x(t)$ or $x(n)$ is observed, is introduced early. All **first-** and **second-order time averages** are visualized statistically. The topics covered with some highlights are as follows:

1. *Discrete probability theory.* The definitions of probability and conditional probability and their evaluation for outcomes of a **random phenomenon** using relative frequency and *axiomatic* formulas are given. The *theorem of total probability*, which elegantly subdivides compound problems into weighted subproblems, is highlighted along with Bayes' theorem.
2. *The theory of one random variable.* The use of **random variables** allows for the solution of probabilistic problems using integration and summations for sequences. The bridge from theory to applications is very short as in the relationship between finding the density or mass function of a function of a random variable and the communication problem of finding statistics of a system's random output given statistics of its random input.
3. *The theory of two random variables.* **Joint distribution, density, and mass functions** give complete probabilistic information about two random variables and are the key to solving most

applications involving continuous and discrete random waveforms and later random processes. Again, the bridge from theory to application is short as in the relationship between finding the density or mass function of a function of two random variables and the communication problem of finding statistics of a system's random output when two random inputs are combined by operations such as summing or multiplying (modulation).

4.2 Discrete Probability Theory

Probability theory concerns itself with assumed random phenomena, which are characterized by outcomes which occur with statistical regularity. These statistical regularities define the probability $p(A)$ of any outcome A and the conditional probability of A given B, $p(A/B)$ by

$$p(A) = \lim_{N\to\infty} \frac{N_A}{N} \qquad \text{and} \qquad p(A/B) = \lim_{N\to\infty} \frac{N_{AB}}{N_B} \tag{4.1}$$

where N, N_A, N_{AB}, and N_B denote the total number of trials, those for which A occurs, those for which both A and B occur, and those for which B occurs, respectively. In order to use these definitions, it is required to know *counting theory* or *permutations and combinations*.

Counting Formulas

A permutation of size k is an ordered array of k objects and a combination of size k is a group or set of size k where order of inclusion is immaterial. From four basic counting formulas, which are inductively developed, plus an understanding of counting factors, probabilities of complex outcomes may be found. These counting formulas are:

$$P_{n,k} = (n)_k = n(n-1)\times\cdots\times(n-k+1) \tag{4.2}$$

which is the number of permutations of size k from n different objects,

$$P_{n,n(n_1,n_2,\ldots,n_k)} = \frac{n!}{n_1!n_2!\times\cdots\times n_k!} \tag{4.3}$$

which is the number of permutations of size n from a total of n objects of which n_1 are identical of type 1 and so on until n_k are identical of type k, and

$$P_{\infty,k(a\text{ types})} = (a)^k \tag{4.4}$$

which is the number of permutations of size k from an infinite supply of a different types.

For example, there are $(10)_3 = 720$ different possible results from a race with 10 competitors where only the first three positions are listed, there are 56 different 8-bit words that can be formed from five 1s and three 0s, and there are $6^5 = 7776$ different results from rolling a die five times and stating the result of each roll.

The only formula used for combinations, which is the most versatile formula in counting, is

$$C_{n,k} = \binom{n}{k} = \frac{(n)_k}{k!} \tag{4.5}$$

for the number of combinations of size k from n different objects.

For example, given 52 balls which come in four colors with each color numbered from 1 to 13, we can form

$$\binom{13}{1}\binom{12}{1}\binom{4}{3}\binom{4}{2}$$

combinations of size 5 where each contains three of one number and two of another number and

$$\binom{13}{3}\binom{4}{2}\binom{4}{2}\binom{4}{2}$$

combinations of size 6 where each contains two each of three different numbers. Also, the number of permutations of size 5 where each contains three of one number and two of another number is

$$\binom{13}{1}\binom{12}{1}\binom{4}{3}\binom{4}{2}5!$$

A sample problem will be solved using the relative frequency definitions of probabilities and counting formulas.

Problem 1

Consider an urn with 50 balls of which 20 are red and 30 have other colors. If a sample of size 6 is drawn from the urn, find the following probabilities:

1. Exactly four red balls are drawn.
2. Exactly four red balls given at least three are red.

Consider the case drawing the sample with and without replacement.

Solution

1. Case without replacement between draws:

$$P[\text{exactly 4 red}] = \frac{\binom{20}{4}\binom{30}{2}}{\binom{50}{6}}$$

or

$$= \frac{(20)_4(30)_2\frac{6!}{4!2!}}{(50)_6} = 0.13$$

$$P[(\text{exactly 4 red})/(\text{at least 3 red})] = \frac{\binom{20}{4}\binom{30}{2}}{\binom{50}{6}-\left[\binom{30}{6}+\binom{30}{5}\binom{20}{1}+\binom{30}{4}\binom{20}{2}\right]}$$

or

$$= \frac{(20)_4(30)_2\frac{6!}{4!2!}}{(50)_6-\left[(30)_6+(30)_5(20)\frac{6!}{5!1!}+(30)_4(20)_2\frac{6!}{4!2!}\right]}$$

$$= 0.291$$

2. Case with replacement between draws:

$$P[\text{exactly 4 red}] = \frac{20^4 30^2 \dfrac{6!}{4!2!}}{(50)^6} = (0.4)^4(0.6)^2\frac{6!}{4!2!} = 0.138$$

$$P[(\text{exactly 4 red})/(\text{at least 3 red})]$$

$$= \frac{(20)^4(30)^2\dfrac{6!}{4!2!}}{(50)^6 - \left[(30)^6 + (30)^5(20)\dfrac{6!}{5!2!} + (30)^4(20)^2\dfrac{6!}{4!2!}\right]}$$

$$= 0.303$$

On reflection, much philosophy is involved in this problem. In order to use the relative frequency formulas, we imagine all of the balls are different and, from symmetry, know the answer is correct. In the case without replacement, we can consider drawing a sample of six balls at once or drawing one at a time and noting the result.

Axiomatic Formulas of Probability Theory

The axioms of probability state that a probability is between zero and one and that the probability of the union of mutually exclusive outcomes is the sum of their probabilities. The axiomatic formulas are:

$$P(A/B) = \frac{P(AB)}{P(B)} \qquad \text{if } P(B) \neq 0 \tag{4.6}$$

and

$$p(AB) = P(A)P(B/A) \tag{4.7}$$

Equation (4.7) may be extended for the intersection of n outcomes to give

$$P\left[\bigcap_{i=1}^{n} A_i\right] = p(A_1)p(A_2/A_1) \times \cdots \times p\left(A_n \Big/ \bigcap_{i=1}^{n-1} A_i\right) \tag{4.8}$$

Also, two outcomes are independent if

$$P(AB) = P(A)P(B) \qquad \text{or} \qquad P(A/B) = P(A) \tag{4.9}$$

and this may be extended to n outcomes. Let us now resolve part of Problem 1 using the axiomatic formulas.

Solution

1. Without replacement:

$$P[(\text{exactly 4 red})] = \left(\frac{20}{50} \times \frac{19}{49} \times \frac{18}{48} \times \frac{17}{47} \times \frac{30}{46} \times \frac{29}{45}\right)\frac{6!}{4!2!}$$

The outcome of exactly 4 red is the union of 6!/4! 2! mutually exclusive outcomes, each with the same probability.

2. With replacement:

$$P[(\text{exactly 4 red})] = [(0.4)^4(0.6)^2]\frac{6!}{4!2!}$$

as before.

The Theorem of Total Probability and Bayes' Theorem

An historically elegant theorem which simplifies compound problems is *Bayes' Theorem*, and we will state it in a casual manner and note some applications.

Bayes' Theorem(s)

Consider a random phenomenon where one trial consists in performing a trial of one of m random phenomena $B_1, B_2, \ldots, B_m$ where the probability of performing a trial of B_1 is $P(B_1)$ and so on until $P(B_m)$ is the probability of performing a trial of B_m. Then, for any outcome A,

$$P(A) = P(B_1)P(A/B_1) + P(B_2)P(A/B_2) + \cdots + P(B_m)P(A/B_m) \tag{4.10}$$

This is the *theorem of total probability*. Using the axiomatic formula, we obtain

$$P(B_k/A) = \frac{P(B_k)P(A/B_k)}{\sum_{i=1}^{m} P(B_i)P(A/B_i)} \tag{4.11}$$

This is Bayes' theorem.

Problem 2

Consider the random phenomenon of sampling the periodic waveform

$$x(t) = \sum_n g(t-5n)$$

which is shown plotted in Fig. 4.1. A value of time is uniformly chosen over a long time (or over exactly one period) and $x(t)$ is observed. First find

$$P[-0.2 < x(t) \le 2.3]$$

then find

$$P\left[\left(x(t) \text{ from } \frac{1}{t} \text{ part}\right)\Big/(-0.2 < x(t) \le 2.3)\right]$$

Solution

Let us define the Bayesian space (B_1, B_2, B_3) to denote sampling the $\frac{1}{t}$ portion, the $x(t) = -1$ portion, or the $2t - 8$ portion, respectively, with probabilities $p(B_1) = 0.4$, $p(B_2) = 0.2$, and $p(B_3) = 0.4$. If we let A be the event $(-0.2 < x(t) \le 2.3)$, then, using the theorem of total probability,

$$\begin{aligned} P(A) &= P(B_1)P(A/B_1) + P(B_2)P(A/B_2) + P(B_3)P(A/B_3) \\ &= 0.4\frac{(2-0.43)}{2} + 0.2(0) + 0.4\frac{5-3.9}{5-3} = 0.534 \end{aligned}$$

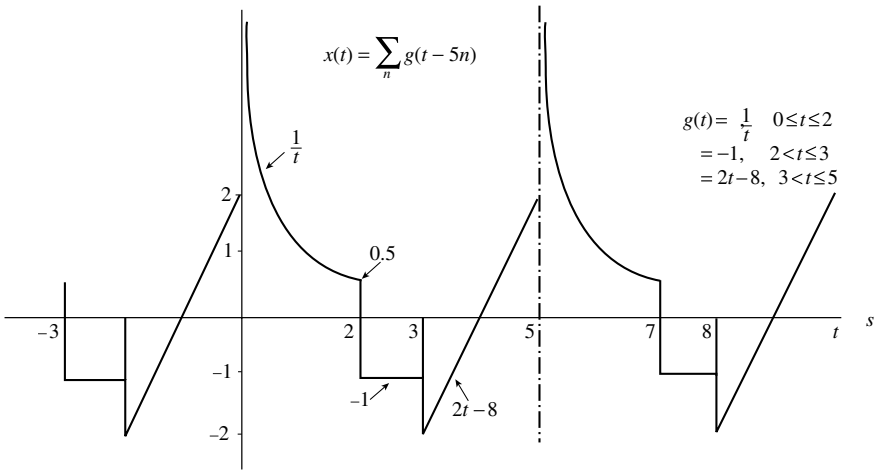


FIGURE 4.1 The periodic function for Problem 2.

and

$$p(B_1/A) = \frac{0.2(2-0.43)}{0.534} = 0.59$$

using Bayes' theorem.

Structured Problem Solving

Whenever many probabilistic questions are asked about a random phenomenon, a structured outline is given for the solution in three steps.

In step 1, an appropriate *event space* or the **sample description space** of the phenomenon is listed. The sample description space is the set of finest grain, mutually exclusive, collectively exhaustive outcomes, whereas an event space is a listing of mutually exclusive, collectively exhaustive outcomes. In step 2, a probability is assigned to each outcome of the chosen space using the relative frequency or axiomatic formulas. In step 3, any desired probabilities are found by using the axiomatic formulas.

For example, consider a seven-game playoff series between two teams A and B, which terminates when a team wins four games. Assume that for any game, $P(A \text{ wins}) = 0.6$ and $P(B \text{ wins}) = 0.4$. If our interest is in how long the series lasts and who wins in how many games, then an appropriate event space is

$$E = \{e_{A4}, e_{A5}, e_{A6}, e_{A7}, e_{B4}, e_{B5}, e_{B6}, e_{B7}\}$$

where, for example, e_{B6} is the outcome B wins in exactly six games and

$$p(e_{B6}) = (0.4)^4(0.6)^2\frac{5!}{3!2!} = 0.092$$

Similarly, we can find the probabilities for the other seven outcomes. Some problems which now can be easily solved are

$$P[(\text{series lasts at least 6 games})] = p(e_{A6}) + p(e_{A7}) + p(e_{B6}) + p(e_{B7})$$

and

$$P[(A \text{ wins in at most 6 games})/(\text{series lasts at least 5 games})]$$
$$= (p[e_{A5}] + p[e_{A6}]) \div \left(\sum_{i=5}^{7} p(e_{Ai}) + \sum_{i=5}^{7} p(e_{Bi})\right)$$

4.3 The Theory of One Random Variable

A random variable X assigns to every outcome of the sample description space of a random phenomenon a real number $X(s_i)$. Associated with X are three very important functions:

1. The **cumulative distribution function** $F_X(\alpha)$ defined by

$$F_X(\alpha) = P[X \le \alpha], \quad -\infty < \alpha < \infty \tag{4.12}$$

2. The density function $f_X(\alpha)$ defined by

$$f_X(\alpha) = \frac{\mathrm{d}}{\mathrm{d}\alpha} F_X(\alpha)$$

3. If appropriate, the probability mass function $p_X(\alpha_i)$ defined by $p_X(\alpha_i) = P[X = \alpha_i]$.

Given a random phenomenon, finding $F_X(\alpha)$, $f_X(\alpha)\Delta\alpha$, or $p_X(\alpha_i)$ are problems in discrete probability.

Problem 3

Given the periodic waveform

$$x(t) = \sum_n g(t - 5n)$$

of Fig. 4.1, let the **random variable *X*** describe sampling $x(t)$ with a one-to-one mapping. Find and plot $F_X(\alpha)$ and $f_X(\alpha)$.

Solution

Observing $x(t)$ in Fig. 4.1, we will define X_1 as sampling $x(t) = 1/t$, X_2 as sampling $x(t) = -1$, and X_3 as sampling $x(t) = 2t - 8$. Figure 4.2(a) shows the solution for $F_{X_i}(\alpha)$ and $f_{X_i}(\alpha)$ and Fig. 4.2(b) uses the theorem of total probability to find $F_X(\alpha)$ as $0.4F_{X_1}(\alpha) + 0.2F_{X_2}(\alpha) + 0.4F_{X_3}(\alpha)$ and, similarly, $f_X(\alpha)$ is found. In this case, X is said to be a mixed random variable as $F_X(\alpha)$ changes both continuously and in jumps. The delta function allows for the inclusion of a point taking on a specific probability in $f_X(\alpha)$.

Figure 4.3 shows density or mass functions for four of the most commonly occurring random variables: the continuous uniform and Gaussian density functions and the discrete binomial and Poisson mass functions. Also included are some statistics to be encountered later.

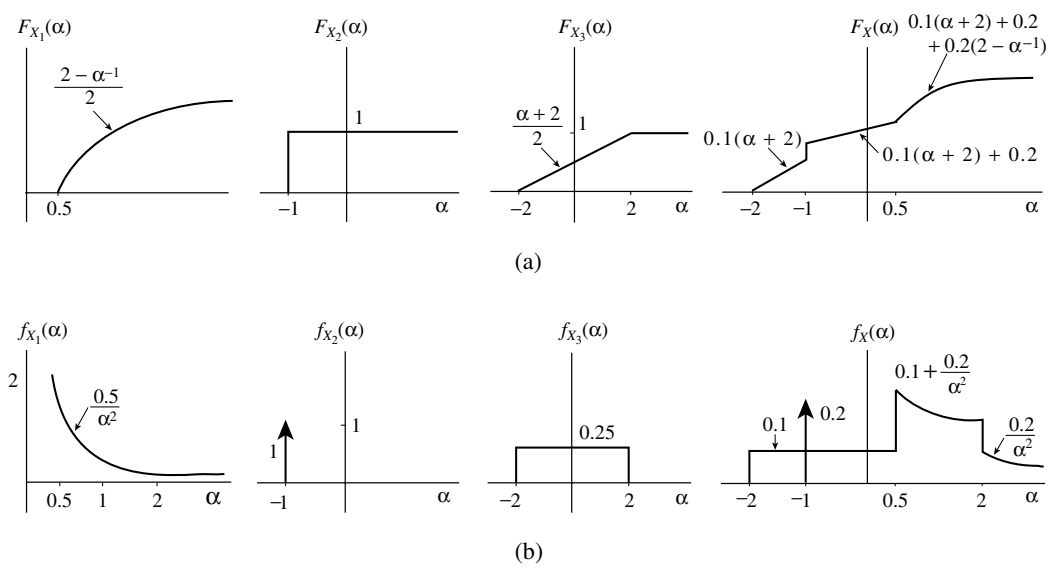


FIGURE 4.2 (a) $F_X(\alpha)$ for $x(t)$ in Fig. 4.1 for Problem 3 and (b) $f_X(\alpha)$ for $x(t)$ in Fig. 4.1.

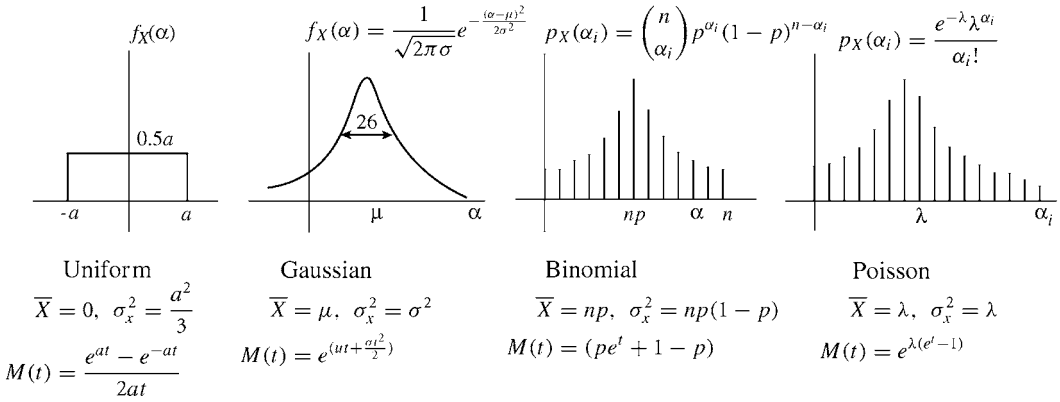


FIGURE 4.3 Important density and mass functions.

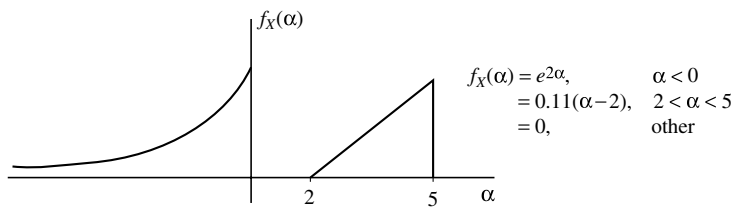


FIGURE 4.4 Density function $f_X(\alpha)$ for Problem 4.

Finding Probabilities from Density and Mass Functions

Density and mass functions allow for the use of integration and summations to answer probabilistic questions about a random variable. This is illustrated with two problems.

Problem 4

Consider the density function $f_X(\alpha) = e^{2\alpha}$, $-\infty < \alpha \le 0$; $f_X(\alpha) = 0.11(\alpha - 2)$, $2 \le \alpha \le 5$; and $f_X(\alpha) = 0$ otherwise. Find $P[(-1.0 < X < 4.0)/(-X^2 < -5)]$.

Solution

A sketch of $f_X(\alpha)$ is shown in Fig. 4.4. Using the axiomatic formula,

$$P[(-1.0 < X < 4.0)/(-X^2 < -5)] = \frac{0 + \int_{2.24}^{4} 0.11(\alpha - 2)\,\mathrm{d}\alpha}{\int_{-\infty}^{-2.24} 0.5e^{\alpha}\,\mathrm{d}\alpha + \int_{2.24}^{5} 0.11(\alpha - 2)\,\mathrm{d}\alpha} = 0.398$$

Problem 5

Given the probability mass function $p_X(\alpha_i) = 0.01$, $-40 \le \alpha_1 \le -10$, α_i integer; $p_X(\alpha_i) = 0.10(0.9)^{\alpha_i}$, $\alpha_i \ge 5$, α_i integer; and $p_X(\alpha_i) = 0$ otherwise. Find (1) $P[(|X| \le 20)/(X^2 > 200)]$ and (2) $p_X[\alpha_i/\ (X^2 > 200)]$.

Solution

1. A sketch of $p_X(\alpha_i)$ is shown in Fig. 4.5(a). Using the axiomatic formula,

$$P[(|X| \le 20)/(X^2 > 200)] = \frac{\sum_{-20}^{-15} 0.01 + \sum_{15}^{20} 0.10(0.9)^{\alpha_i}}{\sum_{-40}^{-15} 0.01 + \sum_{15}^{\infty} 0.10(0.9)^{\alpha_i}} = \frac{0.156}{0.467} = 0.34$$

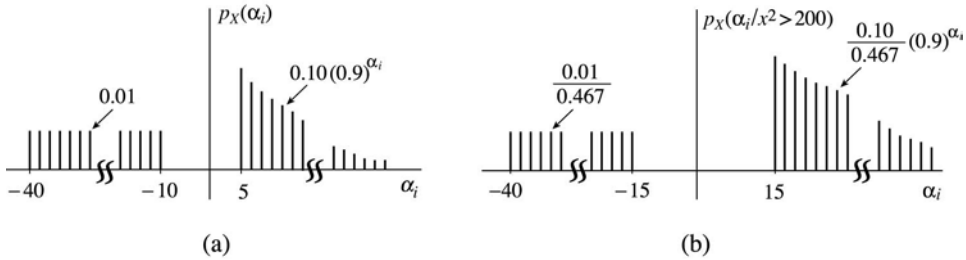


FIGURE 4.5 Probability mass function $p_X(\alpha_i)$ and $p_X(\alpha_i/X^2 > 200)$ for Problem 5.

The formulas for a finite geometric progression

$$\sum_{p=0}^{n} \alpha^p = (1-\alpha^{n+1}) \div (1-\alpha)$$

and for a geometric progression

$$\sum_{p=0}^{\infty} \alpha^p = (1-\alpha)^{-1}, \quad |\alpha| < 1$$

were used in the calculation.

2. The second solution is

$$\begin{aligned} p_X(\alpha_i/(X^2>200)) &= \frac{p_X(\alpha_i)}{P[X^2>200]}, \quad \alpha_i^2 > 200 \\ &= 0, \qquad \text{otherwise} \end{aligned}$$

and is shown plotted in Fig. 4.5(b).

The Density or Mass Function of a Function of a Known Random Variable

Continuous Case

Consider the problem: Given $Y = g(X)$ where $f_X(\alpha)$ is known, find $f_Y(\beta)$.

$$F_Y(\beta) = P[g(X) \le \beta] = \int_{\alpha} f_X(\alpha)\,\mathrm{d}\alpha \quad \text{where } g(\alpha) \le \beta$$

In order to find $f_Y(\beta)$ directly, Leibnitz's rule is used. Leibnitz's rule states

$$\frac{\mathrm{d}}{\mathrm{d}\beta}\int_{k_1(\beta)}^{k_2(\beta)} f_X(\alpha)\,\mathrm{d}\alpha = f_X(k_2(\beta))k_2'(\beta) - f_X(k_1(\beta))k_1'(\beta) \tag{4.13}$$

For example, if

$$f_X(\alpha) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\alpha-\mu)^2}{2\sigma^2}}$$

and $Y = aX + b$, where $a > 0$,

$$F_Y(\beta) = P[aX + b \le \beta] = \int_{-\infty}^{\frac{\beta-b}{a}} f_X(\alpha)\,\mathrm{d}\alpha$$

Using Leibnitz's rule,

$$f_Y(\beta) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\beta-b-a\mu)^2}{2a^2\sigma^2}} \frac{1}{a} = N(a\mu + b,\ (a\sigma)^2)$$

which means Y is also Gaussian with $\mu_Y = a\mu + b$ and $\sigma_Y^2 = (a\sigma)^2$.

Discrete Case

The discrete case is much more direct. If $Y = g(x)$ where $p_X(\alpha_i)$ is known, then $p_Y(\beta_j) = p_X(g^{-1}(\beta))$. For example, if $Y = X^2$, then $p_Y(\beta_j) = p_X(\sqrt{\beta_j}) + p_X(-\sqrt{\beta_j})$.

Problem 6

Given the probability mass function $p_X(\alpha_i) = 0.096(0.9)^{\alpha_i+5}$, $-5 \le \alpha_i \le 30$, α_i integer, $p_X(\alpha_i) = 0$ otherwise, and $Y = |2X|$, find $p_Y(\beta_j)$.

Solution

$$\begin{aligned} Y = |2X| &= 2X, \quad X \ge 0 \\ &= -2X, \quad X < 0 \end{aligned}$$

$$p_Y(0) = 0.096(0.9)^5$$

Range 1: $\beta_j = 2, 4, 6, 8, 10$ or $\beta_j = |2\alpha_i|$, $\alpha_i = \pm 1, \ldots, \pm 5$,

$$p_Y(\beta_j) = 0.096\left[(0.9)^{\frac{\beta_j+10}{2}} + (0.9)^{\frac{-\beta_j+10}{2}}\right]$$

Range 2: $\beta_j = 12, 14, \ldots, 60$, $\quad \beta_j = 2\alpha_i$, $\alpha_i = 6, 30$

$$p_Y(\beta_j) = 0.096\left[(0.9)^{\frac{\beta_j+10}{2}}\right]$$

Statistics of a Random Variable

The three most widely used statistics of a random variable are

$$\begin{aligned} \bar{X} &= \int_{-\infty}^{\infty} \alpha f_X(\alpha)\,\mathrm{d}\alpha \\ &\text{or} \\ &= \sum_{\alpha_i} \alpha_i p_X(\alpha_i) \\ \overline{X^2} &= \int_{-\infty}^{\infty} \alpha^2 f_X(\alpha)\,\mathrm{d}\alpha \\ &\text{or} \\ &= \sum_{\alpha_i} \alpha_i^2 p_\alpha(\alpha_i) \end{aligned}$$

and

$$\sigma_X^2 = \overline{(X-\bar{X})^2} = \int_{-\infty}^{\infty} (\alpha-\bar{X})^2 f_X(\alpha)\,\mathrm{d}\alpha = \overline{X^2} - (\bar{X})^2$$

These are the mean or average value, the mean squared value, and the variance, respectively. The formulas follow from a relative frequency interpretation based on a large number of trials of a phenomenon. Higher order statistics are defined by

$$m_n = \overline{X^n} = \int_{\infty}^{\infty} \alpha^n f_X(\alpha)\,\mathrm{d}\alpha$$

or

$$= \sum_{\alpha_i} \alpha_i^n p_X(\alpha_i)$$

and

$$\overline{(X-\bar{X})^n} = \int_{-\infty}^{\infty} (\alpha-\bar{X})^n f_X(\alpha)\,\mathrm{d}\alpha$$

or

$$= \sum_{\alpha_i} (\alpha-\bar{X})^n p_X(\alpha_i)$$

These are called moments and central moments, respectively. General expressions for higher order statistics may be found using the moment generation function $M(t)$ defined by

$$M(t) = \overline{e^{tX}} = \sum_{\alpha_i} e^{t\alpha_i} p_X(\alpha_i)$$

or

$$= \int_{-\infty}^{\infty} e^{t\alpha} f_X(\alpha)\,\mathrm{d}\alpha \tag{4.14}$$

where $\overline{X^n} = M^n(0)$ for $n \geq 1$. In Fig. 4.3, moment generating functions plus $\bar{X}$ and σ_X^2 are given for the uniform, Gaussian, binomial, and Poisson random variables.

Time Averages as Statistical Averages for Finite Power Waveforms

Continuous Waveforms

A finite power waveform is defined as one for which

$$P_{av} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} x^2(t)\,\mathrm{d}t$$

exists.

The time average definition for some first-order time averages are

$$\widetilde{x(t)} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} x(t)\,\mathrm{d}t \qquad \text{and} \qquad \widetilde{x^2 t} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} x^2(t)\,\mathrm{d}t \tag{4.15}$$

called the average and mean squared value, respectively.

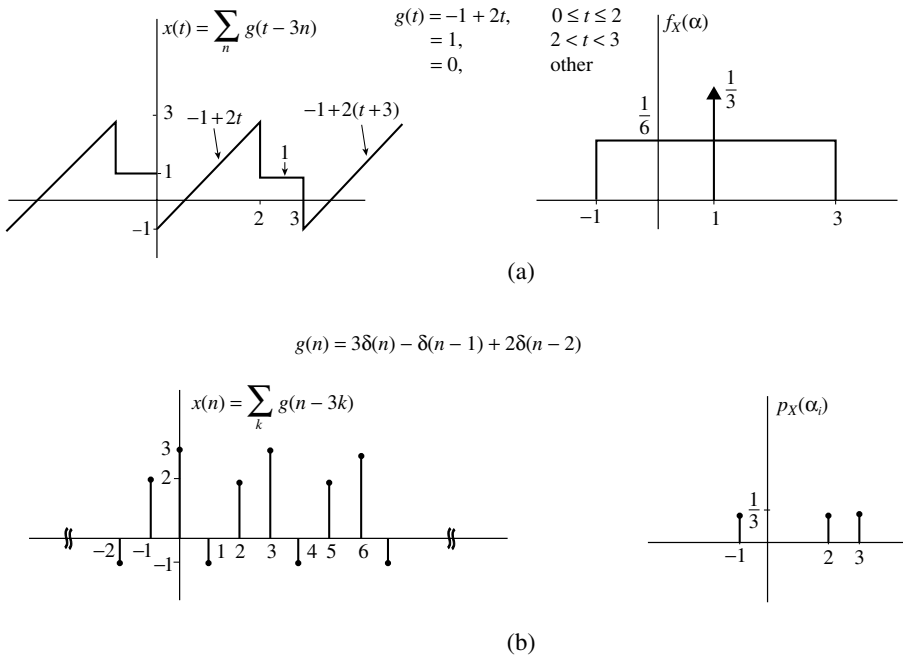


FIGURE 4.6 (a) Waveform $x(t) = \sum_n g(t-3n)$ and $f_X(\alpha)$ for sampling $x(t)$, and (b) waveform $x(n) = \sum_k g(n-3k)$ and $p_X(\alpha_i)$ for sampling $x(n)$.

Applying these definitions to the periodic waveform

$$x(t) = \sum_n g(t-3n)$$

shown in Fig. 4.6(a), it is easy to find the average and mean squared values,

$$\widetilde{x(t)} = \frac{1}{3}\int_0^3 x(t)\,\mathrm{d}t = \frac{1}{3}\left[\int_0^2 (-1+2t)\,\mathrm{d}t + \int_2^3 1\,\mathrm{d}t\right] = 1.0$$

$$\widetilde{x^2(t)} = \frac{1}{3}\left[\int_0^2 (-1+2t)^2\,\mathrm{d}t + \int_2^3 1\,\mathrm{d}t\right] = 1.89 \qquad \text{and} \qquad \sigma^2_{x(t)} = 0.89$$

We will now interpret these averages statistically. Let the random variable X describe sampling t over a long time and observing $x(t)$. The density function $f_X(\alpha)$ is shown plotted in Fig. 4.6(a), and it is easy to show

$$\widetilde{x(t)} = \overline{X} = \left[\int_{-1}^3 \alpha\frac{1}{6}\,\mathrm{d}\alpha + \int_{-\infty}^{\infty} \alpha\frac{1}{3}\delta(\alpha-1)\,\mathrm{d}\alpha\right] = 1.0$$

and

$$\widetilde{x^2(t)} = \overline{X^2} = \left[\int_{-1}^3 \alpha^2\frac{1}{6}\,\mathrm{d}\alpha + \int_{-\infty}^{\infty} \alpha^2\frac{1}{3}\delta(\alpha-1)\,\mathrm{d}\alpha\right] = 1.89$$

The general formula for any first-order statistic is

$$\widetilde{g(x(t))} = \int_{-\infty}^{\infty} g(\alpha) f_X(\alpha) \mathrm{d}\alpha \tag{4.16}$$

Discrete Waveforms

A finite power discrete waveform is one for which

$$P_{\text{av}} = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} x^2(n)$$

exists.

The time average definitions for some first-order time averages are

$$\widetilde{x(n)} = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{-N}^{N} x(n) \qquad \text{and} \qquad \widetilde{x^2(n)} = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{-N}^{N} x^2(n) \tag{4.17}$$

called the average and mean square value, respectively. Applying these definitions to the periodic waveform

$$x(n) = \sum_{k=-\infty}^{\infty} g(n-3k)$$

shown in Fig. 4.6(b), it is easy to find the average, mean square, and variance of $x(n)$,

$$\widetilde{x(n)} = \frac{1}{3}[3 + (-1) + 2] = 1.33$$

$$\widetilde{x^2(n)} = \frac{1}{3}[9 + 1 + 4] = 4.67 \qquad \text{and} \qquad \sigma^2_{x(n)} = 4.67 - 1.78 = 2.89$$

If the random variable X describes uniformly choosing an integer n, $-N \le n \le N$ and observing $x(n)$, then the statistical definition for any first-order time average $\widetilde{g(x(n))}$ is

$$\widetilde{g[x(n)]} = \overline{g(X)} = \int g(\alpha) f_X(\alpha)\, \mathrm{d}\alpha$$

or

$$= \sum g(\alpha_i) p_X(\alpha_i)$$

which is identical to Eq. (4.16). $P_X(\alpha_i)$ is shown in Fig. 4.6(b), $\widetilde{x(n)} = \overline{X} = 3(\frac{1}{3}) + (-1)(\frac{1}{3}) + 2(\frac{1}{3}) = 1.33$, and $\widetilde{x^2(n)} = \overline{X^2} = 9(\frac{1}{3}) + 1(\frac{1}{3}) + 4(\frac{1}{3}) = 4.67$ as before.

Random Waveforms

A finite-power random or noise waveform is defined as being nondeterministic and such that statistics of a large section or of the whole waveform exist. Often, $x(t)$ or $x(n)$ is represented by a formula involving time and random variables,

$$x(t) = f(t, \text{r. vs}) \qquad \text{or} \qquad x(n) = f(n, \text{r. vs})$$

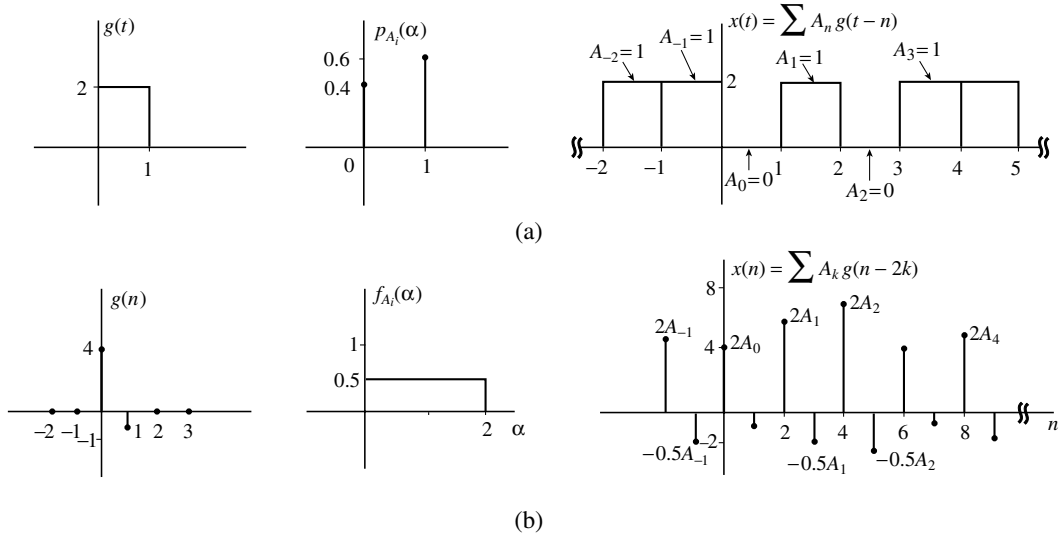


FIGURE 4.7 The random waveforms: (a) $x(t) = \sum_n A_n\, g(t-n)$ and (b) $x(n) = \sum_k A_k\, g(n-2k)$.

where "r. vs" denotes "random variables." For example,

$$x(t) = \sum_{n=-\infty}^{\infty} A_n g(t - nT_0)$$

is a periodic function with each period weighted by a value statistically determined by the density or joint density functions for the As. Consider when $g(t) = 2, 0 \le t \le 1$, and $T_0 = 1$ and the As are independent, each with $p_{A_i}(1) = 0.6$ and $p_{A_i}(0) = 0.4$.

A section of $x(t)$ may be as shown in Fig. 4.7(a). Consider

$$x(n) = \sum_k A_k g(n - 2k)$$

where $g(n) = 4\delta(n) - \delta(n-1)$ and the As are independent random variables each with $f_{A_i}(\alpha) = 0.5$, $0 \le \alpha \le 2$; $f_{A_i}(\alpha) = 0$ otherwise. A typical section of $x(n)$ may appear as in Fig. 4.7(b).

If X describes sampling $x(t)$ or $x(n)$ where t is uniformly chosen $-T < t < T$ and $x(t)$ observed or n is discretely uniformly chosen $-N \le n \le N$ and $x(n)$ observed, then, as T or N approach infinity, $f_X(\alpha)$ or $p_X(\alpha_i)$ may be found.

Any time average may be found by the statistical formula (4.16)

$$\widetilde{g(x(t))} \quad \text{or} \quad \widetilde{g[x(n)]} = \int_{-\infty}^{\infty} g(\alpha) f_X(\alpha)\, d\alpha$$

or

$$= \sum_{\alpha_i} g(\alpha_i) p_X(\alpha_i)$$

Two simple examples will be used for demonstration.

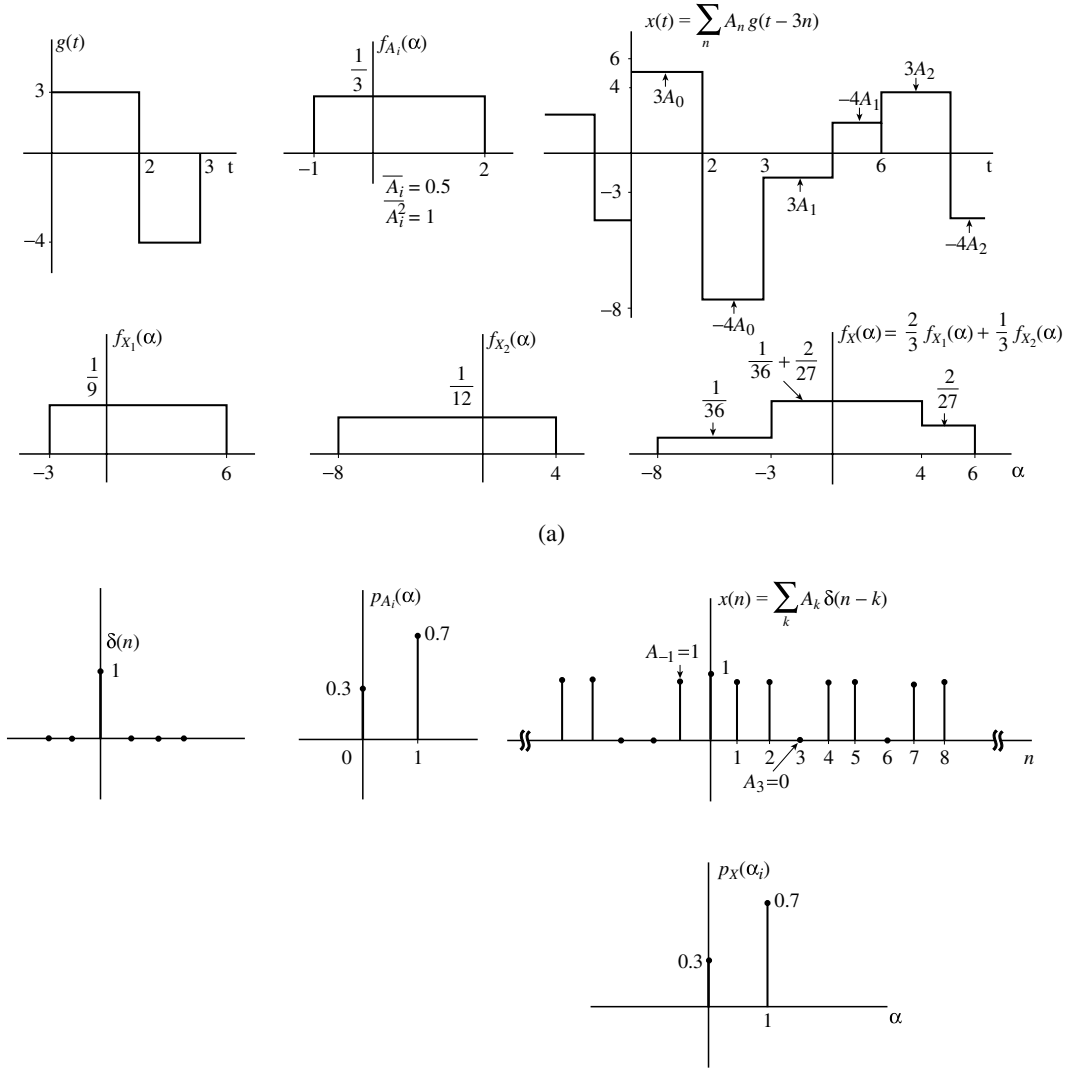


FIGURE 4.8 The random waveform of (a) Problem 7 and $f_X(\alpha)$ and (b) Problem 8 and $p_X(\alpha_i)$.

Problem 7

Given the random waveform

$$x(t) = \sum_n A_n g(t-3n)$$

where $g(t) = 3$, $0 \le t \le 2$; $g(t) = -4$, $2 < t \le 3$; $g(t) = 0$ otherwise and the As are independent random variables with $f_{A_i}(\alpha) = \frac{1}{3}$, $-1 \le \alpha \le 2$; $f_{A_i}(\alpha) = 0$ otherwise, find $f_X(\alpha)$, $\overline{x(t)}$, and $\overline{x^2(t)}$.

Solution

Figure 4.8(a) shows a plot of a typical section of $x(t)$. We notice the front two units of a pulse take on values uniformly from −3 to 6 and the last unit takes on that value multiplied by −1.33. Using the theorem of total probability, $f_X(\alpha) = \frac{2}{3} f_{X_1}(\alpha) + \frac{1}{3} f_{X_2}(\alpha)$ where X_1 and X_2 describe sampling the first two and last

unit of a pulse, respectively. The density function $f_X(\alpha)$ is shown in Fig. 4.8(a) and it is easy to find; $\widetilde{x(t)} = \overline{X} = \frac{2}{3}(3\overline{A}_i) + \frac{1}{3}(-4\overline{A}_i) = 0.33$ and $\widetilde{x^2(t)} = \overline{X}^2 = \frac{2}{3}(9\overline{A_i^2} + \frac{1}{3}(16\overline{A_i^2}) = 11.33$.

Problem 8

Given the random waveform

$$x(n) = \sum_k A_k \delta(n-k)$$

where the A_s are independent random variables with $p_{A_i}(1) = 0.7$ and $p_{A_i}(0) = 0.3$, find $p_X(\alpha_i)$, $\widetilde{x(t)}$ and $\widetilde{x^2(t)}$.

Solution

Figure 4.8(b) shows a plot for a typical section of $x(n)$ and $p_X(\alpha_i)$. $\widetilde{x(n)} = \overline{X} = (1)0.7$ and $\widetilde{x^2(t)} = \overline{X^2} = 0.7$.

Finally, let us consider an application for the density function of a known random variable.

The Square Wave Detector

Consider a noise waveform input to a system where the output and input are related by $y(t) = x^2(t)$. If the density function for sampling the input $x(t)$ is

$$f_X(\alpha) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(\alpha-\mu)^2}{2\sigma^2}}$$

find $f_Y(\beta)$ for sampling the output $y(t)$.

Solution

$$F_Y(\beta) = 2\int_0^{\sqrt{\beta}} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(\alpha-\mu)^2}{2\sigma^2}} \, d\alpha$$

and by Leibnitz's rule

$$f_Y(\beta) = \frac{2}{\sqrt{2\pi}\,\sigma} e^{-\frac{(\sqrt{\beta}-\mu)^2}{2\sigma}} 0.5(\beta)^{-0.5}$$

which is the famous chi-squared density function.

4.4 The Theory of Two Random Variables

Definitions of Joint Distribution, Density, and Mass Functions

Given a random phenomenon, two random variables X and Y map every point of the sample description space onto an α–β plane via $X(s_i)$ and $Y(s_i)$. The following important functions are associated with X and Y:

$$F_{XY}(\alpha, \beta) = P[(X \le \alpha) \cap (Y \le \beta)], \qquad -\infty < \alpha \le \infty, \quad -\infty < \beta < \infty \tag{4.18}$$

This is the joint cumulative distribution function.

$$f_{XY}(\alpha, \beta) = \frac{\partial^2}{\partial\alpha\partial\beta} F_{XY}(\alpha, \beta) \tag{4.19}$$

is the joint density function and

$$p_{XY}(\alpha_i, \beta_j) = P[(X = \alpha_i) \cap (Y = \beta_j)] \tag{4.20}$$

is the joint mass function when appropriate.

Finding $F_{XY}(\alpha, \beta)$, $f_{XY}(\alpha, \beta)\,\Delta\alpha\,\Delta\beta$, or $p_{XY}(\alpha_i, \beta_j)$ are problems in discrete probability theory. Knowing $f_{XY}(\alpha, \beta)$ and $p_{XY}(\alpha_i, \beta_j)$ or $F_{XY}(\alpha, \beta)$, it is possible to answer any probabilistic question about X and Y, X given a condition on Y, or Y given a condition on X. The following are some formulas which are easy to derive starting from a statement with probabilistic meaning:

$$f_X(\alpha) = \int_{-\infty}^{\infty} f_{XY}(\alpha, \beta)\,\mathrm{d}\beta, \quad p_Y(\beta_j) = \sum_{\beta_j} p_{XY}(\alpha_i, \beta_j) \tag{4.21}$$

$$f_Y(\beta/g(X) < C_1) = \frac{\int f_{XY}(\alpha, \beta)\,\mathrm{d}\alpha}{\int_{-\infty}^{\infty}\int f_{XY}(\alpha, \beta)\,\mathrm{d}\alpha\,\mathrm{d}\beta} \tag{4.22}$$

where the two missing limits of integration are: all α such that $g(\alpha) < C_1$,

$$p_X(\alpha_i/Y^2 \le 25) = \frac{\sum_{\beta_j=-5}^{5} p_{XY}(\alpha_i, \beta_j)}{\sum_{\alpha_i}\sum_{\beta_j=-5}^{5} p_{XY}(\alpha_i, \beta_j)} \tag{4.23}$$

and

$$f_Y(\beta/X = \alpha) = \frac{f_{XY}(\alpha, \beta)}{f_X(\alpha)} \qquad \text{all } \beta \text{ when } X = \alpha \tag{4.24}$$

The use of these formulas involves detailed attention and care with different ranges.

Finding Density and Mass Functions from Joint Functions

Problem 9

Given $f_{XY}(\alpha, \beta) = -0.0010\beta$, $0.2\beta \le \alpha \le -0.2\beta$, $-20 \le \beta \le -8$; $f_{XY}(\alpha, \beta) = 0$ otherwise. Find and plot $f_Y(\beta)$ and $f_X(\alpha/(Y = \beta))$.

Solution

The region for which $f_{XY}(\alpha, \beta)$ is defined is shown in Fig. 4.9(a). Using

$$f_Y(\beta) = \int_{-\infty}^{\infty} f_{XY}(\alpha, \beta)\,\mathrm{d}\alpha$$

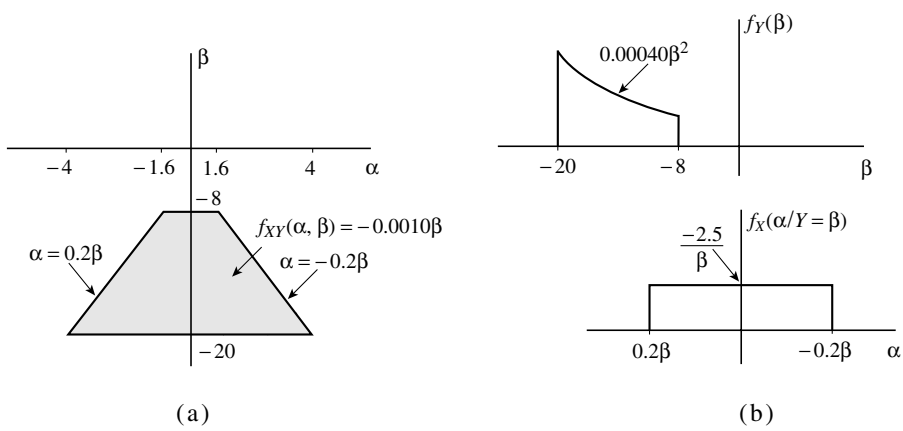


FIGURE 4.9 (a) Region for $f_{XY}(\alpha, \beta)$ for Problem 9 and (b) $f_Y(\beta)$ and $f_X(\alpha/(Y = \beta))$.

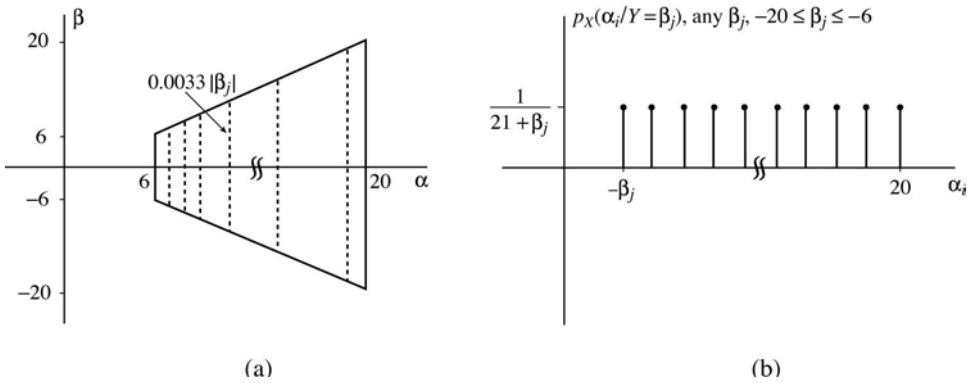


FIGURE 4.10 (a) Region for $p_{XY}(\alpha_i, \beta_j)$ for Problem 10 and (b) $p_X(\alpha_i/(Y = \beta_j))$.

we obtain the following.

Range 1: $-20 \le \beta \le -8$

$$\begin{aligned} f_Y(\beta) &= \int_{0.2\beta}^{-0.2\beta} -0.0010\beta \, d\alpha = 0.0004\beta^2, \quad -20 \le \beta \le -8 \\ &= 0 \qquad \text{otherwise,} \qquad \text{and} \\ f_X(\alpha/Y = \beta) &= \frac{f_{XY}(\alpha, \beta)}{f_Y(\beta)} \\ &= \frac{-0.0010\beta}{0.0004\beta^2} = \frac{-2.5}{\beta}, \quad 0.2\beta < \alpha < -0.2\beta \\ &\quad \text{for any } \beta, \quad -20 \le \beta \le -8 \\ &= 0 \qquad \text{otherwise} \end{aligned}$$

The density functions are plotted in Fig. 4.9(b).

Problem 10

Given $p_{XY}(\alpha_i, \beta_j) = 0.0033|\beta_j|$, $-\alpha_i \le \beta_j \le \alpha_i$, $6 \le \alpha_i \le 20$, both integer, $p_{XY}(\alpha_i, \beta_j) = 0$ otherwise, find $p_X(\alpha_i)$ and $p_X(\alpha_i/(Y = \beta_j))$ for $-20 \le \beta_j \le -6$. Are X and Y independent?

Solution

The region for which $p_{XY}(\alpha_i, \beta_j)$ is defined is shown in Fig. 4.10(a).

Range 1: $6 \le \alpha_i \le 20$

$$\begin{aligned} p_X(\alpha_i) &= \sum_{\beta_j=-\alpha_i}^{\alpha_i} 0.0033|\beta_j| = 2\sum_{1}^{\alpha_i} 0.0033\beta_j \\ &= 0.0033(2)\frac{(\alpha_i + 1)\alpha_i}{2} = 0.0033(\alpha_i + 1)\alpha_i \end{aligned}$$

using the formula for an arithmetic progression. Otherwise, $p_X(\alpha_i) = 0$.

$$p_X[\alpha_i/Y = \beta_j] = \frac{p_{XY}(\alpha_i, \beta_j)}{p_Y(\beta_j)}$$

Looking at Fig. 4.10(a), for the region when $-20 \le \beta_j \le -6$, the formula

$$p_Y(\beta_j) = \sum_{\alpha_i} p_{XY}(\alpha_i, \beta_j)$$

becomes

$$p_Y(\beta_j) = \sum_{\alpha_i=-\beta_j}^{20} p_{XY}(\alpha_i, \beta_j) = 0.0033(21+\beta_j)(-\beta_j)$$

and

$$p_X(\alpha_i/Y=\beta_j) = \frac{0.0033(-\beta_j)}{0.0033(21+\beta_j)(-\beta_j)} = \frac{1}{21+\beta_j}, \quad -\beta_j \le \alpha_i \le 20, \text{ any } \beta_j, -20 \le \beta_j \le -6$$
$$= 0 \qquad \text{otherwise.}$$

This conditional mass function is plotted in Fig. 4.10(b). Since $p_X(\alpha_i/(Y=\beta_j))$ does not equal $p_X(\alpha_i)$ for all α_i, then X and Y are not independent.

The Density (or Mass) Function of a Function of Two Random Variables

Continuous Case

Given $Z = g(X, Y)$ where $f_{XY}(\alpha, \beta)$ is known, $f_Z(\gamma)$ is found as

$$f_Z(\gamma) = \frac{\mathrm{d}}{\mathrm{d}\gamma} F_Z(\gamma) = \frac{\mathrm{d}}{\mathrm{d}\gamma} \iint f_{XY}(\alpha, \beta)\,\mathrm{d}\alpha\,\mathrm{d}\beta \tag{4.25a}$$

where the limits on the double integral satisfy $g(\alpha, \beta) \le \gamma$.

By Leibnitz's rule, Eq. (4.25a) becomes, for the case $Z = aX + bY$,

$$f_Z(\gamma) = \int_{-\infty}^{\infty} C_1 f_{XY}(\alpha, g_1(\alpha, \gamma))\,\mathrm{d}\alpha$$

or

$$= \int_{-\infty}^{\infty} C_2 f_{XY}(g_2(\beta, \gamma), \beta)\,\mathrm{d}\beta \tag{4.25b}$$

Problem 11

Given $f_{XY}(\alpha, \beta) = 0.25$; $0 \le \beta \le 2\alpha$, $0 < \alpha < 2$; $f_{XY}(\alpha, \beta) = 0$ otherwise, find $f_Z(\gamma)$ when $Z = Y - 2X$.

Solution

Figure 4.11 shows the region for which $f_{XY}(\alpha, \beta)$ is defined. If we visualize the family of curves $\beta - 2\alpha = \gamma$ as γ varies, $-\infty < \gamma < \infty$, we note $F_Z(\gamma)$ is the weighed integral of all points to the right of $\beta - 2\alpha = \gamma$. $F_Z(\gamma) = 0$, $\gamma < -4$, and $F_Z(\gamma) = 1$, $\gamma > 0$.

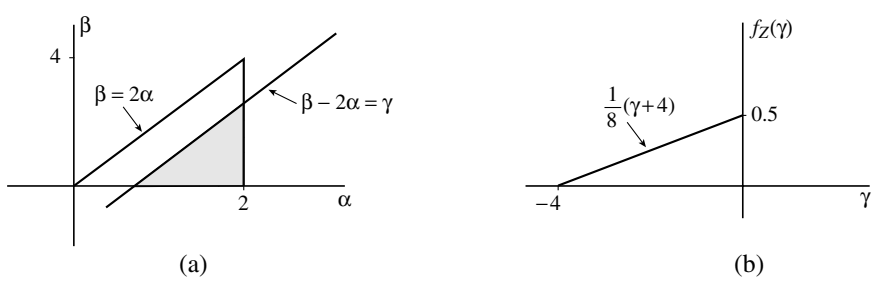


FIGURE 4.11 (a) Region of $f_{XY}(\alpha, \beta)$ for Problem 11 and (b) $f_Z(\gamma)$, $Z = Y - 2X$.

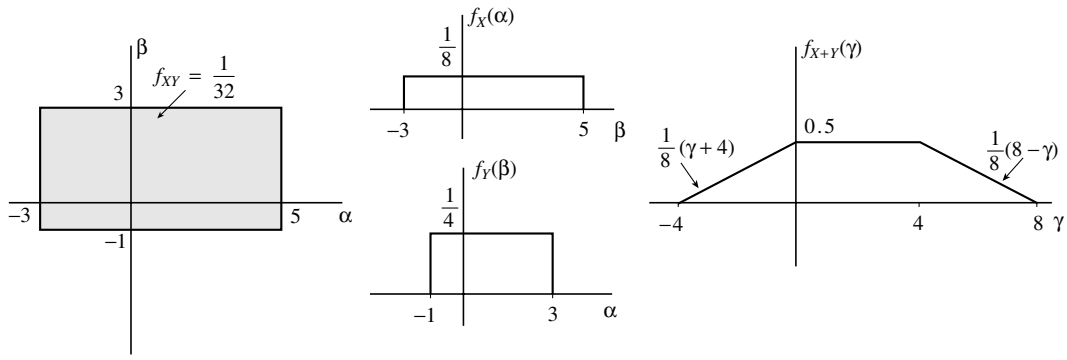


FIGURE 4.12 Function $f_Z(\gamma)$ by convolution when $f_{XY}(\alpha, \beta) = \frac{1}{32}$; $-3 \le \alpha \le 5$, $-1 \le \beta \le 3$, and $Z = X + Y$.

In general, $F_Z(\gamma) = \int_{-\infty}^{\infty}(\int_{-\infty}^{2\alpha+\gamma} f_{XY}(\alpha, \beta)\,\mathrm{d}\beta)\mathrm{d}\alpha$ and using Leibnitz's rule

$$f_Z(\gamma) = \int_{-\infty}^{\infty} f_{XY}(\alpha, 2\alpha+\gamma)\,\mathrm{d}\alpha$$

Observing Fig. 4.11(a),

$$\begin{aligned} f_Z(\gamma) &= \int_{-\frac{\gamma}{2}}^{2} 0.25\,\mathrm{d}\alpha \\ &= \frac{1}{8}(\gamma+4), \qquad -4 < \gamma < 0 \end{aligned}$$

which is shown in Fig. 4.11(b).

An important situation occurs when X and Y are independent random variables or when $f_{XY}(\alpha, \beta) = f_X(\alpha)\, f_Y(\beta)$ for all α, β. In this case, if $Z = X + Y$, then

$$\begin{aligned} f_Z(\gamma) = f_X(\gamma) * f_Y(\gamma) &= \int_{-\infty}^{\infty} f_X(\gamma-\beta) f_Y(\beta)\,\mathrm{d}\beta \\ &\text{or} \\ &= \int_{-\infty}^{\infty} f_X(\alpha) f_Y(\gamma-\alpha)\,\mathrm{d}\alpha \end{aligned} \tag{4.26}$$

which is the convolution of the two density functions. For example, if $f_{XY}(\alpha, \beta) = \frac{1}{32}$, $-3 \le \alpha \le 5$, $-1 \le \beta \le 3$, $f_{XY}(\alpha, \beta) = 0$ otherwise, then it is easy to find $f_X(\alpha) = \frac{1}{8}$, $-3 \le \alpha \le 5$, $f_X(\alpha) = 0$ otherwise and $f_Y(\beta) = \frac{1}{4}$, $-1 \le \beta \le 3$, $f_Y(\beta) = 0$ otherwise. Also, X and Y are independent. Figure 4.12 shows $f_{XY}(\alpha, \beta)$, $f_X(\alpha)$, and $f_Y(\beta)$ and gives the convolution of $f_X(\gamma)$ with $f_Y(\gamma)$.

Discrete Case

Given $Z = g(X, Y)$ where $p_{XY}(\alpha_i, \beta_j)$ is known,

$$p_Z(\gamma_i) = \sum_{\beta_j}\sum_{\alpha_i} p_{XY}(\alpha_i, \beta_j)$$

for all α_i, β_j such that $g(\alpha_i, \beta_j) = \gamma_i$. For example, if $Z = XY$,

$$p_Z(\gamma_i) = \sum_{\alpha_i} p_{XY}\left(\alpha_i, \frac{\gamma_i}{\alpha_i}\right) = \sum_{\beta_j} p_{XY}\left(\frac{\gamma_i}{\beta_j}, \beta_j\right)$$

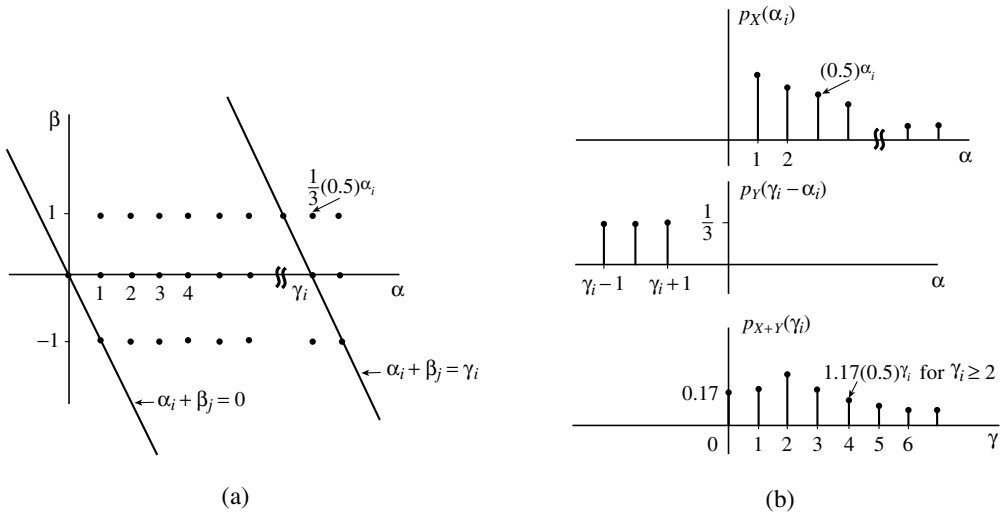


FIGURE 4.13 (a) Region of $p_{XY}(\alpha_i)$ for Problem 12 and (b) $p_X(\alpha_i)$, $p_Y(\gamma_i - \alpha_i)$, $p_{X+Y}(\gamma_i)$, $Z = X + Y$.

Problem 12

Given $p_{XY}(\alpha_i, \beta_j) = \frac{1}{3}(0.5)^{\alpha_i}$, $1 \leq \alpha_i \leq \infty$, $-1 \leq \beta_j \leq 1$, both integers, $p_{XY}(\alpha_i, \beta_j) = 0$ otherwise, and $Z = X + Y$, find (1) $p_Z(\gamma_i)$ directly and (2) if X and Y are independent, find $p_Z(\gamma_i)$ using discrete convolution.

Solution

1. A sketch of the region for which $p_{XY}(\alpha_i, \beta_j)$ exists is shown in Fig. 4.13. Since $Z = X + Y$, then, when $Z = \gamma_i$,

$$p_Z(\gamma_i) = \sum_{\alpha_i} p_{XY}(\alpha_i, \gamma_i - \alpha_i) = \sum_{\beta_j} p_{XY}(\gamma_j - \beta_j, \beta_j)$$

We will use the latter formula as β_j has only three values:

$$p_Z(0) = p_{XY}(1, -1) = \frac{1}{6}$$

$$p_Z(1) = \frac{1}{3}\left[\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)\right] = \frac{1}{4}$$

and for the range $\gamma_i \geq 2$,

$$p_Z(\gamma_i) = \sum_{\beta_j=-1}^{1} p_{XY}(\gamma_i - \beta_j, \beta_j) = \frac{1}{3}\left[\left(\frac{1}{2}\right)^{\gamma_i - 1} + \left(\frac{1}{2}\right)^{\gamma_i} + \left(\frac{1}{2}\right)^{\gamma_i + 1}\right]$$

$$= \frac{7}{6}\left(\frac{1}{2}\right)^{\gamma_i} = 1.17(0.5)^{\gamma_i}$$

Figure 4.13(b) shows $p_Z(\gamma_i)$ plotted.

2. It is easy to show that $p_X(\alpha_i) = (0.5)^{\alpha_i}$, $\alpha_i \geq 1$ and is zero otherwise and

$$p_Y(\beta_j) = \frac{1}{3}\sum_{1}^{\infty}(0.5)^{\alpha_i} = \frac{1}{3}, \quad -1 \leq \beta_j \leq 1;$$

$$= 0, \quad \text{otherwise}$$

Since $p_{XY}(\alpha_i, \beta_j) = p_X(\alpha_i)p_Y(\beta_j)$, then X and Y are independent and

$$p_Z(\gamma_i) = \sum_{\beta_j} p_X(\gamma_i - \beta_j)p_Y(\beta_j) \qquad \text{or} \qquad \sum_{\alpha_i} p_X(\alpha_i)p_Y(\gamma_i - \alpha_i) \tag{4.27}$$

which is the discrete convolution of $p_X(\gamma_i)$ with $p_Y(\gamma_i)$. We will use

$$p_Z(\gamma_i) = \sum_{\alpha_i} p_X(\alpha_i)p_Y(\gamma_i - \alpha_i)$$

For the range $\gamma_i < 0$, $p_Z(\gamma_i) = 0$ as $p_Y(\gamma_i - \alpha_i)$ has not reached $\alpha_i = 1$. $p_Z(0) = \frac{1}{3}(\frac{1}{2}) = \frac{1}{6}$, $p_Z(1) = \frac{1}{3}(\frac{1}{4}) + \frac{1}{3}(\frac{1}{2}) = \frac{1}{4}$, and

$$p_Z(\gamma_i) = \sum_{\gamma_i - 1}^{\gamma_i + 1} (0.5)^{\alpha_i} \frac{1}{3} = \frac{7}{6}(0.5)^{\gamma_i} = 1.17(0.5)^{\gamma_i}, \qquad \gamma_i \geq 2.$$

Statistics of Two Random Variables

Associated with two random variables X and Y, the general second-order moment is

$$M_{p,q} = \overline{X^p Y^q} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \alpha^p \beta^q f_{XY}(\alpha, \beta)\, \mathrm{d}\alpha\, \mathrm{d}\beta$$

or

$$= \sum\sum \alpha_i^p \beta_j^q p_{XY}(\alpha_i, \beta_j) \tag{4.28}$$

The general second-order central moment $L_{XY} = \overline{(X - \overline{X})^p (Y - \overline{Y})^q}$ is similarly defined. The two most widely used of these are the correlation $R_{XY} = \overline{XY}$ and the covariance $L_{XY} = \overline{(X - \overline{X})(Y - \overline{Y})}$. Associated with these statistics, a two-dimensional moment generating function is defined:

$$M(s, t) = e^{\overline{sX + tY}} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{s\alpha + t\beta} f_{XY}(\alpha, \beta)\, \mathrm{d}\alpha\, \mathrm{d}\beta$$

and

$$\overline{X^p Y^q} = \frac{\delta^{p+q}}{(\delta s)^p (\delta l)^q} M(s, t)\big|_{s=t=0}$$

The moment generating function is particularly useful for obtaining results for jointly Gaussian random variables.

Second-Order Time Averages for Finite Power Waveforms (FPW)

The time autocorrelation function for a deterministic or random finite power waveform (FPW) is

$$\tilde{R}_{xx}(\tau) = \widetilde{x(t)x(t+\tau)} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} x(t)x(t+\tau)\, \mathrm{d}t \tag{4.29}$$

or

$$\tilde{R}_{xx}(k) = \widetilde{x(n)x(n+k)} = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} x(n)x(n+k) \tag{4.30}$$

The ~ symbol is used to distinguish them from the fundamental ensemble definition for $R_{xx}(\tau)$ or $R_{xx}(k)$ used for random processes.

These time-average definitions may be assigned an equivalent statistical interpretation. For a continuous FPW $x(t)$, X and Y describe uniformly sampling $x(t)$ at t and $t+\tau$ and observing $X = x(t)$ and $Y = x(t+\tau)$. For N trials as $N \to \infty$,

$$\tilde{R}_{xx}(\tau) = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} x(t_i)x(t_i+\tau) \tag{4.31}$$

where $x(t_i)$ is the value of $x(t)$ on the ith trial, or

$$\begin{aligned}\tilde{R}_{xx}(\tau) &= \widetilde{x(t)x(t+\tau)} = \overline{XY} \\ &= \iint \alpha\beta f_{XY}(\alpha, \beta)\, d\alpha\, d\beta \\ &\quad \text{or} \\ &= \sum\sum \alpha_i\beta_j p_{XY}(\alpha_i, \beta_j)\end{aligned} \tag{4.32}$$

For a discrete FPW $x(n)$, X and Y describe uniformly sampling $x(n)$ at n and $n+k$ and observing $X = x(n)$ and $Y = x(n+k)$. For N trials as $N \to \infty$,

$$\tilde{R}_{xx}(k) = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} x(n_i)x(n_i+k) \tag{4.33}$$

where $x(n_i)$ is the value of $x(n)$ on the ith trial, or

$$\begin{aligned}\tilde{R}_{xx}(k) &= \widetilde{x(n)x(n+k)} = \overline{XY} \\ &= \iint \alpha\beta f_{xy}(\alpha, \beta)\, d\alpha\, d\beta \\ &\quad \text{or} \\ &= \sum\sum \alpha_i\beta_j p_{XY}(\alpha_i, \beta_j)\end{aligned} \tag{4.34}$$

To clarify the relative frequency notation of Eq. (4.31) or (4.33), for example, on the sixth trial, t is chosen uniformly to obtain t_6 and $x(t_6)x(t_6+\tau)$ and n is chosen uniformly to obtain n_6 and $x(n_6)x(n_6+\mathrm{k})$. We also note Eqs. (4.32) and (4.34) are identical.

Two illustrative problems will be solved.

Problem 13

Consider the discrete random waveform

$$x(n) = \sum_k A_k g(n-2k)$$

where $g(n) = 3\delta(n) - \delta(n-1)$ and the As are independent random variables, each with mass function $p_{A_i}(1) = p_{A_i}(-1) = 0.5$. Find $\tilde{R}_{xx}(k)$.

Solution

A typical section of $x(n)$ is shown in Fig. 4.14(a). Using formula (4.34) for $\tilde{R}_{xx}(k)$, $\tilde{R}_{xx}(0) = [(3)^2 + (-3)^2 + (1)^2 + (-1)^2]\frac{1}{4} = 5.0$, $\tilde{R}_{xx}(\pm 1) = [3(-1) + (-3)(1)]\frac{1}{4} + [(1)(-3) + (1)(3)]\frac{1}{8}$ + other terms adding to zero $= -1.5$, and $\tilde{R}_{xx}(k) = 0$ for all $|k| > 1$. $\tilde{R}_{xx}(k)$ is shown plotted in Fig. 4.14(b).

Problem 14

Consider the continuous random waveform

$$x(t) = \sum_n A_n g(t-n)$$

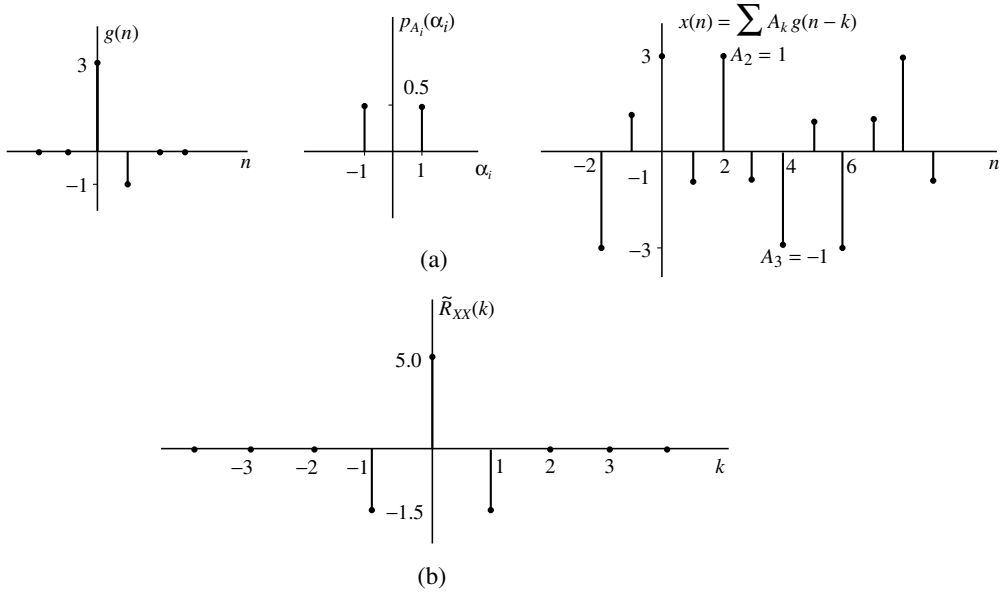


FIGURE 4.14 (a) A typical section of $x(n)$ for Problem 13 and (b) $\tilde{R}_{XX} = \widetilde{x(n)x(n+k)} = \overline{XY}$.

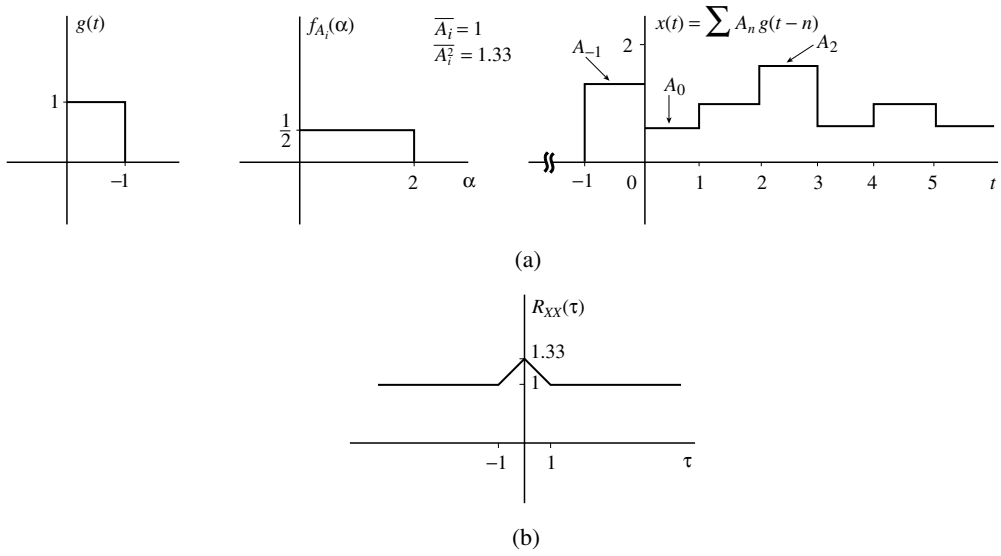


FIGURE 4.15 (a) A typical section of $x(t)$ for Problem 14 and (b) $\tilde{R}_{XX}(\tau)$.

where $g(t) = 1$, $0 \le t \le 1$; $g(t) = 0$ otherwise and the As are independent random variables with $f_{A_i}(\alpha) = \frac{1}{2}$, $0 \le \alpha \le 2$, $f_{A_i}(\alpha) = 0$ otherwise. Find $\tilde{R}_{xx}(\tau)$.

Solution

A typical section of $x(t)$ is shown in Fig. 4.15(a). Since the As are independent, $\overline{A_i} = 1$, $\overline{A_i^2} = 1.33$, and $\overline{A_iA_j} = 1$ for all $i \ne j$. Observing Fig. 4.15(a), we obtain the following.

For $0 < \tau < 1$:

$$\begin{aligned}\tilde{R}_{xx}(\tau) &= \overline{A_i^2}(1-\tau) + \overline{A_iA_{i+1}}\,\tau \\ &= 1.33 - 0.33\tau\end{aligned}$$

For $\tau > 1$:

$$\tilde{R}_{xx}(\tau) = \overline{A_i A_j} = 1$$

and $\tilde{R}_{xx}(-\tau) = \tilde{R}_{xx}(\tau)$ by symmetry. $\tilde{R}_{xx}(\tau)$ is shown plotted in Fig. 4.15(b). We also note that the same result is obtained when the As are independent and $f_{A_i}(\alpha)$ or $p_{A_i}(\alpha)$ are such $\overline{A_i} = 1$ and $\overline{A_i^2} = 1.33$. Thus, many different waveforms have the same autocorrelation function.

4.5 Summary and Future Study

The elements of probability and random variable theory were presented as a prerequisite to the study of random processes and communication theory. Highlighted were the concepts of finding first- and second-order time averages statistically, by uniformly sampling a continuous random finite power waveform at t or at t and $t + \tau$ or discretely uniformly sampling a random finite power discrete waveform at n or at n and $n + k$. The illustrative problems chosen were simple. When ergodic random processes are later encountered, the time averages for any one member of the ensemble will be equivalent to corresponding ensemble averages. A future task is to approximate time averages on a computer. Here, from a section of a waveform T with its values sampled every τ seconds, an estimate for the autocorrelation function and its Fourier transom $S_{XX}(f)$ called the power spectral density will be found.

Defining Terms

Cumulative distribution, density and mass function: $F_X(\alpha) = P[X \le \alpha]$, $f_X(\alpha) = \mathrm{d}/\mathrm{d}\alpha\, F_X(\alpha)$, $p_X(\alpha_i) = P[X = \alpha_i]$. Other noted distribution or density functions are: total signal energy or total signal power vs. frequency, corresponding to $F_X(\alpha)$; energy or power spectral density vs. frequency, corresponding to $f_X(\alpha)$; total signal power vs. frequency for a periodic waveform, corresponding to $F_X(\alpha)$ for a discrete random variable.

Joint distribution, density and mass function: $F_{XY}(\alpha, \beta) = P[(X \le \alpha) \cap (Y \le \beta)]$, $f_{XY}(\alpha, \beta) = \partial^2 / \partial\alpha\, \partial\beta\, F_{XY}(\alpha, \beta)$, $p_{XY}(\alpha_i, \beta_j) = P[(X = \alpha_i) \cap (Y = \beta_j)]$.

Joint random variables *X* and *Y*: $X(s_i)$ and $Y(s_i)$ jointly map each outcome of the SDS onto an α–β plane.

Probability: $P(A)$ is the statistical regularity of A, $P(A/B)$ is the statistical regularity of A based only on trials when B occurs, axiomatically $P(A) = \sum_{i=1}^{n} p(s_i)$ if $A = \bigcup_{i=1}^{n} s_i$ and $s_i s_j = 0$, $i \ne j$. $P(A/B) = P(AB)/P(B)$ if $P(B) \ne 0$.

Random phenomenon: An experiment which, when performed many times N, yields outcomes that occur with statistical regularity as $N \to \infty$.

Random variable *X*: $X(s_i)$ maps the outcomes of the SDS onto a real axis.

Sample description space (SDS): A listing of mutually exclusive, collectively exhaustive, finest grain outcomes.

Special Terminology for Random Finite Power Waveforms

Finite power waveform:

$$P_{\text{av}} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} x^2(t)\,\mathrm{d}t$$

or

$$= \lim_{N\to\infty} \frac{1}{2N+1}\sum_{n=-N}^{N} x^2(n) \qquad \text{exists.}$$

First-order time averages:

$$\widetilde{g(x(t))} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} g(x(t))\,\mathrm{d}t = \int_{-\infty}^{\infty} g(\alpha) fx(\alpha)\,\mathrm{d}\alpha$$

and

$$\widetilde{g(x(n))} = \lim_{N\to\infty} \frac{1}{2N+1}\sum_{-N}^{N} g(x(n)) = \int_{-\infty}^{\infty} g(\alpha) fx(\alpha)\,\mathrm{d}\alpha$$

The most noted first-order statistics are $\widetilde{x(t)}$ or $\widetilde{x(n)} = \overline{X}$, $\widetilde{x^2(t)}$ or $\widetilde{x^2(n)} = \overline{X^2}$ and $\sigma_X^2 = \overline{(X-\overline{X})^2}$.

Jointly sampling a waveform: Uniformly choosing t and observing $X = x(t)$ and $Y = x(t+\tau)$ or discretely uniformly choosing n and observing $X = x(n)$ and $Y = x(n+k)$.

Sampling a waveform: For a continuous waveform uniformly choosing a value of time over a long time T and observing $X = x(t)$; for a discrete waveform discretely uniformly choosing n, $-N \le n \le N$ over a long time $2N+1$ and observing $x(n)$. The random variable X describes $x(t)$ or $x(n)$.

Second-order time averages:

$$\widetilde{g[x(t),\, x(t+\tau)]} = \overline{g(X,Y)} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} g[x(t),\, x(t+\tau)]\,\mathrm{d}t$$

$$\widetilde{g[x(n),\, x(n+k)]} = \overline{g(X,Y)} = \lim_{N\to\infty} \frac{1}{2N+1}\sum_{N} g[x(n),\, x(n+k)]$$

and statistically

$$\overline{g(X,\,Y)} = \iint g(\alpha,\,\beta) f_{XY}(\alpha,\,\beta)\,\mathrm{d}\alpha\,\mathrm{d}\beta$$

or

$$= \sum\sum g(\alpha_i,\,\beta_j) p_{XY}(\alpha_i,\,\beta_j)$$

The most noted second-order statistics are $\tilde{R}_{xx}(\tau)$ or $\tilde{R}_{xx}(k) = \overline{XY}$, the time autocorrelation function, and/or $\tilde{L}_{xx}(\tau)$ or $\tilde{L}_{xx}(k) = \overline{(X-\overline{X})(Y-\overline{Y})} = \overline{XY} - \overline{X}\,\overline{Y}$, the covariance function.

References

Olkin, Gleser, and Derman. 1994. *Probability Models and Applications,* Macmillan, New York.
O'Flynn, M. 1982. *Probabilities, Random Variables and Random Processes,* Wiley, New York.
Papoulis, A. 1965. *Probability, Random Variables and Stochastic Processes,* McGraw-Hill, New York.
Parzen, E. 1960. *Modern Probability Theory and Its Applications,* Wiley, New York.
Ross, S. 1994. *A First Course in Probability,* Macmillan, New York.

Further Information

Some important topics were omitted or only briefly mentioned here. The bivariate Gaussian or normal joint density function is given an excellent treatment in Olkin, Gleser, and Derman [1994, Chap. 12]. The topic of finding the density function of a function of two known random variables can be extended. The case of finding the joint density function for U and V, each defined as a function of two known random variables X and Y, is well handled in Olkin, Gleser, and Derman [1994, Chap. 13]. Finally, the most general transformation problem of finding the joint density function of m random variables, each defined as a linear combination of n known jointly Gaussly random variables, is given in Chap. 9 of O'Flynn [1982].

5

Random Processes, Autocorrelation, and Spectral Densities

Lew E. Franks
University of Massachusetts

5.1 Introduction

The modeling of signals and noise in communication systems as random processes is a well-established and proven tradition. Much of the system performance characterization and evaluation can be done in terms of a few low-order statistical moments of these processes. Hence, a system designer often needs to be able to evaluate mean values and the **autocorrelation** of a process and the **cross correlation** of two processes. It is especially true in communication system applications that a frequency-domain-based evaluation is often easier to perform and interpret. Therefore, a knowledge of spectral properties of random processes is valuable.

More recently, engineers are finding that modeling of typical communication signals as **cyclostationary** (CS) processes, rather than stationary processes, more accurately captures their true nature and permits the design of more effective signal processing operations. In the book edited by Gardner [1994], several authors describe applications that have benefited from the exploitation of cyclostationarity. Although these processes are nonstationary, frequency-domain methods prove to be very effective, as in the stationary case.

Another notion that is becoming well accepted by communication engineers is the utility of complex-valued representations of signals. Accordingly, all of the results and relations presented here will be valid for complex-valued random processes.

5.2 Basic Definitions

We consider the random process, $x(t)$, defined as a set of jointly distributed random variables $\{x(t)\}$ indexed by a time parameter t. If the set of values of t is the real line, we say the process is a *continuous-time* process. If the set of values is the integers, we call the process a *discrete-time* process. The *mean* and *autocorrelation* of the process are given by the expected values $E[x(t)]$ and $E[x(t_1)x*(t_2)]$, respectively.

In dealing with cyclostationary processes, it is often more convenient to define an autocorrelation function in terms of a simple transformation on the t_1 and t_2 variables; namely, let $t = (t_1 + t_2)/2$ and $\tau = t_1 - t_2$ denote the mean value and difference of t_1 and t_2. Accordingly, we define the autocorrelation function for the process as

$$R_{xx}(t, \tau) = E[x(t + \tau/2)x^*(t - \tau/2)] \tag{5.1}$$

A process is *cyclostationary*, in the wide sense, with period T if both the mean and autocorrelation are periodic in t with period T. Hence, for a cyclostationary process CS(T), we can write*

$$m_x(t) = E[x(t)] = \sum_m M_x(m)e^{j2\pi mt/T} \tag{5.2}$$

$$\begin{aligned} R_{xx}(t, \tau) &= \sum_k Tc_{xx}(t - kT, \tau) \\ &= \sum_m \tilde{c}_{xx}\left(\frac{m}{T}, \tau\right)e^{j2\pi mt/T} \end{aligned} \tag{5.3}$$

Equation (5.3) results from an application of the **Poisson sum formula** (see Appendix), where $\tilde{c}_{xx}(\nu, \tau)$ is the Fourier transform of $c_{xx}(t, \tau)$ with respect to the t variable; that is,

$$\tilde{c}_{xx}(\nu, \tau) = \int c_{xx}(t, \tau)e^{-j2\pi\nu t}\,\mathrm{d}t \tag{5.4}$$

hence,

$$\tilde{c}_{xx}(m/T, \tau) = \frac{1}{T}\int_0^T R_{xx}(t, \tau)e^{-j2\pi mt/T}\,\mathrm{d}t \tag{5.5}$$

is the mth Fourier coefficient for the t-variation in $R_{xx}(t, \tau)$. It is called the **cyclic autocorrelation function** for $x(t)$ with **cycle frequency** m/T [Gardner, 1989]. Note that the $c_{xx}(t, \tau)$ function in Eq. (5.3) is not uniquely determined by the autocorrelation; only a discrete set of values of its Fourier transform are specified, as indicated in Eq. (5.3). Nevertheless, the notation is useful because the $c_{xx}(t, \tau)$ function arises naturally in modeling typical communication signal time-sequential formats that use sampling, scanning, or multiplexing. The **pulse amplitude modulation** (**PAM**) signal discussed in Section 5.4 is a good example of this.

A frequency-domain characterization for the CS(T) process is given by the double Fourier transform of $R_{xx}(t, \tau)$.

$$\begin{aligned} S_{xx}(\nu, f) &= \iint R_{xx}(t, \tau)e^{-j2\pi(\nu t + f\tau)}\,\mathrm{d}t\,\mathrm{d}\tau \\ &= \sum_m C_{xx}\left(\frac{m}{T}, f\right)\delta\left(\nu - \frac{m}{T}\right) \end{aligned} \tag{5.6}$$

where $C_{xx}(\nu, f)$ is the double Fourier transform of $c_{xx}(t, \tau)$; hence, it is the Fourier transform, with respect to the τ-variation, of the cyclic autocorrelation. It is called the *cyclic spectral density* for the cycle frequency m/T. Notice that Eq. (5.6) is a set of *impulse fences* located at discrete values of ν in the dual-frequency (ν, f) plane. The impulse fence at $\nu = 0$ has a special significance as the (one-dimensional) **power spectral**

*All summation indices will run from $-\infty$ to ∞ unless otherwise indicated. The same is true for integration limits.

density of the process. The other impulse fences characterize the correlation between signal components in different spectral regions. The physical significance of these concepts is discussed in the next section.

A process is *wide-sense stationary* (WSS) if the mean and autocorrelation are independent of the t variable, that is, if only the $m = 0$ terms in Eqs. (5.2) and (5.3) are nonzero. For the WSS process, we write the autocorrelation function as

$$R_{xx}(\tau) = R_{xx}(0, \tau) = \tilde{c}_{xx}(0, \tau) \tag{5.7}$$

since the cyclic autocorrelation is nonzero only for a cycle frequency of $\nu = 0$. The *power spectral density* of the process is simply the Fourier transform of $R_{xx}(\tau)$:

$$S_{xx}(f) = \int R_{xx}(\tau) e^{-j2\pi f\tau}\, \mathrm{d}\tau \tag{5.8}$$

A CS(T) process can be converted to a WSS process by phase randomization. Let $x'(t) = x(t - \theta)$ where the time shift θ is a random variable which is independent of the other parameters of the process. Then the mean and autocorrelation of the new process are obtained by convolution, with respect to the t variable, with the probability density function $p_\theta(\cdot)$ for the θ variable. Letting $p_\theta(\cdot)$ denote the Fourier transform of $p_\theta(\cdot)$, we have

$$m_{x'}(t) = \sum_m P_\theta\left(\frac{m}{T}\right) M_x(m) e^{j2\pi mt/T}; \quad P_\theta(0) = 1 \tag{5.9a}$$

$$R_{x'x'}(t, \tau) = \sum_m P_\theta\left(\frac{m}{T}\right) \tilde{c}_{xx}\left(\frac{m}{T}, \tau\right) e^{j2\pi mt/T} \tag{5.9b}$$

which indicates that phase randomization tends to suppress the cyclic components with cycle frequencies different from zero. In fact, for a uniform distribution of θ over a T-second interval, only the $m = 0$ terms in Eq. (5.9) remain; hence, $x'(t)$ is WSS. Using the Poisson sum formula, we see that any $p_\theta(\sigma)$ satisfying $\sum_k p_\theta(\sigma - kT) = 1/T$ makes $P_\theta(m/T) = 0$ for $m \neq 0$, hence making $x'(t)$ WSS.

It is important to note [from Eq. (5.5)] that the autocorrelation of the phase-randomized process is the same as the time average of the autocorrelation of the original CS(T) process [Franks, 1969].

For a WSS discrete-time process $a(k)$, where k is an integer, we write the mean and autocorrelation as

$$\bar{a} = E[a(k)] \tag{5.10a}$$

$$R_{aa}(m) = E[a(k+m)a^*(k)] \tag{5.10b}$$

The power spectral density of the discrete-time process is the *discrete-time Fourier transform* (DTFT) of the autocorrelation function; that is,

$$S_{aa}(f) = \sum_m R_{aa}(m) e^{-j2\pi mTf} \tag{5.11}$$

which is always a periodic function of f with a period of $1/T$. A case of frequent interest is when the $a(k)$ are uniformly spaced sample values of a continuous-time WSS random process. Suppose $a(k) = y(kT)$, then using the alternate form of the Poisson sum formula (Appendix), we have $R_{aa}(m) = R_{yy}(mT)$ and

$$S_{aa}(f) = \sum_m R_{yy}(mT) e^{-j2\pi mTf} = \frac{1}{T} \sum_n S_{yy}\left(f - \frac{n}{T}\right) \tag{5.12}$$

In communication system applications, the $a(k)$ sequence usually represents digital information to be transmitted, and the WSS assumption is reasonable. In some cases, such as when synchronizing pulses or other header information are added to the data stream, the $a(k)$ sequence is more appropriately modeled as a cyclostationary process also. The examples we consider here, however, will assume that $a(k)$ is a WSS process

5.3 Properties and Interpretation

Many interesting features of a random process are revealed when the effects of a linear transformation on the first- and second-moment statistics are examined. For this purpose, we use linear transformations which are a combination of frequency shifting and time-invariant filtering, as shown in Fig. 5.1. The exponential multipliers (modulators) provide the frequency shifts, and the filters with transfer functions $H_1(f)$ and $H_2(f)$ provide frequency selective filtering. If $H_1(f)$ and $H_2(f)$ are narrowband low-pass filters, each path of the structure in Fig. 5.1 can be regarded as a sort of heterodyne technique for spectrum analysis. The cross-correlation between the outputs of the two paths provides further useful characterization of the input process.

The cross-correlation function for a pair of jointly random complex processes is defined as

$$R_{yz}(t, \tau) = E[y(t+\tau/2)z^*(t-\tau/2)] \tag{5.13}$$

Substituting the indicated expressions for $y(t)$ and $z(t)$ in Fig. 5.1 and using Eq. (5.3) for the autocorrelation of the $x(t)$ process, we can derive the following general expression for the cross-correlation:

$$\begin{aligned} R_{yz}(t, \tau) = \sum_m e^{j\pi(f_1+f_2)\tau} \cdot e^{j2\pi(f_1-f_2)\tau} \cdot e^{j2\pi mt/T} \cdot \int H_1\left(f+f_1+\frac{m}{2T}\right) \\ \times H_2^*\left(f+f_2-\frac{m}{2T}\right) C_{xx}\left(\frac{m}{T}, f\right) e^{j2\pi f\tau}\, \mathrm{d}f \end{aligned} \tag{5.14}$$

First, we consider the case where $H_1(f) = H_2(f) = B(f)$, which is a very narrowband low-pass analysis filter. Let $B(0) = 1$ and let Δ denote the *noise bandwidth* of the filter given by

$$\Delta = \int \left|\frac{B(f)}{B(0)}\right|^2 \mathrm{d}f \tag{5.15}$$

Hence, an ideal low-pass filter with a rectangular gain shape and a cutoff frequency at $f = \Delta/2$ has a noise bandwidth of Δ. On the other hand, a low-pass filter resulting from averaging a signal over the last T_0 s, that is, one with an impulse response $b(t) = 1/T_0$ in $0 < t < T_0$ and zero elsewhere, has a noise bandwidth

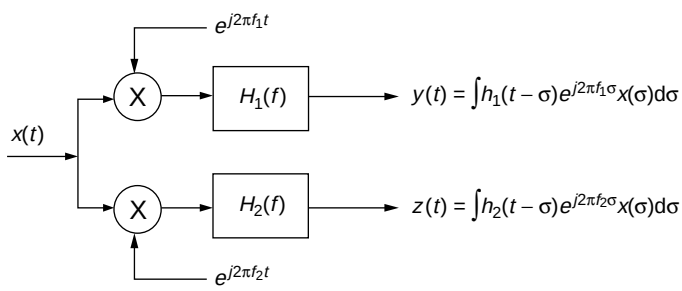


FIGURE 5.1 Conceptual apparatus for analysis of random processes.

of $\Delta = 1/T_0$. Whatever kind of analysis filter we want to use, we shall consider the results as Δ approaches zero so that we can assume that factors of the form $B(f + m/2T)B^*(f - m/2T)$ which will appear in Eq. (5.14) will vanish except for $m = 0$.

For example, let $f_1 = f_2 = -f_0$. Then we see that both paths in the figure are identical, so that $z(t) = y(t)$ and

$$R_{yy}(t, \tau) = e^{-j2\pi f_0 \tau} \int |B(f - f_0)|^2 C_{xx}(0, f) e^{j2\pi f \tau} \, \mathrm{d}f \tag{5.16}$$

because only the $m = 0$ term in Eq. (5.14) is nonzero under these conditions. Note also that $y(t)$ is WSS, and its mean-squared value (also called the power of the process) is given by

$$\begin{aligned} E[|y(t)|^2] &= R_{yy}(0) = \int |B(f - f_0)|^2 C_{xx}(0, f) \, \mathrm{d}f \\ &\cong \Delta C_{xx}(0, f_0) = \Delta S_{xx}(f_0) \qquad \text{as } \Delta \to 0 \end{aligned} \tag{5.17}$$

This expression, valid for small Δ, gives the power in $x(t)$ due to components in a narrow band centered at a frequency of f_0. This provides the motivation for calling $S_{xx}(f) = C_{xx}(0, f)$ the power spectral density of the $x(t)$ process. Note that our spectrum analyzer gives the same results for a CS(T) process and its phase-randomized WSS version because of the time averaging inherent in the $B(f)$ analysis filter.

Now, keeping the same low-pass filters, $H_1(f) = H_2(f) = B(f)$, let $f_0 = -(f_1 + f_2)/2$ and $n/T = (f_2 - f_1)$, so that the mean value of the two analysis frequencies is f_0 and their difference is an integer multiple of $1/T$. Then, only the $m = n$ term in Eq. (5.14) is nonzero and we have the following expression for the cross correlation of the $y(t)$ and $z(t)$ processes:

$$\begin{aligned} R_{yz}(t, 0) &= \int |B(f - f_0)|^2 C_{xx}\left(\frac{n}{T}, f\right) \mathrm{d}f \\ &\cong \Delta C_{xx}(n/T, f_0) \end{aligned} \tag{5.18}$$

for small Δ. Thus, a physical significance of the cyclic spectral density, defined implicity in Eq. (5.6), is that it represents the correlation that exists between spectral components separated by the cycle frequency. Note that no such spectral correlation exists for a WSS process. This spectral correlation is used to advantage in communication systems to extract timing or synchronization information or to eliminate certain kinds of interference by means of periodically time-varying filtering of signals [Gardner, 1994].

Equation (5.14) is useful to establish many other useful relationships. For example, let $x(t)$ be WSS and let $f_1 = f_2 = 0$, then $y(t)$ and $z(t)$ are jointly WSS and the cross-spectral density is

$$S_{yz}(f) = H_1(f) H_2^*(f) S_{xx}(f) \tag{5.19}$$

and so, if $H_1(f) = H_2(f)$, then $z(t) = y(t)$ and the power spectral density of a filtered WSS process is given by

$$S_{yy}(f) = |H_1(f)|^2 S_{xx}(f) \tag{5.20}$$

whereas letting $H_2(f) = 1$ makes $z(t) = x(t)$ so that the cross-spectral density of the input and output of a filtered WSS process is

$$S_{yx}(f) = H_1(f) S_{xx}(f) \tag{5.21}$$

If $H_1(f) H_2^*(f) = 0$ for all f, then $S_{yz}(f) = 0$, indicating that spectral components in nonoverlapping frequency bands are always uncorrelated. This is a fundamental property of all WSS processes.

5.4 Baseband Digital Data Signals

Perhaps the most basic form of mapping from a discrete-time data sequence $\{a(k)\}$ into a continuous-time signal $x(t)$ for transmission over a communication channel is baseband pulse amplitude modulation. The mapping uses a single carrier pulse shape, $g(t)$, and is a linear function of the data, which we will assume is a WSS discrete-time process [Eq. (5.10)].

$$x(t) = \sum_k a(k)g(t-kT) \tag{5.22}$$

The $x(t)$ process is CS(T) with

$$m_x(t) = \bar{a}\sum_k g(t-kT) = (\bar{a}/T)\sum_m G\left(\frac{m}{T}\right)e^{j2\pi mt/T} \tag{5.23a}$$

$$R_{xx}(t,\tau) = \sum_k\sum_m R_{aa}(m)g\left(t+\frac{\tau}{2}-kT-mT\right)g^*\left(t-\frac{\tau}{2}-kT\right) \tag{5.23b}$$

By inspection, the $c_{xx}(t,\tau)$ function in Eq. (5.3) is

$$c_{xx}(t,\tau) = \frac{1}{T}\sum_m R_{aa}(m)g\left(t+\frac{\tau}{2}-mT\right)g^*\left(t-\frac{\tau}{2}\right) \tag{5.24}$$

so that the cyclic autocorrelation is

$$\tilde{c}_{xx}(n/T,\tau) = \frac{1}{T}\int S_{aa}(f)G\left(\frac{n}{T}+f\right)G^*(f)e^{j2\pi f\tau}\,\mathrm{d}f\cdot e^{j\pi n\tau/T} \tag{5.25}$$

Note that the strength of the cyclic autocorrelations (equivalently, the amount of spectral correlation) depends strongly on the amount of overlap of $G(f)$ and its frequency-translated versions. Data carrier pulses are often sharply bandlimited, with **excess bandwidth** factors not exceeding 100%. A pulse with 100% excess bandwidth is bandlimited to the frequency interval $|f| < 1/T$. In this case, $x(t)$ has cycle frequencies of 0 and $\pm 1/T$ only. Note also that an excess bandwidth of less than 100% means that $G(n/T) = 0$ for $n \neq 0$ so that the mean of the PAM signal [Eq. (5.23a)] is constant.

Using time averaging to get the power spectrum of the PAM signal, we can write

$$R_{xx}(\tau) = \frac{1}{T}\sum_m R_{aa}(m)r_g(\tau-mT) \tag{5.26}$$

where

$$r_g(\tau) = \int g(t+\tau)g^*(t)\,\mathrm{d}t \tag{5.27}$$

is the **time-ambiguity function** for the data pulse, $g(t)$ [Franks, 1969]. Its Fourier transform is $|G(f)|^2$, called the *energy spectral density* of the pulse. Hence, taking the Fourier transform of Eq. (5.26), the power spectral density of the PAM signal is

$$S_{xx}(f) = \frac{1}{T}S_{aa}(f)|G(f)|^2 \tag{5.28}$$

which nicely separates the influence of the pulse shape from the correlation structure of the data.

Although it may not be apparent from Eq. (5.28), the power spectrum expression may contain δ-function terms, which represent nonzero concentrations of power at a discrete set of frequencies. In order to make evident the discrete part of the power spectrum, we express the continuous part of the spectrum as the Fourier transform of the autocovariance function, which is the autocorrelation of the zero-mean process, $x'(t) = x(t) - m_x(t)$. Then

$$R_{xx}(t, \tau) = R'_{xx}(t, \tau) + m_x(t + \tau/2)m_x^*(t - \tau/2) \tag{5.29}$$

where

$$R'_{xx}(t, \tau) = R_{x'x'}(t, \tau) = \frac{1}{T}\sum_m R'_{aa}(m)r_g(\tau - mT) \tag{5.30}$$

and

$$R'_{aa}(m) = R_{aa}(m) - |\bar{a}|^2 \tag{5.31}$$

Now, performing the time averaging, the power spectrum is expressed as

$$S_{xx}(f) = \frac{1}{T}S'_{aa}(f)|G(f)|^2 + \left|\frac{\bar{a}}{T}\right|^2 \sum_n \left|G\left(\frac{n}{T}\right)\right|^2 \delta\left(f - \frac{n}{T}\right) \tag{5.32}$$

The second term in Eq. (5.32) is the discrete part of the spectrum. Notice that there are no discrete components for zero-mean data ($\bar{a} = 0$) or if the data pulse has less than 100% excess bandwidth ($G(n/T) = 0$ for $n \neq 0$). For zero-mean independent data, pulse amplitudes are uncorrelated and the power spectrum is simply

$$S_{xx}(f) = \frac{\sigma_a^2}{T}|G(f)|^2 \tag{5.33}$$

where σ_a^2 is the data variance.

A more general format for baseband data signals is to use a separate waveform for each data symbol from the symbol alphabet. Familiar examples are *frequency-shift keying* (FSK), *phase-shift keying* (PSK), and *pulse-position modulation* (PPM). For the binary alphabet case, $[a(k) \in \{0, 1\}]$, the format is basically PAM. Using waveforms $g_0(t)$ and $g_1(t)$ to transmit a zero or one, respectively, we have

$$x(t) = \sum_k [1 - a(k)]g_0(t - kT) + a(k)g_1(t - kT) \tag{5.34}$$

which can be rewritten, using $d(t) = g_1(t) - g_0(t)$, as

$$x(t) = \sum_k g_0(t - kT) + \sum_k a(k)d(t - kT) \tag{5.35}$$

which is a periodic fixed term plus a binary PAM signal. Let $p = Pr[a(k) = 1] = \bar{a}$ and let $P_{11}(m)$ denote the conditional probability, $Pr[a(k + m) = 1|a(k) = 1]$, then with $a'(k) = a(k) - \bar{a}$,

$$\begin{aligned} R'_{aa}(m) &= E[a'(k + m)a'(k)] \\ &= R_{aa}(m) - p^2 = pP_{11}(m) - p^2 \end{aligned} \tag{5.36}$$

and carrying out the calculations, the power spectral density is

$$S_{xx}(f) = \frac{1}{T} s'_{aa}(f)\left|G_1(f) - G_0(f)\right|^2 + \left(\frac{1}{T}\right)^2 \sum_n \left|(1-p)G_0\left(\frac{n}{T}\right) + pG_1\left(\frac{n}{T}\right)\right|^2 \delta\left(f - \frac{n}{T}\right) \tag{5.37}$$

where the second term gives the discrete components. Note that the strength of the discrete components depends only on the mean value of the data and not the correlation between data symbols.

To generalize the preceding result to an M-ary symbol alphabet for the data, we write

$$x(t) = \sum_k \sum_{m=0}^{M-1} a_m(k) g_m(t - kT) \tag{5.38}$$

with $a_m(k) \in \{0, 1\}$ and, for each k, only one of the $a_m(k)$ is nonzero for each realization of the process. Then, calculating the autocorrelation function, we find that $c_{xx}(t, \tau)$ in Eq. (5.3) can be expressed as

$$c_{xx}(t, \tau) = \frac{1}{T} \sum_i \sum_{m,n=0}^{M-1} E[a_m(i+j)a_n(i)] g_m\left(t + \frac{\tau}{2} - iT\right) g_n^*\left(t - \frac{\tau}{2}\right) \tag{5.39}$$

We assume that the data process can be modeled as a *Markov chain* characterized by a probability transition matrix $P_{mn}(i)$ giving the probability that $a_m(i+j) = 1$ under the condition that $a_n(j) = 1$. We further assume the process is in a steady-state condition so that it is WSS. This means that the symbol probabilities, $p_n = Pr[a_n(j) = 1]$, must satisfy the homogeneous equation

$$p_m = \sum_{n=0}^{M-1} P_{mn}(1) p_n \tag{5.40}$$

and all of the transition matrices can be generated by recursion:

$$P_{mn}(i+1) = \sum_{k=0}^{M-1} P_{mk}(i) P_{kn}(1); \quad i \geq 1 \tag{5.41}$$

Now, calculation of the autocorrelation can proceed using $E[a_m(i+j)a_n(j)] = q_{mn}(i)$ in Eq. (5.39), where

$$\begin{aligned} q_{mn}(i) &= Pr[a_m(i+j) = 1 \text{ and } a_n(j) = 1] \\ &= P_{mn}(i)p_n && \text{for } i > 0 \\ &= \delta_{mn} p_n && \text{for } i = 0 \\ &= P_{mn}(-i)p_m && \text{for } i < 0 \end{aligned} \tag{5.42}$$

Taking the DTFT of the $q_{mn}(i)$ sequence

$$Q_{mn}(f) = \sum_i q_{mn}(i) e^{-j2\pi iTf} \tag{5.43}$$

we arrive at a remarkably compact expression for the power spectral density of this general form for the baseband data signal

$$S_{xx}(f) = \frac{1}{T} \sum_{m,n=0}^{M-1} Q_{mn}(f) G_m(f) G_n^*(f) \tag{5.44}$$

The power spectrum may contain discrete components. The continuous part of the spectrum can be extracted by replacing the $q_{mn}(i)$ in Eqs. (5.42–5.44) by $q_{mn}(i) - p_m p_n$. Then, the remaining part is the discrete spectrum given by

$$S_{xx}(f)_{\text{discrete}} = \left(\frac{1}{T}\right)^2 \sum_k \left| \sum_{m=0}^{M-1} p_m G_m\left(\frac{k}{T}\right) \right|^2 \delta\left(f - \frac{k}{T}\right) \tag{5.45}$$

For the special case of independent data from symbol to symbol, the probability transition matrix takes on the special form $P_{mn}(i) = p_m$ for each m and each $i > 1$; that is, all columns are identical and consist of the set of steady-state symbol probabilities. In this case, the continuous part of the power spectrum is

$$S_{xx}(f)_{\text{continuous}} = \frac{1}{T}\left[\sum_{m=0}^{M-1} p_m |G_m(f)|^2 - \left|\sum_{m=0}^{M-1} p_m G_m(f)\right|^2\right] \tag{5.46}$$

and the discrete part is the same as in Eq. (5.45).

As an example of signalling with M different waveforms and with independent, equiprobable data, consider the case of pulse position modulation where

$$g_m(t) = g\left(t - m\frac{T}{M}\right); \quad p_m = 1/M; \quad m = 0, 1, 2, \ldots, M-1 \tag{5.47}$$

The resulting power spectrum is

$$S_{xx}(f) = \left(\frac{1}{T}\right) Q(f)|G(f)|^2 + \left(\frac{1}{T}\right)^2 \sum_k \left| G\left(\frac{kM}{T}\right)\right|^2 \delta\left(f - \frac{kM}{T}\right) \tag{5.48}$$

where

$$Q(f) = \left[1 - \left(\frac{1}{M}\right)\frac{\sin^2(\pi T f)}{\sin^2(\pi T f/M)}\right]$$

The continuous part of the spectrum will be broad because pulse widths will be narrow (on the order of T/M s duration) for a reliable determination of which symbol was transmitted. Note, however, that the spectrum has nulls at multiples of $1/T$ due to the shape of the $Q(f)$ factor in Eq. (5.48).

5.5 Coding for Power Spectrum Control

From Eq. (5.28) or (5.44), we see that that the shape of the continuous part of the power spectral density of data signals can be controlled either by choice of pulse shapes or by manipulating the correlation structure of the data sequence. Often the latter method is more convenient to implement. This can be accomplished by mapping the data sequence into a new sequence by a procedure called *line coding* [Lee and Messerschmitt, 1988]. The simpler types of line codes are easily handled by Markov chain models. As an example, we consider the *differential binary* (also called NRZI) coding scheme applied to a binary data sequence. Let $a(k) \in \{0, 1\}$ represent the data sequence and $b(k)$ the coded binary sequence. The coding rule is that $b(k)$ differs from $b(k-1)$ if $a(k) = 1$, otherwise $b(k) = b(k-1)$. This can also be

expressed as $b(k) = [b(k-1) + a(k)]\text{mod}2$. Letting $q = Pr[a(k) = 1]$, the probability transition matrix can be expressed as

$$P(1) = \begin{bmatrix} 1-q & q \\ q & 1-q \end{bmatrix} \tag{5.49}$$

from which it follows that the steady-state probabilities for $b(k) = 1$ or 0 [Eq. (5.40)] are given by $p = 1 - p = 1/2$, regardless of the uncoded symbol probability q. Because the columns of a transition matrix sum to 1, the evolution of the transition probability [Eq. (5.41)] can be written as a single scalar equation which, in this case, is a first-order difference equation with an initial condition of $P_{11}(0) = 1$.

$$P_{11}(i+1) = (1-2q)P_{11}(i) + q; \quad i \geq 0 \tag{5.50}$$

Using $P_{11}(i) = P_{11}(-i)$, the solution is

$$P_{11}(i) = \frac{1}{2}[(1-2q)^{|i|} + 1] \tag{5.51}$$

Supposing that the transmitted signal is binary PAM using the $b(k)$ sequence, that is,

$$x(t) = \sum_k b(k)g(t-kT) \tag{5.52}$$

then, taking the DTFT of Eq. (5.51) and using Eq. (5.32), the power spectral density becomes [Franks, 1969]

$$S_{xx}(f) = \left(\frac{1}{4T}\right)Q(f)|G(f)|^2 + \left(\frac{1}{2T}\right)^2 \sum_n \left|G\left(\frac{n}{T}\right)\right|^2 \delta\left(f - \frac{n}{T}\right) \tag{5.53}$$

where the spectrum shaping function is

$$Q(f) = q(1-q)[q^2 + (1-2q)\sin^2(\pi T f)]^{-1} \tag{5.54}$$

It is sometimes stated that differential binary (NRZI) coding does not alter the shape of the power spectrum. We see from Eq. (5.54) that this is true only in the case that the original data are equiprobable ($q = 1/2$). As shown in Fig. 5.2, the shaping factor $Q(f)$ differs greatly from a constant as q approaches 0 or 1. Note, however, that the spectrum of the coded signal depends only on the symbol probability q and not the specific correlation structure of the original data sequence.

A more practical line coding scheme is *bipolar coding*, also called *alternate mark inversion* (AMI). It is easily derived from the differential binary code previously discussed. The bipolar line signal is a ternary PAM signal

$$x(t) = \sum_k c(k)g(t-kT) \tag{5.55}$$

where $c(k) = 0$ if $a(k) = 0$ and $c(k) = +1$ or -1 if $a(k) = 1$, with the two pulse polarities forming an alternating sequence; that is, a positive pulse will always be followed by a negative pulse and vice versa. Hence, the bipolar code can be expressed in terms of the differential binary code as

$$c(k) = b(k) - b(k-1) \tag{5.56}$$

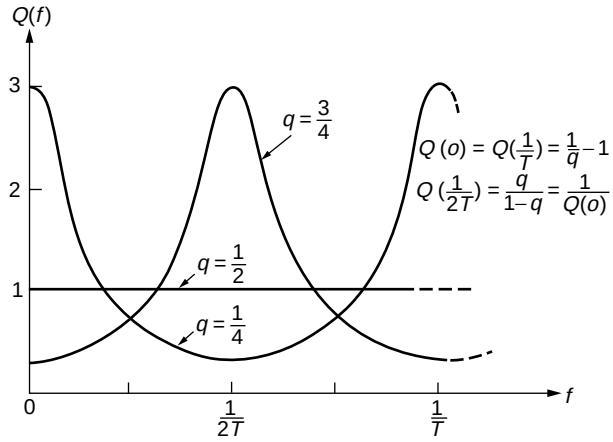


FIGURE 5.2 Spectral density shaping function for differential binary coding.

which itself is a form of *partial response coding* [Lee and Messerschmitt, 1988; Proakis and Salehi, 1994]. The autocorrelation function for $c(k)$ is easily expressed in terms of the autocorrelation for $b(k)$,

$$R_{cc}(m) = 2R_{bb}(m) - R_{bb}(m-1) - R_{bb}(m+1) \tag{5.57}$$

and since

$$R_{bb}(m) = \frac{1}{2}P_{11}(m) = \frac{1}{4}[(1-2q)^{|m|} + 1] \tag{5.58}$$

we get

$$\begin{aligned} R_{cc}(0) &= q \\ R_{cc}(m) &= -q^2(1-2q)^{|m-1|}; \quad |m| \geq 1 \end{aligned} \tag{5.59}$$

and, taking the DTFT of $R_{cc}(m)$, the power spectral density for the bipolar signal is

$$S_{xx}(f) = \frac{1}{T}Q(f)|G(f)|^2 \tag{5.60}$$

where

$$Q(f) = [q(1-q)\sin^2(\pi Tf)][q^2 + (1-2q)\sin^2(\pi Tf)]^{-1}$$

Since $E[c(k)] = \bar{c} = \bar{b} - \bar{b} = 0$, the spectrum has no discrete components, which can be a substantial practical advantage. The spectral shaping factor $Q(f)$, shown in Fig. 5.3, produces spectral nulls at $f = 0$ and $f = 1/T$. This was used to advantage in early wire-cable PCM systems to mitigate the effects of no low-frequency transmission and the effects of increased crosstalk between wire pairs at the higher frequencies. Note that the spectrum shape is still dependent on data symbol probability but not on the correlation structure of the data.

Another simple line code, found useful in optical and magnetic recording systems, is the *Miller code*, whose power spectrum can be determined with a four-state Markov chain [Proakis and Salehi, 1994].

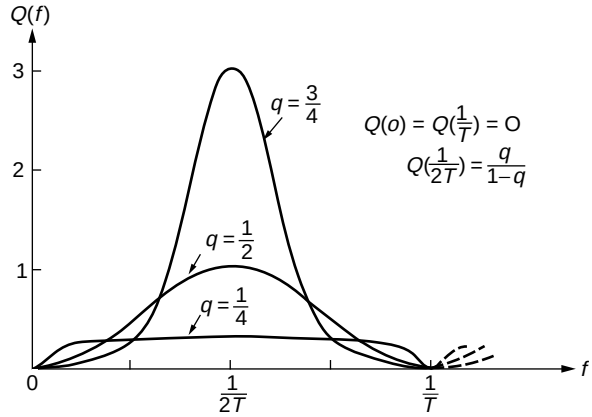


FIGURE 5.3 Spectral density shaping function for bipolar coding.

5.6 Bandpass Digital Data Signals

In most communication systems, signals are positioned in a frequency band which excludes low frequencies. These are called bandpass signals. For digital data signals, we could simply consider the baseband formats discussed previously, but with data pulse shapes, $g(t)$, which have a bandpass nature. However, it is more common to produce bandpass signals by modulation of a sinusoidal carrier with low-pass (baseband) data signals. Hence, we examine here the correlation and spectral properties of **quadrature amplitude modulation** (**QAM**) carrier signals. Let $u(t)$ and $v(t)$ be real baseband signal processes, then the QAM signal with a carrier frequency of f_0 is

$$x(t) = u(t)\cos(2\pi f_0 t) - v(t)\sin(2\pi f_0 t) \tag{5.61}$$

which can be more conveniently handled in terms of complex signals

$$x(t) = \Re[w(t)e^{j2\pi f_0 t}]; \qquad w(t) = u(t) + jv(t) \tag{5.62}$$

The signals $u(t)$ and $v(t)$ are often called the *in-phase* and *quadrature* components, respectively, of the QAM signal. The autocorrelation function for $x(t)$ becomes

$$\begin{aligned} R_{xx}(t, \tau) &= \frac{1}{2}\Re[R_{ww}(t, \tau)e^{j2\pi f_0 \tau}] \\ &\quad + \frac{1}{2}\Re[R_{ww^*}(t, \tau)e^{j2\pi(2f_0)t}] \end{aligned} \tag{5.63}$$

First we consider the situation where $u(t)$ and $v(t)$ are jointly WSS processes, which might be a good model for some types of analog systems. In this case, there are three cycle frequencies $(0, +2f_0, -2f_0)$. We obtain the cyclic autocorrelation functions from Eq. (5.5) by time averaging over a period of $T = 1/f_0$ with the result that

$$\tilde{c}_{xx}(0, \tau) = \frac{1}{2}\Re[R_{ww}(\tau)e^{j2\pi f_0 \tau}] \tag{5.64a}$$

$$\tilde{c}_{xx}(2f_0, \tau) = \frac{1}{4}R_{ww^*}(\tau) = \tilde{c}^*_{xx}(-2f_0, \tau) \tag{5.64b}$$

From the second term in Eq. (5.63), in the WSS case, it is evident that a period of $T = 1/2f_0$ would also work; however, we choose to regard the QAM signal as CS($1/f_0$) since the mean value could have a component at f_0. The autocorrelation result is the same in either case.

We note that $x(t)$ itself would be WSS under the condition that the cross-correlation, $R_{ww*}(\tau)$, vanishes, making the second term in Eq. (5.63) disappear. This condition is equivalent to $R_{uu}(\tau) = R_{vv}(\tau)$ and $R_{vu}(\tau) = -R_{uv}(\tau) = -R_{vu}(-\tau)$; in other words, $x(t)$ is WSS if the in-phase and quadrature components have the same autocorrelation and their cross-correlation function is odd. If $u(t)$ and $v(t)$ are independent zero-mean processes, then we require only that their autocorrelations match. This is called *balanced* QAM and its power spectral density is

$$S_{xx}(f) = C_{xx}(0, f) = \frac{1}{2}S_{uu}(f-f_0) + \frac{1}{2}S_{uu}(f+f_0) \tag{5.65}$$

For the case of digital data signals, we examine the QAM/PAM signal where the complex baseband signal is a pair of PAM signals,

$$w(t) = \sum_k c(k)g(t-kT); \qquad c(k) = a(k) + jb(k) \tag{5.66}$$

Here we encounter a problem regarding the period of the cyclostationarity, because the PAM signal exhibits a period of T while the carrier modulation introduces a period of $1/f_0$, and these periods are not necessarily commensurate. The problem is easily handled by methods used for almost periodic functions, by extending the integration interval indefinitely

$$R_{xx}(t, \tau) = \sum_i \tilde{c}_{xx}(\nu_i, \tau)e^{j2\pi\nu_i t} \tag{5.67a}$$

$$\tilde{c}_{xx}(\nu_i, \tau) = \lim_{T_0\to\infty} \frac{1}{2T_0}\int_{-T_0}^{T_0} R_{xx}(t, \tau)e^{-j2\pi\nu_i t}dt \tag{5.67b}$$

The values of ν_i for which Eq. (5.67b) are nonzero are the cycle frequencies. The process is sometimes called *polycyclostationary*. Now, using Eq. (5.25) for the cyclic autocorrelations of $w(t)$ and Eq. (5.63) for the effect of the modulation, then taking Fourier transforms, we get the following results for the cyclic spectral density functions for $x(t)$:

$$\begin{aligned} C_{xx}\left(\frac{m}{T}, f\right) &= \left(\frac{1}{4T}\right)S_{cc}\left(f-f_0-\frac{m}{2T}\right)K(f-f_0; m) \\ &\quad + \left(\frac{1}{4T}\right)S_{cc}\left(-f-f_0+\frac{m}{2T}\right)K(f+f_0; m) \end{aligned} \tag{5.68a}$$

$$C_{xx}\left(\frac{m}{T}+2f_0, f\right) = \left(\frac{1}{4T}\right)S_{cc*}\left(f-\frac{m}{2T}\right)K(f; m) \tag{5.68b}$$

$$C_{xx}\left(\frac{m}{T}-2f_0, f\right) = \left(\frac{1}{4T}\right)S^*_{cc*}\left(-f+\frac{m}{2T}\right)K(f; m) \tag{5.68c}$$

where, for convenience, we have defined the function

$$K(f; m) = G\left(f+\frac{m}{2T}\right)G^*\left(f-\frac{m}{2T}\right) \tag{5.69}$$

The cycle frequencies are $\{m/T, m/T + 2f_0, m/T - 2f_0\}$. For example, if the data pulse $g(t)$ had less than 100% excess bandwidth, then $K(f; m)$ vanishes for $|m| > 2$ and there are a total of nine cycle frequencies. If the QAM is balanced, for example, by making the $a(k)$ and $b(k)$ identically distributed and independent, then the cross-spectral density $S_{cc*}(f)$ vanishes and only three cycle frequencies remain.

As an illustration, let us assume that the $a(k)$ and $b(k)$ sequences are independent and each sequence has zero-mean, independent data, but with variances, σ_a^2 and σ_b^2, for in-phase and quadrature data, respectively. Then $S_{cc}(f) = \sigma_a^2 + \sigma_b^2$ and $S_{cc^*}(f) = \sigma_a^2 = \sigma_b^2$, giving the following simple expressions for the cyclic spectral densities:

$$C_{xx}\left(\frac{m}{T}, f\right) = \frac{\sigma_a^2 + \sigma_b^2}{4T}[K(f - f_0; m) + K(f + f_0; m)] \tag{5.70a}$$

$$C_{xx}\left(\frac{m}{T} + 2f_0, f\right) = C_{xx}\left(\frac{m}{T} - 2f_0, f\right) = \frac{\sigma_a^2 - \sigma_b^2}{4T}K(f; m) \tag{5.70b}$$

and the power spectral density is

$$S_{xx}(f) = C_{xx}(0, f) = \frac{\sigma_a^2 + \sigma_b^2}{4T}[|G(f - f_0)|^2 + |G(f + f_0)|^2] \tag{5.71}$$

which is simply a pair of frequency-translated versions of the corresponding baseband PAM spectrum.

The case of *double-sideband amplitude modulation* (DSB-AM/PAM) is covered in the preceding equations simply by making all of the $b(k)$ equal to zero. This can be regarded as extremely unbalanced QAM; hence, the nonzero cyclic components will tend to be greater than in the QAM case. Other modulation formats can be handled by essentially the same methods. For example, *vestigial-sideband modulation* (VSB-AM/PAM) uses a single data stream and a complex data pulse; $w(t) = \sum a(k)[g(t - kT) + j\tilde{g}(t - kT)]$. A more general version of QAM uses a different pulse shape in the in-phase and quadrature channels. Staggered QAM (SQAM) uses time-shifted versions of the same pulse; for example, $w(t) = \Sigma a(k)g(t - kT) + jb(k)g(t - T/2 - kT)$. *Quaternary phase-shift keying* (QPSK) is a special case of QAM with $a(k), b(k) \in \{+1, -1\}$.

5.7 Appendix: The Poisson Sum Formula

For any time function $g(t)$ having Fourier transform $G(f)$, the following relation is often useful in system analysis:

$$\sum_k g(t - kT) = \frac{1}{T}\sum_m G\left(\frac{m}{T}\right)e^{j2\pi mt/T} \tag{A.5.1}$$

The relation is easily verified by making a Fourier series expansion of the periodic function on the left-hand side of Eq. (A.5.1). The mth Fourier coefficient is

$$c(m) = \frac{1}{T}\int_0^T \sum_k g(t - kT)e^{-j2\pi mt/T}\,\mathrm{d}t \tag{A.5.2}$$

and, by a change of integration variable,

$$c(m) = \frac{1}{T}\int_{-\infty}^{\infty} g(\sigma)e^{-j2\pi m\sigma/T}\,\mathrm{d}\sigma = \frac{1}{T}G\left(\frac{m}{T}\right) \tag{A.5.3}$$

Another useful version of the Poisson sum formula is the time/frequency dual of Eq. (A.5.1).

$$\sum_k g(kT)e^{-j2\pi kTf} = \frac{1}{T}\sum_m G\left(f - \frac{m}{T}\right) \tag{A.5.4}$$

Defining Terms

Autocorrelation: The expected value of the product of two elements (time samples) of a single random process. The variables of the autocorrelation function are the two time values indexing the random variables in the product.

Cross correlation: The expected value of the product of elements from two distinct random processes.

Cycle frequency: The frequency parameter associated with a sinusoidal variation of a component of the autocorrelation of a cyclostationary process.

Cyclic autocorrelation: A parameter characterizing the amount of contribution to the overall autocorrelation at a particular cycle frequency.

Cyclostationary random process (wide-sense): A random process whose mean and autocorrelation functions vary periodically with time.

Excess bandwidth: The amount of bandwidth of a PAM signal that exceeds the minimum bandwidth for no intersymbol interference. It is usually expressed as a fraction of the Nyquist frequency of $1/2T$, where $1/T$ is the PAM symbol rate.

Poisson sum formula: A useful identity relating the Fourier series and sum-of-translates representations of a periodic function.

Power spectral density: A real, nonnegative function of frequency that characterizes the contribution to the overall mean-squared value (power) of a random process due to components in any specified frequency interval.

Pulse amplitude modulation (PAM): A signal format related to a discrete-time sequence by superposing time-translated versions of a single waveform, each scaled by the corresponding element of the sequence.

Quadrature amplitude modulation (QAM): A signal format obtained by separately modulating the amplitude of two components of a sinusoidal carrier differing in phase by 90°. In digital communications, the two components (referred to as in-phase and quadrature components) can each be regarded as a PAM signal.

Time-ambiguity function: A property of a waveform, similar to convolution of the waveform with itself, which characterizes the degree to which the waveform can be localized in time. The Fourier transform of the time-ambiguity function is the energy spectral density of the waveform.

References

Franks, L.E. 1969. *Signal Theory,* Prentice–Hall, Englewood Cliffs, NJ; rev. ed. 1981. Dowden and Culver.

Gardner, W.A. 1989. *Introduction to Random Processes with Applications to Signals and Systems*, 2nd ed., McGraw–Hill, New York.

Gardner, W.A., Ed. 1994. *Cyclostationarity in Communications and Signal Processing*, IEEE Press.

Lee, E.A. and Messerschmitt, D.G. 1988. *Digital Communication*, Kluwer Academic, Dordrecht, the Netherlands.

Papoulis, A. 1991. *Probability, Random Variables, and Stochastic Processes*, 3rd ed., McGraw–Hill, New York.

Proakis, J.G. 1989. *Digital Communications*, McGraw–Hill, New York.

Proakis, J.G. and Salehi, M. 1994. *Communication Systems Engineering*, Prentice–Hall, Englewood Cliff, NJ.

Stark, H. and Woods, J.W. 1986. *Probability Random Processes, and Estimation Theory for Engineers*, 2nd ed., Prentice–Hall, Englewood Cliffs, NJ.

Further Information

The text by Papoulis [1991] is the third edition of a well known and referenced work on probability and random processes that first appeared in 1965. It is a concise and comprehensive treatment of the topic and is widely regarded as an ideal reference source. Gardner [1989] provides the most complete textbook treatment of cyclostationary processes. He presents a careful development of the theory of such processes as well as several examples of applications relevant to the communications field. Stark and Woods [1986] is a somewhat more advanced textbook. It uses modern notation and viewpoints and gives the reader a good introduction to the related topics of parameter estimation and decision theory.

6

Queuing*

Richard H. Williams
University of New Mexico

6.1 Introduction

Everyone has experienced being in a **queue**: waiting in line at a checkout counter, waiting in traffic as we enter a highway's tollgate, or waiting for a time-sharing computer to run our program. Examples of queues from everyday life and technical experiences are almost endless.

A study of queues is important for the following reasons:

1. From the user's point of view, a queue, that is, waiting in line, is a nuisance and will be tolerated only at some minimal level of inconvenience.
2. From the point of view of a provider of some service, queues are necessary. It is extremely uneconomical to provide a large number of servers to handle the biggest possible rush, only to have many of them idle most of the time.

One motivation to study queues is to find some balance between these two conflicting viewpoints. Probability theory is the tool we use for a study of queues because arrivals and departures from a **queuing system** occur randomly. In Fig 6.1, a **customer** arrives at the input to the queuing system, and if no **server** is free, the customer must wait in a queue until it is his, her, or its turn to be served. A **queue discipline** is the name given to a rule by which a customer is selected from a queue for service.

6.2 Little's Formula

Let the average rate at which customers arrive at a queuing system be the constant λ. Departures from the queuing system occur when a server finishes with a customer. In the steady state, the average rate at which customers depart the queuing system is also λ. Arrivals to a queuing system may be described as a counting process $A(t)$ such as shown in Fig 6.2. We estimate an average rate of arrivals with

$$\bar{\lambda} = \frac{A(t)}{t} \quad \text{(per unit time)} \tag{6.1}$$

*Portions are reprinted by permission. Williams, R.H. 1991. *Electrical Engineering Probability,* Copyright © by West, St. Paul MN. All Rights Reserved.

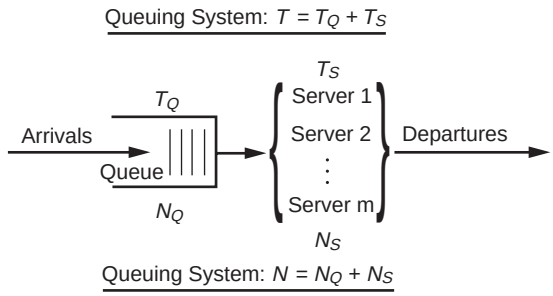


FIGURE 6.1 Illustration of queuing system: Customers arrive at a queue and, after some time, depart from a server that serves the queue.

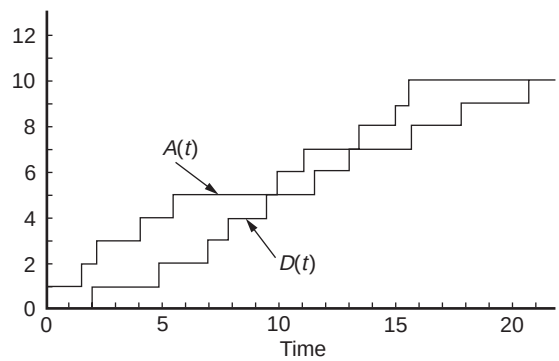


FIGURE 6.2 Illustration of the arrival and departure counting processes for a queuing system.

where $A(t)$ is the number of arrivals in $[0, t]$. Departures from the system $D(t)$ are also described by a counting process, and an example is also shown in Fig. 6.2. Note that $A(t) \geq D(t)$ must be true at all times. The time that each customer spends in the queuing system is the time difference between the customer's arrival and subsequent departure from a server. The time spent in the queuing system for arrival i is denoted t_i, $i = 1, 2, 3, \ldots$.

The average time a customer spends waiting in a queuing system is

$$\bar{t} = \frac{1}{A(t)} \sum_{i=1}^{A(t)} t_i \quad \text{(units of time)} \tag{6.2}$$

The average number of customers in a queuing system $\bar{n}$ is, from Eqs. (6.1) and (6.2),

$$\bar{n} = \frac{1}{t} \sum_{i=1}^{A(t)} t_i = \bar{\lambda}\bar{t} \quad \text{(dimensionless)} \tag{6.3}$$

If we assume that, with enough data, sample averages approximate their expected values, then $\bar{\lambda} \to \lambda$, $\bar{t} \to E[T]$, and $\bar{n} \to E[N]$, where N and T are random variables. Also, Eq. (6.3) becomes

$$E[N] = \lambda E[T] \tag{6.4}$$

This relation is known as Little's formula [Williams, 1991]; it has a generality much beyond what one might suppose, given the heuristic demonstration leading to Eq. (6.4).

TABLE 6.1 Data from a FIFO Queue Used with Example 1

i	t_A	t_Q	t_S	t_i	t_D
1	0.0	0.0	2.0	2.0	2.0
2	1.5	0.5	2.9	3.4	4.9
3	2.2	2.7	2.1	4.8	7.0
4	4.1	2.9	0.8	3.7	7.8
5	5.5	2.3	1.7	4.0	9.5
6	10.0	0.0	1.5	1.5	11.5
7	11.3	0.2	1.6	1.8	13.1
8	13.4	0.0	2.3	2.3	15.7
9	15.0	0.7	2.0	2.7	17.7
10	15.5	2.2	3.0	5.2	20.7

Example 1

The data in Table 6.1 are the data used to plot $A(t)$ and $D(t)$ in Fig. 6.1. In Table 6.1, the times, given in the units of minutes, are: arrival time t_A, time spent in the queue t_Q, time spent with the server t_S, total time in the queuing system for the ith arrival t_i, and the departure time t_D. For each i we see in Table 6.1, that $t_i = t_Q + t_S$ and that $t_D = t_A + t_i$. These data describe a first-in–first-out (FIFO) queuing discipline: If a customer arrives before the previous customer departs, then the customer must wait his turn in the queue until the server is available. If the previous customer has departed prior to a customer's arrival, then the waiting time in the queue is $t_Q = 0$.

According to the data in Table 6.1, when $t = 20.7$ then $A(t) = i = 10$. From Eq. (6.1),

$$\bar{\lambda} = \frac{10}{20.7} = 0.48 \quad (\text{min}^{-1})$$

The sum of all the t_i in Table 6.1 is 31.4 min. Therefore, from Eq. (6.2),

$$\bar{t} = \frac{31.4}{10} = 3.14 \quad (\text{min})$$

Finally, using Little's formula in Eq. (6.3), the sample average number of customers in the queuing system is

$$\bar{n} = \frac{10}{20.7} \times \frac{31.4}{10} = 1.52$$

6.3 The M/M/1 Queuing System: State Probabilities

A generally accepted notation used to name a queuing system is A/S/m: A is a description of the dynamics by which arrivals come to the queuing system; S is a description of the dynamics by which a server processes a customer taken from the queue; and m is the number of servers. The arrivals and the server are independent of each other except that a server cannot serve a customer who has not yet arrived into the queue.

We define the **state** of a queuing system as the total number of customers within it, that is, either within the queue or being served. The only queuing system that we discuss in this chapter is M/M/1: The arrival dynamics are Markov (with rate constant λ), the server dynamics are Markov (with rate constant λ_S), and there is only one server.

There can be, at most, only *one* arrival and, at most, *one* departure in a time interval Δt when $\Delta t \to 0$. As far as arrivals are concerned, when Δt is a small increment of time, the probability of an arrival into the queuing system in the interval Δt is

$$P(A) = \lambda \Delta t \tag{6.5}$$

Correspondingly, the probability of no arrival in Δt is

$$P(\bar{A}) = 1 - \lambda \Delta t \tag{6.6}$$

The probability of serving (discharging) a customer $P(D)$ is actually conditioned by the event that there is at least one customer in the queuing system. Thus, *given* that the state is at least one,

$$P(D) = \lambda_S \Delta t \tag{6.7}$$

The probability of continuing to serve a customer who is already being served is

$$P(\bar{D}) = 1 - \lambda_S \Delta t \tag{6.8}$$

Suppose that the state of a queuing system is zero at time $t + \Delta t$. This can occur in only one of three mutually exclusive ways. If the state at time t is zero, then either there is no arrival in the increment Δt, or, if there is an arrival in Δt, then there is also a departure within the same Δt. If the state at t is one, then there must be one departure and no arrival within Δt. Let $P_n(t)$ denote the probability of a queuing system being in the state n at time t. Then, with this notation, a summary of the preceding statement is

$$P_0(t + \Delta t) = P_0(t)\{P(\bar{A}) + P(A)P(D)\} + P_1(t)P(\bar{A})P(D)$$

Using Eqs. (6.5–6.8), this may be arranged into

$$\frac{P_0(t + \Delta t) - P_0(t)}{\Delta t} = -\lambda P_0(t)\{1 - \lambda_S \Delta t\} + \lambda_S P_1(t)\{1 - \lambda \Delta t\}$$

Then, taking the limit as $\Delta t \to 0$, we have the derivative

$$\frac{\mathrm{d}P_0(t)}{\mathrm{d}t} = \lambda_S P_1(t) - \lambda P_0(t)$$

When a queuing system is in the steady state, the rate of change of the probability $P_0(t)$ with respect to time is zero, and we have

$$\lambda P_0(t) = \lambda_S P_1(t) \tag{6.9}$$

If at time $t + \Delta t$ the state of the queuing system is greater than zero, we have

$$\begin{aligned} P_n(t + \Delta t) = {} & P_n(t)\{P(\bar{A})P(\bar{D}) + P(A)P(D)\} \\ & + P_{n+1}(t)P(\bar{A})P(D) + P_{n-1}(t)P(A)P(\bar{D}) \end{aligned}$$

The first term on the right of this equation describes two ways for the state to remain the same when time increases from t to $t + \Delta t$: either no arrival and no departure occur, or exactly one arrival occurs and exactly one departure occurs. The second term describes the state $n + 1$ changing to n by exactly one departure with no arrival. The third term describes the change from state $n - 1$ to n by exactly one arrival and no departure. No other combinations of arrivals and departures can happen using our assumptions when $n > 0$. Then, as before, finding the derivative of $P_n(t)$ and setting it equal to zero in the steady-state condition, we find

$$(\lambda + \lambda_S)P_n(t) = \lambda_S P_{n+1}(t) + \lambda P_{n-1}(t) \tag{6.10}$$

We introduce the ratio

$$\rho = \frac{\lambda}{\lambda_S} \tag{6.11}$$

Then, when $n = 0$, we have from Eq. (6.9),

$$P_1(t) = \rho P_0(t)$$

When $n = 1$, Eq. (6.10) becomes

$$(\lambda + \lambda_S)P_1(t) = \lambda_S P_2(t) + \lambda P_0(t)$$

Combining this with $P_1(t)$, we have an expression which, when simplified, becomes

$$P_2(t) = \rho^2 P_0(t)$$

Continuing in this way, it follows that

$$P_n(t) = \rho^n P_0(t) \tag{6.12}$$

Since the queuing system can theoretically be in any state $n = 0, 1, 2, \ldots$ (that is, there can be any nonnegative number of customers in the queuing system), we must require that

$$\sum_{n=0}^{\infty} P_n(t) = 1 \tag{6.13}$$

Combining Eqs. (6.12) and (6.13), we have

$$P_0(t)\sum_{n=0}^{\infty} \rho^n = 1$$

By stipulating that $\rho < 1$, we can sum the geometric progression [Beyer, 1987, p. 8]:

$$\sum_{n=0}^{\infty} \rho^n = \frac{1}{1-\rho} \tag{6.14}$$

Thus, when $n = 0$, we have

$$P_0(t) = 1 - \rho \tag{6.15}$$

Combining Eq. (6.15) with Eq. (6.12) gives the probability of any state n for the M/M/1 queuing system:

$$P_n(t) = (1-\rho)\rho^n \tag{6.16}$$

As long as the state of a queuing system is greater than zero, then one of the customers must be with the server. Therefore, the probability that the server is busy is $P(N > 0)$. Since $P(N > 0) + P_0(t) = 1$, it follows, using Eq. (6.15), that

$$P(N > 0) = \rho \tag{6.17}$$

The result in Eq. (6.17) is why the ratio ρ is called the **utilization** of an M/M/1 queuing system.

Example 2

A manager of a queuing system wants to know the probability that k or more customers are in the queue waiting to be served by the server; $k = 1, 2, 3, \ldots$. If even one customer is in the queue, then there must also be one customer with the server. Therefore, what the manager wants to know is the probability that the state of the queuing system is $k + 1$ or more. Then, with Eq. (6.16),

$$P(N_Q \geq k) = P(N \geq k+1) = \sum_{n=k+1}^{\infty} (1-\rho)\rho^n$$

If we use the change of variables $m = n - (k + 1)$, and Eq. (6.14), we then answer the manager's question:

$$P(N_Q \geq k) = (1-\rho)\rho^{k+1} \sum_{m=0}^{\infty} \rho^m = \rho^{k+1}$$

6.4 The M/M/1 Queuing System: Averages and Variances

The purpose of this section is to find the statistical parameters of mean and variance for both random variables N and T shown in Fig. 6.1 for an M/M/1 queue. In the following, the lower case n and t stand for realizations of the random variables N and T.

If the geometric progression in Eq. (6.14) is differentiated once with respect to ρ, we have

$$\sum_{n=0}^{\infty} n\rho^n = \frac{\rho}{(1-\rho)^2}$$

Then, differentiating this once again,

$$\sum_{n=0}^{\infty} n^2\rho^n = \frac{\rho(1+\rho)}{(1-\rho)^3}$$

The first of these two derivatives can be used along with Eq. (6.16) to show that the mean of N is

$$E[N] = \sum_{n=0}^{\infty} nP_n(t) = \sum_{n=0}^{\infty} n(1-\rho)\rho^n = \frac{\rho}{1-\rho} \tag{6.18}$$

The second of these two derivatives is used to express the mean of the square of N,

$$E[N^2] = \sum_{n=0}^{\infty} n^2P_n(t) = \sum_{n=0}^{\infty} n^2(1-\rho)\rho^n = \frac{\rho(1+\rho)}{(1-\rho)^2} \tag{6.19}$$

Equations (6.18) and (6.19) are used to find the variance of N,

$$\mathrm{Var}[N] = E[N^2] - (E[N])^2 = \frac{\rho}{(1-\rho)^2} \tag{6.20}$$

One notes that as $\lambda_S \to \lambda$, the utilization approaches unity, and both the mean and the variance of N for the M/M/1 queue become unbounded.

Assume that a customer arrives at an M/M/1 queuing system at time $t = 0$. The customer finds that the queuing system is in state n where n could be 0, 1, 2,.... We use the notation T_n to denote the random variable for the total time in the queuing system *given* that the state found by a customer is n. The probability of that customer leaving the queuing system in the interval between t and $t + \Delta t$ is

$$P(t < T_n \le t + \Delta t) = F_{T_n}(t + \Delta t) - F_{T_n}(t)$$

where $F_{T_n}(t)$ is the probability cumulative distribution function (cdf) for T_n. Another way of expressing this probability is

$$\begin{aligned} P(t < T_n \le t + \Delta t) &= P(\text{exactly } n \text{ departures from the server between 0 and } t) \\ &\quad \times P(\text{the customer departs from the server in } \Delta t) \end{aligned}$$

The first of the probabilities is Poisson [Williams, 1991] with rate parameter λ_S and time interval t. The second probability is obtained from Eq. (6.7). Thus,

$$P(t < T_n \le t + \Delta t) = \left\{ \frac{(\lambda_S t)^n}{n!} e^{-\lambda_S t} \right\} \lambda_S \Delta t$$

Combining these two expressions for $P(t < T_n \le t + \Delta t)$, we have the following:

$$F_{T_n}(t + \Delta t) - F_{T_n}(t) = \left\{ \frac{(\lambda_S t)^n}{n!} e^{-\lambda_S t} \right\} \lambda_S \Delta t$$

We can use this to construct the following derivative:

$$\lim_{\Delta t \to 0} \frac{F_{T_n}(t + \Delta t) - F_{T_n}(t)}{\Delta t} = \frac{\mathrm{d}}{\mathrm{d}t} F_{T_n}(t) = f_{T_n}(t)$$

where $f_{T_n}(t)$ is the probability density function (pdf) for T_n,

$$f_{T_n}(t) = \frac{\lambda_S^{n+1} t^n}{n!} e^{-\lambda_S t}, \quad t \ge 0 \tag{6.21}$$

The random variable T is the time in the queuing system considering all possible states. Then, using Eqs. (6.16), (6.21), and the series expansion for e^x [Beyer, 1987, p. 299],

$$\begin{aligned} f_T(t) &= \sum_{n=0}^{\infty} f_{T_n}(t) P_n(0) \\ &= \sum_{n=0}^{\infty} \frac{\lambda_S^{n+1} t^n}{n!} e^{-\lambda_S t} (1 - \rho) \rho^n \\ &= (1 - \rho) \lambda_S e^{-\lambda_S t} \sum_{n=0}^{\infty} \frac{\lambda_S^n t^n \rho^n}{n!} \\ &= (1 - \rho) \lambda_S e^{-\lambda_S t} e^{\rho \lambda_S t} \end{aligned}$$

Using Eq. (6.11), we finally have the result

$$f_T(t) = (\lambda_S - \lambda)e^{(\lambda_S - \lambda)t}, \quad t \geq 0 \tag{6.22}$$

This is recognized as the pdf for an exponential random variable. It is therefore known that [Williams, 1991]

$$E[T] = \frac{1}{\lambda_S - \lambda} = \frac{1/\lambda_S}{1-\rho} \tag{6.23}$$

$$\mathrm{Var}[T] = \frac{1}{(\lambda_S - \lambda)^2} = \frac{1/\lambda_S^2}{(1-\rho)^2} \tag{6.24}$$

Example 3

Assume that an M/M/1 queuing system has a utilization of $\rho = 0.6$; this means that the server is idle 40% of the time. The average number of customers in the queuing system is, from Eq. (6.18),

$$E[N] = \frac{\rho}{1-\rho} = 1.50$$

The standard deviation for this value is, from Eq. (6.20),

$$\sqrt{\mathrm{Var}[N]} = \frac{\sqrt{\rho}}{1-\rho} = 1.94$$

From Eqs. (6.23) and (6.24),

$$E[T] = \sqrt{\mathrm{Var}[T]} = \left(\frac{1}{1-\rho}\right)\frac{1}{\lambda_S} = \frac{2.5}{\lambda_S}$$

Thus, the mean and the standard deviation of the time a customer spends in the M/M/1 queuing system are the same.

Realizations of either N and T in an M/M/1 queuing system with $\rho = 0.6$ should be expected to fluctuate markedly about their means because of their relatively large standard deviations. A similar behavior may occur with other values of the utilization.

6.5 Averages for the Queue and the Server

The average number of customers in the server can be calculated using Eq. (6.17):

$$E[N_S] = (1)P(N > 0) + (0)P_0(t) = P(N > 0) = \rho \tag{6.25}$$

Combining Eq. (6.18) with Eq. (6.25), we find the average number of customers in the queue:

$$\begin{aligned} E[N_Q] &= E[N] - E[N_S] \\ &= \frac{\rho}{1-\rho} - \rho \\ &= \frac{\rho^2}{1-\rho} \end{aligned} \tag{6.26}$$

The average time a customer spends with the server is the reciprocal of the server's average rate,

$$E[T_S] = 1/\lambda_S \tag{6.27}$$

Therefore, using Eqs. (6.23) and (6.27), the average time spent waiting in the queue before being served is

$$\begin{aligned} E[T_Q] &= E[T] - E[T_S] \\ &= \frac{1/\lambda_S}{1-\rho} - \frac{1}{\lambda_S} \\ &= \frac{\rho/\lambda_S}{1-\rho} \end{aligned} \tag{6.28}$$

Example 4

An M/M/1 queuing system has a utilization of $\rho = 0.6$ and a rate at which the server can work of $\lambda_S = \frac{1}{45}(\mathrm{s}^{-1})$. Using Eq. (6.28), customers should expect to spend the following time in the queue before being served:

$$E[T_Q] = \frac{45\rho}{1-\rho} = 1.13 \quad (\mathrm{min})$$

At any time, the expected number of customers in the queue waiting to be served is, using Eq. (6.26),

$$E[N_Q] = \frac{\rho^2}{1-\rho} = 0.90$$

And, the expected number of customers with the server is its utilization, Eq. (6.25),

$$E[N_S] = \rho = 0.6$$

Defining Terms

Customer: A person or object that is directed to a server.
Queue: A line of people or objects waiting to be served by a server.
Queue discipline: A rule by which a customer is selected from a queue for service by a server.
Queuing system: A combination of a queue, a server, and a queue discipline.
Server: A person, process, or subsystem designed to perform some specific operation on or for a customer.
State: The total number of customers within a queuing system.
Utilization: The percentage of time that the server is busy.

References

Beyer, W.H. 1987. *CRC Standard Mathematical Tables,* 28th ed., CRC Press, Boca Raton, FL.
Williams, R.H. 1991. *Electrical Engineering Probability,* West, St. Paul, MN.

Further Information

Asmussen, S. 1987. *Applied Probability and Queues*, Wiley, Chichester, UK.
Kleinrock, L. 1975. *Queuing Systems*, Vol. 1, Wiley, New York.
Leon-Garcia, A. 1994. *Probability and Random Processes for Electrical Engineering*, 2nd ed., Addison–Wesley, Reading, MA.
Payne, J.A. 1982. *Introduction to Simulation*, McGraw–Hill, New York.

7

Multiplexing

Martin S. Roden
California State University

7.1 Introduction

A *channel* is the link between a source and a sink (receiver). In the early days of electrical wire communication, each channel was used to transmit only one signal. If you were to view an early photograph of the streets of New York, you would note that utility poles dominate the scene. Literally hundreds of wires formed a web above the city streets. Some people considered this beautiful, particularly on an icy winter morning. Of course, aside from aesthetic considerations, the proliferation of separate channels could not be permitted to continue as communications expanded.

In many contemporary applications, a variety of signals must be transmitted on a single channel. The channel must, therefore, be shared among the various applications. The process of combining multiple signals for transmission through a single channel is called **multiplexing**.

When multiple signals simultaneously occupy the same channel, the various signals must be separable from each other. In a mathematical sense, the various signals must be *orthogonal*.

The form of multiplexing used in everyday conversation requires time separation. In such situations, the participants try not to speak at the same time. But there are more efficient ways that signals can share a channel without overlapping. This section examines frequency multiplexing, wavelength multiplexing, time multiplexing, space multiplexing, and several *multiple access* schemes.

7.2 Frequency Multiplexing

Frequency-division multiplexing (FDM) is the technique used in traditional analog broadcast systems including AM radio, FM radio, and television. As a simple example, suppose two people in a room wanted to talk simultaneously. A listener would hear a superposition of both speech waveforms and would have difficulty separating one from the other. If, however, one of the speakers spoke in a baritone and the other in soprano, the listener could resolve the two voice signals using an acoustic filter. For example, a lowpass filter would only admit the baritone's voice, thereby blocking the other signal.

Of course, we do not want to restrict the various sources in an unrealistic manner. FDM takes advantage of the observation that all frequencies of a particular message waveform can be easily shifted by a constant amount. The shifting is performed using a sinusoidal carrier signal. The original message

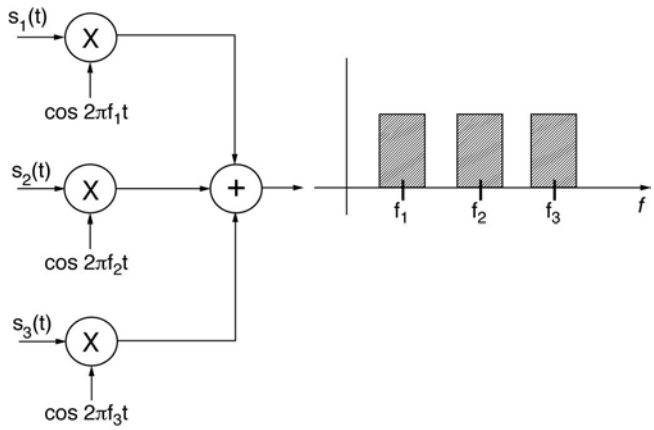


FIGURE 7.1 Frequency-division multiplexing.

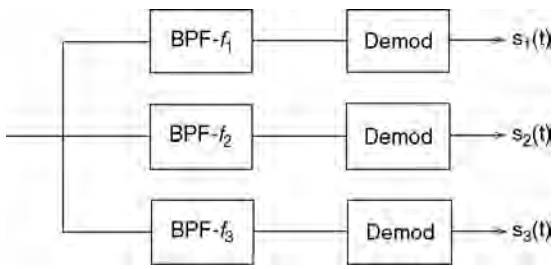


FIGURE 7.2 Demultiplexing of FDM.

signal is multiplied by the carrier. This shifts the frequencies of the message signal by the frequency of the carrier. If, for example, the original signal is baseband (composed of relatively low frequencies around DC), multiplication of the signal by a sinusoidal carrier shifts all frequencies to a range centered about the sinusoidal frequency. We are describing *amplitude modulation*. By using multiple sinusoids of differing frequencies, signals can be stacked in frequency in a non-overlapping manner. This is illustrated in Fig. 7.1. The right portion of the figure illustrates typical frequency spectra (Fourier transforms). To be able to separate the signals, the various carriers must be spaced by at least twice the highest frequency of the message signal.

The multiplexed signals can be extracted using bandpass filters (i.e., frequency *gates*). The frequencies can then be shifted back to baseband using demodulators. This is shown in Fig. 7.2.

The demodulator block in Fig. 7.2 takes on various forms. The coherent form simply shifts the signal back down to baseband frequencies using an exact replica of the original sinusoid. If the original modulated signal is modified to explicitly contain the carrier sinusoid, incoherent demodulators (e.g., the envelope detector) may obviate the need for reproducing the original sinusoid.

In designing FDM systems, attention must be given to the minimum separation between carriers. Although the minimum spacing is twice the highest signal frequency, using this exact value would place unrealistic constraints on the bandpass filters of Fig. 7.2. For this reason, a *guard slot* is often included between multiplexed frequency components. The larger the guard band, the easier it is to design the separation filters. However, the price being paid is a reduction in the number of channels that can be multiplexed within a given bandwidth. Note that this is not the approach used by broadcast multiplex systems. Instead, there is no guard slot and the FCC avoids assigning adjacent frequency slots to two strong signal sources.

Frequency-division multiplexing is sometimes used to create a composite baseband signal. For example, in FM stereo, the two audio signals are frequency multiplexed to produce a new baseband signal. One of the audio signals occupies the band of frequencies between DC and 15 kHz, whereas the second audio signal is shifted by 38 kHz. It then occupies a band between 23 and 53 kHz. The composite sum signal can then be considered as a new baseband waveform with frequencies between DC and 53 kHz. The composite baseband waveform frequency modulates a carrier and is frequency multiplexed a second time, this time with other modulated composite signals.

When frequency-division multiplexing is used on fiber-optic channels, it is sometimes referred to as **wavelength-division multiplexing (WDM)**. WDM systems encode each signal on a different carrier frequency. For relatively short distances, WDM is often more complex (i.e., expensive) than running multiple fiber cables, but this may change in time.

7.3 Time Multiplexing

There is a complete duality between the time domain and the frequency domain. Therefore, the discussion of FDM of the preceding section can be extended to **time-division multiplexing (TDM)**. This form of multiplexing is common to pulse modulation systems.

An analog signal can be transmitted by first sampling the waveform. In accordance with the *sampling theorem*, the number of samples required each second is at least twice the highest frequency of the waveform. For example, a speech waveform with an upper frequency of 4 kHz can be fully represented by samples taken at a rate of 8000 samples/s. If each sample value modulates the height of a pulse, as in PAM, the signal can be sent using 8000 varying-height pulses each second. If each pulse occupies only a fraction of the time spacing between samples, the time axis can be shared with other sampled signals. Figure 7.3 shows an example for three signals. The notation used designates s_{ij} as the ith time sample of the jth signal. Thus, for example, s_{23} is the second sample of the third signal. The three signals are said to be time-division multiplexed. If the pulses are made narrower, additional signals can be multiplexed. Of course, the narrower the pulses, the wider the bandwidth of the multiplexed signal. Once you are given the bandwidth of the channel, the minimum pulse width is set. In using TDM, one must also be sensitive to channel distortion, which widens the pulses before they arrive at the receiver. Pulses that do not overlap at the transmitter may, therefore, overlap at the receiver, thus causing *intersymbol interference (ISI)*. Equalizers can be used to reduce the ISI.

If the transmission is (binary) digital rather than analog, each sample is coded into a binary number. Instead of sending a single pulse, multiple binary pulses are sent for each sample. This leads to *pulse code modulation*, or PCM. TDM can still be used, but the number of pulses in each sampling interval increases by a factor equal to the number of bits of A/D conversion.

The **T-1 carrier** transmission system has historical significance and also forms the basis of many contemporary multiplexing systems. It was developed by Bell Systems (before divestiture) in the early 1960s.

The T-1 carrier system develops a 1.544-Mb/s pulsed digital signal by multiplexing 24 voice channels. Each channel is sampled at a rate of 8000 samples/s. Each sample is then converted into 8 bits using companded PCM. Although all 8 bits corresponding to a particular sample can be available for sending that sample value, one of the bits is sometimes devoted to signaling, which includes bookkeeping needed

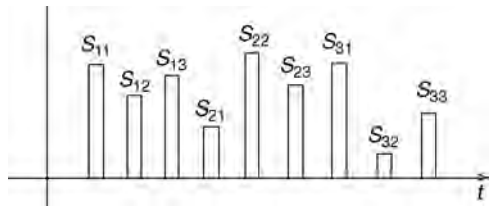


FIGURE 7.3 Time-division multiplexing of PAM signals.

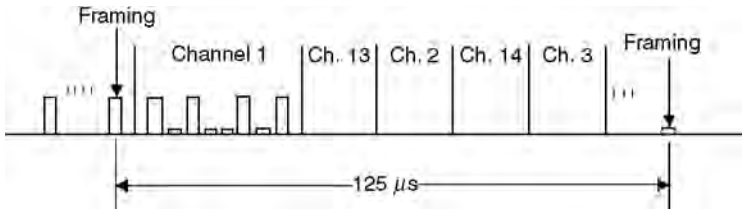


FIGURE 7.4 Multiplexing structure of T-1 carrier.

to set up the call and keep track of billing. The 24 channels are then time-division multiplexed, as shown in Fig. 7.4.

Each frame, which corresponds to one sampling period, contains 192 bits (8 times 24), plus a 193rd bit for **frame synchronization**. Since frames occur 8000 times/s, the system transmits 1.544 Mb/s.

Need for Frame Synchronization

When a receiver captures and demodulates a TDM signal, the result is a continuous sequence of numbers (usually binary). These numbers must be sorted to associate the correct numbers with the correct signal. This sorting process is known as demultiplexing. To perform this function, the receiver needs to know when each frame begins. If nothing is known about the value of the signals, frame synchronization must be performed using overhead. A known synchronizing signal is transmitted along with the multiplexed signals. This synchronizing signal can be sent as one of the multiplexed signals, or it can be a framing bit as in the case of T-1 carrier. The receiver searches for the known pattern and locks on to it. Of course, there is some probability that one of the actual signals will resemble the synchronizing signal, and false sync lock can result. The probability of false lock decreases as the synchronizing signal is made longer.

Dissimilar Channels

The examples given so far in this section assume that the signals to be multiplexed are of the same form. For example, in the T-1 carrier system, each of the 24 signals is sampled at the same rate (8 kHz) and converted to the same number of bits.

Dissimilar signals must sometimes be multiplexed. In the general case, we refer to this as *asynchronous multiplexing*. We present two examples. In the first example, the data rates of the various channels are rationally related. In the second example, they are not rationally related.

Suppose first that we had to multiplex three channels, A, B, and C. Assume further that the data rates for both channels B and C are one-half that of A. That is, A is producing symbols at twice the rate of B or C. A simple multiplexing scheme would include two symbols of signal A for each symbol of B or C. Thus, the transmitted sequence could be

$$A\,B\,A\,C\,A\,B\,A\,C\,A\,B\,A\,C\,A\,B\,A\,C^{\circ}$$

This is known as *supercommutation*. It can be visualized as a rotating switch with one contact each for signal B and C and two contacts for signal A. This is shown in Fig. 7.5.

Supercommutation can be combined with *subcommutation* to yield greater flexibility. Suppose, in the preceding example, the signal C were replaced by four signals, C1, C2, C3, and C4, each of which produces symbols at 1/4 the rate of C. Thus, every time the **commutator** of Fig. 7.5 reaches the C contact, we wish to feed a sample of one of the lower rate signals. This is accomplished using subcommutation, as shown in Fig. 7.6.

If the rational relationships needed to do sub- and supercommutation are not present, we need to use a completely asynchronous multiplexing approach. Symbols from the various sources are stored in a buffer and the buffer sorts things out and transmits the interleaved symbols at the appropriate rate.

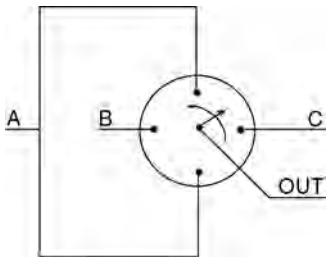


FIGURE 7.5 Supercommutation.

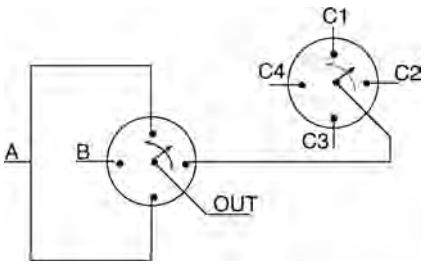


FIGURE 7.6 Combination of sub- and supercommutation.

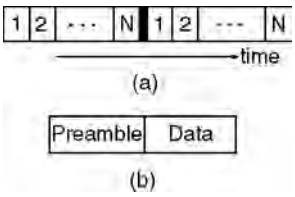


FIGURE 7.7 Typical frame for fixed-assignment TDMA.

This is a variation of the **statistical multiplexer (stat-MUX)** commonly used in packet switching and other asynchronous applications. If the individual channels are not synchronous, the symbols arrive at varying rates and times. Consider, for example, the output of a workstation connected to a network. The human operator certainly does not create symbols at a synchronous rate. The buffer in the multiplexer, therefore, must be carefully sized. If it is too small, it could become overloaded and lose information. At the other extreme, if symbols exit the MUX too rapidly, then the buffer could become empty and stuffing symbols must be added.

Time Multiplexing at Higher Layers

We have been talking about interleaving at the sample or symbol (bit) level. There are many contemporary systems that share a channel at the signal level. Thus, each source has the resources of the channel for a segment of time. This is particularly suited to bursty signals.

In **time-division multiple access (TDMA)**, the channel is shared among multiple users by assigning time slots to each user. Each source, therefore, must transmit its signal in bursts that occur only during the time allocated to that source.

In *fixed-assignment TDMA*, the time axis is divided into frames. Each frame is composed of time slots assigned to various users. Figure 7.7(a) shows a typical frame structure for N users. Two consecutive frames are illustrated. If a particular user generates a signal with a high symbol rate, that user could be assigned more than one time slot within the frame. Figure 7.7(b) shows a representative time slot.

Note that the transmission begins with a preamble which contains addressing information. Each slot must be self-synchronizing (i.e., an individual receiver needs to be able to establish symbol synchronization), and so the preamble might also contain a unique waveform to help establish synchronization. In some systems, the preamble also contains error control data (e.g., parity bits).

Sources sharing the channel need not be in the same geographical location. Therefore, as the signals line up in the channel as a frame, they experience varying time delays depending on the length of the transmission path. For this reason, guard time slots are included in the frame structure to assure that the delayed signals do not overlap in time.

Interleaving at the signal burst level includes such concepts as packet switching and asynchronous transfer mode (ATM).

7.4 Space Multiplexing

Time and frequency are the traditional domains used for multiplexing. However, there are other forms of multiplexing that force the signals to be orthogonal to each other. Among these are space and polarization multiplexing.

When terrestrial multipoint communication became popular, the general approach was to maximize the distance over which signals could be transmitted. In broadcast commercial systems, advertising revenue is related to the number of households the signal reaches. Obviously, higher powers and greater distances translate into more revenue. However, with the proliferation of high-power signal sources, frequency and time separation became increasingly difficult. Mobile radio was the driving force for a reversal of philosophy. The **cellular radio** concept is based on intentionally reducing the signal coverage area. With signal power low enough to limit range, signals that overlap in all other domains (e.g., time and frequency) can be transmitted simultaneously and separated by concentrating on the limited geographical area of coverage. To achieve communication over a distance, these low-frequency wireless techniques must be combined with more traditional long distance communication systems. The practicality of this approach had to wait for developments in electronics and computing that permitted reliable handing off among antennas as the location of the transmitter and/or receiver changes.

Space-division multiplexing can also be accomplished with highly directional antennas. Spot beam antennas can focus a signal along a particular direction, and multiple signals can be transmitted using slightly different headings. Some satellite systems (e.g., INTELSAT) divide the earth into regions and do simultaneous transmission of different signals to these regions using directional antennas.

Polarization multiplexing is yet another way to simultaneously send multiple signals on the same channel. If two signals occupy the same frequency band on the same channel yet are polarized in different planes, they can be separated at the receiver. The receiving antenna must be polarized in the same direction as the desired signal.

7.5 Techniques for Multiplexing in Spread Spectrum

We have examined three areas of nonoverlapping signals: time, frequency, and space. There are hybrid variations of these in which the signals appear to overlap in all three domains, yet are still separable. The proliferation of spread spectrum systems has opened up the domain of **code-division multiple access (CDMA)**. Spread spectrum is a transmission method that intentionally widens the signal bandwidth by modulating the bandlimited message with a wideband noiselike signal. Since the spreading function is pseudorandom (i.e., completely known although possessing noiselike properties), different widening functions can be made orthogonal to each other.

One way to visualize this is by considering the **frequency hopping** form of spread spectrum. This is a form of spread spectrum where the bandlimited message modulates a carrier whose frequency jumps around over a very wide range. The pattern of this jumping follows a noiselike sequence. However, since the jumping is completely specified, a second signal could occupy the same frequency band but jump across carrier frequencies that are always different from the first. The signals can be demodulated using

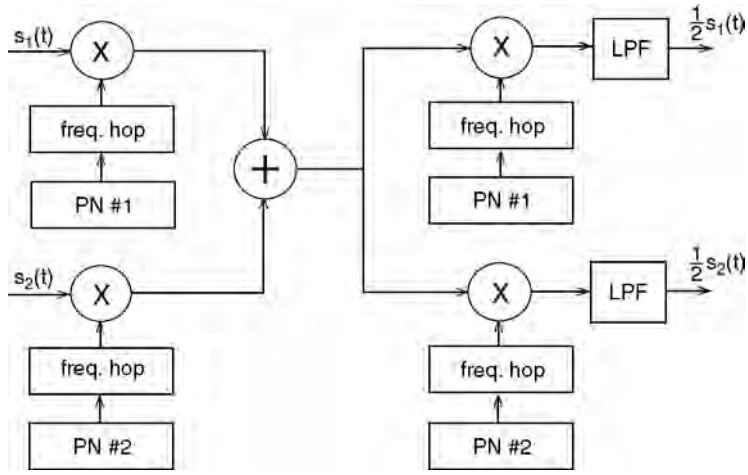


FIGURE 7.8 Code-division multiple access.

the same jumped carrier that was used in the transmitter. This is illustrated in Fig. 7.8. A pseudonoise sequence generator (labeled PN in the figure) produces a noiselike sequence. This sequence controls the frequency of an oscillator (known as a frequency hopper). The two pseudonoise sequences, PN1 and PN2, result from different initiating sequences, so their resulting frequency patterns appear unrelated.

7.6 Concluding Remarks

We have seen that signals can be forced to be nonoverlapping in a variety of domains. As it becomes possible to implement increasingly complex hardware and/or software systems, we can expect additional forms of multiplexing to be developed. The expanding applications of wireless communication are significantly increasing the volume of signals being transmitted. As competition for channel use expands at an accelerating pace, the ingenuity and imagination of engineers will be continually challenged.

Defining Terms

Asynchronous multiplexing: A technique for combining the output of various sources when those sources are producing information at unrelated rates.

Cellular radio: Each transmitting antenna uses a sufficiently low power so that the transmission range is relatively short (within a cell).

Code-division multiple access (CDMA): A technique for simultaneous transmission of wideband signals (spread spectrum) that occupy the same band of frequencies.

Commutation: Interspersing of signal pulses using the equivalent of a rotating multi-contact switch.

Frame synchronization: In time-division multiplexing, frame synchronization allows the receiver to associate pulses with the correct original sources.

Frequency-division multiplexing (FDM): Multiplexing technique that requires that signals be confined to assigned, nonoverlapping frequency bands.

Frequency hopping: Spread spectrum technique that uses carriers whose frequencies hop around in a prescribed way to create a wideband modulated signal.

Interleaving: Putting signals together in a defined sequence. Similar to shuffling a deck of cards.

Multiplexing: The process of combining signals for transmission on a single channel.

Polarization multiplexing: More than one signal can share a channel if the signals are polarized to different planes.

Space-division multiplexing: Sharing of a channel by concentrating individual signals in non-overlapping narrow beams.

Statistical multiplexer (Stat-MUX): Multiplexer that can accept asynchronous inputs and produce a synchronous output signal.

T-1 carrier: Time-division multiplexing technique for combining 24 voice channels.

Time-division multiple access (TDMA): A channel is shared among various users by assigning time slots to each user. Transmissions are bursts within the assigned time slots.

Time-division multiplexing (TDM): Multiplexing technique which requires that signal symbols be confined to assigned, non-overlapping portions of the time axis.

Wavelength-division multiplexing: Optical multiplexing systems where each signal is modulated on a different carrier frequency.

References

Couch, L. W., II, 2001, *Digital and Analog Communication Systems,* 6th ed., Prentice-Hall, Upper Saddle River, NJ.

Haykin, S., 2000, *Communication Systems,* 4th ed., John Wiley & Sons, New York.

Proakis, J. G., 2000, *Digital Communications,* McGraw-Hill, New York.

Roden, M.S., 2000, *Analog and Digital Communication Systems,* 4th ed., Discovery Press, Los Angeles, CA.

Sklar, B., 2001, *Digital Communications, Fundamentals and Applications,* 2nd ed., Prentice-Hall, Upper Saddle River, NJ.

Further Information

The basics of multiplexing are covered in any general communications textbook. The more recent state-of-the-art information can be found in the technical literature, primarily that of the IEEE. In particular, the *IEEE Transactions on Communications* and *IEEE Selected Papers on Communications* contain articles on advances in research areas. Other *IEEE Transactions* often contain related articles (e.g., the *IEEE Transactions on Vehicular Technology* contain articles related to multiplexing in mobile radio systems). The *IEEE Communications Magazine* is a highly readable periodical that covers the broad spectrum of communications in the form of tutorial papers. *IEEE Spectrum Magazine* occasionally contains articles of interest related to communications, and special issues of *IEEE Proceedings* cover communications topics.

8

Pseudonoise Sequences

Tor Helleseth
University of Bergen

P. Vijay Kumar
University of Southern California

8.1 Introduction

Pseudonoise sequences (PN sequences), also referred to as pseudorandom sequences, are sequences that are deterministically generated and yet possess some properties that one would expect to find in randomly generated sequences. Applications of PN sequences include signal synchronization, navigation, radar ranging, random number generation, spread-spectrum communications, multipath resolution, cryptography, and signal identification in multiple-access communication systems. The *correlation* between two sequences $\{x(t)\}$ and $\{y(t)\}$ is the complex inner product of the first sequence with a shifted version of the second sequence. The correlation is called (1) an autocorrelation if the two sequences are the same, (2) a crosscorrelation if they are distinct, (3) a periodic correlation if the shift is a cyclic shift, (4) an aperiodic correlation if the shift is not cyclic, and (5) a partial-period correlation if the inner product involves only a partial segment of the two sequences. More precise definitions are given subsequently.

Binary ***m* sequences**, defined in the next section, are perhaps the best-known family of PN sequences. The balance, run-distribution, and autocorrelation properties of these sequences mimic those of random sequences. It is perhaps the random-like correlation properties of PN sequences that make them most attractive in a communications system, and it is common to refer to any collection of low-correlation sequences as a family of PN sequences.

Section 8.2 begins by discussing *m* sequences. Thereafter, the discussion continues with a description of sequences satisfying various correlation constraints along the lines of the accompanying self-explanatory figure, Fig. 8.1. Expanded tutorial discussions on pseudorandom sequences may be found in [14], in [15, Chapter 5], and in [6].

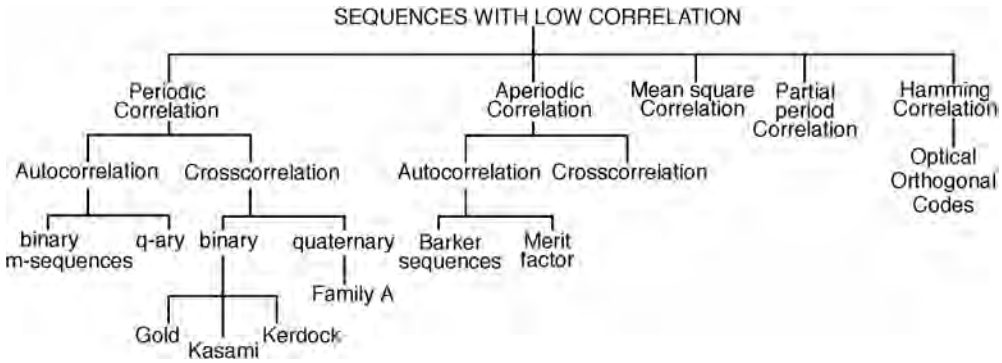


FIGURE 8.1 Overview of pseudonoise sequences.

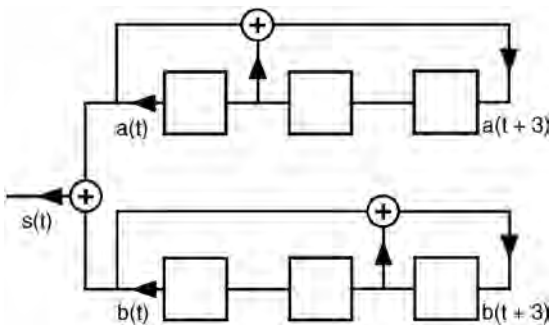


FIGURE 8.2 An example Gold sequence generator. Here $\{a(t)\}$ and $\{b(t)\}$ are m sequences of length 7.

8.2 *m* Sequences

A binary $\{0, 1\}$ **shift-register sequence** $\{s(t)\}$ is a sequence that satisfies a linear recurrence relation of the form

$$\sum_{i=0}^{r} f_i s(t+i) = 0, \quad \text{for all } t \geq 0 \tag{8.1}$$

where $r \geq 1$ is the *degree* of the recursion; the coefficients f_i belong to the finite field $GF(2) = \{0, 1\}$ where the leading coefficient $f_r = 1$. Thus, both sequences $\{a(t)\}$ and $\{b(t)\}$ appearing in Fig. 8.2 are shift-register sequences. A sequence satisfying a recursion of the form in Eq. (8.1) is said to have *characteristic polynomial* $f(x) = \sum_{i=0}^{r} f_i x^i$. Thus, $\{a(t)\}$ and $\{b(t)\}$ have characteristic polynomials given by $f(x) = x^3 + x + 1$ and $f(x) = x^3 + x^2 + 1$, respectively.

Since an r-bit binary shift register can assume a maximum of 2^r different states, it follows that every shift-register sequence $\{s(t)\}$ is eventually periodic with period $n \leq 2^r$, i.e.,

$$s(t) = s(t+n), \quad \text{for all } t \geq N$$

for some integer N. In fact, the maximum period of a shift-register sequence is $2^r - 1$, since a shift register that enters the all-zero state will remain forever in that state. The upper shift register in Fig. 8.2, when initialized with starting state 0 0 1, generates the periodic sequence $\{a(t)\}$ given by

$$0010111 \quad 0010111 \quad 0010111 \quad \cdots \tag{8.2}$$

of period $n = 7$. It follows then that this shift register generates sequences of maximal period starting from any nonzero initial state.

An m sequence is simply a binary shift-register sequence having maximal period. For every $r \geq 1$, m sequences are known to exist. The periodic **autocorrelation** function θ_s of a binary $\{0, 1\}$ sequence $\{s(t)\}$ of period n is defined by

$$\theta_s(\tau) = \sum_{t=0}^{n-1} (-1)^{s(t+\tau)-s(t)}, \quad 0 \leq \tau \leq n-1$$

An m sequence of length $2^r - 1$ has the following attributes: (1) *Balance property:* in each period of the m sequence there are 2^{r-1} ones and $2^{r-1} - 1$ zeros, (2) *Run property:* every nonzero binary s-tuple, $s \leq r$ occurs 2^{r-s} times, the all-zero s-tuple occurs $2^{r-s} - 1$ times, and (3) *Two-level autocorrelation function:*

$$\theta_s(\tau) = \begin{cases} n & \text{if } \tau = 0 \\ -1 & \text{if } \tau \neq 0 \end{cases} \tag{8.3}$$

The first two properties follow immediately from the observation that every nonzero r-tuple occurs precisely once in each period of the m sequence. For the third property, consider the difference sequence $\{s(t + \tau) - s(\tau)\}$ for $\tau \neq 0$. This sequence satisfies the same recursion as the m sequence $\{s(t)\}$ and is clearly not the all-zero sequence. It follows, therefore, that $\{s(t + \tau) - s(t) \equiv \{s(t + \tau')\}$ for some τ', $0 \leq \tau' \leq n - 1$, i.e., is a different cyclic shift of the m sequence $\{s(t)\}$. The balance property of the sequence $\{s(t + \tau')\}$ then gives us attribute 3. The m sequence $\{a(t)\}$ in Eq. (8.2) can be seen to have the three listed properties.

If $\{s(t)\}$ is any sequence of period n and d is an integer, $1 \leq d \leq n$, then the mapping $\{s(t)\} \rightarrow \{s(dt)\}$ is referred to as a *decimation* of $\{s(t)\}$ by the integer d. If $\{s(t)\}$ is an m sequence of period $n = 2^r - 1$ and d is an integer relatively prime to $2^r - 1$, then the decimated sequence $\{s(dt)\}$ clearly also has period n. Interestingly, it turns out that the sequence $\{s(dt)\}$ is always also an m sequence of the same period. For example, when $\{a(t)\}$ is the sequence in Eq. (8.2), then

$$a(3t) = 0011101 \quad 0011101 \quad 0011101 \quad \cdots \tag{8.4}$$

and

$$a(2t) = 0111001 \quad 0111001 \quad 0111001 \quad \cdots \tag{8.5}$$

The sequence $\{a(3t)\}$ is also an m sequence of period 7, since it satisfies the recursion

$$s(t + 3) + s(t + 2) + s(t) = 0 \qquad \text{for all } t$$

of degree $r = 3$. In fact $\{a(3t)\}$ is precisely the sequence labeled $\{b(t)\}$ in Fig. 8.2. The sequence $\{a(2t)\}$ is simply a cyclically shifted version of $\{a(t)\}$ itself; this property holds in general. If $\{s(t)\}$ is any m sequence of period $2^r - 1$, then $\{s(2t)\}$ will always be a shifted version of the same m sequence. Clearly, the same is true for decimations by any power of 2.

Starting from an m sequence of period $2^r - 1$, one can generate all m sequences of the same period through decimations by integers d relatively prime to $2^r - 1$. The set of integers d, $1 \leq d \leq 2^r - 1$, satisfying $(d, 2^r - 1) = 1$ forms a group under multiplication modulo $2^r - 1$, with the powers $\{2^i | \, 0 \leq i \leq r - 1\}$ of 2 forming a subgroup of order r. Since decimation by a power of 2 yields a shifted version of the same m sequence, it follows that the number of distinct m sequences of period $2^r - 1$ is $[\phi(2^r - 1)/r]$ where $\phi(n)$ denotes the number of integers d, $1 \leq d \leq n$, relatively prime to n. For example, when $r = 3$, there are just two cyclically distinct m sequences of period 7, and these are precisely the sequences $\{a(t)\}$ and $\{b(t)\}$

discussed in the preceding paragraph. Tables provided in [12] can be used to determine the characteristic polynomial of the various m sequences obtainable through the decimation of a single given m sequence. The classical reference on m sequences is [4].

If one obtains a sequence of some large length n by repeatedly tossing an unbiased coin, then such a sequence will very likely satisfy the balance, run, and autocorrelation properties of an m sequence of comparable length. For this reason, it is customary to regard the extent to which a given sequence possesses these properties as a measure of randomness of the sequence. Quite apart from this, in many applications such as signal synchronization and radar ranging, it is desirable to have sequences $\{s(t)\}$ with low autocorrelation sidelobes, i.e., $|\theta_s(\tau)|$ is small for $\tau \neq 0$. Whereas m sequences are a prime example, there exist other methods of constructing binary sequences with low out-of-phase autocorrelation.

Sequences $\{s(t)\}$ of period n having an autocorrelation function identical to that of an m sequence, i.e., having θ_s satisfying Eq. (8.3), correspond to well-studied combinatorial objects known as *cyclic Hadamard difference sets*. Known infinite families fall into three classes: (1) Singer and Gordon, Mills and Welch, (2) quadratic residue, and (3) twin-prime difference sets. These correspond, respectively, to sequences of period n of the form $n = 2^r - 1$, $r \geq 1$; n prime; and $n = p(p + 2)$ with both p and $p + 2$ being prime in the last case. For a detailed treatment of cyclic difference sets, see [2]. A recent observation by Maschietti in [9] provides additional families of cyclic Hadamard difference sets that also correspond to sequences of period $n = 2^r - 1$.

8.3 The *q*-ary Sequences with Low Autocorrelation

As defined earlier, the autocorrelation of a binary $\{0, 1\}$ sequence $\{s(t)\}$ leads to the computation of the inner product of an $\{-1, +1\}$ sequence $\{(-1)^{s(t)}\}$ with a cyclically shifted version $\{(-1)^{s(t+\tau)}\}$ of itself. The $\{-1, +1\}$ sequence is transmitted as a phase shift by either 0° and 180° of a radio-frequency carrier, i.e., using binary phase-shift keying (PSK) modulation. If the modulation is q-ary PSK, then one is led to consider sequences $\{s(t)\}$ with symbols in the set Z_q, i.e., the set of integers modulo q. The relevant autocorrelation function $\theta_s(\tau)$ is now defined by

$$\theta_s(\tau) = \sum_{t=0}^{n-1} \omega^{s(t+\tau)-s(t)}$$

where n is the period of $\{s(t)\}$ and ω is a complex primitive qth root of unity. It is possible to construct sequences $\{s(t)\}$ over Z_q whose autocorrelation function satisfies

$$\theta_s(\tau) = \begin{cases} n & \text{if} \quad \tau = 0 \\ 0 & \text{if} \quad \tau \neq 0 \end{cases}$$

For obvious reasons, such sequences are said to have an *ideal autocorrelation function*.

We provide without proof two sample constructions. The sequences in the first construction are given by

$$s(t) = \begin{cases} t^2/2 & (\text{mod } n) \quad \text{when } n \text{ is even} \\ t(t+1)/2 & (\text{mod } n) \quad \text{when } n \text{ is odd} \end{cases}$$

Thus, this construction provides sequences with ideal autocorrelation for any period n. Note that the size q of the sequence symbol alphabet equals n when n is odd and $2n$ when n is even.

The second construction also provides sequences over Z_q of period n but requires that n be a perfect square. Let $n = r^2$ and let π be an arbitrary permutation of the elements in the subset $\{0, 1, 2, \ldots, (r-1)\}$

of Z_n: Let g be an arbitrary function defined on the subset $\{0, 1, 2, \ldots, r-1\}$ of Z_n. Then any sequence of the form

$$s(t) = rt_1\pi(t_2) + g(t_2) \pmod{n}$$

where $t = rt_1 + t_2$ with $0 \le t_1, t_2 \le r-1$ is the base-r decomposition of t and has an ideal autocorrelation function. When the alphabet size q equals or divides the period n of the sequence, ideal-autocorrelation sequences also go by the name *generalized bent functions*. For details, see [6].

8.4 Families of Sequences with Low Crosscorrelation

Given two sequences $\{s_1(t)\}$ and $\{s_2(t)\}$ over Z_q of period n, their **crosscorrelation** function $\theta_{1,2}(\tau)$ is defined by

$$\theta_{1,2}(\tau) = \sum_{t=0}^{n-1} \omega^{s_1(t+\tau)-s_2(t)}$$

where ω is a primitive qth root of unity. The crosscorrelation function is important in code-division multiple-access (CDMA) communication systems. Here, each user is assigned a distinct signature sequence, and to minimize interference due to the other users, it is desirable that the signature sequences have pairwise, low values of crosscorrelation function. To provide the system in addition with a self-synchronizing capability, it is desirable that the signature sequences have low values of the autocorrelation function as well.

Let $\mathcal{F} = \{\{s_i(t)\} | 1 \le i \le M\}$ be a family of M sequences $\{s_i(t)\}$ over Z_q, each of period n. Let $\theta_{i,j}(\tau)$ denote the crosscorrelation between the ith and jth sequence at shift τ, i.e.,

$$\theta_{i,j}(\tau) = \sum_{t=0}^{n-1} \omega^{s_i(t+\tau)-s_j(t)}, \quad 0 \le \tau \le n-1$$

The classical goal in sequence design for CDMA systems has been minimization of the parameter

$$\theta_{\max} = \max\{|\theta_{i,j}(\tau)| \quad \text{either } i \ne j \text{ or } \tau \ne 0\}$$

for fixed n and M. It should be noted though that, in practice, because of data modulation, the correlations that one runs into are typically of an aperiodic rather than a periodic nature (see Section 8.5). The problem of designing for low aperiodic correlation, however, is a more difficult one. A typical approach, therefore, has been to design based on periodic correlation and then analyze the resulting design for its aperiodic correlation properties. Again, in many practical systems, the mean square correlation properties are of greater interest than the worst-case correlation represented by a parameter such as $\theta_{\max}$. The mean square correlation is discussed in Section 8.6.

Bounds on the minimum possible value of $\theta_{\max}$ for given period n, family size M, and alphabet size q are available that can be used to judge the merits of a particular sequence design. The most efficient bounds are those due to Welch, Sidelnikov, and Levenshtein, see [6]. In CDMA systems, there is greatest interest in designs in which the parameter $\theta_{\max}$ is in the range $\sqrt{n} \le \theta_{\max} \le 2\sqrt{n}$. Accordingly, Table 8.1 uses the Welch, Sidelnikov, and Levenshtein bounds to provide an order-of-magnitude upper bound on the family size M for certain $\theta_{\max}$ in the cited range.

Practical considerations dictate that q be small. The bit-oriented nature of electronic hardware makes it preferable to have q a power of 2. With this in mind, a description of some efficient sequence families having low auto- and crosscorrelation values and alphabet sizes $q = 2$ and $q = 4$ are described next.

TABLE 8.1 Bounds on Family Size M for Given n, $\theta_{\max}$

$\theta_{\max}$	Upper bound on M $q = 2$	Upper Bound on M $q > 2$
$\sqrt{n}$	$n/2$	n
$\sqrt{2n}$	n	$n^2/2$
$2\sqrt{n}$	$3n^2/10$	$n^3/2$

Gold and Kasami Sequences

Given the low autocorrelation sidelobes of an m sequence, it is natural to attempt to construct families of low correlation sequences starting from m sequences. Two of the better known constructions of this type are the families of Gold and Kasami sequences.

Let r be odd and $d = 2^k + 1$, where k, $1 \le k \le r - 1$, is an integer satisfying $(k, r) = 1$. Let $\{s(t)\}$ be a cyclic shift of an m sequence of period $n = 2^r - 1$ that satisfies $S(dt) \not\equiv 0$ and let $\mathcal{G}$ be the *Gold* family of $2^r + 1$ sequences given by

$$\mathcal{G} = \{s(t)\} \cup \{s(dt)\} \cup \{\{s(t) + s(d[t + \tau])\} | 0 \le \tau \le n - 1\}$$

Then, each sequence in $\mathcal{G}$ has period $2^r - 1$ and the maximum-correlation parameter $\theta_{\max}$ of $\mathcal{G}$ satisfies

$$\theta_{\max} \le \sqrt{2^{r+1}} + 1$$

An application of the Sidelnikov bound coupled with the information that $\theta_{\max}$ must be an odd integer yields that for the family $\mathcal{G}$, $\theta_{\max}$ is as small as it can possibly be. In this sense, the family $\mathcal{G}$ is an optimal family. We remark that these comments remain true even when d is replaced by the integer $d = 2^{2k} - 2^k + 1$ with the conditions on k remaining unchanged.

The Gold family remains the best-known family of m sequences having low crosscorrelation. Applications include the Navstar Global Positioning System whose signals are based on Gold sequences.

The family of Kasami sequences has a similar description. Let $r = 2v$ and $d = 2^v + 1$. Let $\{s(t)\}$ be a cyclic shift of an m sequence of period $n = 2^r - 1$ that satisfies $s(dt) \not\equiv 0$, and consider the family of Kasami sequences given by

$$\mathcal{K} = \{s(t)\} \cup \{\{s(t) + s(d[t + \tau])\} | 0 \le \tau \le 2^v - 2\}$$

Then the Kasami family $\mathcal{K}$ contains 2^v sequences of period $2^r - 1$. It can be shown that in this case

$$\theta_{\max} = 1 + 2^v$$

This time an application of the Welch bound and the fact that $\theta_{\max}$ is an integer shows that the Kasami family is optimal in terms of having the smallest possible value of $\theta_{\max}$ for given n and M.

Quaternary Sequences with Low Crosscorrelation

The entries in Table 8.1 suggest that nonbinary (i.e., $q > 2$) designs may be used for improved performance. A family of quaternary sequences that outperform the Gold and Kasami sequences is discussed next.

Let $f(x)$ be the characteristic polynomial of a binary m sequence of length $2^r - 1$ for some integer r. The coefficients of $f(x)$ are either 0 or 1. Now, regard $f(x)$ as a polynomial over Z_4 and form the product $(-1)^r f(x) f(-x)$. This can be seen to be a polynomial in x^2. Define the polynomial $g(x)$ of degree r by

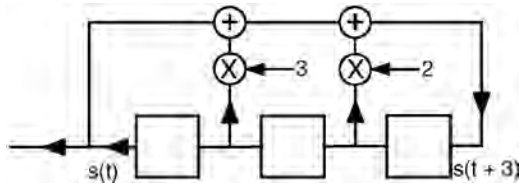


FIGURE 8.3 Shift register that generates family A quaternary sequences $\{s(t)\}$ of period 7.

setting $g(x^2) = (-1)^r f(x)f(-x)$. Let $g(x) = \sum_{i=0}^{r} g_i x^i$ and consider the set of all quaternary sequences $\{a(t)\}$ satisfying the recursion $\sum_{i=0}^{r} g_i a(t+i) = 0$ for all t.

It turns out that with the exception of the all-zero sequence, all of the sequences generated in this way have period $2^r - 1$. Thus, the recursion generates a family $\mathcal{A}$ of $2^r + 1$ cyclically distinct quaternary sequences. Closer study reveals that the maximum correlation parameter $\theta_{\max}$ of this family satisfies $\theta_{\max} \le 1 + \sqrt{2^r}$. Thus, in comparison to the family of Gold sequences, the family $\mathcal{A}$ offers a lower value of $\theta_{\max}$ (by a factor of $\sqrt{2}$) for the same family size. In comparison to the set of Kasami sequences, it offers a much larger family size for the same bound on $\theta_{\max}$. Family $\mathcal{A}$ sequences may be found discussed in [16, 3].

We illustrate with an example. Let $f(x) = x^3 + x + 1$ be the characteristic polynomial of the m sequence $\{a(t)\}$ in Eq. (8.1). Then over Z_4

$$g(x^2) = (-1)^3 f(x)f(-x) = x^6 + 2x^4 + x^2 + 3$$

so that $g(x) = x^3 + 2x^2 + x + 3$. Thus, the sequences in family $\mathcal{A}$ are generated by the recursion $s(t+3) + 2s(t+2) + s(t+1) + 3s(t) = 0 \bmod 4$. The corresponding shift register is shown in Fig. 8.3. By varying initial conditions, this shift register can be made to generate nine cyclically distinct sequences, each of length 7. In this case $\theta_{\max} \le 1 + \sqrt{8}$.

Binary Kerdock Sequences

The Gold and Kasami families of sequences are closely related to binary linear cyclic codes. It is well known in coding theory that there exist nonlinear binary codes whose performance exceeds that of the best possible linear code. Surprisingly, some of these examples come from binary codes, which are images of *linear quaternary* ($q = 4$) *codes* under the Gray map: $0 \to 00$, $1 \to 01$, $2 \to 11$, $3 \to 10$. A prime example of this is the Kerdock code, which recently has been shown to be the Gray image of a quaternary linear code. Thus, it is not surprising that the Kerdock code yields binary sequences that significantly outperform the family of Kasami sequences.

The Kerdock sequences may be constructed as follows: let $f(x)$ be the characteristic polynomial of an m sequence of period $2^r - 1$, r odd. As before, regarding $f(x)$ as a polynomial over Z_4 (which happens to have $\{0, 1\}$ coefficients), let the polynomial $g(x)$ over Z_4 be defined via $g(x^2) = -f(x)f(-x)$. [Thus, $g(x)$ is the characteristic polynomial of a family $\mathcal{A}$ sequence set of period $2^r - 1$.] Set $h(x) = -g(-x) = \sum_{i=0}^{r} h_i x^i$, and let S be the set of all Z_4 sequences satisfying the recursion $\sum_{i=0}^{r} h_i s(t+i) = 0$. Then S contains 4^r-distinct sequences corresponding to all possible distinct initializations of the shift register.

Let T denote the subset S of size 2^r consisting of those sequences corresponding to initializations of the shift register only using the symbols 0 and 2 in Z_4. Then the set $S - T$ of size $4^r - 2^r$ contains a set $\mathcal{U}$ of 2^{r-1} cyclically distinct sequences each of period $2(2^r - 1)$. Given $x = a + 2b \in Z_4$ with $a, b \in \{0, 1\}$, let μ denote the most significant bit (MSB) map $\mu(x) = b$. Let $\mathcal{K}_E$ denote the family of 2^{r-1} binary sequences obtained by applying the map μ to each sequence in $\mathcal{U}$. It turns out that each sequence in $\mathcal{U}$ also has period $2(2^r - 1)$ and that, furthermore, for the family $\mathcal{K}_E$, $\theta_{\max} \le 2 + \sqrt{2^{r+1}}$. Thus, $\mathcal{K}_E$ is a much larger family than the Kasami family, while having almost exactly the same value of $\theta_{\max}$.

For example, taking $r = 3$ and $f(x) = x^3 + x + 1$, we have from the previous family $\mathcal{A}$ example that $g(x) = x^3 + 2x^2 + x + 3$, so that $h(x) = -g(-x) = x^3 + 2x^2 + x + 1$. Applying the MSB map to the head of

the shift register and discarding initializations of the shift register involving only 0's and 2's yields a family of four cyclically distinct binary sequences of period 14. Kerdock sequences are discussed in [6,11,1,17].

8.5 Aperiodic Correlation

Let $\{x(t)\}$ and $\{y(t)\}$ be complex-valued sequences of length (or period) n, not necessarily distinct. Their *aperiodic correlation* values $\{\rho_{x,y}(\tau) \mid -(n-1) \le \tau \le n-1\}$ are given by

$$\rho_{x,y}(\tau) = \sum_{t=\max\{0,-\tau\}}^{\min\{n-1,n-1-\tau\}} x(t+\tau)y^*(t)$$

where $y^*(t)$ denotes the complex conjugate of $y(t)$. When $x \equiv y$, we will abbreviate and write ρ_x in place of $\rho_{x,y}$. The sequences described next are perhaps the most famous example of sequences with low-aperiodic autocorrelation values.

Barker Sequences

A binary $\{-1,+1\}$ sequence $\{s(t)\}$ of length n is said to be a *Barker sequence* if the aperiodic autocorrelation values $\rho_s(\tau)$ satisfy $|\rho_s(\tau)| \le 1$ for all τ, $-(n-1) \le \tau \le n-1$. The Barker property is preserved under the following transformations:

$$s(t) \to -s(t), \qquad s(t) \to (-1)^t s(t) \quad \text{and} \quad s(t) \to s(n-1-t)$$

as well as under compositions of the preceding transformations. Only the following Barker sequences are known:

$$\begin{aligned}
n = 2 &\quad ++ \\
n = 3 &\quad ++- \\
n = 4 &\quad +++- \\
n = 5 &\quad +++-+ \\
n = 7 &\quad +++--+- \\
n = 11 &\quad +++---+--+- \\
n = 13 &\quad +++++--++-+-+
\end{aligned}$$

where + denotes +1 and − denotes −1 and sequences are generated from these via the transformations already discussed. It is known that if any other Barker sequence exists, it must have length $n > 1{,}898{,}884$, a multiple of 4.

For an upper bound to the maximum out-of-phase aperiodic autocorrelation of an m sequence, see [13].

Sequences with High Merit Factor

The *merit factor* F of a $\{-1, +1\}$ sequence $\{s(t)\}$ is defined by

$$F = \frac{n^2}{2\sum_{\tau=1}^{n-1}\rho_s^2(\tau)}$$

Since $\rho_s(\tau) = \rho_s(-\tau)$ for $1 \le |\tau| \le n-1$ and $\rho_s(0) = n$, factor F may be regarded as the ratio of the square of the in-phase autocorrelation to the sum of the squares of the out-of-phase aperiodic autocorrelation values. Thus, the merit factor is one measure of the aperiodic autocorrelation properties of a binary $\{-1,+1\}$ sequence. It is also closely connected with the signal to self-generated noise ratio of a communication system in which coded pulses are transmitted and received.

Let F_n denote the largest merit factor of any binary $\{-1, +1\}$ sequence of length n. For example, at length $n = 13$, the Barker sequence of length 13 has a merit factor $F = F_{13} = 14.08$. Assuming a certain ergodicity postulate it was established by Golay that $\lim_{n\to\infty} F_n = 12.32$. Exhaustive computer searches carried out for $n \le 40$ have revealed the following.

1. For $1 \le n \le 40$, $n \ne 11, 13$,

$$3.3 \le F_n \le 9.85,$$

2. $F_{11} = 12.1$, $F_{13} = 14.08$.

The value F_{11} is also achieved by a Barker sequence. From partial searches, for lengths up to 117, the highest known merit factor is between 8 and 9.56; for lengths from 118 to 200, the best-known factor is close to 6. For lengths >200, statistical search methods have failed to yield a sequence having merit factor exceeding 5.

An *offset sequence* is one in which a fraction θ of the elements of a sequence of length n are chopped off at one end and appended to the other end, i.e., an offset sequence is a cyclic shift of the original sequence by $n\theta$ symbols. It turns out that the asymptotic merit factor of m sequences is equal to 3 and is independent of the particular offset of the m sequence. There exist offsets of sequences associated with quadratic-residue and twin-prime difference sets that achieve a larger merit factor of 6. Details may be found in [7].

Sequences with Low Aperiodic Crosscorrelation

If $\{u(t)\}$ and $\{v(t)\}$ are sequences of length $2n-1$ defined by

$$u(t) = \begin{cases} x(t) & \text{if} \quad 0 \le t \le n-1 \\ 0 & \text{if} \quad n \le t \le 2n-2 \end{cases}$$

and

$$v(t) = \begin{cases} y(t) & \text{if} \quad 0 \le t \le n-1 \\ 0 & \text{if} \quad n \le t \le 2n-2 \end{cases}$$

then

$$\{\rho_{x,y}(\tau) \mid -(n-1) \le \tau \le n-1\} = \{\theta_{u,v}(\tau) \mid 0 \le \tau \le 2n-2\} \tag{8.6}$$

Given a collection

$$U = \{\{x_i(t)\} \mid 1 \le i \le M\}$$

of sequences of length n over Z_q, let us define

$$\rho_{\max} = \max\{|\rho_{a,b}(\tau)| \mid a,b \in U, \quad \text{either } a \ne b \text{ or } \tau \ne 0\}$$

It is clear from Eq. (8.6) how bounds on the *periodic* correlation parameter $\theta_{\max}$ can be adapted to give bounds on $\rho_{\max}$. Translation of the Welch bound gives that for every integer $k \geq 1$,

$$\rho_{\max}^{2k} \geq \left(\frac{n^{2k}}{M(2n-1)-1}\right)\left\{\frac{M(2n-1)}{\binom{2n+k-2}{k}} - 1\right\}$$

Setting $k = 1$ in the preceding bound gives

$$\rho_{\max} \geq n\sqrt{\frac{M-1}{M(2n-1)-1}}$$

Thus, for fixed M and large n, Welch's bound gives

$$\rho_{\max} \geq \mathcal{O}(n^{1/2})$$

There exist sequence families which asymptotically achieve $\rho_{\max} \approx \mathcal{O}(n^{1/2})$, [10].

8.6 Other Correlation Measures

Partial-Period Correlation

The *partial-period* (*p-p*) *correlation* between the sequences $\{u(t)\}$ and $\{v(t)\}$ is the collection $\{\Delta_{u,v}(l, \tau, t_0) | 1 \leq l \leq n, 0 \leq \tau \leq n-1, 0 \leq t_0 \leq n-1\}$ of inner products

$$\Delta_{u,v}(l, \tau, t_0) = \sum_{t=t_0}^{t=t_0+l-1} u(t+\tau)v^*(t)$$

where l is the length of the partial period and the sum $t + \tau$ is again computed modulo n.

In direct-sequence CDMA systems, the pseudorandom signature sequences used by the various users are often very long for reasons of data security. In such situations, to minimize receiver hardware complexity, correlation over a partial period of the signature sequence is often used to demodulate data, as well as to achieve synchronization. For this reason, the p-p correlation properties of a sequence are of interest.

Researchers have attempted to determine the moments of the p-p correlation. Here the main tool is the application of the Pless power-moment identities of coding theory [8]. The identities often allow the first and second p-p correlation moments to be completely determined. For example, this is true in the case of m sequences (the remaining moments turn out to depend upon the specific characteristic polynomial of the m sequence). Further details may be found in [15].

Mean Square Correlation

Frequently in practice, there is a greater interest in the mean-square correlation distribution of a sequence family than in the parameter $\theta_{\max}$. Quite often in sequence design, the sequence family is derived from a linear, binary cyclic code of length n by picking a set of cyclically distinct sequences of period n. The families of Gold and Kasami sequences are so constructed. In this case, as pointed out by Massey, the mean square correlation of the family can be shown to be either optimum or close to optimum, under certain easily satisfied conditions, imposed on the minimum distance of the dual code. A similar situation holds even when the sequence family does not come from a linear cyclic code. In this sense, mean square

correlation is not a very discriminating measure of the correlation properties of a family of sequences. An expanded discussion of this issue may be found in [5].

Optical Orthogonal Codes

Given a pair of $\{0, 1\}$ sequences $\{s_1(t)\}$ and $\{s_2(t)\}$ each having period n, we define the *Hamming correlation* function $\theta_{12}(\tau)$, $0 \le \tau \le n-1$, by

$$\theta_{12}(\tau) = \sum_{t=0}^{n-1} s_1(t+\tau)s_2(t)$$

Such correlations are of interest, for instance, in optical communication systems where the 1s and 0s in a sequence correspond to the presence or absence of pulses of transmitted light.

An (n, w, λ) optical orthogonal code (OOC) is a family $\mathcal{F} = \{\{s_i(t)\} \mid i = 1, 2, \ldots, M\}$, of M $\{0, 1\}$ sequences of period n and constant Hamming weight w, where w is an integer lying between 1 and $n-1$ satisfying $\theta_{ij}(\tau) \le \lambda$ whenever either $i \ne j$ or $\tau \ne 0$.

Note that the Hamming distance $d_{a,b}$ between a period of the corresponding codewords $\{a(t)\}$, $\{b(t)\}$, $0 \le t \le n-1$, in an (n, w, λ) OOC having Hamming correlation ρ, $0 \le \rho \le \lambda$, is given by $d_{a,b} = 2(w-\rho)$, and, thus, OOCs are closely related to constant-weight error correcting codes. Given an (n, w, λ) OOC, by enlarging the OOC to include every cyclic shift of each sequence in the code, one obtains a constant-weight, minimum distance $d_{\min} \ge 2(w-\lambda)$ code. Conversely, given a constant-weight cyclic code of length n, weight w, and minimum distance $d_{\min}$, one can derive an (n, w, λ) OOC code with $\lambda \le w - d_{\min}/2$ by partitioning the code into cyclic equivalence classes and then picking precisely one representative from each equivalence class of size n.

By making use of this connection, one can derive bounds on the size of an OOC from known bounds on the size of constant-weight codes. The bound given next follows directly from the Johnson bound for constant weight codes [8]. The number $M(n, w, \lambda)$ of codewords in a (n, w, λ) OOC satisfies

$$M(n, w, \lambda) \le \frac{1}{w}\left\lfloor \frac{n-1}{w-1} \cdots \left\lfloor \frac{n-\lambda+1}{w-\lambda+1}\left\lfloor \frac{n-\lambda}{w-\lambda}\right\rfloor\right\rfloor \cdots \right\rfloor$$

An OOC code that achieves the Johnson bound is said to be optimal. A family $\{\mathcal{F}_n\}$ of OOCs indexed by the parameter n and arising from a common construction is said to be asymptotically optimum if

$$\lim_{n\to\infty} \frac{|\mathcal{F}_n|}{M(n, w, \lambda)} = 1$$

Constructions for optical orthogonal codes are available for the cases when $\lambda = 1$ and $\lambda = 2$. For larger values of λ, there exist constructions which are asymptotically optimum. Further details may be found in [6].

Defining Terms

Autocorrelation of a sequence: The complex inner product of the sequence with a shifted version itself.

Crosscorrelation of two sequences: The complex inner product of the first sequence with a shifted version of the second sequence.

***m* Sequence:** A periodic binary $\{0, 1\}$ sequence that is generated by a shift register with linear feedback and which has maximal possible period given the number of stages in the shift register.

Pseudonoise sequences: Also referred to as pseudorandom sequences (PN), these are sequences that are deterministically generated and yet possess some properties that one would expect to find in randomly generated sequences.

Shift-register sequence: A sequence with symbols drawn from a field, which satisfies a linear-recurrence relation and which can be implemented using a shift register.

References

1. Barg, A., On small families of sequences with low periodic correlation, *Lecture Notes in Computer Science,* 781, 154–158, Springer-Verlag, Berlin, 1994.
2. Baumert, L.D., *Cyclic Difference Sets,* Lecture Notes in Mathematics 182, Springer–Verlag, New York, 1971.
3. Boztaş, S., Hammons, R., and Kumar, P.V. 4-phase sequences with near-optimum correlation properties, *IEEE Trans. Inform. Theory,* IT-38, 1101–1113, 1992.
4. Golomb, S.W., *Shift Register Sequences,* Aegean Park Press, San Francisco, CA, 1982.
5. Hammons, A.R., Jr. and Kumar, P.V., On a recent 4-phase sequence design for CDMA, *IEICE Trans. Commun.,* E76-B(8), 1993.
6. Helleseth, T. and Kumar, P.V. (planned), Sequences with low correlation, In *Handbook of Coding Theory,* V.S. Pless and W.C. Huffman, Eds., Elsevier Science Publishers, Amsterdam, 1998.
7. Jensen, J.M., Jensen, H.E., and Høholdt, T., The merit factor of binary sequences related to difference sets, *IEEE Trans. Inform. Theory,* IT-37(May), 617–626, 1991.
8. MacWilliams, F.J. and Sloane, N.J.A., *The Theory of Error-Correcting Codes,* North-Holland, Amsterdam, 1977.
9. Maschietti, A., Difference sets and hyperovals, *Designs, Codes and Cryptography,* 14, 89–98, 1998.
10. Mow, W.H., On McEliece's open problem on minimax aperiodic correlation, In *Proc. IEEE Intern. Symp. Inform. Theory,* 75, 1994.
11. Nechaev, A., The Kerdock code in a cyclic form, *Discrete Math. Appl.,* 1, 365–384, 1991.
12. Peterson, W.W. and Weldon, E.J., Jr., *Error-Correcting Codes,* 2nd ed., MIT Press, Cambridge, MA, 1972.
13. Sarwate, D.V., An upper bound on the aperiodic autocorrelation function for a maximal-length sequence, *IEEE Trans. Inform. Theory,* IT-30(July), 685–687, 1984.
14. Sarwate, D.V. and Pursley, M.B., Crosscorrelation properties of pseudorandom and related sequences, *Proc. IEEE,* 68(May), 593–619, 1980.
15. Simon, M.K.,Omura, J.K., Scholtz, R.A., and Levitt, B.K., *Spread Spectrum Communications Handbook,* revised ed., McGraw-Hill, New York, 1994.
16. Solé, P., A quaternary cyclic code and a family of quadriphase sequences with low correlation properties, *Coding Theory and Applications, Lecture Notes in Computer Science,* 388, 193–201, Berlin, Springer-Verlag, 1989.
17. Udaya, P. and Siddiqi, M., Optimal biphase sequences with large linear complexity derived from sequences over Z_4, *IEEE Trans. Inform. Theory,* IT-42 (Jan), 206–216, 1996.

Further Information

A more in-depth treatment of pseudonoise sequences may be found in the following.

1. Golomb, S.W., *Shift Register Sequences,* Aegean Park Press, San Francisco, 1982.
2. Helleseth, T. and Kumar, P.V., Sequences with Low Correlation, in *Handbook of Coding Theory,* edited by V.S. Pless and W.C. Huffman, Elsevier Science Publishers, Amsterdam, 1998 (planned).
3. Sarwate, D.V. and Pursley, M.B., Crosscorrelation Properties of Pseudorandom and Related Sequences, *Proc. IEEE,* 68, May, 593–619, 1980.
4. Simon, M.K., Omura, J.K., Scholtz, R.A., and Levitt, B.K., *Spread Spectrum Communications Handbook,* revised ed., McGraw Hill, New York, 1994.

9

D/A and A/D Converters

Susan A.R. Garrod
Purdue University

Digital-to-analog (D/A) conversion is the process of converting digital codes into a continuous range of analog signals. *Analog-to-digital (A/D) conversion* is the complementary process of converting a continuous range of analog signals into digital codes. Such conversion processes are necessary to interface real-world systems, which typically monitor continuously varying analog signals, with digital systems that process, store, interpret, and manipulate the analog values.

D/A and A/D applications have evolved from predominately military-driven applications to consumer-oriented applications. Up to the mid-1980s, the military applications determined the design of many D/A and A/D devices. The military applications required very high performance coupled with hermetic packaging, radiation hardening, shock and vibration testing, and military specification and record keeping. Cost was of little concern, and "low power" applications required approximately 2.8 W. The major applications up to the mid-1980s included military radar warning and guidance systems, digital oscilloscopes, medical imaging, infrared systems, and professional video.

The applications currently requiring D/A and A/D circuits have different performance criteria from those of earlier years. In particular, low power and high speed applications are driving the development of D/A and A/D circuits, as the devices are used extensively in battery-operated consumer products. The predominant applications include cellular telephones, hand-held camcorders, portable computers, and set-top cable TV boxes. These applications generally have low power and long battery life requirements, or they may have high speed and high resolution requirements, as is the case with the set-top cable TV boxes.

9.1 D/A and A/D Circuits

D/A and A/D conversion circuits are available as integrated circuits (ICs) from many manufacturers. A huge array of ICs exists, consisting of not only the D/A or A/D conversion circuits, but also closely related circuits such as sample-and-hold amplifiers, analog multiplexers, voltage-to-frequency and frequency-to-voltage converters, voltage references, calibrators, operation amplifiers, isolation amplifiers, instrumentation amplifiers, active filters, DC-to-DC converters, analog interfaces to digital signal processing systems, and data acquisition subsystems. Data books from the IC manufacturers contain an enormous amount of information about these devices and their applications to assist the design engineer.

TABLE 9.1 D/A and A/D Integrated Circuits

D/A Converter ICs	Resolution, b	Multiplying vs Fixed Reference	Settling Time, ms	Input Data Format
Analog devices AD558	8	Fixed reference	3	Parallel
Analog devices AD7524	8	Multiplying	0.400	Parallel
Analog devices AD390	Quad, 12	Fixed reference	8	Parallel
Analog devices AD1856	16	Fixed reference	1.5	Serial
Burr–Brown DAC729	18	Fixed reference	8	Parallel
DATEL DACHF8	8	Multiplying	0.025	Parallel
National DAC0800	8	Multiplying	0.1	Parallel
A/D Converter ICs	**Resolution, b**	**Signal Inputs**	**Conversion Speed, ms**	**Output Data Format**
Analog devices AD572	12	1	25	Serial and Parallel
Burr–Brown ADC803	12	1	1.5	Parallel
Burr–Brown ADC701	16	1	1.5	Parallel
National ADC1005B	10	1	50	Parallel
TI, National ADC0808	8	8	100	Parallel
TI, National ADC0834	8	4	32	Serial
TI TLC0820	8	1	1	Parallel
TI TLC1540	10	11	21	Serial
Interface ICs A/D and D/A	**Resolution, b**	**Onboard Filters**	**Sampling Rate, kHz**	**Data Format**
TI TLC32040	14	Yes	19.2 (programmable)	Serial
TI 2914 PCM codec and filter	8	Yes	8	Serial

The ICs discussed in this chapter will be strictly the D/A and A/D conversion circuits. Table 9.1 lists a small sample of the variety of the D/A and A/D converters currently available. The ICs usually perform either D/A or A/D conversion. There are serial interface ICs, however, typically for high-performance audio and digital signal processing applications, that perform both A/D and D/A processes.

D/A and A/D Converter Performance Criteria

The major factors that determine the quality of performance of D/A and A/D converters are *resolution, sampling rate, speed*, and *linearity*.

The *resolution* of a D/A circuit is the smallest change in the output analog signal. In an A/D system, the resolution is the smallest change in voltage that can be detected by the system and that can produce a change in the digital code. The resolution determines the total number of digital codes, or *quantization levels,* that will be recognized or produced by the circuit.

The *resolution* of a D/A or A/D IC is usually specified in terms of the bits in the digital code or in terms of the least significant bit (LSB) of the system. An *n*-bit code allows for 2^n quantization levels, or $2^n - 1$ steps between quantization levels. As the number of bits increases, the step size between quantization levels decreases, therefore increasing the accuracy of the system when a conversion is made between an analog and digital signal. The system resolution can be specified also as the voltage step size between quantization levels. For A/D circuits, the resolution is the smallest input voltage that is detected by the system.

The *speed* of a D/A or A/D converter is determined by the time it takes to perform the conversion process. For D/A converters, the speed is specified as the *settling time*. For A/D converters, the speed is specified as the *conversion time*. The settling time for D/A converters will vary with supply voltage and transition in the digital code; thus, it is specified in the data sheet with the appropriate conditions stated.

A/D converters have a maximum *sampling rate* that limits the speed at which they can perform continuous conversions. The sampling rate is the number of times per second that the analog signal can be sampled and converted into a digital code. For proper A/D conversion, the minimum sampling rate must be at least two times the highest frequency of the analog signal being sampled to satisfy the Nyquist sampling criterion. The conversion speed and other timing factors must be taken into consideration to determine the maximum sampling rate of an A/D converter. **Nyquist A/D converters** use a sampling rate that is slightly more than twice the highest frequency in the analog signal. **Oversampling A/D converters** use sampling rates of *N* times rate, where *N* typically ranges from 2 to 64.

Both D/A and A/D converters require a voltage reference in order to achieve absolute conversion accuracy. Some conversion ICs have internal voltage references, whereas others accept external voltage references. For high-performance systems, an external precision reference is needed to ensure long-term stability, load regulation, and control over temperature fluctuations. External precision voltage reference ICs can be found in manufacturers' data books.

Measurement accuracy is specified by the converter's *linearity*. *Integral linearity* is a measure of linearity over the entire conversion range. It is often defined as the deviation from a straight line drawn between the endpoints and through zero (or the offset value) of the conversion range. Integral linearity is also referred to as *relative accuracy*. The *offset* value is the reference level required to establish the zero or midpoint of the conversion range. *Differential linearity* is the linearity between code transitions. Differential linearity is a measure of the *monotonicity* of the converter. A converter is said to be monotonic if increasing input values result in increasing output values.

The accuracy and linearity values of a converter are specified in the data sheet in units of the LSB of the code. The linearity can vary with temperature, and so the values are often specified at +25°C as well as over the entire temperature range of the device.

D/A Conversion Processes

Digital codes are typically converted to analog voltages by assigning a voltage weight to each bit in the digital code and then summing the voltage weights of the entire code. A general D/A converter consists of a network of precision resistors, input switches, and level shifters to activate the switches to convert a digital code to an analog current or voltage. D/A ICs that produce an analog current output usually have a faster settling time and better linearity than those that produce a voltage output. When the output current is available, the designer can convert this to a voltage through the selection of an appropriate output amplifier to achieve the necessary response speed for the given application.

D/A converters commonly have a fixed or variable reference level. The reference level determines the switching threshold of the precision switches that form a controlled impedance network, which in turn controls the value of the output signal. **Fixed reference D/A converters** produce an output signal that is proportional to the digital input. **Multiplying D/A converters** produce an output signal that is proportional to the product of a varying reference level times a digital code.

D/A converters can produce bipolar, positive, or negative polarity signals. A four-quadrant multiplying D/A converter allows both the reference signal and the value of the binary code to have a positive or negative polarity. The four-quadrant multiplying D/A converter produces bipolar output signals.

D/A Converter ICs

Most D/A converters are designed for general-purpose control applications. Some D/A converters, however, are designed for special applications, such as video or graphic outputs, high-definition video displays, ultra high-speed signal processing, digital video tape recording, digital attenuators, or high-speed function generators.

D/A converter ICs often include special features that enable them to be interfaced easily to microprocessors or other systems. Microprocessor control inputs, input latches, buffers, input registers, and compatibility to standard logic families are features that are readily available in D/A ICs. In addition, the ICs usually have laser-trimmed precision resistors to eliminate the need for user trimming to achieve full-scale performance.

A/D Conversion Processes

Analog signals can be converted to digital codes by many methods, including integration, **successive approximation**, parallel (flash) conversion, **delta modulation**, **pulse code modulation**, and **sigma–delta conversion**. Two of the most common A/D conversion processes are successive approximation A/D conversion and parallel or **flash A/D** conversion. Very high-resolution digital audio or video systems require specialized A/D techniques that often incorporate one of these general techniques as well as specialized A/D conversion processes. Examples of specialized A/D conversion techniques are pulse code modulation (PCM) and sigma–delta conversion. PCM is a common voice encoding scheme used not only by the audio industry in digital audio recordings but also by the telecommunications industry for voice encoding and multiplexing. Sigma–delta conversion is an oversampling A/D conversion where signals are sampled at very high frequencies. It has very high resolution and low distortion and is being used in the digital audio recording industry.

Successive approximation A/D conversion is a technique that is commonly used in medium- to high-speed data acquisition applications. It is one of the fastest A/D conversion techniques that requires a minimum amount of circuitry. The conversion times for successive approximation A/D conversion typically range from 10 to 300 ms for 8-b systems.

The successive approximation A/D converter can approximate the analog signal to form an n-bit digital code in n steps. The successive approximation register (SAR) individually compares an analog input voltage to the midpoint of one of n ranges to determine the value of 1 b. This process is repeated a total of n times, using n ranges, to determine the n bits in the code. The comparison is accomplished as follows. The SAR determines if the analog input is above or below the midpoint and sets the bit of the digital code accordingly. The SAR assigns the bits beginning with the most significant bit. The bit is set to a 1 if the analog input is greater than the midpoint voltage, or it is set to a 0 if it is less than the midpoint voltage. The SAR then moves to the next bit and sets it to a 1 or a 0 based on the results of comparing the analog input with the midpoint of the next allowed range. Because the SAR must perform one approximation for each bit in the digital code, an n-bit code requires n approximations.

A successive approximation A/D converter consists of four functional blocks, as shown in Fig. 9.1: the SAR, the analog comparator, a D/A converter, and a clock.

Parallel or flash A/D conversion is used in high-speed applications such as video signal processing, medical imaging, and radar detection systems. A flash A/D converter simultaneously compares the input analog voltage to $2^n - 1$ threshold voltages to produce an n-bit digital code representing the analog voltage. Typical flash A/D converters with 8-b resolution operate at 20–100 MHz.

The functional blocks of a flash A/D converter are shown in Fig. 9.2. The circuitry consists of a precision resistor ladder network, 2^n-1 analog comparators, and a digital priority encoder. The resistor network establishes threshold voltages for each allowed quantization level. The analog comparators indicate whether the input analog voltage is above or below the threshold at each level. The output of the analog comparators is input to the digital priority encoder. The priority encoder produces the final digital output code that is stored in an output latch.

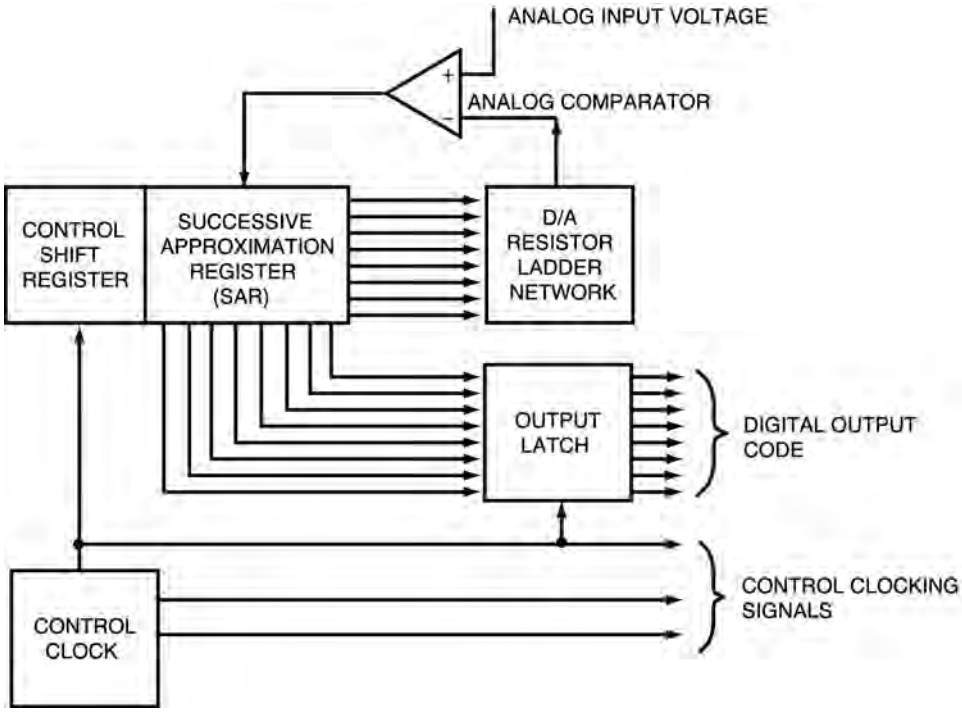


FIGURE 9.1 Successive approximation A/D converter block diagram. (*Source:* Garrod, S. and Borns, R. 1991, *Digital Logic: Analysis, Application, and Design,* p. 919 Copyright " 1991 by Saunders College Publishing, Philadelphia, PA. Reprinted with permission of the publisher.)

An 8-b flash A/D converter requires 255 comparators. The cost of high-resolution A/D comparators escalates as the circuit complexity increases and as the number of analog converters rises by 2^n-1. As a low-cost alternative, some manufacturers produce modified flash A/D converters that perform the A/D conversion in two steps to reduce the amount of circuitry required. These modified flash A/D converters are also referred to as *half-flash* A/D converters, since they perform only half of the conversion simultaneously.

A/D Converter ICs

A/D converter ICs can be classified as general-purpose, high-speed, flash, and sampling A/D converters. The *general-purpose A/D converters* are typically low speed and low cost, with conversion times ranging from 2 ms to 33 ms. A/D conversion techniques used by these devices typically include successive approximation, tracking, and integrating. The general-purpose A/D converters often have control signals for simplified microprocessor interfacing. These ICs are appropriate for many process control, industrial, and instrumentation applications, as well as for environmental monitoring such as seismology, oceanography, meteorology, and pollution monitoring.

High-speed A/D converters have conversion times typically ranging from 400 ns to 3 ms. The higher speed performance of these devices is achieved by using the successive approximation technique, modified flash techniques, and statistically derived A/D conversion techniques. Applications appropriate for these A/D ICs include fast Fourier transform (FFT) analysis, radar digitization, medical instrumentation, and multiplexed data acquisition. Some ICs have been manufactured with an extremely high degree of linearity to be appropriate for specialized applications in digital spectrum analysis, vibration analysis, geological research, sonar digitizing, and medical imaging.

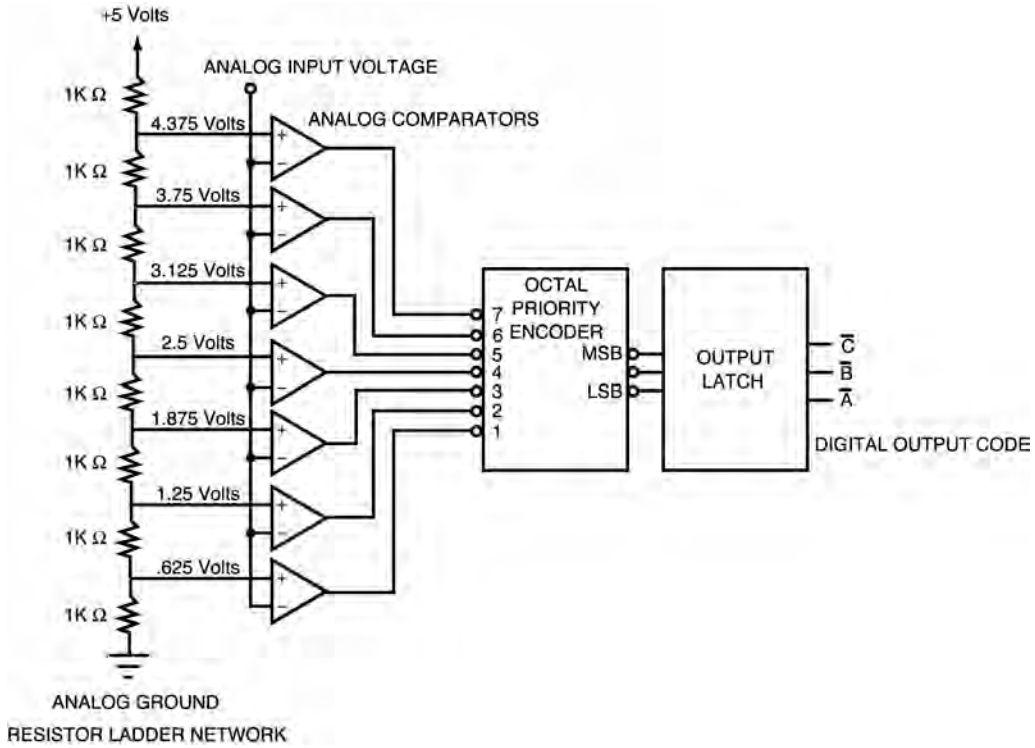


FIGURE. 9.2 Flash A/D converter block diagram. (*Source:* Garrod, S. and Borns, R., *Digital Logic: Analysis, Application, and Design,* p. 928. Copyright " 1991 by Saunders College Publishing, Philadelphia, PA. Reprinted with permission of the publisher.)

Flash A/D converters have conversion times ranging typically from 10 to 50 ns. Flash A/D conversion techniques enable these ICs to be used in many specialized high-speed data acquisition applications such as TV-video digitizing (encoding), radar analysis, transient analysis, high-speed digital oscilloscopes, medical ultrasound imaging, high-energy physics, and robotic vision applications.

Sampling A/D converters have a sample-and-hold amplifier circuit built into the IC. This eliminates the need for an external sample-and-hold circuit. The throughput of these A/D converter ICs ranges typically from 35 kHz to 100 MHz. The speed of the system is dependent on the A/D technique used by the sampling A/D converter.

A/D converter ICs produce digital codes in a serial or parallel format, and some ICs offer the designer both formats. The digital outputs are compatible with standard logic families to facilitate interfacing to other digital systems. In addition, some A/D converter ICs have a built-in analog multiplexer and therefore can accept more than one analog input signal.

Pulse code modulation (PCM) ICs are high-precision A/D converters. The PCM IC is often referred to as a PCM *codec* with both encoder and decoder functions. The encoder portion of the codec performs the A/D conversion, and the decoder portion of the codec performs the D/A conversion. The digital code is usually formatted as a serial data stream for ease of interfacing to digital transmission and multiplexing systems.

PCM is a technique whereby an analog signal is sampled, quantized, and then encoded as a digital word. The PCM IC can include successive approximation techniques or other techniques to accomplish the PCM encoding. In addition, the PCM codec may employ nonlinear data compression techniques, such as companding, if it is necessary to minimize the number of bits in the output digital code. Companding is a logarithmic technique used to compress a code to fewer bits before transmission. The inverse

logarithmic function is then used to expand the code to its original number of bits before converting it to the analog signal. Companding is typically used in telecommunications transmission systems to minimize data transmission rates without degrading the resolution of low-amplitude signals. Two standardized companding techniques are used extensively: A-law and m-law. The A-law companding is used in Europe, whereas the m-law is used predominantly in the United States and Japan. Linear PCM conversion is used in high-fidelity audio systems to preserve the integrity of the audio signal throughout the entire analog range.

Digital signal processing (DSP) techniques provide another type of A/D conversion ICs. Specialized A/D conversion such as *adaptive differential pulse code modulation* (ADPCM), sigma–delta modulation, *speech subband encoding, adaptive predictive speech encoding*, and *speech recognition* can be accomplished through the use of DSP systems. Some DSP systems require analog front ends that employ traditional PCM codec ICs or DSP interface ICs. These ICs can interface to a digital signal processor for advanced A/D applications. Some manufacturers have incorporated DSP techniques on board the single-chip A/D IC, as in the case of the DSP56ACD16 sigma–delta modulation IC by Motorola.

Integrating A/D converters are used for conversions that must take place over a long period of time, such as digital voltmeter applications or sensor applications such as thermocouples. The integrating A/D converter produces a digital code that represents the average of the signal over time. Noise is reduced by means of signal averaging, or integration. Dual-lope integration is accomplished by a counter that advances while an input voltage charges a capacitor in a specified time interval, *T*. This is compared to another count sequence that advances while a reference voltage is discharging across the same capacitor in a time interval, *delta t*. The ratio of the charging count value to the discharging count value is proportional to the ratio of the input voltage to the reference voltage. Hence, the integrating converter provides a digital code that is a measure of the input voltage averaged over time. The conversion accuracy is independent of the capacitor and the clock frequency since they affect both the charging and discharging operations. The charging period, T, is selected to be the period of the fundamental frequency to be rejected. The maximum conversion rate is slightly less than 1/(2T) conversions per second. While this limits the conversion rate to be too slow for high-speed data acquisition applications, it is appropriate for long-duration applications of slowly varying input signals.

Grounding and Bypassing on D/A and A/D ICs

D/A and A/D converter ICs require correct grounding and capacitive bypassing in order to operate according to performance specifications. The digital signals can severely impair analog signals. To combat the electromagnetic interference induced by the digital signals, the analog and digital grounds should be kept separate and should have only one common point on the circuit board. If possible, this common point should be the connection to the power supply.

Bypass capacitors are required at the power connections to the IC, the reference signal inputs, and the analog inputs to minimize noise that is induced by the digital signals. Each manufacturer specifies the recommended bypass capacitor locations and values in the data sheet. The 1-mF tantalum capacitors are commonly recommended, with additional high-frequency power supply decoupling sometimes being recommended through the use of ceramic disc shunt capacitors. The manufacturers' recommendations should be followed to ensure proper performance.

Selection Criteria for D/A and A/D Converter ICs

Hundreds of D/A and A/D converter ICs are available, with prices ranging from a few dollars to several hundred dollars each. The selection of the appropriate type of converter is based on the application requirements of the system, the performance requirements, and cost. The following issues should be considered in order to select the appropriate converter.

1. What are the input and output requirements of the system? Specify all signal current and voltage ranges, logic levels, input and output impedances, digital codes, data rates, and data formats.

2. What level of accuracy is required? Determine the resolution needed throughout the analog voltage range, the dynamic response, the degree of linearity, and the number of bits encoding.
3. What speed is required? Determine the maximum analog input frequency for sampling in an A/D system, the number of bits for encoding each analog signal, and the rate of change of input digital codes in a D/A system.
4. What is the operating environment of the system? Obtain information on the temperature range and power supply to select a converter that is accurate over the operating range.

Final selection of D/A and A/D converter ICs should be made by consulting manufacturers to obtain their technical specifications of the devices. Major manufacturers of D/A and A/D converters include Analog Devices, Burr–Brown, DATEL, Maxim, National, Phillips Components, Precision Monolithics, Signetics, Sony, Texas Instruments, Ultra Analog, and Yamaha. Information on contacting these manufacturers and others can be found in an IC *Master Catalog*.

Defining Terms

Companding: A process designed to minimize the transmission bit rate of a signal by compressing it prior to transmission and expanding it upon reception. It is a rudimentary "data compression" technique that requires minimal processing.

Delta modulation: An A/D conversion process where the digital output code represents the change, or slope, of the analog input signal, rather than the absolute value of the analog input signal. A 1 indicates a rising slope of the input signal. A 0 indicates a falling slope of the input signal. The sampling rate is dependent on the derivative of the signal, since a rapidly changing signal would require a rapid sampling rate for acceptable performance.

Fixed reference D/A converter: The analog output is proportional to a fixed (nonvarying) reference signal.

Flash A/D: The fastest A/D conversion process available to date, also referred to as parallel A/D conversion. The analog signal is simultaneously evaluated by $2^n - 1$ comparators to produce an n-bit digital code in one step. Because of the large number of comparators required, the circuitry for flash A/D converters can be very expensive. This technique is commonly used in digital video systems.

Integrating A/D: The analog input signal is integrated over time to produce a digital signal that represents the area under the curve, or the integral.

Multiplying D/A: A D/A conversion process where the output signal is the product of a digital code multiplied by an analog input reference signal. This allows the analog reference signal to be scaled by a digital code.

Nyquist A/D converters: A/D converters that sample analog signals that have a maximum frequency that is less than the Nyquist frequency. The Nyquist frequency is defined as one-half of the sampling frequency. If a signal has frequencies above the Nyquist frequency, a distortion called *aliasing* occurs. To prevent aliasing, an *antialiasing filter* with a flat passband and very sharp rolloff is required.

Oversampling converters: A/D converters that sample frequencies at a rate much higher than the Nyquist frequency. Typical oversampling rates are 32 and 64 times the sampling rate that would be required with the Nyquist converters.

Pulse code modulation (PCM): An A/D conversion process requiring three steps: the analog signal is sampled, quantized, and encoded into a fixed length digital code. This technique is used in many digital voice and audio systems. The reverse process reconstructs an analog signal from the PCM code. The operation is very similar to other A/D techniques, but specific PCM circuits are optimized for the particular voice or audio application.

Sigma–delta A/D conversion: An *oversampling* A/D conversion process where the analog signal is sampled at rates much higher (typically 64 times) than the sampling rates that would be required with a Nyquist converter. Sigma–delta modulators integrate the analog signal before performing

the delta modulation. The integral of the analog signal is encoded rather than the change in the analog signal, as is the case for traditional delta modulation. A digital sample rate reduction filter (also called a digital decimation filter) is used to provide an output sampling rate at twice the Nyquist frequency of the signal. The overall result of oversampling and digital sample rate reduction is greater resolution and less distortion compared to a Nyquist converter process.

Successive approximation: An A/D conversion process that systematically evaluates the analog signal in *n* steps to produce an *n*-bit digital code. The analog signal is successively compared to determine the digital code, beginning with the determination of the most significant bit of the code.

References

Analog Devices. 1989. *Analog Devices Data Conversion Products Data Book.* Analog Devices, Inc., Norwood, MA.

Burr–Brown. 1989. *Burr–Brown Integrated Circuits Data Book*. Burr–Brown, Tucson, AZ.

DATEL. 1988. *DATEL Data Conversion Catalog*. DATEL, Inc., Mansfield, MA.

Drachler, W. and Murphy, B. 1995. New High-Speed, Low-Power Data-Acquisition ICs, *Analog Dialogue* 29(2):3–6. Analog Devices, Inc., Norwood, MA.

Garrod, S. and Borns, R. 1991. *Digital Logic: Analysis, Application and Design,* Chap. 16. Saunders College Publishing, Philadelphia PA.

Jacob, J.M. 1989. *Industrial Control Electronics,* Chap. 6. Prentice–Hall, Englewood Cliffs, NJ.

Keiser, B. and Strange, E. 1995. *Digital Telephony and Network Integration,* 2nd ed. Van Nostrand Reinhold, New York.

Motorola. 1989. *Motorola Telecommunications Data Book.* Motorola, Inc., Phoenix, AZ.

National Semiconductor. 1989. *National Semiconductor Data Acquisition Linear Devices Data Book.* National Semiconductor Corp., Santa Clara, CA.

Park, S. 1990. *Principles of Sigma–Delta Modulation for Analog-to-Digital Converters.* Motorola, Inc., Phoenix, AZ.

Texas Instruments. 1986. *Texas Instruments Digital Signal Processing Applications with the TMS320 Family.* Texas Instruments, Dallas, TX.

Texas Instruments. 1989. *Texas Instruments Linear Circuits Data Acquisition and Conversion Data Book.* Texas Instruments, Dallas, TX.

Further Information

Analog Devices, Inc. has edited or published several technical handbooks to assist design engineers with their data acquisition system requirements. These references should be consulted for extensive technical information and depth. The publications include *Analog-Digital Conversion Handbook,* by the engineering staff of Analog Devices, published by Prentice–Hall, Englewood Cliffs, NJ, 1986; *Nonlinear Circuits Handbook, Transducer Interfacing Handbook*, and *Synchro and Resolver Conversion*, all published by Analog Devices Inc., Norwood, MA.

Engineering trade journals and design publications often have articles describing recent A/D and D/A circuits and their applications. These publications include *EDN Magazine, EE Times,* and *IEEE Spectrum.* Research-related topics are covered in *IEEE Transactions on Circuits and Systems* and *IEEE Transactions on Instrumentation and Measurement.*

10

Signal Space

Roger E. Ziemer
University of Colorado at Colorado Springs

10.1 Introduction

Signal space concepts have their roots in the mathematical theory of inner product spaces known as **Hilbert spaces** [Stakgold, 1967]. Many books on linear systems touch on the subject of signal spaces in the context of Fourier series and transforms [Ziemer, Tranter, and Fannin, 1998]. The applications of signal space concepts in communication theory find their power in the representation of signal detection and estimation problems in geometrical terms, which provides much insight into signalling techniques and communication system design. The first person to have apparently exploited the power of signal space concepts in communication theory was the Russian Kotel'nikov [1968], who presented his doctoral dissertation in January, 1947. Wozencraft and Jacobs [1965] expanded on this approach and their work is still today widely referenced. Arthurs and Dym [1962] made use of signal space concepts in the performance analysis of several digital modulation schemes. A one-chapter summary of the use of signal space methods in signal detection and estimation is provided in Ziemer and Tranter [2002]. Another application of signal space concepts is in signal and image compression. Wavelet theory [Rioul and Vetterli, 1991] is currently finding use in these application areas. In the next section, the fundamentals of generalized vector spaces are summarized, followed by an overview of several applications to signal representations.

10.2 Fundamentals

A linear space or **vector space** (signal space) [Stakgold, 1967] is a collection of elements (called **vectors**) $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \ldots$, for which the following axioms are satisfied:

1. To every pair of vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ there corresponds a vector $\boldsymbol{x} + \boldsymbol{y}$, with the properties:
   a. $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$
   b. $\boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z}) = (\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z}$
   c. There exists a unique element $\mathbf{0}$ such that $\boldsymbol{x} + \mathbf{0} = \boldsymbol{x}$ for every $\boldsymbol{x}$
   d. To every $\boldsymbol{x}$, there exists a unique vector labeled $-\boldsymbol{x}$ such that $\boldsymbol{x} + (-\boldsymbol{x}) = \mathbf{0}$
2. To all vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, and all numbers a and b (in general, complex), the following commutative and associative rules hold:

a. $\alpha(\beta \boldsymbol{x}) = (\alpha\beta)\boldsymbol{x}$
b. $(\alpha + \beta)\boldsymbol{x} = \alpha\boldsymbol{x} + \beta\boldsymbol{y}$
c. $\alpha(\boldsymbol{x} + \boldsymbol{y}) = \alpha\boldsymbol{x} + \alpha\boldsymbol{y}$
d. $1\boldsymbol{x} = \boldsymbol{x}$, where 1 is the **identity element**

A vector is said to be a **linear combination** of the vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ in a vector space if there exist numbers (in general, complex) $\alpha_1, \alpha_2, \ldots, \alpha_k$ such that

$$\boldsymbol{x} = \sum_{i=1}^{k} \alpha_i \boldsymbol{x}_i \tag{10.1}$$

The vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ are said to be **linearly dependent** (or form a dependent set) if there exist complex numbers $\alpha_1, \alpha_2, \ldots, \alpha_k$, not all zero, such that

$$\alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2 + \cdots + \alpha_k \boldsymbol{x}_k = 0 \tag{10.2}$$

If Eq. (10.2) can be satisfied only for $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$, the vectors are **linearly independent**.

One is tempted to use the infinite-sum version of Eq. (10.2) in defining the notion of independence for an infinite set of vectors. This is not true in general; one needs the notion of convergence, which is based on the concept of distance between vectors.

With the idea of linear independence firmly in mind, the concept of **dimension** of a vector space readily follows. A vector space is *n-dimensional* if it possesses a set of n **independent vectors**, but every set of $n + 1$ vectors is a dependent set. If for every positive integer k, a set of k independent vectors in the space can be found, the space is said to be infinite dimensional. A **basis** for a vector space means a finite set of vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k$ with the following attributes:

1. They are linearly independent.
2. Every vector $\boldsymbol{x}$ in the space can be written as a linear combination of the basis vectors; that is

$$\boldsymbol{x} = \sum_{i=1}^{k} \xi_i \boldsymbol{e}_i \tag{10.3}$$

It can be proved that the representation (10.3) is unique and that if the space is n-dimensional, any set of n independent vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ forms a basis.

The next concept to be developed is that of a **metric space**. In addition to the addition of vectors and multiplication of vectors by scalars, as is true of ordinary three-dimensional vectors, it is important to have the notions of length and direction of a vector imposed. In other words, a metric structure must be added to the algebraic structure already defined. A collection of elements $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \ldots$ in a space will be called a metric space if to each pair of elements $\boldsymbol{x}, \boldsymbol{y}$ there corresponds a real number $d(\boldsymbol{x}, \boldsymbol{y})$ satisfying the properties:

1. $d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x})$
2. $d(\boldsymbol{x}, \boldsymbol{y}) \geq 0$ with equality if and only if $\boldsymbol{x} = \boldsymbol{y}$
3. $d(\boldsymbol{x}, \boldsymbol{z}) \leq d(\boldsymbol{x}, \boldsymbol{y}) + d(\boldsymbol{x}, \boldsymbol{z})$ (called the **triangle inequality**)

The function $d(\boldsymbol{x}, \boldsymbol{y})$ is called a **metric** (or distance function). Note that the definition of $d(x, y)$ *does not require that the elements be vectors*; there may not be any way of adding elements or multiplying them by scalars as required for a vector space.

With the definition of a metric, one can now discuss the idea of convergence of a sequence $\{\boldsymbol{x}_k\}$ of elements in the space. Note that $d(\boldsymbol{x}_k, \boldsymbol{x})$ is a sequence of real numbers. Therefore, it is sensible to write that

$$0 \lim_{k \to \infty} \boldsymbol{x}_k = \boldsymbol{x} \tag{10.4}$$

if the sequence of numbers $d(\boldsymbol{x}_k, \boldsymbol{x})$ converges to 0 in the ordinary sense of convergence of sequences of real numbers. If

$$\lim_{m,p\to\infty} d(\boldsymbol{x}_m, \boldsymbol{x}_p) = 0 \tag{10.5}$$

the sequence $\{\boldsymbol{x}_k\}$ is said to be a **Cauchy sequence**. It can be shown that if a sequence $\{\boldsymbol{x}_k\}$ converges, it is a Cauchy sequence. The converse is not necessarily true, for the limit may have carelessly been excluded from the space. If the converse is true, then the metric space is said to be **complete**.

The next vector space concept to be defined is that of length or norm of a vector. A **normed vector space** (or linear space) is a vector space in which a real-valued function $\|\boldsymbol{x}\|$ (known as the *norm* of $\boldsymbol{x}$) is defined, with the properties

1. $\|\boldsymbol{x}\| \geq 0$ with equality if and only if $\boldsymbol{x} = 0$
2. $\|\alpha\boldsymbol{x}\| = |\alpha|\,\|\boldsymbol{x}\|$
3. $\|\boldsymbol{x}_1 + \boldsymbol{x}_2\| \leq \|\boldsymbol{x}_1\| + \|\boldsymbol{x}_2\|$

A normed vector space is automatically a metric space if the metric is defined as

$$d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\| = \langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y}\rangle^{1/2} \tag{10.6}$$

(see definition below) which is called the natural metric for the space. A normed vector space may be viewed either as a linear space, a metric space, or both. Its elements may be interpreted as *vectors* or *points*.

The structure of a normed vector space will now be refined further with the definition of the notion of angle between two vectors. In particular, it will be possible to tell whether two vectors are perpendicular. The notion of angle between two vectors will be obtained by defining the **inner product** (also known as a scalar or dot product). In general, an inner product in a vector space is a complex-valued function of ordered pairs $\boldsymbol{x}, \boldsymbol{y}$ with the properties

1. $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \langle \boldsymbol{y}, \boldsymbol{x}\rangle^*$ (the asterisk denotes complex conjugate)
2. $\langle \alpha\boldsymbol{x}, \boldsymbol{y}\rangle = \alpha\langle \boldsymbol{x}, \boldsymbol{y}\rangle$
3. $\langle \boldsymbol{x}_1 + \boldsymbol{x}_2, \boldsymbol{y}\rangle = \langle \boldsymbol{x}_1, \boldsymbol{y}\rangle + \langle \boldsymbol{x}_2, \boldsymbol{y}\rangle$
4. $\langle \boldsymbol{x}, \boldsymbol{x}\rangle \geq 0$ with equality if and only if $\boldsymbol{x} = 0$

From the first two properties, it follows that

$$\langle \boldsymbol{x}, \alpha\boldsymbol{y}\rangle = \alpha^* \langle \boldsymbol{x}, \boldsymbol{y}\rangle \tag{10.7}$$

Also, **Schwarz's inequality** can be proved and is given by

$$|\langle \boldsymbol{x}, \boldsymbol{y}\rangle| \leq \langle \boldsymbol{x}, \boldsymbol{x}\rangle^{1/2} \langle \boldsymbol{y}, \boldsymbol{y}\rangle^{1/2} \tag{10.8}$$

with equality if and only if $\boldsymbol{x} = \alpha\boldsymbol{y}$. The real, nonnegative quantity $\langle \boldsymbol{x}, \boldsymbol{x}\rangle^{1/2}$ satisfies all of the properties of a norm. Therefore, it is adopted as the definition of the norm, and Schwartz's inequality assumes the form

$$|\langle \boldsymbol{x}, \boldsymbol{y}\rangle| \leq \|\boldsymbol{x}\|\,\|\boldsymbol{y}\| \tag{10.9}$$

The natural metric in the space is given by Eq. (10.6). An inner product space, which is complete in its natural metric, is called a Hilbert space.

Example 10.1

Consider the space of all complex-valued functions $\boldsymbol{x}(t)$ defined on $a \leq t \leq b$ for which the integral

$$E_x = \int_a^b |\boldsymbol{x}(t)|^2 \,\mathrm{d}t \tag{10.10}$$

exists (i.e., the space of all finite-energy signals in the interval $[a, b]$). The inner product may be defined as

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \int_a^b \boldsymbol{x}(t)\boldsymbol{y}^*(t)\,\mathrm{d}t \tag{10.11}$$

The natural norm is

$$\|\boldsymbol{x}\| = \left[\int_a^b |\boldsymbol{x}(t)|^2\,\mathrm{d}t\right]^{1/2} \tag{10.12}$$

and the metric is

$$d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\| = \left[\int_a^b |\boldsymbol{x}(t) - \boldsymbol{y}(t)|^2\,\mathrm{d}t\right]^{1/2} \tag{10.13}$$

respectively. Schwarz's inequality becomes

$$\left|\int_a^b |\boldsymbol{x}(t)\boldsymbol{y}^*(t)|\,\mathrm{d}t\right| \le \left[\int_a^b |\boldsymbol{x}(t)|^2\,\mathrm{d}t\right]^{1/2}\left[\int_a^b |\boldsymbol{y}(t)|^2\,\mathrm{d}t\right]^{1/2} \tag{10.14}$$

It can be shown that this space is complete and, hence, is a Hilbert space.

An additional requirement that can be imposed on a Hilbert space is separability, which, roughly speaking, restricts the number of elements in the space. A Hilbert space $\mathcal{H}$ is **separable** if there exists a **countable** (i.e., can be put in one-to-one correspondence with the positive integers) set of elements ($\boldsymbol{f}_1$, $\boldsymbol{f}_2, \ldots, \boldsymbol{f}_n, \ldots$) whose finite linear combinations are such that for any element f in $\mathcal{H}$ there exist an index N and constants $\alpha_1, \alpha_2, \ldots, \alpha_N$ such that

$$\left\|\boldsymbol{f} - \sum_{k=1}^{N} \alpha_k \boldsymbol{f}_k\right\| < \epsilon, \quad \epsilon > 0 \tag{10.15}$$

The set ($\boldsymbol{f}_1$, $\boldsymbol{f}_2, \ldots, \boldsymbol{f}_n, \ldots$) is called a **spanning set**. The discussions from here on are limited to separable Hilbert spaces.

Any finite-dimensional Hilbert space E_n is separable. In fact, there exists a set of n vectors ($\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n$) such that each vector $\boldsymbol{x}$ in E_n has the representation

$$\boldsymbol{x} = \sum_{k=1}^{n} \alpha_k f_k \tag{10.16}$$

It can be shown that the spaces consisting of square-integrable functions on the intervals $[a, b]$, $(-\infty, b]$, $[a, \infty)$, and $(-\infty, \infty)$ are all separable, where a and b are finite, and spanning sets exist for each of these spaces. For example, a spanning set for the space of square-integrable functions on $[a, b]$ is the set $(1, t, t^2, \ldots)$, which is clearly countable.

The concepts of convergence and Cauchy sequences carry over to Hilbert spaces, with *convergence in the mean* defined as

$$\lim_{k\to\infty} \int_a^b |\boldsymbol{x}_k(t) - \boldsymbol{x}(t)|^2 = 0 \tag{10.17}$$

Similarly, the ideas of independence and basis sets apply to Hilbert spaces in infinite-dimensional form.

It is necessary to distinguish between the concepts of a *basis* and a *spanning* set consisting of independent vectors. As ϵ is reduced in Eq. (10.15), it is expected that N must be increased, and it may also be

necessary to change the previously found coefficients $\alpha_1, \ldots, \alpha_N$. Hence, there might not exist a fixed sequence of constants $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ with the property

$$\boldsymbol{x} = \sum_{k=1}^{\infty} \xi_k f_k \tag{10.18}$$

as would be required if the set $\{\boldsymbol{f}_k\}$ were a basis. For example, on the space of square-integrable functions on $[-1, 1]$, the independent set $\boldsymbol{f}_0 = 1, \boldsymbol{f}_1 = t, \boldsymbol{f}_2 = t^2, \ldots$, is a spanning set, but not a basis, since there are many square-integrable functions on $[-1, 1]$ that cannot be expanded in a series like Eq. (10.18) (an example is $|t|$). Odd as it may seem at first, it is possible if the powers of t in this spanning set are regrouped into the set of polynomials known as the Legendre polynomials, $P_k(t)$.

Two vectors $\boldsymbol{x}, \boldsymbol{y}$ are **orthogonal** or perpendicular if $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$. A finite or countably infinite set of vectors $\{\phi_1, \phi_2, \ldots, \phi_k, \ldots\}$ is said to be an *orthogonal set* if $\langle \phi_i, \phi_j \rangle = 0$, $i \neq j$. A *proper orthogonal set* is an orthogonal set none of whose elements is the zero vector. A proper orthogonal set is an independent set. A set is **orthonormal** if

$$\langle \phi_i, \phi_j \rangle = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} \tag{10.19}$$

An important concept is that of a **linear manifold** in a Hilbert space. A set M is said to be a linear manifold if, for $\boldsymbol{x}$ and $\boldsymbol{y}$ belonging to M, so does $\alpha\boldsymbol{x} + \beta\boldsymbol{y}$ for arbitrary complex numbers α and β; thus, M is itself a linear space. If a linear manifold is a closed set, it is called a closed linear manifold (i.e., every Cauchy sequence has a limit in the space) and is itself a Hilbert space. In three-dimensional Euclidean space, linear manifolds are simply lines and planes containing the origin. In a finite-dimensional space, every linear manifold is necessarily closed.

Let M be a linear manifold, closed or not. Consider the set $M^{\perp}$ of all vectors which are orthogonal to every vector in M. It is a linear manifold, which can be shown to be closed. If M is closed, M and $M^{\perp}$ are known as **orthogonal complements**. Given a linear manifold M, each vector in the space can be decomposed in a unique manner, as a sum $\boldsymbol{x}_p + \boldsymbol{z}$, where $\boldsymbol{x}_p$ is in M and $\boldsymbol{z}$ is in $M^{\perp}$.

Given an infinite orthonormal set $\{\phi_1, \phi_2, \ldots, \phi_n, \ldots\}$ in the space of all square-integrable functions on the interval $[a, b]$, let $\{a_n\}$ be a sequence of complex numbers. The **Riesz–Fischer theorem** tells how to represent an element in the space and states:

1. If

$$\sum_{n=1}^{\infty} |a_n|^2 \tag{10.20}$$

   diverges, then

$$\sum_{n=1}^{\infty} a_n \phi_n \tag{10.21}$$

   diverges.
2. If Eq. (10.20) converges, then Eq. (10.21) also converges to some element $\boldsymbol{g}$ in the space and

$$a_n = \langle \boldsymbol{g}, \phi_n \rangle \tag{10.22}$$

The next question that arises is how to construct an orthonormal set from an independent set $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k, \ldots\}$. A way to do this is known as the **Gram–Schmidt procedure**. The construction is as follows:

1. Pick a function from the set $\{\boldsymbol{e}_k\}$, say $\boldsymbol{e}_1$. Let

$$\phi_1 = \frac{\boldsymbol{e}_1}{\langle \boldsymbol{e}_1, \boldsymbol{e}_1 \rangle^{1/2}} \tag{10.23}$$

2. Remove from a second function in the set $\{\boldsymbol{e}_k\}$, say $\boldsymbol{e}_2$, its projection on ϕ_1. This yields

$$\boldsymbol{g}_2 = \boldsymbol{e}_2 - \langle \boldsymbol{e}_2, \phi_1 \rangle \phi_1 \tag{10.24}$$

   The vector $\boldsymbol{g}_2$ is a linear combination of $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$, and is orthogonal to ϕ_1. To normalize it, form

$$\phi_2 = \frac{\boldsymbol{g}_2}{\langle \boldsymbol{g}_2, \boldsymbol{g}_2 \rangle^{1/2}} \tag{10.25}$$

3. Pick another function from the set $\{e_k\}$, say e_3, and form

$$\boldsymbol{g}_3 = \boldsymbol{e}_3 - \langle \boldsymbol{e}_3, \phi_2 \rangle \phi_2 - \langle e_3, \phi_1 \rangle \phi_1 \tag{10.26}$$

   Normalize $\boldsymbol{g}_3$ in a manner similar to that used for $\boldsymbol{g}_2$.
4. Continue until all functions in the set $\{\boldsymbol{e}_k\}$ have been used.

Note that the sets $\{\boldsymbol{e}_k\}$, $\{\boldsymbol{g}_k\}$, and $\{\phi_k\}$ all generate the same linear manifold.

A basis consisting of orthonormal vectors is known as an **orthonormal basis**. If the basis vectors are not normalized, it is simply an *orthogonal basis*.

Example 10.2

Consider the interval $[-1, 1]$ and the independent set $\boldsymbol{e}_0 = 1$, $\boldsymbol{e}_2 = t, \ldots, \boldsymbol{e}_k = t^k, \ldots$. The Gram–Schmidt procedure applied to this set without normalization, but with the requirement that all orthogonal functions take on the value 1 at $t = 1$, gives the set of *Lengendre polynomials*, which is

$$\psi_0(t) = 1, \quad \psi_1(t) = t, \quad \psi_2(t) = \frac{1}{2}(3t^2 - 1), \quad \psi_3(t) = \frac{1}{2}(5t^3 - 3t), \ldots \tag{10.27}$$

It is next desired to approximate an arbitrary vector $\boldsymbol{x}$ in a Hilbert space in terms of a linear combination of the independent set $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k\}$, where $k \le n$ if the space is an n-dimensional Euclidean space, and k is an arbitrary integer if the space is infinite-dimensional. First, the orthonormal set $\{\phi_1, \ldots, \phi_k\}$ is constructed from $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k\}$. The unique, best approximation to $\boldsymbol{x}$ is the **Fourier sum**

$$\sum_{i=1}^{k} \langle \boldsymbol{x}, \phi_i \rangle \phi_i \tag{10.28}$$

which is geometrically the projection of $\boldsymbol{x}$ onto the linear manifold generated by $\{\phi_1, \ldots, \phi_k\}$, or equivalently, the sum of the projections along the individual axes defined by $\phi_1, \phi_2, \ldots, \phi_k$. The square of the distance between $\boldsymbol{x}$ and its projection is

$$\left\| \boldsymbol{x} - \sum_{i=1}^{k} \langle \boldsymbol{x}, \phi_i \rangle \phi_i \right\|^2 = \|\boldsymbol{x}\|^2 - \sum_{i=1}^{k} |\langle \boldsymbol{x}, \phi_i \rangle|^2 \tag{10.29}$$

Since the left-hand side is nonnegative, Eq. (10.29) gives **Bessel's inequality**, which is

$$\|\boldsymbol{x}\|^2 \geq \sum_{i=1}^{k} |\langle \boldsymbol{x}, \phi_i \rangle|^2 \tag{10.30}$$

A convenient feature of the Fourier sum is the following: If another vector ϕ_{k+1} is added to the orthonormal approximating set of vectors, the best approximation now becomes

$$\sum_{i=1}^{k+1} \langle \boldsymbol{x}, \phi_i \rangle \phi_i \tag{10.31}$$

Thus, an additional term is added to the series expansion without changing previously computed coefficients, which makes the extension to a countably infinite orthonormal approximating set simple to envision. In the case of an infinite orthonormal approximating set, Bessel's inequality (10.30) now has an infinite limit on the sum. Does the approximating sum (10.31) converge to $\boldsymbol{x}$? The answer is that convergence can be guaranteed only if the set $\{\phi_k\}$ is extensive enough, that is, if it is a *basis or a complete orthonormal set*. In such cases, Eq. (10.30) becomes an equality. In fact, a number of equivalent criteria can be stated to determine whether an orthonormal set $\{\phi_k\}$ is a basis or not [Stakgold, 1967]. These are:

1. In finite n-dimensional Euclidean space, $\{\phi_k\}$ has exactly n elements for completeness.
2. For every $\boldsymbol{x}$ in the space of square-integrable functions,

$$\boldsymbol{x} = \sum_{i} \langle \boldsymbol{x}, \phi_i \rangle \phi_i \tag{10.32}$$

3. For every $\boldsymbol{x}$ in the space of square-integrable functions,

$$\|\boldsymbol{x}\|^2 = \sum_{i=1}^{k} |\langle \boldsymbol{x}, \phi_i \rangle|^2 \tag{10.33}$$

   (known as **Parseval's equality**).
4. The only $\boldsymbol{x}$ in the space of square-integrable functions for which all of the Fourier coefficients vanish is the 0 function.
5. There exists no function $\phi(t)$ in the space of square-integrable functions such that $\{\phi_1, \phi_2, \ldots, \phi_k, \ldots\}$ is an orthonormal set.

Examples of complete orthonormal sets of trigonometric functions over the interval $[0, T]$ are as follows: (1) The complex exponentials with frequencies equal to the harmonics of the fundamental frequency $\omega_0 = 2\pi/T$, or

$$\frac{1}{\sqrt{T}}, \quad \frac{e^{j\omega_0 t}}{\sqrt{T}}, \quad \frac{e^{-j\omega_0 t}}{\sqrt{T}}, \quad \frac{e^{j2\omega_0 t}}{\sqrt{T}}, \quad \frac{e^{-j2\omega_0 t}}{\sqrt{T}}, \ldots \tag{10.34}$$

The factor $(T)^{1/2}$ in the denominator is necessary to normalize the functions and is often not included in the definition of the complex exponential Fourier series. (2) The sines and cosines with frequencies equal to harmonics of the fundamental frequency $\omega_0 = 2\pi/T$, or

$$\frac{1}{\sqrt{T}}, \quad \frac{(\cos \omega_0 t)}{\sqrt{T/2}}, \quad \frac{(\sin \omega_0 t)}{\sqrt{T/2}}, \quad \frac{(\cos 2\omega_0 t)}{\sqrt{T/2}}, \quad \frac{(\sin 2\omega_0 t)}{\sqrt{T/2}}, \ldots \tag{10.35}$$

Note that if any function is left out of these sets, the basis is incomplete.

10.3 Application of Signal Space Representation to Signal Detection

The M-ary signal detection problem is as follows: given M signals, $s_0(t)$, $s_1(t),\ldots,s_M(t)$, defined over $0 \le t \le T$. One is chosen at random and sent each T-second interval through a channel that adds white, Gaussian noise of power spectral density $N_0/2$ to it. The challenge is to design a receiver that will decide which signal was sent through the channel during each T-second interval with minimum probability of making an error.

An approach to this problem, as expanded upon in greater detail by Wozencraft and Jacobs [1965, Chap. 4] and Ziemer and Tranter [2002, Chap. 9], is to construct a linear manifold, called the signal space, using the Gram–Schmidt procedure on the M signals. Suppose that this results in the orthonormal basis set $\{\phi_1, \phi_2,\ldots, \phi_K\}$ where $K \le M$. The received signal plus noise is represented in this signal space as vectors with coordinates (note that they depend on the signal transmitted)

$$Z_{ij} = A_{ij} + N_j, \qquad \begin{aligned} j &= 1, 2, \ldots, K, \\ i &= 1, 2, \ldots, M, \end{aligned} \tag{10.36}$$

where

$$A_{ij} = \int_0^T s_i(t)\phi_j(t)dt \tag{10.37}$$

The numbers Z_{ij} are components of vectors referred to as the **signal vectors**, and the space of all signal vectors is called the **observation space**. An apparent problem with this approach is that not all possible noise waveforms added to the signal can be represented as vectors in this K-dimensional observation space. The part of the noise that is represented is

$$n_|(t) = \sum_{j=1}^{K} N_j\phi_j(t) \tag{10.38}$$

where

$$N_j = \int_0^T n(t)\phi_j(t)\,\mathrm{d}t \tag{10.39}$$

In terms of Hilbert space terminology, Eq. (10.38) is the projection of the noise waveform onto the observation space (i.e., a linear manifold). The unrepresented part of the noise is

$$n_\perp(t) = n(t) - n_|(t) \tag{10.40}$$

and is the part of the noise that must be represented in the orthogonal complement of the observation space. The question is, will the decision process be harmed by ignoring this part of the noise? It can be shown that $n_\perp(t)$ is uncorrelated with $n_|(t)$. Thus, since $n(t)$ is Gaussian they are statistically independent and $n_\perp(t)$ has no bearing on the decision process; nothing is lost by ignoring $n_\perp(t)$.

The decision process can be shown to reduce to choosing the signal $s_\ell(t)$, which minimizes the distance to the data vector; that is,

$$d(\boldsymbol{z}, s_\ell) = \|z - s_\ell\| = \left[\sum_{j=1}^{K}(Z_{ij} - A_{\ell j})^2\right]^{1/2} = \text{minimum}, \quad \ell = 1, 2, \ldots, M \tag{10.41}$$

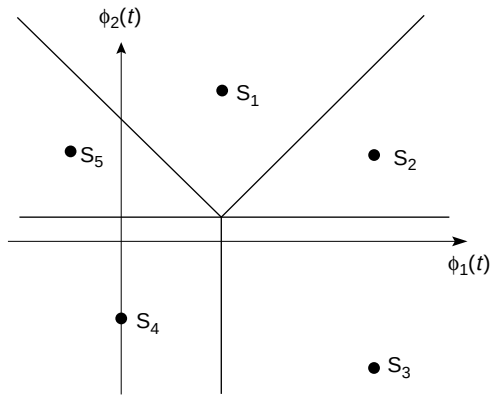


FIGURE 10.1 Observation space in two dimensions, showing five signal points with boundaries for decision regions providing the minimum distance decision in each case.

where

$$z(t) = \sum_{j=1}^{K} Z_{ij}\phi_j(t) \tag{10.42}$$

and

$$s_\ell(t) = \sum_{j=1}^{K} A_{\ell j}\phi_j(t) \tag{10.43}$$

Thus, the signal detection problem is reduced to a geometrical one, where the observation space is subdivided into decision regions in order to make a decision, as shown in Fig. 10.1.

10.4 Application of Signal Space Representation to Parameter Estimation

The procedure used in applying signal space concepts to estimation is similar to that used for signal detection. Consider the observed waveform consisting of additive signal and noise of the form

$$y(t) = s(t, A) + n(t), \quad 0 \le t \le T \tag{10.44}$$

where A is a parameter to be estimated and the noise is white as before. Let $\{\phi_k(t)\}$, $k = 1, 2, \ldots$, be a complete orthonormal basis set. The observed waveform can be represented as

$$y(t) = \sum_{k=1}^{\infty} S_k(A)\phi_k(t) + \sum_{k=1}^{\infty} N_k\phi_k(t) \tag{10.45}$$

where

$$S_k(A) = \langle s, \phi_k \rangle = \int_0^T s(t, A)\phi_k(t)\,\mathrm{d}t \tag{10.46}$$

and N_j is defined by Eq. (10.39). Hence, an estimate can be made on the basis of the set of coefficients

$$Z_k = S_k(A) + N_k \tag{10.47}$$

or on the basis of a vector in the signal space with these coordinates. A reasonable criterion for estimating A is to maximize the likelihood ratio, or a monotonic function thereof. Its logarithm, for $n(t)$ Gaussian, can be shown to reduce to

$$\ell(A) = \lim_{K\to\infty} L_K(A) = \lim_{K\to\infty}\left[\frac{2}{N_0}\sum_{k=1}^{K} Z_k S_k(A) - \frac{1}{N_0}\sum_{k=1}^{K} S_k^2(A)\right] \tag{10.48}$$

In the limit as $k \to \infty$, this becomes

$$\ell(A) = \frac{2}{N_0}\int_0^T z(t)s(t,\ A)\,\mathrm{d}t - \frac{1}{N_0}\int_0^T s^2(t,A)\,\mathrm{d}t \tag{10.49}$$

A necessary condition for the value of A that maximizes Eq. (10.49) is

$$\frac{\partial \ell(A)}{\partial A} = \frac{2}{N_0}\int_0^T [z(t) - s(t,\ A)]\frac{\partial s(t,A)}{\partial A}\mathrm{d}t\bigg|_{A=\hat{A}} = 0 \tag{10.50}$$

The value of A that maximizes Eq. (10.49), denoted $\hat{A}$, is called the **maximum likelihood estimate**.

Wavelet Transforms

Wavelet transforms can be continuous time or discrete time. They find applications in speech and image compression, signal and image classification, and pattern recognition. Wavelet representations have been adopted by the U.S. Federal Bureau of Investigation for fingerprint compression.

The continuous-time **wavelet transform** of a signal $x(t)$ takes the form [Rioul and Vetterli, 1991]

$$W_x(\tau, a) = \int_{-\infty}^{\infty} x(t)h^*_{a,\tau}(t)\,\mathrm{d}t \tag{10.51}$$

where

$$h_{a,\tau}(t) = \frac{1}{\sqrt{a}}h\left(\frac{t-\tau}{a}\right) \tag{10.52}$$

are basis functions called **wavelets**. Thus, the wavelets defined in Eq. (10.52) are scaled and translated versions of the basic wavelet prototype $h(t)$ (also known as the *mother wavelet*), and the wavelet transform is seen to be a convolution of the conjugate of a wavelet with the signal $x(t)$. Substitution of Eq. (10.52) into Eq. (10.51) yields

$$W_x(\tau, a) = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} \boldsymbol{x}(t)h^*\left(\frac{t-\tau}{a}\right)\mathrm{d}t \tag{10.53}$$

Note that $h(t/a)$ is contracted if $a < 1$ and expanded if $a > 1$. Thus, an interpretation of Eq. (10.53) is that as a increases, the function $h(t/a)$ becomes spread out over time and takes the long-term behavior

of $\boldsymbol{x}(t)$ into account; as a decreases the short-time behavior of $\boldsymbol{x}(t)$ is taken into account. A change of variables in Eq. (10.53) gives

$$W_x(\tau, a) = \sqrt{a}\int_{-\infty}^{\infty} x(at)h^*\left(t - \frac{\tau}{a}\right)\mathrm{d}t \tag{10.54}$$

Now the interpretation of Eq. (10.54) is as the scale increases ($a < 1$), an increasingly contracted version of the signal is seen through a constant-length sifting function, $h(t)$. This is only the barest of introductions to wavelets, and the reader is urged to consult the references to learn more about wavelets and their applications, particularly their discrete-time implementation.

Mean Square Estimation—The Orthogonality Principle

Given n random variables, $X_1, X_2, \ldots, X_n$, it is desired to find n constants $a_1, a_2, \ldots, a_n$ such that when another random variable S is estimated by the sum

$$\hat{S} = \sum_{i=1}^{n} a_i X_i \tag{10.55}$$

then the mean square error

$$\mathrm{MSE} = E\left\{\left|S - \sum_{i=1}^{n} a_i X_i\right|^2\right\} \tag{10.56}$$

is a minimum, where $E\{\}$ denotes expectation or statistical average. It is shown in [Papoulis, 1984] that the **mean square estimate** (**MSE**) is minimized when the error is orthogonal to the data, or when

$$E\left\{\left[S - \sum_{i=1}^{n} a_i X_i\right] X_j^*\right\} = 0, \quad j = 1, 2, \ldots, n \tag{10.57}$$

This is known as the **orthogonality principle** or **projection theorem** and can be interpreted as stating that the MSE is minimized when the error vector $\boldsymbol{S} - \hat{\boldsymbol{S}}$ is orthogonal to the subspace (linear manifold) spanned by the vectors $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$, as shown in Fig. 10.2. The projection theorem has many applications including MSE filtering of noisy signals, known as Wiener filtering.

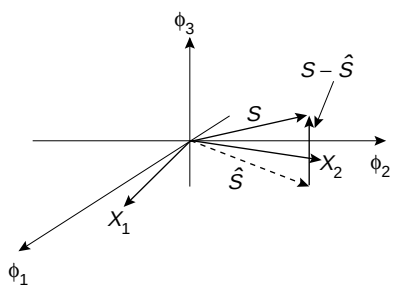


FIGURE 10.2 Schematic representation of mean square estimation in three dimensions with the estimate expressed in terms of a linear combination of two random variables, $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, which lie in the $\phi_1 - \phi_2$ plane.

Defining Terms

Basis (basis functions): An independent set of vectors in a linear space in terms of which any vector in the space can be represented as a linear combination.

Bessel's inequality: The statement that the norm squared of any vector in a vector space is greater or equal to the sum of the squares of the projections onto a set of orthonormal basis vectors.

Cauchy sequence: A convergent sequence whose members become progressively closer as the limit is approached.

Complete: The idea that a basis set is extensive enough to represent any vector in the space as a Fourier sum.

Countable: A set that can be put in one-to-one correspondence with the positive integers.

Dimension: A vector space is n-dimensional if it possesses a set of n independent vectors, but every set of $n + 1$ vectors is linearly dependent.

Fourier sum: An approximation of a vector in a Hilbert space as a linear combination of orthonormal basis functions where the coefficient of each basis function is the projection of the vector onto the respective basis function.

Gram–Schmidt procedure: An algorithm that produces a set of orthonormal vectors from an arbitrary set of vectors.

Hilbert space: A normed linear space complete in its natural metric.

Identity element: An element in a vector space that when multiplied by any vector reproduces that vector.

Independent vectors: A set of vectors is independent if any one of them cannot be expressed as a linear combination of the others.

Inner product: A function of ordered pairs of vectors in a vector space, which is analogous to the dot product in ordinary Euclidean vector space.

Linear combination: A linear sum of a set of vectors in a vector space with each member in the sum, in general, multiplied by a different scalar.

Linear manifold: A subspace of a Hilbert space that contains the origin.

Linearly dependent (dependent): A set of vectors that is not linearly independent.

Linearly independent (independent): A set of vectors, none of which can be expressed as a linear combination of the others.

Maximum likelihood estimate: An estimate for a parameter that maximizes the likelihood function.

Mean square estimation: Estimation of a random variable by a linear combination of n other random variables that minimizes the expectation of the squared difference between the random variable to be estimated and the approximating linear combination.

Metric: A real-valued function of two vectors in a vector space that is analogous to the distance between them in the ordinary Euclidean sense.

Metric space: A vector space in which a metric is defined.

Norm: A function of a single vector in a vector space that is analogous to the length of a vector in ordinary Euclidean vector space.

Normed vector space: A vector space in which a norm has been defined.

Observation space: The space of all possible received data vectors in a signal detection or estimation problem.

Orthogonal: The property of two vectors expressed by their inner product being zero.

Orthogonal complement: The space of vectors that are orthogonal to a linear manifold or subspace.

Orthogonality principle: The theorem that states that the minimum mean-square estimate of a random variable in terms of a linear combination of n other random variables requires the difference between the random variable and linear combination, or error, to be statistically orthogonal to each random variable in the linear combination (i.e., the expectation of the product of the error and each random variable is zero). Also called the *projection theorem*.

Orthonormal: Orthogonal vectors that have unit norm.

Orthonormal basis: A basis set for which the basis vectors are orthonormal.

Parseval's equality: The statement that the norm squared of a vector in a complete Hilbert space equals the sum of the squares of the vector's projections onto an orthonormal basis set in the space.

Projection theorem: See **orthogonality principle**.

Riesz–Fischer theorem: A theorem that states the conditions under which an element of a space of square-integrable functions can be represented in terms of an infinite orthonormal set.

Schwarz's inequality: An inequality expressing the fact that the absolute value of the inner product of any pair of vectors in a Hilbert space is less than or equal to the product of their respective norms.

Separable: A Hilbert space in which a countable set of elements exists that can be used to represent any element in the space to any degree of accuracy desired as a linear combination of the members of the set.

Signal vector: A vector representing a received signal in a signal detection or estimation problem.

Spanning set: The set of elements of a separable Hilbert space used to represent an arbitrary element with any degree of accuracy desired.

Triangle inequality: An inequality of a normed linear space that is analogous to the fact that the length of the sum of any two sides of a triangle is less than or equal to the sum of their respective lengths.

Vector: An element of a linear, or vector, space.

Vector space: A space of elements, called vectors, which obey certain laws of associativity and commutativity and have identity elements for scalar multiplication and element addition defined.

Wavelet: See **wavelet transform**.

Wavelet transform: Resolution of a signal into a set of basis functions called wavelets. As a parameter of the wavelet is changed, the behavior of the signal over progressively shorter time intervals is resolved.

References

Arthurs, E. and Dym, H. 1962. On the optimum detection of digital signals in the presence of white gaussian noise—A geometric approach and a study of three basic data transmission systems. *IRE Trans. Commn. Syst.*, CS-10 (Dec.):336–372.

Kotel'nikov, V.A. 1968. *The Theory of Optimum Noise Immunity* (trans. R.A. Silverman), Dover, New York.

Papoulis, A. 1984. *Probability, Random Variables, and Stochastic Processes,* 2nd ed., McGraw–Hill, New York.

Rioul, O. and Vetterli, M. 1991. Wavelets in signal processing. *IEEE Signal Proc. Mag.,* 8(Oct.):14–38.

Stakgold, I. 1967. *Boundary Value Problems of Mathematical Physics,* Vol. 1, Macmillan, Collier–Macmillan Ltd., London.

Wozencraft, J.M. and Jacobs, I.M. 1965. *Principles of Communication Engineering,* Wiley, New York. (Available from Waveland Press, Prospect Heights, IL.)

Ziemer, R.E., Tranter, W.H., and Fannin, D.R. 1998. *Signals and Systems: Continuous and Discrete,* 4th ed., Macmillan, New York.

Ziemer, R.E. and Tranter, W.H. 2002. *Principles of Communications: Systems, Modulation, and Noise,* 5th ed., Houghton Mifflin, Boston, MA.

Further Information

Very readable expositions of signal space concepts as applied to signal detection and estimation are found in Chapter 4 of the classic book by Wozencraft and Jacobs. The paper by Arthurs and Dym listed in the references is also very readable. A treatment of signal space concepts and applications to signal detection and parameter estimation is given in Chapter 9 of the book by Ziemer and Tranter. For the mathematical theory behind signal space concepts, the book by Stakgold is recommended. For those interested in wavelet transforms and their relationship to signal spaces, the April 1992 issue of the *IEEE Signal Processing Magazine* provides a tutorial article on wavelet transforms.

11

Channel Models

David R. Smith
George Washington University

11.1 Introduction

Early channel models used in conjunction with analysis of communication systems assumed a binary symmetric channel (BSC) [Wozencraft and Jacobs, 1967]. In this case, the probability of bit error is assumed constant. This channel is memoryless, that is, the presence or absence of an error for a particular symbol has no influence on past or future symbols. Few transmission paths can be characterized as a BSC; most channels of interest are subject to some degree of fading and dispersion which introduce memory (correlation) between symbol errors. Therefore, it is necessary that a fading dispersive channel model be described to allow complete characterization of communication systems for typical applications.

Modeling of the fading dispersive channel is based on the use of mathematics to describe the physical or observed properties of the channel. Stein, Schwartz, and Bennett [1966] and Kennedy [1969] first used a deterministic characterization of the channel and then introduced dynamics into the model to account for the time-varying nature of a fading channel. Bello [1963] described a channel model using a tapped delay line representation, which is based on knowledge of the correlation properties of the channel.

The fading dispersive channel is described in Section 11.2 as a linear time-varying filter. Mobile and scatter channels produce a received signal consisting of the sum of multiple time-variant paths. Such channels are assumed to impart zero mean Gaussian statistics on a transmitted signal. Characterization of the channel then reduces to the specification of a correlation function or power spectral density. Many radio channels are adequately modeled by making these assumptions, including tropospheric scatter, high-frequency (HF) skywave, VHF/UHF mobile channels, and lunar reflection channels. Line-of-sight (LOS) channels, however, consist of a direct ray plus multiple indirect rays. The LOS channel, therefore, is assumed to impart nonzero mean Gaussian statistics on the transmitted signal. Modeling of the LOS channel can be somewhat simplified as compared to the general fading dispersive channel model, as shown in Section 11.3.

Digital communication channels introduce errors in bursts, which invalidates the use of the BSC. To characterize error clustering effects, various error models have been proposed, which can be divided into two basic categories, *descriptive* and *generative*. Descriptive models attempt to characterize channels by use of statistics that describe the general behavior of the channel. The Neyman type A distribution [Neyman, 1939] is an example of a descriptive model; here channel errors are modeled as a compound Poisson distribution in which error clusters have a Poisson distribution and errors within a cluster also

have a Poisson distribution. Such models have been shown to apply to channels that have relatively few error-causing mechanisms, such as cable transmission [Becam et al., 1984], but have not been effectively used with more complex channels such as digital radio. Generative models, however, are able to match channel data by iteratively expanding the size of the model to fit the data. The generative model allows error statistics to be generated from a model made up of a finite-state Markov chain, using the states and transitions to represent the error behavior in the communications channel. These error models are usually developed to represent empirical results, but extensions are often made to more general cases.

11.2 Fading Dispersive Channel Model

Fading arises from destructive interference among multiple propagation paths, hence the term *multipath fading*. These paths arise due to reflection, refraction, or diffraction in the channel. The amplitude and phase of each path vary in time due to changes in the structure of the medium. The resultant signal at a receiver will experience fading, defined as changes in the received signal level in time, where resulting destructive or constructive interference depends on the relative amplitudes and phases of the multiple paths. The received signal may also experience dispersion, defined as spreading of the signal in time or frequency. Channels may be characterized as having multiple, randomly behaving paths or as having a dominant path in addition to multiple secondary paths. Forward scatter and mobile channels tend to have a large number of indirect paths, whereas line-of-sight channels tend to have a direct path that predominates over a few indirect paths.

The fading dispersive channel model may be best presented by using complex representation of the transmitted and received waveforms. Let us define $s(t)$ to be the transmitted signal and $y(t)$ to be the received signal. Assuming both $s(t)$ and $y(t)$ are narrowband signals, we can represent them as

$$s(t) = \mathrm{Re}[z(t)e^{j2\pi f_c t}] \tag{11.1}$$

$$y(t) = \mathrm{Re}[w(t)e^{j2\pi f_c t}] \tag{11.2}$$

where f_c is the carrier frequency and $z(t)$ and $w(t)$ are the complex envelopes of $s(t)$ and $y(t)$, respectively. The complex envelope notation carries information about the amplitude and phase of the signal waveforms. The magnitude of the complex envelope is the conventional envelope of the signal, and the angle of the complex envelope is the phase of the signal with respect to the carrier frequency f_c.

The first assumption made in characterizing the fading dispersive channel is that a large number of paths exist, and so the central limit theorem can be applied. We then can represent the channel impulse response as a complex Gaussian process, $g(t, \xi)$, which displays the time-varying nature of the channel. Assuming a linear channel,

$$w(t) = \int_{-\infty}^{\infty} z(t-\xi)g(t, \xi)\,\mathrm{d}\xi \tag{11.3}$$

We also need to define the time-varying transfer function of the channel, which is the Fourier transform dual of the channel impulse response

$$G(f, t) = \int_{-\infty}^{\infty} g(t, \xi)e^{-j2\pi f\xi}\,\mathrm{d}\xi \tag{11.4}$$

In this section, we will further assume that no direct or dominant path is present in the channel and that the scatterers behave as a random process, so that the statistics are Gaussian with zero mean. The received envelope then displays Rayleigh law statistics. Mobile and scatter communication systems are examples of such channels.

The final assumption required for characterization of the channel model involves the stationarity of the channel impulse response. Although the fluctuations in the channel are due to non-stationary statistical phenomena, on a short enough time scale and for small enough bandwidth, the fluctuations in time and frequency can be characterized as approximately stationary. This approximate stationarity is often called quasistationarity. Since our interest here is in short-term fading, it is reasonable to assume that $g(t, \xi)$ is stationary in a time sense. Channels displaying stationarity in time are called *wide-sense stationary* (WSS). For most scatter and mobile channels, the channel may be modeled as a continuum of uncorrelated scatters. If the channel impulse response $g(t, \xi)$ is independent for different values of delay (ξ), the channel is said to exhibit *uncorrelated scattering* (US). When the time-varying impulse response is assumed to have stationary fluctuations in time and frequency, the channel is said to be wide-sense stationary with uncorrelated scattering (WSSUS).

A WSSUS complex Gaussian process is completely determined statistically by specifying its autocorrelation function or its power spectral density. Because of the dual (time and frequency) nature of the channel, there exist several correlation functions and power spectral densities that are used to characterize the channel, and the most applicable of these are defined and briefly described as follows:

1. The *time-frequency correlation function* (sometimes called the two-frequency correlation function) is defined as

$$R(\Omega, \tau) = E[G^*(f, t)G(f + \Omega, t + \tau)] \tag{11.5}$$

   Note that the complex character of $G(f, t)$ requires a slightly altered definition of autocorrelation, with the asterisk denoting the complex conjugate operation. $R(\Omega, \tau)$ represents the cross-correlation function between the complex envelopes of received carriers transmitted Ω Hz apart. Since the transfer function is assumed stationary with uncorrelated scattering, $R(\Omega, \tau)$ is dependent only on the frequency separation Ω and time delay τ.
2. The *tap gain correlation function* (sometimes called the path gain correlation function) is defined as

$$Q(\tau, \xi)\delta(\Delta\xi) = E[g^*(t, \xi)g(t + \tau, \xi + \Delta\xi)] \tag{11.6}$$

   where $\delta(\bullet)$ is the unit impulse function and $Q(\tau, \xi)$ is the Fourier transform of $R(\Omega, \tau)$ on the Ω variable. $Q(\tau, \xi)$ represents the autocorrelation function of the gain fluctuations for paths providing delays in the interval $(\xi, \xi + \Delta\xi)$. The form of Eq. (11.6) implies that the fluctuations of the complex gains at different delays are uncorrelated, which is a result of the uncorrelated scattering assumption.
3. The *scattering function* is defined as the power spectrum of the complex gain fluctuations at delay ξ and is obtained from $Q(\tau, \xi)$ by applying the Fourier transform on the τ variable,

$$S(\xi, \nu) = \int_{-\infty}^{\infty} Q(\tau, \xi)e^{-j2\pi f\nu}\,\mathrm{d}\tau \tag{11.7}$$

   The scattering function directly exhibits the delay and Doppler spreading characteristics of the dispersive channel. The power spectral density associated with values of time delay ξ and Doppler frequency ν in this interval ξ to $\xi + \Delta\xi$ and ν to $\nu + \Delta\nu$, respectively, is then $S(\xi, \nu)\,\Delta\xi\,\Delta\nu$. The scattering function is also seen to be the double Fourier transform of the time-frequency correlation function.

Special cases of the described channel model that are of interest include the frequency-selective channel, time-selective channel, time- and frequency-selective channel, and nonselective channel. These cases are described by first showing how the general model is simplified and then showing the physical conditions for which these special cases arise.

Frequency-Selective Channel

The frequency-selective channel is characterized by a transfer function that varies in a more or less random fashion with frequency. An approximately constant amplitude and phase can only be observed over a frequency interval sufficiently small. The fading, which occurs at each frequency, will be very nearly the same, that is, highly correlated, for closely spaced frequencies and essentially independent, that is, no correlation for frequencies sufficiently far apart. The lack of correlation in fading of spaced frequencies is thus called **frequency-selective fading**.

A fading channel which is selective only in frequency is then implicitly also a nontime selective or *time-flat* channel. This time-flat channel is characterized by a channel impulse response that is time invariant. When the fading rate is so slow that little change in the channel impulse response occurs over the duration of a transmitted pulse, that channel may be approximated as a time-flat channel. Assuming the channel impulse response to be time invariant, the input–output relationship (11.3) may be simplified to the usual convolution integral

$$w(t) = \int_{-\infty}^{\infty} z(t-\xi)g(\xi)\,\mathrm{d}\xi \tag{11.8}$$

in which $g(\xi)$ is now the time-invariant channel impulse response. With the time-invariance assumption, the complex correlation function of $g(\xi)$ becomes

$$E[g^*(\xi)g(\xi+\Delta\xi)] = Q(\xi)\delta(\Delta\xi) \tag{11.9}$$

where the function

$$Q(\xi) = Q(O, \xi) \tag{11.10}$$

is called the *delay power spectrum*. The *frequency correlation function*, which is the Fourier transform of the delay power spectrum, is given by

$$q(\Omega) = R(\Omega, \xi) = E[G^*(f)G(f+\Omega)] \tag{11.11}$$

where $G(f)$ is the time-invariant channel transfer function. A sufficient characterization of this channel is provided by either the delay power spectrum $Q(\xi)$ or the frequency correlation function $q(\Omega)$.

The frequency correlation function $q(\Omega)$ would be determined experimentally by transmitting two sine waves of different frequency and measuring the correlation between the complex envelopes as a function of frequency separation of the sine waves. When the frequency separation Ω is such that the correlation function $q(\Omega)$ is very near the maximum value $q(O)$ for all $\Omega < B_c$, all transmitted frequencies less than B_c will be received fading in a highly correlated fashion. For this reason B_c is called the **coherence bandwidth**, where B_c has been defined as the frequency separation for which $q(\Omega)$ equals $1/e$, or alternatively, 1/2.

The delay power spectrum $Q(\xi)$ is the average power at time delay ξ in the multipath structure of the channel. One may define a **multipath** or **delay spread** parameter L as the width of $Q(\xi)$. Two measures of L that occur frequently are *total* and 2σ. The total delay spread L_T is meant to define the spread of $Q(\xi)$ for values of ξ where $Q(\xi)$ is significantly different from zero. The 2σ delay spread $L_{2\sigma}$ is defined as twice the standard deviation of $Q(\xi)$ when it has been normalized to unit area and regarded as a probability density function. Note that the coherence bandwidth and delay spread are reciprocals of one another, due to the Fourier transform relationship between $q(\Omega)$ and $Q(\xi)$.

If we assume that the received signal consists of scattered components ranging over a delay interval L, then the received signal corresponding to a transmitted symbol of duration T will be spread over $T + L$. As long as $L > T$, the received signal is spread in time, and the corresponding channel is said to be *time dispersive*. In a digital communications system, the effect of time dispersion is to cause intersymbol interference and a corresponding degradation in probability of symbol error.

A channel exhibits frequency selectivity when the transmitted bandwidth of the signal is sufficiently larger than the coherence bandwidth of the channel. Stated another way, frequency selectivity is observed if the delay spread L of the channel is sufficiently larger than the transmitted signal duration T. This channel is simultaneously time flat since the duration of a transmitted pulse is sufficiently small compared to the fading correlation time. For example, a high-speed troposcatter channel (in the range of megabits per second) often exhibits frequency selectivity since the transmitted bandwidth is larger than the channel coherence bandwidth.

Time-Selective Channel

The time-selective channel has the property that the impulse response is time variant. An approximately constant amplitude and phase can only be observed over a time interval sufficiently small. The fading which occurs at some frequency will be very nearly the same, that is, highly correlated, for closely spaced times, and essentially independent for times sufficiently far apart. The lack of correlation in fading of a given frequency over time intervals is thus called **time-selective fading**.

A fading channel that is selective only in time is then implicitly also a nonfrequency selective or *frequency-flat* channel. Such a channel has the property that the amplitude and phase fluctuations observed on the channel response to a given frequency are the same for any frequency. The frequency-flat fading channel is observed when the signal bandwidth is smaller than the correlation bandwidth of the channel. Low-speed data transmission over an HF channel (in the range of kilobits per second) is an example of time selectivity wherein the transmitted pulse width is greater than the average fade duration.

For the frequency-flat fading channel the channel impulse response is independent of the frequency variable, so that the input–output relationship (11.3) becomes

$$w(t) = z(t)g(t) \tag{11.12}$$

in which $g(t)$ is the time-varying, frequency-invariant, impulse response. The complex correlation function of $g(t)$ then becomes

$$E[g^*(t)g(t+\tau)] = p(\tau) = R(0, \tau) \tag{11.13}$$

where the function $p(\tau)$ is the *time* (*fading*) *correlation function*. A sufficient characterization of this channel is provided by either the time correlation function $p(\tau)$, or its Fourier transform $P(\nu)$, the *Doppler power spectrum*.

To make an experimental determination of $p(\tau)$, one could simply transmit a sine wave and evaluate the autocorrelation function of the received process. Here one may define a **coherence time** parameter (also called *fading time constant*) T_c in terms of $p(\tau)$ in the same way as B_c was defined in terms of $q(\Omega)$. The channel parameter T_c is then a measure of the average fade duration and is particularly useful in predicting the occurrence of time-selective fading in the channel, just as B_c is useful in predicting the occurrence of frequency-selective fading.

The Fourier transform of the time-correlation function is the Doppler power spectrum $P(\nu)$, which yields signal power as a function of Doppler shift. The total **Doppler spread** B_T and 2σ Doppler spread $B_{2\sigma}$ parameters are defined analogous to delay spread L_T and $L_{2\sigma}$, respectively, and have analogous utilities. Note that the coherence time and Doppler spread are reciprocals of one another, due to the Fourier transform relationship between $p(\tau)$ and $P(\nu)$. If the Doppler spread B is greater than the transmitted signal bandwidth W, then the received signal will be spread over $W + B$. For the case $B > W$, the received signal is spread in frequency and the corresponding channel is said to be *frequency dispersive*.

Observation of Doppler effects is evidence that the channel is behaving as a time-varying filter. These time variations in the channel can be related directly to motion in the medium or motion by the communications device. In the case of a troposcatter channel, fluctuations caused by motion in the scatterers within the medium cause Doppler effects. Motion of the user or vehicle in a mobile communications channel likewise gives rise to Doppler effects.

Time- and Frequency-Selective Channel

Channels that exhibit both time and frequency selectivity are called doubly dispersive. Such channels are neither time flat nor frequency flat. In general, real channels do not exhibit time and frequency selectivity simultaneously. To simultaneously exhibit time and frequency selectivity, the coherence bandwidth must simultaneously be both larger and smaller than the transmitted bandwidth, or alternatively, the multipath spread must simultaneously be both larger and smaller than the signal duration. In view of this contradiction, we assume that either time or frequency selectivity effects can occur with this channel but not simultaneously. For this channel we then can make the often used approximation

$$S(\xi, \nu) = P(\nu)Q(\xi) \tag{11.14}$$

which thus separates the frequency and time selectivity effects on the channel. Equation (11.14) points out the fact that time and frequency selectivity are independent phenomena.

Nonselective Channel

The nonselective fading channel exhibits neither frequency nor time selectivity. This channel is also called *time flat, frequency flat*, or just flat-flat. This channel is then the random-phase Rayleigh fading channel. For a transmitted signal given by Eq. (11.1), the received signal is

$$y(t) = A \operatorname{Re}[z(t)e^{j(2\pi f_c t-\theta)}] \tag{11.15}$$

where A and θ are statistically independent random variables that are time invariant, the former being Rayleigh distributed and the latter being uniformly distributed in the interval $-\pi$ to π.

The flat-flat fading model can closely approximate the behavior of a channel if the bandwidth of the transmitted signal is much less than the correlation bandwidth of the channel and if the transmitted pulse width is very much less than the correlation time of the fading. However, these two requirements are usually incompatible since the time-bandwidth product of the transmitted signal has a lower bound.

11.3 Line-of-Sight Channel Models

Multipath channel models for line-of-sight (LOS) radio systems assume a direct path in addition to one or more secondary paths observed at the receiver. The channel impulse response is modeled as a nonzero mean, complex Gaussian process, where the envelope has a Ricean distribution and the channel is known as a *Ricean fading channel*. Both multipath and polynomial transfer functions have been used to model the LOS channel.

Several multipath transfer function models have been developed, usually based on the presence of two [Jakes, 1978] or three [Rummler, 1979] rays. In general, the multipath channel transfer function can be written as

$$H(\omega) = 1 + \sum_{i=1}^{n} \beta_i e^{j\omega\tau_i} \tag{11.16}$$

where the direct ray has been normalized to unity and the β_i and τ_i are amplitude and delay of the interfering rays relative to the direct ray. The two-ray model can thus be characterized by two parameters, β and τ. In this case, the amplitude of the resultant signal is

$$R = (1 + \beta^2 + 2\beta\cos\omega\tau)^{1/2} \tag{11.17}$$

and the phase of the resultant is

$$\phi = \arctan\left(\frac{\beta \sin \omega\tau}{1 + \beta \cos \omega\tau}\right) \tag{11.18}$$

The group delay is then

$$T(\omega) = \frac{\mathrm{d}\phi}{\mathrm{d}\omega} = \beta\tau\left(\frac{\beta + \cos \omega\tau}{1 + 2\beta \cos \omega\tau + \beta^2}\right) \tag{11.19}$$

The deepest fade occurs with

$$\omega_d \tau = \pi(2n - 1) \qquad (n = 1, 2, 3, \ldots) \tag{11.20}$$

where both R and T are at a minimum, with

$$R_{\min} = 1 - \beta \tag{11.21}$$

$$T_{\min} = \frac{\beta\tau}{1 - \beta} \tag{11.22}$$

The frequency defined by ω_d is known as the notch frequency and is related to the carrier frequency ω_c by

$$\omega_d = \omega_c + \omega_0 \tag{11.23}$$

where ω_0 is referred to as the offset frequency.

Although the two-ray model described here is easy to understand and apply, most multipath propagation research points toward the presence of three (or more) rays during fading conditions. Out of this research, Rummler's three-ray model is the most widely accepted [Rummler, 1979].

Experiments conducted to develop or verify multipath channel models are limited by the narrowband channels available in existing radio systems. These limited channel bandwidths are generally insufficient to resolve parameters of a two- or three-ray multipath model. One alternative is to use polynomial functions of frequency to analytically fit measured data. Such polynomials have been used to describe amplitude and group delay distortion. Curve fits can be made to individual records of amplitude and group delay distortion using an Mth-order polynomial,

$$p(\omega) = C_0 + C_1\omega + C_2\omega^2 + \cdots + C_M\omega^M \tag{11.24}$$

To obtain the coefficients $\{C_1, C_2, \ldots, C_M\}$, a least squares fit has been used with polynomials typically of order $M = 2$, 4, and 6. For fades that are frequency selective, the most suitable order has been found to be $M = 4$. During periods of nonfading or flat fading, polynomials of order $M = 0$ or 2 have provided acceptable accuracy [Smith and Cormack, 1982]. As shown by Greenstein and Czekaj [1980] these polynomial coefficients can be related to a power series model of the channel transfer function

$$H(\omega) = A_0 + \sum_{k=1}^{n} (A_k + jB_k)(j\omega)^k \tag{11.25}$$

Studies have shown that at least a first-order polynomial, with three coefficients $\{A_0, A_1, \text{and } B_1\}$, is required for LOS channels, and that a second-order polynomial with five coefficients may be required for highly dispersive channels [Greenstein and Czekaj, 1980].

11.4 Digital Channel Models

For digital communications channels, one must be able to relate probability of (symbol, bit) error to the behavior of real channels, which tend to introduce errors in bursts. Two approaches may be taken. The classical method relates signal level variation to error rate statistics. Here, an equation for probability of error for a particular modulation technique is used to develop error statistics such as outage probability, outage rate, and outage duration for a given probability density function of signal level representing a certain channel. The drawback to this approach is the limitation of the results to only classical channels, such as Gaussian or Rayleigh, for which the signal statistics are known. As shown earlier, real channels tend to exhibit complex and time-varying behavior not easily represented by known statistics. For example, radio channels may experience time-selective fading, frequency-selective fading, rain attenuation, interference, and other forms of degradation at one time or another.

The second approach uses error models for channels with memory, that is, for channels where errors occur in bursts. If based on actual channel data, error models have the advantage of exactly reflecting the vagaries of the channel and, depending on the model, allowing various error statistics to be represented as a function of the model's parameters. The chief problem with error modeling is the selection of the model and its parameters to provide a good fit to the data and to the channel(s) in general. The Gilbert model focused on the use of a Markov chain composed of a good state G, which is error free, and a bad state B, with a certain probability of error [Gilbert, 1960]. Subsequent applications used a larger number of states in the Markov chain to better simulate error distributions in real channels. Fritchman [1967] first introduced the simple partitioned Markov chain with k error-free states and $N - k$ error states, which has shown good agreement with experimental results [Knowles and Drukarev, 1988; Semmar et al., 1991].

The single-error partitioned Fritchman model, in which the N states are divided into $N - 1$ error-free states and one error state, with transition probabilities P_{ij}, $i, j = 1$ to N, has been popularly used to characterize sequences of error-free and errored events (e.g., bits or symbols), known as gaps and clusters, respectively. For a binary error sequence where the presence or absence of a bit error is represented by 1 and 0, the error gap distribution $P(0^m|1)$ is the probability that at least m error-free bits will follow given that an errored bit has occurred. Similarly, the error cluster distribution $P(1^m|0)$ is the probability that at least m errored bits will follow given that an error-free bit has occurred. For the single-error Fritchman model [1967],

$$P(0^m|1) = \sum_{i=1}^{k} \frac{P_{N,i}}{P_{i,i}} (P_{i,i})^m \tag{11.26}$$

or

$$P(0^m|1) = \sum_{i=1}^{k} \alpha_i \beta_i^m \tag{11.27}$$

where the values α_i or β_i are the Fritchman model parameters. For $i = 1$ to k and $N - k = 1$, the transition probabilities are related to Fritchman's parameters by

$$P_{i,i} = \beta_i \tag{11.28}$$

$$P_{i,N} = 1 - \beta_i \tag{11.29}$$

$$P_{N,i} = \alpha_i \beta_i \tag{11.30}$$

$$P_{N,N} = 1 - \sum_{i=1}^{k} \alpha_i \beta_i \tag{11.31}$$

After obtaining $P(0^m|1)$ empirically for a given channel, curve fitting techniques can be employed to represent $P(0^m|1)$ as the sum of $(N-1)$ exponentials, with coefficients α_i and β_i. Even simpler, for the single error state model, the error cluster distribution can be described with a single exponential with suitable coefficients α_N and β_N. Fritchman [1967] and others have found that a model with $k = 2$ or 3 exponential functions is generally sufficient. For high error rate sequences, a model with up to four exponential functions is required [Semmar et al., 1991].

Defining Terms

Coherence bandwidth: Bandwidth of a received signal for which fading is highly correlated.

Coherence time: Time interval of a received signal for which fading is highly correlated.

Doppler spread: Range of frequency over which the Doppler power spectrum is significantly different from zero.

Frequency-selective fading: Fading with high correlation for closely spaced frequencies and no correlation for widely spaced frequencies.

Multipath spread or delay spread: Range of time over which the delay power spectrum is significantly different from zero.

Time-selective fading: Fading with high correlation for closely spaced times and no correlation for widely spaced times.

References

Becam, D., Brigant, P., Cohen, R., and Szpirglas, J. 1984. Testing Neyman's model for error performance of 2 and 140 Mb/s line sections. In *1984 International Conference on Communications,* pp. 1362–1365.

Bello, P.A. 1963. Characterization of randomly time-variant linear channels. *IEEE Trans. Commun.,* CS-11(Dec.):360–393.

Fritchman, B.D. 1967. A binary channel characterization using partitioned Markov chains. *IEEE Trans. Commun.,* IT-13(April):221–227.

Gilbert, E.N. 1960. Capacity of a burst noise channel. *Bell Syst. Tech. J.,* 39(Sept.):1253–1265.

Greenstein, L.C. and Czekaj, B.A. 1980. A polynomial model for multipath fading channel responses. *Bell Syst. Tech. J.,* 59(Sept.):1197–1205.

Jakes, W.C. 1978. An approximate method to estimate an upper bound on the effect of multipath delay distortion on digital transmission. In *1978 International Conference on Communications,* pp. 47.1.1–47.1.5.

Kennedy, R.S. 1969. *Fading Dispersive Communication Channels,* Wiley, New York.

Knowles, M.D. and Drukarev, A.I. 1988. Bit error rate estimation for channels with memory. *IEEE Trans. Commun.,* COM-36(June):767–769.

Neyman, J. 1939. On a new class of contagious distribution, applicable in entomology and bacteriology. *Ann. Math. Stat.,* 10:35–57.

Rummler, W.D. 1979. A new selective fading model: Application to propagation data. *Bell Syst. Tech. J.,* 58(May/June):1037–1071.

Semmar, A., Lecours, M., Chouinard, J., and Ahern, J. 1991. Characterization of error sequences in UHF digital mobile radio channels. *IEEE Trans. Veh. Tech.,* 40(Nov.):769–776.

Smith, D.R. and Cormack, J.J. 1982. Measurement and haracterization of a multipath fading channel. In *1982 International Conference on Communications,* pp. 7B.4.1–7B.4.6.

Stein, S., Schwartz, M., and Bennett, W. 1966. *Communication Systems and Techniques,* McGraw–Hill, New York.

Wozencraft, J.M. and Jacobs, I.M. 1967. *Principles of Communication Engineering,* Wiley, New York.

Further Information

A comprehensive treatment of channel models is given in *Fading Dispersive Communication Channels* by R.S. Kennedy. *Data Communications via Fading Channels*, an IEEE volume edited by K. Brayer, provides an excellent collection of papers on the subject.

12

Optimum Receivers

Geoffrey C. Orsak
Southern Methodist University

12.1 Introduction

Every engineer strives for optimality in design. This is particularly true for communications engineers since in many cases implementing suboptimal receivers and sources can result in dramatic losses in performance. As such, this chapter focuses on design principles leading to the implementation of optimum receivers for the most common communication environments.

The main objective in digital communications is to transmit a sequence of bits to a remote location with the highest degree of accuracy. This is accomplished by first representing bits (or more generally short bit sequences) by distinct waveforms of finite time duration. These time-limited waveforms are then transmitted (broadcasted) to the remote sites in accordance with the data sequence.

Unfortunately, because of the nature of the **communication channel**, the remote location receives a corrupted version of the concatenated signal waveforms. The most widely accepted model for the communication channel is the so-called **additive white Gaussian noise**[1] **channel (AWGN channel)**. Mathematical arguments based upon the central limit theorem [7], together with supporting empirical evidence, demonstrate that many common communication channels are accurately modeled by this abstraction. Moreover, from the design perspective, this is quite fortuitous since design and analysis with respect to this channel model is relatively straightforward.

[1] For those unfamiliar with AWGN, a random process (waveform) is formally said to be white Gaussian noise if all collections of instantaneous observations of the process are jointly Gaussian and mutually independent. An important consequence of this property is that the power spectral density of the process is a constant with respect to frequency variation (spectrally flat). For more on AWGN, see Papoulis [4].

12.2 Preliminaries

To better describe the digital communications process, we shall first elaborate on so-called binary communications. In this case, when the source wishes to transmit a bit value of 0, the transmitter broadcasts a specified waveform $s_0(t)$ over the **bit interval** $t \in [0, T]$. Conversely, if the source seeks to transmit the bit value of 1, the transmitter alternatively broadcasts the signal $s_1(t)$ over the same bit interval. The received waveform $R(t)$ corresponding to the first bit is then appropriately described by the following hypotheses testing problem:

$$\begin{aligned} H_0\colon R(t) &= s_0(t) + \eta(t) \qquad 0 \le t \le T \\ H_1\colon R(t) &= s_1(t) + \eta(t) \end{aligned} \tag{12.1}$$

where, as stated previously, $\eta(t)$ corresponds to AWGN with spectral height nominally given by $N_0/2$. It is the objective of the receiver to determine the bit value, i.e., the most accurate hypothesis from the received waveform $R(t)$.

The optimality criterion of choice in digital communication applications is the **total probability of error** normally denoted as P_e. This scalar quantity is expressed as

$$\begin{aligned} P_e = {} & Pr(\text{declaring } 1|0 \text{ transmitted})Pr(0 \text{ transmitted}) \\ & + Pr(\text{declaring } 0|1 \text{ transmitted})Pr(1 \text{ transmitted}) \end{aligned} \tag{12.2}$$

The problem of determining the optimal binary receiver with respect to the probability of error is solved by applying stochastic representation theory [10] to detection theory [5, 9]. The specific waveform representation of relevance in this application is the **Karhunen–Loève (KL) expansion**.

12.3 Karhunen–Loève Expansion

The Karhunen–Loève expansion is a generalization of the Fourier series designed to represent a random process in terms of deterministic basis functions and uncorrelated random variables derived from the process. Whereas the Fourier series allows one to model or represent deterministic time-limited energy signals in terms of linear combinations of complex exponential waveforms, the Karhunen–Loève expansion allows us to represent a second-order random process in terms of a set of **orthonormal** basis functions scaled by a sequence of random variables. The objective in this representation is to choose the basis of time functions so that the coefficients in the expansion are mutually uncorrelated random variables.

To be more precise, if $R(t)$ is a zero mean second-order random process defined over $[0, T]$ with covariance function $K_R(t, s)$, then so long as the basis of deterministic functions satisfy certain integral constraints [9], one may write $R(t)$ as

$$R(t) = \sum_{i=1}^{\infty} R_i \phi_i(t) \qquad 0 \le t \le T \tag{12.3}$$

where

$$R_i = \int_0^T R(t)\phi_i(t)\,\mathrm{d}t$$

In this case the R_i will be mutually uncorrelated random variables with the ϕ_i being deterministic basis functions that are complete in the space of square integrable time functions over $[0, T]$. Importantly, in this case, equality is to be interpreted as **mean-square equivalence**, i.e.,

$$\lim_{N \to \infty} E\left[\left(R(t) - \sum_{i=1}^{N} R_i \phi_i(t)\right)^2\right] = 0$$

for all $0 \le t \le T$.

FACT 12.1 *If $R(t)$ is AWGN, then any basis of the vector space of square integrable signals over $[0, T]$ results in uncorrelated and therefore independent Gaussian random variables.*

The use of Fact 12.1 allows for a conversion of a continuous time detection problem into a finite-dimensional detection problem. Proceeding, to derive the optimal binary receiver, we first construct our set of basis functions as the set of functions defined over $t \in [0, T]$ beginning with the signals of interest $s_0(t)$ and $s_1(t)$. That is,

$$\{s_0(t),\ s_1(t),\ \text{plus a countable number of functions which complete the basis}\}$$

In order to ensure that the basis is orthonormal, we must apply the Gramm–Schmidt procedure* [6] to the full set of functions beginning with $s_0(t)$ and $s_1(t)$ to arrive at our final choice of basis $\{\phi_i(t)\}$.

FACT 12.2 *Let $\{\phi_i(t)\}$ be the resultant set of basis functions.*
Then for all $i > 2$, the $\phi_i(t)$ are orthogonal to $s_0(t)$ and $s_1(t)$. That is,

$$\int_0^T \phi_i(t) s_j(t)\, \mathrm{d}t = 0$$

for all $i > 2$ and $j = 0, 1$.

Using this fact in conjunction with Eq. (12.3), one may recognize that only the coefficients R_1 and R_2 are functions of our signals of interest. Moreover, since the R_i are mutually independent, the optimal receiver will, therefore, only be a function of these two values.

Thus, through the application of the KL expansion, we arrive at an equivalent hypothesis testing problem to that given in Eq. (12.1),

$$\begin{aligned} H_0\colon \boldsymbol{R} &= \begin{bmatrix} \int_0^T \phi_1(t) s_0(t)\ \mathrm{d}t \\ \int_0^T \phi_2(t) s_0(t)\ \mathrm{d}t \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \\ H_1\colon \boldsymbol{R} &= \begin{bmatrix} \int_0^T \phi_1(t) s_1(t)\ \mathrm{d}t \\ \int_0^T \phi_2(t) s_1(t)\ \mathrm{d}t \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \end{aligned} \tag{12.4}$$

where it is easily shown that η_1 and η_2 are mutually independent, zero-mean, Gaussian random variables with variance given by $N_0/2$, and where ϕ_1 and ϕ_2 are the first two functions from our orthonormal set of basis functions. Thus, the design of the optimal binary receiver reduces to a simple two-dimensional detection problem that is readily solved through the application of detection theory.

*The Gramm-Schmidt procedure is a deterministic algorithm that simply converts an arbitrary set of basis functions (vectors) into an equivalent set of orthonormal basis functions (vectors).

12.4 Detection Theory

It is well known from detection theory [5] that under the minimum P_e criterion, the optimal detector is given by the *maximum a posteriori rule* (*MAP*),

$$\text{choose}_i \ \text{ largest } p_{H_i|\boldsymbol{R}}(H_i \,|\, \boldsymbol{R} = \boldsymbol{r}) \tag{12.5}$$

i.e., determine the hypothesis that is most likely, given that our observation vector is $\boldsymbol{r}$. By a simple application of Bayes theorem [4], we immediately arrive at the central result in detection theory: the optimal binary detector is given by the likelihood ratio test (LRT),

$$L(\boldsymbol{R}) = \frac{p_{\boldsymbol{R}|H_1}(\boldsymbol{R})}{p_{\boldsymbol{R}|H_0}(\boldsymbol{R})} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\pi_0}{\pi_1} \tag{12.6}$$

where the π_i are the *a priori* probabilities of the hypotheses H_i being true. Since in this case we have assumed that the noise is white and Gaussian, the LRT can be written as

$$L(\boldsymbol{R}) = \frac{\prod_1^2 \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{1}{2}\frac{(R_i - s_{1,i})^2}{N_0/2}\right)}{\prod_1^2 \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{1}{2}\frac{(R_i - s_{0,i})^2}{N_0/2}\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\pi_0}{\pi_1} \tag{12.7}$$

where

$$s_{j,i} = \int_0^T \phi_i(t) s_j(t)\,\mathrm{d}t$$

By taking the logarithm and cancelling common terms, it is easily shown that the optimum binary receiver can be written as

$$\frac{2}{N_0}\sum_1^2 \boldsymbol{R}_i(s_{1,i} - s_{0,i}) - \frac{1}{N_0}\sum_1^2 (s_{1,i}^2 - s_{0,i}^2) \underset{H_0}{\overset{H_1}{\gtrless}} \ln\frac{\pi_0}{\pi_1} \tag{12.8}$$

This finite-dimensional version of the optimal receiver can be converted back into a continuous time receiver by the direct application of Parseval's theorem [4] where it is easily shown that

$$\begin{aligned} \sum_{i=1}^2 R_i s_{k,i} &= \int_0^T R(t) s_k(t)\,\mathrm{d}t \\ \sum_{i=1}^2 s_{k,i}^2 &= \int_0^T s_k^2(t)\,\mathrm{d}t \end{aligned} \tag{12.9}$$

By applying Eq. (12.9) to Eq. (12.8) the final receiver structure is then given by

$$\int_0^T R(t)[s_1(t) - s_0(t)]\,\mathrm{d}t - \frac{1}{2}(E_1 - E_0) \underset{H_0}{\overset{H_1}{\gtrless}} \frac{N_0}{2}\ln\frac{\pi_0}{\pi_1} \tag{12.10}$$

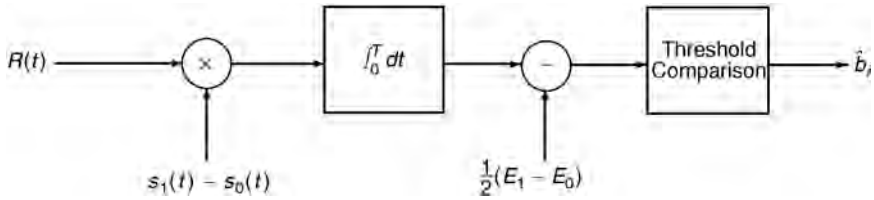


FIGURE 12.1 Optimal correlation receiver structure for binary communications.

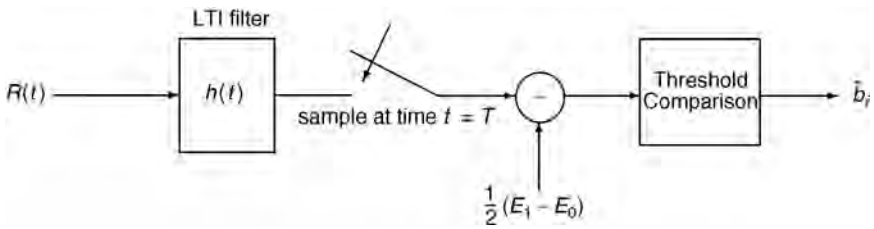


FIGURE 12.2 Optimal matched filter receiver structure for binary communications. In this case $h(t) = s(T - t) - s_0(t - t)$.

where E_1 and E_0 are the energies of signals $s_1(t)$ and $s_0(t)$, respectively. (See Fig. 12.1 for a block diagram.) Importantly, if the signals are equally likely ($\pi_0 = \pi_1$), the optimal receiver is independent of the typically unknown spectral height of the background noise.

One can readily observe that the optimal binary communication receiver correlates the received waveform with the difference signal $s_1(t) - s_0(t)$ and then compares the statistic to a threshold. This operation can be interpreted as identifying the signal waveform $s_i(t)$ that best correlates with the received signal $R(t)$. Based on this interpretation, the receiver is often referred to as the **correlation receiver**.

As an alternate means of implementing the correlation receiver, we may reformulate the computation of the left-hand side of Eq. (12.10) in terms of standard concepts in filtering. Let $h(t)$ be the impulse response of a linear, time-invariant (LTI) system. By letting $h(t) = s_1(T - t) - s_0(T - t)$, then it is easily verified that the output of $R(t)$ to a LTI system with impulse response given by $h(t)$ and then sampled at time $t = T$ gives the desired result. (See Fig. 12.2 for a block diagram.) Since the impulse response is matched to the signal waveforms, this implementation is often referred to as the **matched filter receiver**.

12.5 Performance

Because of the nature of the statistics of the channel and the relative simplicity of the receiver, performance analysis of the optimal binary receiver in AWGN is a straightforward task. Since the conditional statistics of the log likelihood ratio are Gaussian random variables, the probability of error can be computed directly in terms of Marcum Q functions* as

$$P_e = Q\left(\frac{\|\boldsymbol{s}_0 - \boldsymbol{s}_1\|}{\sqrt{2N_0}}\right)$$

where the $\boldsymbol{s}_i$ are the two-dimensional signal vectors obtained from Eq. (12.4), and where $\|\boldsymbol{x}\|$ denotes the Euclidean length of the vector $\boldsymbol{x}$. Thus, $\|\boldsymbol{s}_0 - \boldsymbol{s}_1\|$ is best interpreted as the distance between the respective signal representations. Since the Q function is monotonically decreasing with an increasing argument,

*The Q function is the probability that a standard normal random variable exceeds a specified constant, i.e., $Q(x) = \int_x^\infty 1/\sqrt{2\pi}\exp(-z^2/2)\,dz$.

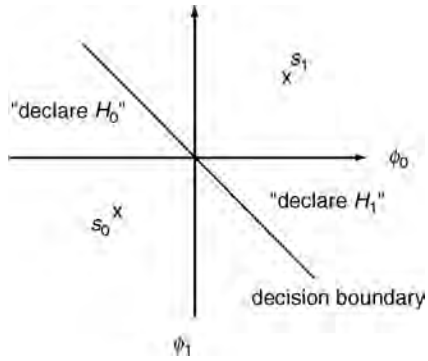


FIGURE 12.3 Signal space and decision boundary for optimal binary receiver.

one may recognize that the probability of error for the optimal receiver decreases with an increasing separation between the signal representations, i.e., the more dissimilar the signals, the lower the P_e.

12.6 Signal Space

The concept of a **signal space** allows one to view the signal classification problem (receiver design) within a geometrical framework. This offers two primary benefits: first, it supplies an often more intuitive perspective on the receiver characteristics (e.g., performance) and second, it allows for a straightforward generalization to standard M-ary signalling schemes.

To demonstrate this, in Fig. 12.3 we have plotted an arbitrary signal space for the binary signal classification problem. The axes are given in terms of the basis functions $\phi_1(t)$ and $\phi_2(t)$. Thus, every point in the signal space is a time function constructed as a linear combination of the two basis functions. By Fact 12.2 we recall that both signals $s_0(t)$ and $s_1(t)$ can be constructed as a linear combination of $\phi_1(t)$ and $\phi_2(t)$ and as such we may identify these two signals in this figure as two points.

Since the decision statistic given in Eq. (12.8) is a linear function of the observed vector $\boldsymbol{R}$ which is also located in the signal space, it is easily shown that the set of vectors under which the receiver declares hypothesis H_i is bounded by a line in the signal space. This so-called **decision boundary** is obtained by solving the equation $\ln[L(\boldsymbol{R})] = 0$. (Here again we have assumed equally likely hypotheses.) In the case under current discussion, this decision boundary is simply the hyperplane separating the two signals in signal space. Because of the generality of this formulation, many problems in communication system design are best cast in terms of the signal space, that is, signal locations and decision boundaries.

12.7 Standard Binary Signalling Schemes

The framework just described allows us to readily analyze the most popular signalling schemes in binary communications: amplitude-shift keying (ASK), frequency-shift keying (FSK), and phase-shift keying (PSK). Each of these examples simply constitutes a different selection for signals $s_0(t)$ and $s_1(t)$.

In the case of ASK, $s_0(t) = 0$, while $s_1(t) = \sqrt{2E/T}\sin(2\pi f_c t)$, where E denotes the energy of the waveform and f_c denotes the frequency of the carrier wave with $f_c T$ being an integer. Because $s_0(t)$ is the null signal, the signal space is a one-dimensional vector space with $\phi_1(t) = \sqrt{2/T}\sin(2\pi f_c t)$. This, in turn, implies that $\|s_0 - s_1\| = \sqrt{E}$. Thus, the corresponding probability of error for ASK is

$$P_e(\text{ASK}) = Q\left(\sqrt{\frac{E}{2N_0}}\right)$$

For FSK, the signals are given by equal amplitude sinusoids with distinct center frequencies, that is, $s_i(t) = \sqrt{2E/T}\sin(2\pi f_i t)$ with $f_i T$ being two distinct integers. In this case, it is easily verified that the signal

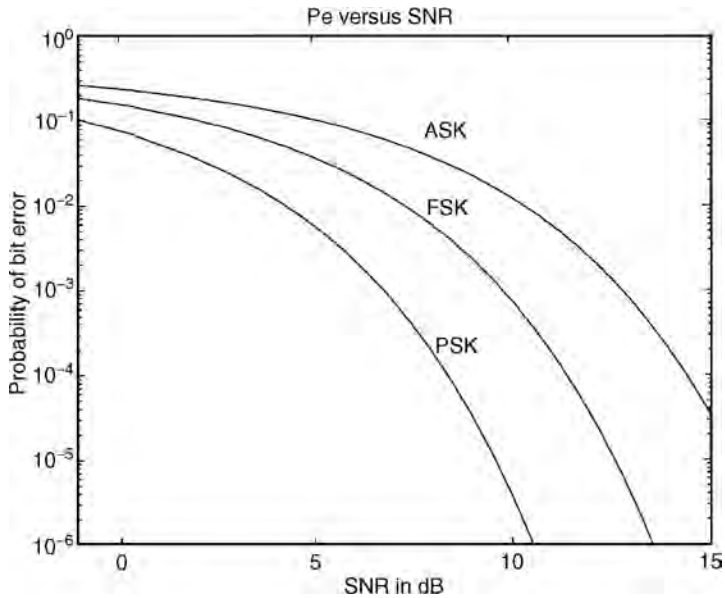


FIGURE 12.4 *Pe* vs. the signal to noise ratio in decibels [dB = $10\log(E/N_0)$] for amplitude-shift keying, frequency-shift keying, and phase-shift keying; note that there is a 3-dB difference in performance from ASK to FSK to PSK.

space is a two-dimensional vector space with $\phi_i(t) = \sqrt{2/T}\sin(2\pi f_i t)$ resulting in $\|\boldsymbol{s}_0 - \boldsymbol{s}_1\| = \sqrt{2\boldsymbol{E}}$. The corresponding error rate is given to be

$$P_e(\text{FSK}) = Q\left(\sqrt{\frac{E}{N_0}}\right)$$

Finally, with regard to PSK signalling, the most frequently utilized binary PSK signal set is an example of an antipodal signal set. Specifically, the antipodal signal set results in the greatest separation between the signals in the signal space subject to an energy constraint on both signals. This, in turn, translates into the energy constrained signal set with the minimum P_e. In this case, the $s_i(t)$ are typically given by $\sqrt{2E/T}\sin[2\pi f_c t + \theta(i)]$, where $\theta(0) = 0$ and $\theta(1) = \pi$. As in the ASK case, this results in a one-dimensional signal space, however, in this case $\|\boldsymbol{s}_0 - \boldsymbol{s}_1\| = 2\sqrt{\boldsymbol{E}}$, resulting in probability of error given by

$$P_e(\text{PSK}) = Q\left(\sqrt{\frac{2E}{N_0}}\right)$$

In all three of the described cases, one can readily observe that the resulting performance is a function of only the signal-to-noise ratio E/N_0. In the more general case, the performance will be a function of the intersignal energy to noise ratio. To gauge the relative difference in performance of the three signalling schemes, in Fig. 12.4, we have plotted the P_e as a function of the SNR. Please note the large variation in performance between the three schemes for even moderate values of SNR.

12.8 *M*-ary Optimal Receivers

In binary signalling schemes, one seeks to transmit a single bit over the bit interval $[0, T]$. This is to be contrasted with M-ary signalling schemes where one transmits multiple bits simultaneously over the so-called symbol interval $[0, T]$. For example, using a signal set with 16 separate waveforms will allow one to transmit a length four-bit sequence per symbol (waveform). Examples of M-ary waveforms are quadrature phase-shift keying (QPSK) and quadrature amplitude modulation (QAM).

The derivation of the optimum receiver structure for M-ary signalling requires the straightforward application of fundamental results in detection theory. As with binary signalling, the Karhunen–Loève expansion is the mechanism utilized to convert a hypotheses testing problem based on continuous waveforms into a vector classification problem. Depending on the complexity of the M waveforms, the signal space can be as large as an M-dimensional vector space.

By extending results from the binary signalling case, it is easily shown that the optimum M-ary receiver computes

$$\xi_i[R(t)] = \int_0^T s_i(t)R(t)\,\mathrm{d}t - \frac{E_i}{2} + \frac{N_0}{2}\ln \pi_i \qquad i = 1, \ldots, M$$

where, as before, the $s_i(t)$ constitute the signal set with the π_i being the corresponding *a priori* probabilities. After computing M separate values of ξ_i, the minimum probability of error receiver simply chooses the largest amongst this set. Thus, the M-ary receiver is implemented with a bank of correlation or matched filters followed by choose-largest decision logic.

In many cases of practical importance, the signal sets are selected so that the resulting signal space is a two-dimensional vector space irrespective of the number of signals. This simplifies the receiver structure in that the sufficient statistics are obtained by implementing only two matched filters. Both QPSK and QAM signal sets fit into this category. As an example, in Fig. 12.5, we have depicted the signal locations for standard 16-QAM signalling with the associated decision boundaries. In this case, we have assumed an equally likely signal set. As can be seen, the optimal decision rule selects the signal representation that is closest to the received signal representation in this two-dimensional signal space.

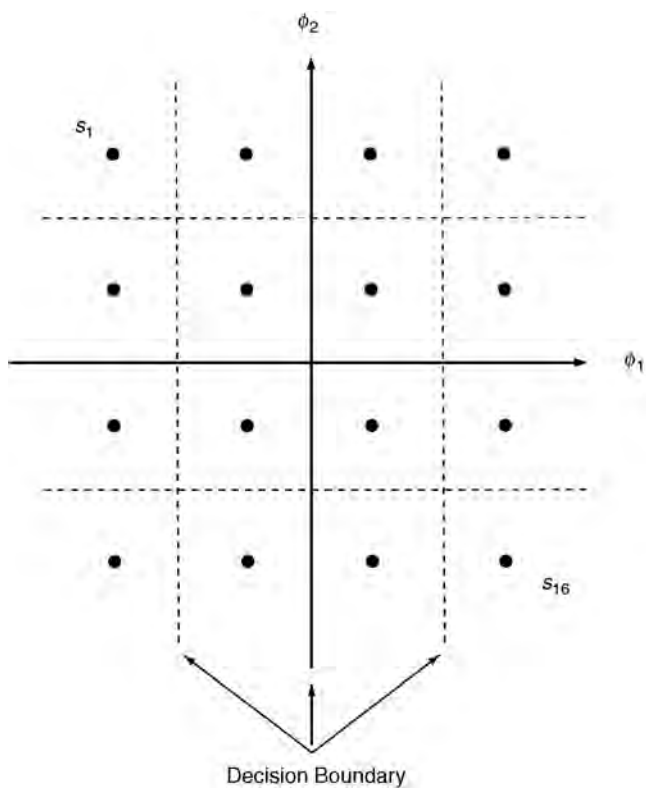


FIGURE 12.5 Signal space representation of 16-QAM signal set. Optimal decision regions for equally likely signals are also noted.

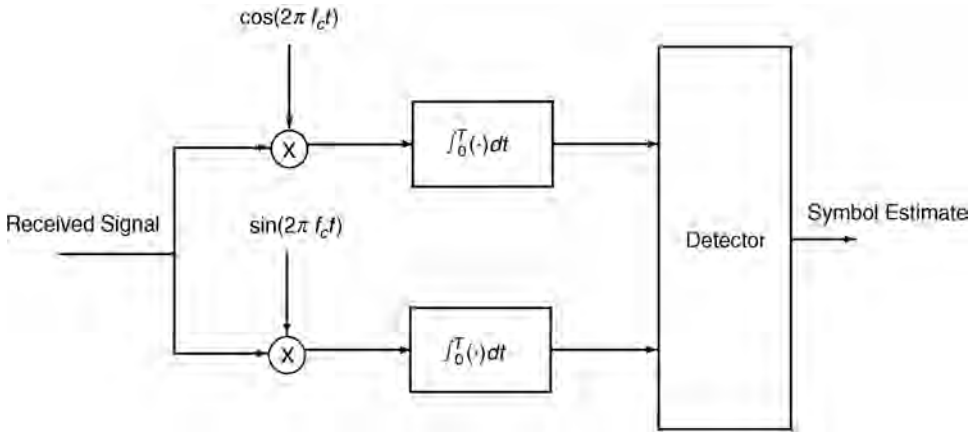


FIGURE 12.6 Optimum receiver structure for noncoherent (random or unknown phase) ASK demodulation.

12.9 More Realistic Channels

As is unfortunately often the case, many channels of practical interest are not accurately modeled as simply an AWGN channel. It is often that these channels impose nonlinear effects on the transmitted signals. The best example of this is channels that impose a random phase and random amplitude onto the signal. This typically occurs in applications such as in mobile communications, where one often experiences rapidly changing path lengths from source to receiver.

Fortunately, by the judicious choice of signal waveforms, it can be shown that the selection of the ϕ_i in the Karhunen–Loève transformation is often independent of these unwanted parameters. In these situations, the random amplitude serves only to scale the signals in signal space, whereas the random phase simply imposes a rotation on the signals in signal space.

Since the Karhunen–Loève basis functions typically do not depend on the unknown parameters, we may again convert the continuous time classification problem to a vector channel problem where the received vector $\boldsymbol{R}$ is computed as in Eq. (12.3). Since this vector is a function of both the unknown parameters (i.e., in this case amplitude A and phase ν), to obtain a likelihood ratio test independent of A and ν, we simply apply Bayes theorem to obtain the following form for the LRT:

$$L(\boldsymbol{R}) = \frac{E[p_{\boldsymbol{R}|H_1,A,\nu}(\boldsymbol{R}|H_1, A, \nu)]}{E[p_{\boldsymbol{R}|H_0,A,\nu}(\boldsymbol{R}|H_0, A, \nu)]} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\pi_0}{\pi_1}$$

where the expectations are taken with respect to A and ν and where $p_{\boldsymbol{R}|H_i,A,\nu}$ are the conditional probability density functions of the signal representations. Assuming that the background noise is AWGN, it can be shown that the LRT simplifies to choosing the largest amongst

$$\xi_i[R(t)] = \pi_i \int_{A,\nu} \exp\left\{\frac{2}{N_0}\int_0^T R(t)s_i(t|A, \nu)\,\mathrm{d}t - \frac{E_i(A, \nu)}{N_0}\right\} p_{A,\nu}(A, \nu)\,\mathrm{d}A\,\mathrm{d}\nu \qquad (12.11)$$

$$i = 1, \ldots, M$$

It should be noted that in Eq. (12.11) we have explicitly shown the dependence of the transmitted signals s_i on the parameters A and ν. The final receiver structures, together with their corresponding performance are, thus, a function of both the choice of signal sets and the probability density functions of the random amplitude and random phase.

Random Phase Channels

If we consider first the special case where the channel simply imposes a uniform random phase on the signal, then it can be easily shown that the so-called in-phase and quadrature statistics obtained from the received signal $R(t)$ (denoted by R_I and R_Q, respectively) are sufficient statistics for the signal classification problem. These quantities are computed as

$$R_I(i) = \int_0^T R(t)\cos[2\pi f_c(i)t]\,\mathrm{d}t$$

and

$$R_Q(i) = \int_0^T R(t)\sin[2\pi f_c(i)t]\,\mathrm{d}t$$

where, in this case, the index i corresponds to the center frequencies of hypotheses H_i, (e.g., FSK signalling). The optimum binary receiver selects the largest from amongst

$$\xi_i[R(t)] = \pi_i \exp\left(-\frac{E_i}{N_0}\right) I_0\left[\frac{2}{N_0}\sqrt{R_I^2(i)+R_Q^2(i)}\right] \qquad i = 1, \ldots, M$$

where I_0 is a zeroth-order, modified Bessel function of the first kind. If the signals have equal energy and are equally likely (e.g., FSK signalling), then the optimum receiver is given by

$$R_I^2(1)+R_Q^2(1) \underset{H_0}{\overset{H_1}{\gtrless}} R_I^2(0)+R_Q^2(0)$$

One may readily observe that the optimum receiver bases its decision on the values of the two envelopes of the received signal $\sqrt{R_I^2(i)+R_Q^2(i)}$ and, as a consequence, is often referred to as an envelope or square-law detector. Moreover, it should be observed that the computation of the envelope is independent of the underlying phase of the signal and is, as such, known as a noncoherent receiver.

The computation of the error rate for this detector is a relatively straightforward exercise resulting in

$$P_e(\text{noncoherent}) = \frac{1}{2}\exp\left(-\frac{E}{2N_0}\right)$$

As before, note that the error rate for the noncoherent receiver is simply a function of the SNR.

Rayleigh Channel

As an important generalization of the described random phase channel, many communication systems are designed under the assumption that the channel introduces both a random amplitude and a random phase on the signal. Specifically, if the original signal sets are of the form $s_i(t) = m_i(t)\cos(2\pi f_c t)$ where $m_i(t)$ is the baseband version of the message (i.e., what distinguishes one signal from another), then the so-called **Rayleigh channel** introduces random distortion in the received signal of the following form:

$$s_i(t) = A m_i(t)\cos(2\pi f_c t + \nu)$$

where the amplitude A is a Rayleigh random variable* and where the random phase ν is a uniformly distributed between zero and 2π.

*The density of a Rayleigh random variable is given by $p_A(a) = a/\sigma^2 \exp(-a^2/2\sigma^2)$ for $a \geq 0$.

To determine the optimal receiver under this distortion, we must first construct an alternate statistical model for $s_i(t)$. To begin, it can be shown from the theory of random variables [4] that if X_I and X_Q are statistically independent, zero mean, Gaussian random variables with variance given by σ^2, then

$$Am_i(t)\cos(2\pi f_c t + \nu) = m_i(t)X_I\cos(2\pi f_c t) + m_i(t)X_Q\sin(2\pi f_c t)$$

Equality here is to be interpreted as implying that both A and ν will be the appropriate random variables. From this, we deduce that the combined uncertainty in the amplitude and phase of the signal is incorporated into the Gaussian random variables X_I and X_Q. The in-phase and quadrature components of the signal $s_i(t)$ are given by $s_{Ii}(t) = m_i(t)\cos(2\pi f_c t)$ and $s_{Q_i}(t) = m_i(t)\sin(2\pi f_c t)$ respectively. By appealing to Eq. (12.11), it can be shown that the optimum receiver selects the largest from

$$\xi_i[R(t)] = \frac{\pi_i}{1 + \frac{2E_i}{N_0}\sigma^2}\exp\left[\frac{\sigma^2}{\frac{1}{2} + \frac{E_i}{N_0}\sigma^2}(\langle R(t), s_{Ii}(t)\rangle^2 + \langle R(t), s_{Q_i}(t)\rangle^2)\right]$$

where the inner product

$$\langle R(t), S_i(t)\rangle = \int_0^T R(t)s_i(t)\,\mathrm{d}t$$

Further, if we impose the conditions that the signals be equally likely with equal energy over the symbol interval, then the optimum receiver selects the largest amongst

$$\xi_i[R(t)] = \sqrt{\langle R(t), s_{Ii}(t)\rangle^2 + \langle R(t), s_{Q_i}(t)\rangle^2}$$

Thus, much like for the random phase channel, the optimum receiver for the Rayleigh channel computes the projection of the received waveform onto the in-phase and quadrature components of the hypothetical signals. From a signal space perspective, this is akin to computing the length of the received vector in the subspace spanned by the hypothetical signal. The optimum receiver then chooses the largest amongst these lengths.

As with the random phase channel, computing the performance is a straightforward task resulting in (for the equally likely equal energy case)

$$P_e(\text{Rayleigh}) = \frac{\frac{1}{2}}{\left(1 + \frac{E\sigma^2}{N_0}\right)}$$

Interestingly, in this case, the performance depends not only on the SNR, but also on the variance (spread) of the Rayleigh amplitude A. Thus, if the amplitude spread is large, we expect to often experience what is known as deep fades in the amplitude of the received waveform and, as such, expect a commensurate loss in performance.

12.10 Dispersive Channels

The **dispersive channel** model assumes that the channel not only introduces AWGN but also distorts the signal through a filtering process. This model incorporates physical realities such as multipath effects and frequency selective fading. In particular, the standard model adopted is depicted in the block diagram given in Fig. 12.7. As can be seen, the receiver observes a filtered version of the signal plus AWGN. If the impulse response of the channel is known, then we arrive at the optimum receiver design by applying the previously presented theory. Unfortunately, the duration of the filtered signal can be a complicating factor.

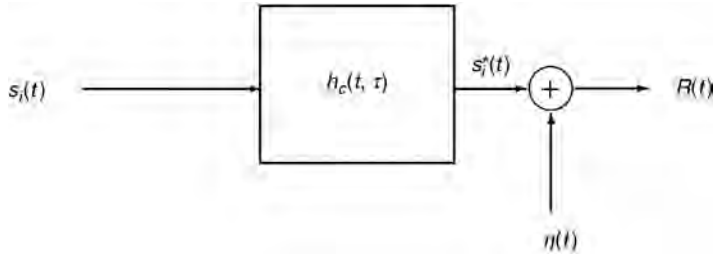


FIGURE 12.7 Standard model for dispersive channel. The time varying impulse response of the channel is denoted by $h_c(t, \tau)$.

More often than not, the channel will increase the duration of the transmitted signals, hence, leading to the description, dispersive channel.

However, if the designers take this into account by shortening the duration of $s_i(t)$ so that the duration of $s_i^*(t)$ is less than T, then the optimum receiver chooses the largest amongst

$$\xi_i(R(t)) = \frac{N_0}{2}\ln \pi_i + \langle R(t), s_i^*(t)\rangle - \frac{1}{2}E_i^*$$

If we limit our consideration to equally likely binary signal sets, then the minimum P_e matches the received waveform to the filtered versions of the signal waveforms. The resulting error rate is given by

$$P_e(\text{dispersive}) = Q\left(\frac{\|\boldsymbol{s}_0^* - \boldsymbol{s}_0^*\|}{\sqrt{2\boldsymbol{N}_0}}\right)$$

Thus, in this case, the minimum P_e is a function of the separation of the filtered version of the signals in the signal space.

The problem becomes substantially more complex if we cannot insure that the filtered signal durations are less than the symbol lengths. In this case, we experience what is known as **intersymbol interference (ISI)**. That is, observations over one symbol interval contain not only the symbol information of interest but also information from previous symbols. In this case, we must appeal to optimum sequence estimation [5] to take full advantage of the information in the waveform. The basis for this procedure is the maximization of the joint likelihood function conditioned on the sequence of symbols. This procedure not only defines the structure of the optimum receiver under ISI but also is critical in the decoding of convolutional codes and coded modulation. Alternate adaptive techniques to solve this problem involve the use of channel equalization.

Defining Terms

Additive white Gaussian noise (AWGN) channel: The channel whose model is that of corrupting a transmitted waveform by the addition of white (i.e., spectrally flat) Gaussian noise.

Bit (symbol) interval: The period of time over which a single symbol is transmitted.

Communication channel: The medium over which communication signals are transmitted. Examples are fiber optic cables, free space, or telephone lines.

Correlation or matched filter receiver: The optimal receiver structure for digital communications in AWGN.

Decision boundary: The boundary in signal space between the various regions where the receiver declares H_i. Typically a hyperplane when dealing with AWGN channels.

Dispersive channel: A channel that elongates and distorts the transmitted signal. Normally modeled as a time-varying linear system.

Intersymbol interference: The ill-effect of one symbol smearing into adjacent symbols thus interfering with the detection process. This is a consequence of the channel filtering the transmitted signals and therefore elongating their duration. See **dispersive channel**.

Karhunen–Loève expansion: A representation for second-order random processes. Allows one to express a random process in terms of a superposition of deterministic waveforms. The scale values are uncorrelated random variables obtained from the waveform.

Mean-square equivalence: Two random vectors or time-limited waveforms are mean-square equivalent if and only if the expected value of their mean-square error is zero.

Orthonormal: The property of two or more vectors or time-limited waveforms being mutually orthogonal and individually having unit length. Orthogonality and length are typically measured by the standard Euclidean inner product.

Rayleigh channel: A channel that randomly scales the transmitted waveform by a Rayleigh random variable while adding an independent uniform phase to the carrier.

Signal space: An abstraction for representing a time limited waveform in a low-dimensional vector space. Usually arrived at through the application of the Karhunen–Loève transformation.

Total probability of error: The probability of classifying the received waveform into any of the symbols that were not transmitted over a particular bit interval.

References

1. Gibson, J.D., *Principles of Digital and Analog Communications,* 2nd ed., MacMillan, New York, 1993.
2. Haykin, S., *Communication Systems,* 3rd ed., John Wiley & Sons, New York, 1994.
3. Lee, E.A. and Messerschmitt, D.G., *Digital Communication,* Kluwer Academic Publishers, Norwell, MA, 1988.
4. Papoulis, A., *Probability, Random Variables, and Stochastic Processes,* 3rd ed., McGraw-Hill, New York, 1991.
5. Poor, H.V., *An Introduction to Signal Detection and Estimation,* Springer-Verlag, New York, 1988.
6. Proakis, J.G., *Digital Communications,* 2nd ed., McGraw-Hill, New York, 1989.
7. Shiryayev, A.N., *Probability,* Springer-Verlag, New York, 1984.
8. Sklar, B., *Digital Communications, Fundamentals and Applications,* Prentice Hall, Englewood Cliffs, NJ, 1988.
9. Van Trees, H.L., *Detection, Estimation, and Modulation Theory, Part I,* John Wiley & Sons, New York, 1968.
10. Wong, E. and Hajek, B., *Stochastic Processes in Engineering Systems,* Springer-Verlag, New York, 1985.
11. Wozencraft, J.M. and Jacobs, I., *Principles of Communication Engineering,* reissue, Waveland Press, Prospect Heights, IL, 1990.
12. Ziemer, R.E. and Peterson, R.L., *Introduction to Digital Communication,* Macmillan, New York, 1992.

Further Information

The fundamentals of receiver design were put in place by Wozencraft and Jacobs in their seminal book. Since that time, there have been many outstanding textbooks in this area. For a sampling see [1, 2, 3, 8, 12]. For a complete treatment on the use and application of detection theory in communications see [5, 9]. For deeper insights into the Karhunen–Loève expansion and its use in communications and signal processing see [10].

13

Forward Error Correction Coding

Vijay K. Bhargava
University of Victoria

Ivan J. Fair
University of Alberta

13.1 Introduction

In 1948, Claude Shannon issued a challenge to communications engineers by proving that communication systems could be made arbitrarily reliable as long as a fixed percentage of the transmitted signal was redundant [Shannon, 1948]. He showed that limits exist only on the rate of communication and not its accuracy, and he went on to prove that errorless transmission could be achieved in an additive white Gaussian noise (AWGN) environment with infinite bandwidth if the ratio of energy per data bit to noise power spectral density exceeds the **Shannon limit** of -1.6 dB. He did not, however, indicate how this could be achieved. Subsequent research has led to a number of techniques that introduce redundancy to allow for correction of errors without retransmission. These techniques, collectively known as forward error correction (FEC) coding techniques, are used in systems where a reverse channel is not available for requesting retransmission, the delay with retransmission would be excessive, the expected number of errors would require a large number of retransmissions, or retransmission would be awkward to implement [Sklar, 1988].

A simplified model of a digital communication system that incorporates FEC coding is shown in Fig. 13.1. The FEC code acts on a **discrete data channel** comprised of all system elements between the encoder output and decoder input. The encoder maps the source data to q-ary code symbols that are modulated and transmitted. During transmission, this signal can be corrupted, causing errors to arise in the demodulated symbol sequence. The FEC decoder attempts to correct these errors and restore the original source data.

A demodulator that outputs only a value for the q-ary symbol received during each symbol interval is said to make **hard decisions**. In the **binary symmetric channel** (BSC), hard decisions are made on binary symbols and the probability of error is independent of the value of the symbol. One example of

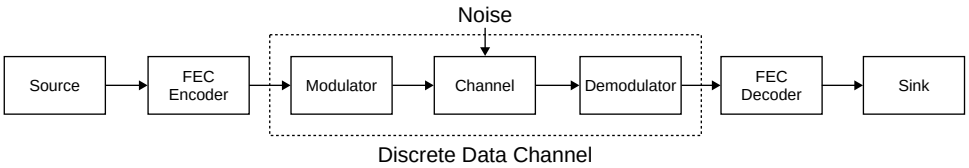


FIGURE 13.1 Block diagram of a digital communication system with forward error correction.

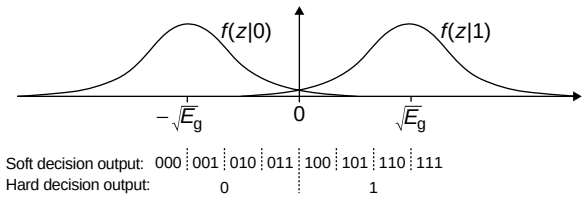


FIGURE 13.2 Hard and soft decision demodulation of a coherently-demodulated BPSK signal corrupted by AWGN. $f(z|1)$ and $f(z|0)$ are the Gaussianly distributed conditional probability density functions at the threshold device.

a BSC is the coherently demodulated BPSK signal corrupted by AWGN. The conditional probability density functions that result with this system are depicted in Fig. 13.2. The probability of error is given by the area under the density functions that lies across the decision threshold and is a function of the symbol energy E_s and the one-sided noise power spectral density N_0.

Alternatively, the demodulator can make **soft decisions**, or output an estimate of the symbol value along with an indication of its confidence in this estimate. For example, if the BPSK demodulator uses three-bit quantization, the two least significant bits can be taken as a confidence measure. Possible soft-decision thresholds for the BPSK signal are depicted in Fig. 13.2. In practice, there is little to be gained by using many soft-decision quantization levels.

Block and convolutional codes introduce redundancy by adding parity symbols to the message data. They map k source symbols to n code symbols and are said to have **code rate** $R = k/n$. With fixed information rates, this redundancy results in increased bandwidth and lower energy per transmitted symbol. At low signal to noise ratios, these codes cannot compensate for these impairments, and performance is degraded. However, at higher ratios of information symbol energy E_b to noise spectral density N_0, there is **coding gain** since the performance improvement offered by coding more than compensates for these impairments. Coding gain is usually defined as the reduction in required E_b/N_0 to achieve a specific error rate in an error-control coded system over one without coding. In contrast to block and convolutional codes, trellis-coded modulation introduces redundancy by expanding the size of the signal set rather than increasing the number of symbols transmitted and so offers the advantages of coding to bandlimited systems.

Each of these coding techniques is considered in turn. Following a discussion of interleaving and concatenated coding, an overview is given of the most recent significant advance in FEC coding, the development of turbo codes. A brief overview of FEC applications concludes this chapter.

13.2 Fundamentals of Block Coding

In block codes there is a one-to-one mapping between k-symbol source words and n-symbol codewords. With q-ary signaling, q^k out of the q^n possible n-tuples are valid code vectors. The set of all n-tuples forms a **vector space** in which the q^k code vectors are distributed. The **Hamming distance** between any two code vectors is the number of symbols in which they differ; the **minimum distance** d_{min} of the code is the smallest Hamming distance between any two codewords.

Contradictory objectives of block codes are:

- to distribute the code vectors in the vector space such that the distance between them is maximized. Then, if the decoder receives a corrupted vector, by evaluating the nearest valid code vector it will decode the correct word with high probability.
- to pack the vector space with as many code vectors as possible to reduce the redundancy in transmission.

When code vectors differ in at least $d_{\min}$ positions, a decoder that evaluates the nearest code vector to each received word is guaranteed to correct up to t random symbol errors per word if

$$d_{\min} \geq 2t + 1 \tag{13.1}$$

Alternatively, all $q^n - q^k$ illegal words can be detected, including all error patterns with $d_{\min} - 1$ or fewer errors. In general, a block code can correct all patterns of t or fewer errors and detect all patterns of u or fewer errors provided that $u \geq t$ and

$$d_{\min} \geq t + u + 1 \tag{13.2}$$

If $q = 2$, knowledge of the positions of the errors is sufficient for their correction; if $q > 2$, the decoder must determine both the positions and values of the errors. If the demodulator indicates positions in which the symbol values are unreliable, the decoder can assume their value unknown and has only to solve for the value of these symbols. These positions are called **erasures**. A block code can correct up to t errors and v erasures in each word if

$$d_{\min} \geq 2t + v + 1 \tag{13.3}$$

13.3 Structure and Decoding of Block Codes

Shannon showed that the performance limit of codes with fixed code rate improves as the block length increases. However, as n and k increase, practical implementation requires that the mapping from message to code vector not be arbitrary but that an underlying structure to the code exist. The structures developed to date limit the error correcting capability of these codes to below what Shannon proved possible, on average, for a code with random codeword assignments. Although turbo codes have made significant strides toward approaching the Shannon limit, the search for good constructive codes continues.

A property which simplifies implementation of the coding operations is that of code linearity. A code is **linear** if the addition of any two code vectors forms another code vector, which implies that the code vectors form a subspace of the vector space of n-tuples. This subspace, which contains the all-zero vector, is spanned by any set of k linearly independent code vectors. Encoding can be described as the multiplication of the information k-tuple by a **generator matrix G**, of dimension $k \times n$, which contains these basis vectors as rows. That is, a message vector $\boldsymbol{m}_i$ is mapped to a code vector $\boldsymbol{c}_i$ according to

$$c_i = \boldsymbol{m}_i\mathbf{G}, \quad i = 0, 1, \ldots, q^k - 1 \tag{13.4}$$

where element-wise arithmetic is defined in the **finite field** GF(q). In general, this encoding procedure results in code vectors with non-systematic form in that the values of the message symbols cannot be determined by inspection of the code vector. However, if **G** has the form $[\mathbf{I}_k, \mathbf{P}]$ where $\mathbf{I}_k$ is the $k \times k$ identity matrix and **P** is a $k \times (n - k)$ matrix of parity checks, then the k most significant symbols of each code vector are identical to the message vector and the code has **systematic** form. This notation assumes

that vectors are written with their most significant or first symbols in time on the left, a convention used throughout this article.

For each generator matrix there is an $(n - k) \times k$ **parity check matrix H** whose rows are orthogonal to the rows in **G**, i.e., $\mathbf{GH}^T = \mathbf{0}$. If the code is systematic, $\mathbf{H} = [\mathbf{P}^T, \mathbf{I}_{n-k}]$. Since all codewords are linear sums of the rows in **G**, it follows that $\boldsymbol{c}_i\mathbf{H}^T = \mathbf{0}$ for all i, $i = 0, 1, \ldots, q^k - 1$, and that the validity of the demodulated vectors can be checked by performing this multiplication. If a codeword $\boldsymbol{c}$ is corrupted during transmission so that the hard-decision demodulator outputs the vector $\hat{\boldsymbol{c}} = \boldsymbol{c} + \boldsymbol{e}$, where $\boldsymbol{e}$ is a non-zero error pattern, the result of this multiplication is an $(n - k)$-tuple that is indicative of the validity of the sequence. This result, called the **syndrome** $\boldsymbol{s}$, is dependent only on the error pattern since

$$\boldsymbol{s} = \hat{\boldsymbol{c}}\mathbf{H}^T = (\boldsymbol{c} + \boldsymbol{e})\mathbf{H}^T = \boldsymbol{c}\mathbf{H}^T + \boldsymbol{e}\mathbf{H}^T = \boldsymbol{e}\mathbf{H}^T \tag{13.5}$$

If the error pattern is a code vector, the errors go undetected. However, for all other error patterns, the syndrome is non-zero. Since there are $q^{n-k} - 1$ non-zero syndromes, $q^{n-k} - 1$ error patterns can be corrected. When these patterns include all those with t or fewer errors and no others, the code is said to be a **perfect code**. Few codes are perfect; most codes are capable of correcting some patterns with more than t errors. **Standard array decoders** use look-up tables to associate each syndrome with an error pattern but become impractical as the block length and number of parity symbols increases. Algebraic decoding algorithms have been developed for codes with stronger structure. These algorithms are simplified with imperfect codes if the patterns corrected are limited to those with t or fewer errors, a simplification called **bounded distance decoding**.

Cyclic codes are a subclass of linear block codes with an algebraic structure that enables encoding to be implemented with a linear feedback shift register and decoding to be implemented without a look-up table. As a result, most block codes in use today are cyclic or are closely related to cyclic codes. These codes are best described if vectors are interpreted as polynomials and arithmetic follows the rules for polynomials where the element-wise operations are defined in GF(q). In a cyclic code, all codeword polynomials are multiples of a **generator polynomial** $g(x)$ of degree $n - k$. This polynomial is chosen to be a divisor of $x^n - 1$ so that a cyclic shift of a code vector yields another code vector, giving this class of codes its name. A message polynomial $m_i(x)$ can be mapped to a codeword polynomial $c_i(x)$ in non-systematic form as

$$c_i(x) = m_i(x)g(x), \quad i = 0, 1, \ldots, q^k - 1 \tag{13.6}$$

In systematic form, codeword polynomials have the form

$$c_i(x) = m_i(x)x^{n-k} - r_i(x), \quad i = 0, 1, \ldots, q^k - 1 \tag{13.7}$$

where $r_i(x)$ is the remainder of $m_i(x)x^{n-k}$ divided by $g(x)$. Polynomial multiplication and division can be easily implemented with shift registers [Blahut, 1983].

The first step in decoding the demodulated word is to determine if the word is a multiple of $g(x)$. This is done by dividing it by $g(x)$ and examining the remainder. Since polynomial division is a linear operation, the resulting syndrome $s(x)$ depends only on the error pattern. If $s(x)$ is the all-zero polynomial, transmission is errorless or an undetectable error pattern has occurred. If $s(x)$ is non-zero, at least one error has occurred. This is the principle of the **cyclic redundancy check** (CRC). It remains to determine the most likely error pattern that could have generated this syndrome.

Single error correcting binary codes can use the syndrome to immediately locate the bit in error. More powerful codes use this information to determine the locations and values of multiple errors. The most prominent approach of doing so is with the iterative technique developed by Berlekamp. This technique, which involves computing an error-locator polynomial and solving for its roots, was subsequently interpreted by Massey in terms of the design of a minimum-length shift register. Once the location and values

of the errors are known, Chien's search algorithm efficiently corrects them. The implementation complexity of these decoders increases only as the square of the number of errors to be corrected [Bhargava, 1983] but does not generalize easily to accommodate soft-decision information. Other decoding techniques, including Chase's algorithm and threshold decoding, are easier to implement with soft-decision input [Clark and Cain, 1981]. Berlekamp's algorithm can be used in conjunction with transform-domain decoding which involves transforming the received block with a finite field Fourier-like transform and solving for errors in the transform domain. Since the implementation complexity of these decoders depends on the block length rather than the number of symbols corrected, this approach results in simpler circuitry for codes with high redundancy [Wu et al., 1987].

Other block codes have also been constructed, including product codes that extend the above ideas to two or more dimensions, codes that are based on transform-domain spectral properties, codes designed specifically for correction of burst errors, and codes that are decodable with straightforward threshold or majority logic decoders [Blahut, 1983; Clark and Cain, 1981; Lin and Costello, 1983].

13.4 Important Classes of Block Codes

When errors occur independently, Bose–Chaudhuri–Hocquenghem (BCH) codes provide among the best performance of known codes for a given block length and code rate. They are cyclic codes with $n = q^m - 1$ where m is any integer greater than two. They are designed to correct up to t errors per word and so have **designed distance** $d = 2t + 1$; the minimum distance may be greater. Generator polynomials for these codes are listed in many texts, including [Clark and Cain, 1981]. These polynomials have degree less than or equal to mt, and so $k \geq n - mt$. BCH codes can be shortened to accommodate system requirements by deleting positions for information symbols.

Some subclasses of these codes are of special interest. Hamming codes are perfect single error correcting binary BCH codes. Full length Hamming codes have $n = 2^m - 1$ and $k = n - m$ for any m greater than two. The duals of these codes are maximal-length codes, with $n = 2^m - 1$, $k = m$, and $d_{\min} = 2^{m-1}$. All $2^m - 1$ non-zero code vectors in these codes are cyclic shifts of a single non-zero code vector. Reed–Solomon (RS) codes are non-binary BCH codes defined over GF(q), where q is often taken as a power of two so that symbols can be represented by a sequence of bits. In these cases, correction of even a single symbol allows for correction of a burst of bit errors. The block length is $n = q - 1$, and the minimum distance $d_{\min} = 2t + 1$ is achieved using only $2t$ parity symbols. Since RS codes meet the Singleton bound of $d_{\min} \leq n - k + 1$, they have the largest possible minimum distance for these values of n and k and are called **maximum distance separable** codes.

The Golay codes are the only non-trivial perfect codes that can correct more than one error. The (11, 6) ternary Golay code has minimum distance five. The (23, 12) binary code is a triple error correcting BCH code with $d_{\min} = 7$. To simplify implementation, it is often extended to a (24, 12) code through addition of an extra parity bit. The extended code has $d_{\min} = 8$.

The (23, 12) Golay code is also a binary quadratic residue code. These cyclic codes have prime length of the form $n = 8m \pm 1$, with $k = (n + 1)/2$ and $d_{\min} \geq \sqrt{n}$. Some of these codes are as good as the best codes known with these values of n and k, but it is unknown if there are good quadratic residue codes with large n [Blahut, 1983].

Reed–Muller codes are equivalent to binary cyclic codes with an additional overall parity bit. For any m, the rth-order Reed–Muller code has $n = 2^m$, $k = \sum_{i=0}^{r} \binom{m}{i}$, and $d_{\min} = 2^{m-r}$. The rth-order and $(m - r - 1)$th-order codes are duals, and the first order codes are similar to maximal-length codes. These codes, and the closely related Euclidean geometry and projective geometry codes, can be decoded with threshold decoding.

The performance of several of these block codes is shown in Fig. 13.3 in terms of decoded bit error probability vs. E_b/N_0 for systems using coherent, hard-decision demodulated BPSK signaling. Many other block codes have also been developed, including Goppa codes, quasi-cyclic codes, burst error correcting Fire codes, and other lesser-known codes.

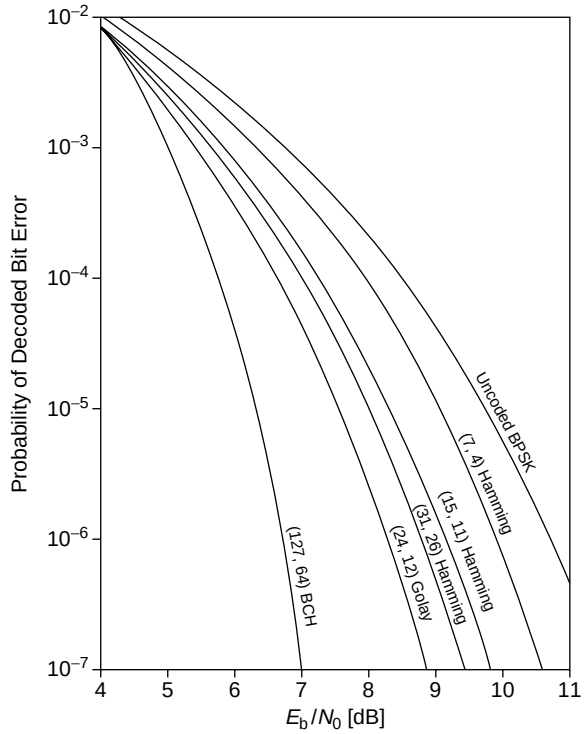


FIGURE 13.3 Block code performance. (From Sklar, B., 1988, *Digital Communications: Fundamentals and Applications,* Prentice-Hall, Englewood Cliffs, NJ. With permission.)

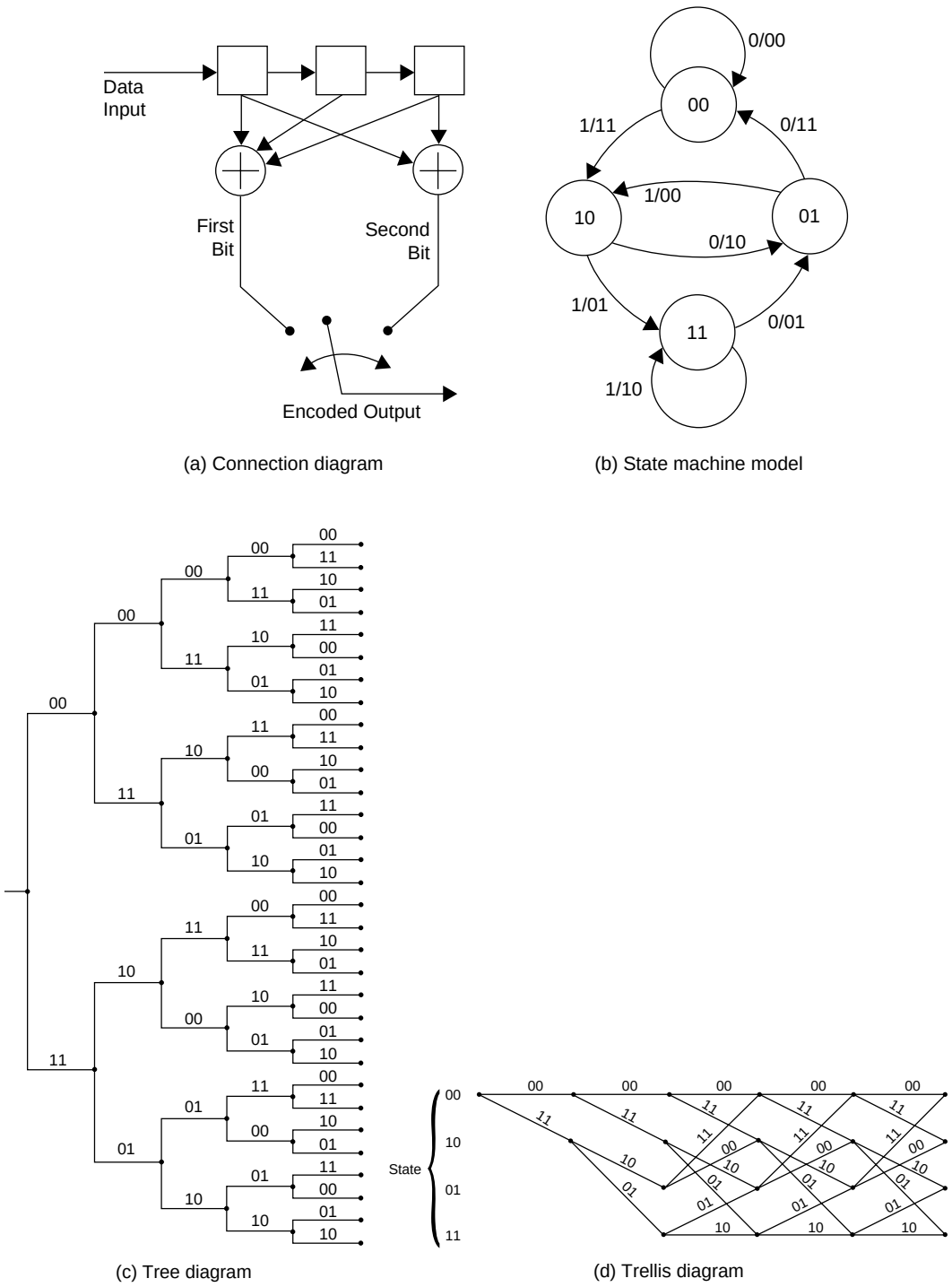
13.5 Principles of Convolutional Coding

Convolutional codes map successive information k-tuples to a series of n-tuples such that the sequence of n-tuples has distance properties that allow for detection and correction of errors. Although these codes can be defined over any alphabet, their implementation has largely been restricted to binary signals, and only binary convolutional codes are considered here.

In addition to the code rate $R = k/n$, the **constraint length** K is an important parameter for these codes. Definitions vary; we will use the definition that K equals the number of k-tuples that affect formation of each n-tuple during encoding. That is, the value of an n-tuple depends on the k-tuple which arrives at the encoder during that encoding interval as well as the $K - 1$ previous information k-tuples.

Binary convolutional encoders can be implemented with $k(K - 1)$-stage shift registers and n modulo-2 adders, an example of which is given in Fig. 13.4(a) for a rate 1/2, constraint length 3 code. The structure depicted in this figure is the encoder for a non-recursive, non-systematic convolutional code because the encoder does not involve feedback and the data sequence is not directly visible in the encoded sequence; recursive and systematic structures are described below. Regardless of their recursive or systematic nature, each convolutional encoder shifts in a new k-tuple during each encoding interval and samples the outputs of the adders sequentially to form the coded output.

Although connection diagrams similar to that of Fig. 13.4(a) completely describe the code, a more concise description can be given by stating the values of n, k, and K and giving the adder connections in the form of vectors or polynomials. For instance, the rate 1/2 code has the generator vectors $\boldsymbol{g}_1 = 111$ and $\boldsymbol{g}_2 = 101$, or equivalently, the generator polynomials $g_1(x) = x^2 + x + 1$ and $g_2(x) = x^2 + 1$. Since feedforward shift registers of this form implement polynomial multiplication [Blahut, 1983], the sequence of bits from each tap can be interpreted as the product of the source sequence and the corresponding



(a) Connection diagram

(b) State machine model

(c) Tree diagram

(d) Trellis diagram

FIGURE 13.4 A rate 1/2, constraint length 3, non-recursive, non-systematic convolutional code.

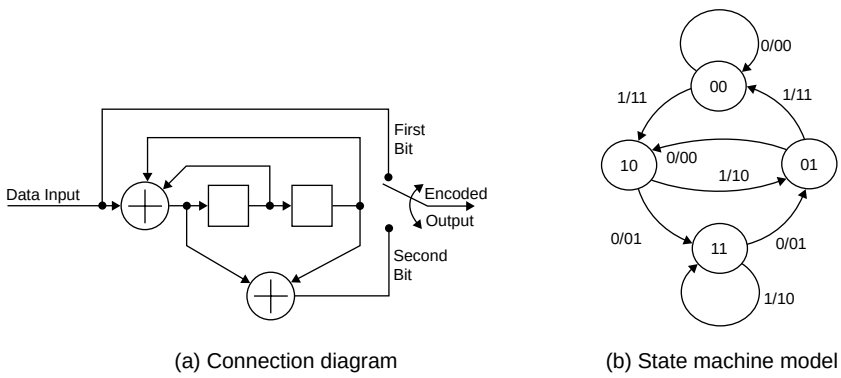
generator polynomial. Alternatively, a convolutional code can be characterized by its impulse response, the coded sequence generated due to input of a single logic-1. It is straightforward to verify that the impulse response of the circuit in Fig. 13.4(a) is 111011 followed by zeros. Since modulo-2 addition is a linear operation, convolutional codes are linear and the coded output can be viewed as the convolution of the input sequence with the impulse response, hence the name of this coding technique. Shifted versions of the impulse response or generator vectors can be combined to form an infinite-order generator matrix that also describes the code.

Shift register circuits can be modeled as finite state machines. A Mealy machine description of a convolutional encoder requires $2^{k(K-1)}$ states, each describing a different value of the $K - 1$ k-tuples in the shift register. Each state has 2^k exit paths that correspond to the value of the incoming k-tuple. A state machine description for the rate 1/2 encoder depicted in Fig. 13.4(a) is given in Fig. 13.4(b). States are labeled with the contents of the shift register; edges are labeled with information bit values followed by their corresponding coded output.

The dimension of time is added to the description of the encoder with tree and trellis diagrams. The tree diagram for the rate 1/2 convolutional code is given in Fig. 13.4(c), assuming the shift register is initially clear. Each node represents an encoding interval, from which the upper branch is taken if the input bit is a 0 and the lower branch is taken if the input bit is a 1. Each branch is labeled with the corresponding output bit sequence. A drawback of the tree representation is that it grows without bound as the length of the input sequence increases. This is overcome with the trellis diagram depicted in Fig. 13.4(d). Again, encoding results in left-to-right movement, where the upper of the two branches is taken whenever the input is a 0, the lower branch is taken when the input is a 1, and the output is the bit sequence that weights the branch taken. Each level of nodes corresponds to a state of the encoder, as shown on the left-hand side of the diagram.

Figure 13.5(a) depicts a recursive systematic convolutional encoder that generates a code that is equivalent to the code depicted in Fig. 13.4. This encoder is systematic because the source bit stream is copied directly into the encoded sequence; it is recursive because it involves feedback. Since a feedback shift register implements polynomial division, the sequence of non-systematic bits can be interpreted as the source sequence divided by the feedback polynomial $g_1(x) = x^2 + x + 1$ and multiplied by the feedforward polynomial $g_2(x) = x^2 + 1$. The codes in Figs. 13.4 and 13.5 are equivalent because their state diagrams, code trees, and trellises are identical with the exception of the mapping from source sequence to encoded sequence. For example, the state diagram in Fig. 13.5(b) for the recursive systematic code exhibits the same structure as the diagram in Fig. 13.4(b); only the input data bits weighting some of the branches differ. Recursive non-systematic codes and non-recursive systematic codes also exist.

If the received sequence contains errors, it may no longer depict a valid path through the tree or trellis. It is the job of the decoder to determine the original path. In doing so, the decoder does not so much correct errors as find the closest valid path to the received sequence. As a result, the error correcting



(a) Connection diagram (b) State machine model

FIGURE 13.5 A rate 1/2, constraint length 3, recursive systematic convolutional code.

capability of a convolutional code is more difficult to quantify than that of a block code; it depends on how valid paths differ. One measure of this difference is the **column distance** $d_c(i)$, the minimum Hamming distance between all coded sequences generated over i encoding intervals that differ in the first interval. The non-decreasing sequence of column distance values is the **distance profile** of the code. The column distance after K intervals is the minimum distance of the code, and is important for evaluating the performance of a code that uses threshold decoding. As i increases, $d_c(i)$ approaches the **free distance** of the code d_{free} which is the minimum Hamming distance in the set of arbitrarily long paths that diverge and then remerge in the trellis.

With maximum likelihood decoding, convolutional codes can generally correct up to t errors within three to five constraint lengths, depending on how the errors are distributed, where

$$d_{\text{free}} \geq 2t + 1 \tag{13.8}$$

The free distance can be calculated by exhaustively searching for the minimum weight path that returns to the all-zero state or by evaluating the term of lowest degree in the generating function of the code.

The objective of a convolutional code is to maximize these distance properties. They generally improve as the constraint length of the code increases, and non-recursive, non-systematic codes generally have better properties than non-recursive systematic ones. As indicated above, non-recursive, non-systematic codes always have an equivalent recursive systematic form that has identical distance properties. Good codes have been found by computer search and are tabulated in many texts, including [Clark and Cain, 1981].

Convolutional codes with high code rate can be constructed by **puncturing** or periodically deleting coded symbols from a low rate code. A list of low rate codes and perforation matrices that result in good high rate codes can be found in many sources, including [Wu et al., 1987]. The performance of good punctured codes approaches that of the best convolutional codes known with similar rate, and decoder implementation is significantly less complex.

Convolutional codes can be **catastrophic**, having the potential to generate an unlimited number of decoded bit errors in response to a finite number of errors in the demodulated bit sequence. Catastrophic error propagation is avoided if the code has generator polynomials with a greatest common divisor of the form x^a for any a, or equivalently, if there are no closed loop paths in the state diagram with all-zero output other than the one taken with all-zero input. Systematic codes are not catastrophic.

13.6 Decoding of Convolutional Codes

Maximum Likelihood Decoding

Maximum likelihood (ML) decoding selects the source sequence $\boldsymbol{m}_i$ that maximizes the probability $P(\boldsymbol{x}|\boldsymbol{m}_i)$, $i \in 0, 1, \ldots, 2^L - 1$, where $\boldsymbol{x}$ is the received sequence and L is the number of symbols in the source sequence. Maximizing this probability is equivalent to selecting the path through the tree or trellis that differs from the received sequence by the smallest amount. In 1967, Viterbi developed a maximum likelihood decoding algorithm that takes advantage of the trellis structure to reduce the complexity of this evaluation. This algorithm has become known as the **Viterbi algorithm**. With each received n-tuple, the decoder computes a **metric** or measure of likelihood for all paths that could have been taken during that interval and discards all but the most likely to terminate on each node. An arbitrary decision is made if path metrics are equal. The metrics can be formed using either hard- or soft-decision information with little difference in implementation complexity.

If the message has finite length and the encoder is subsequently flushed with zeros, a single decoded path remains. With a BSC, this path corresponds to the valid code sequence with minimum Hamming distance from the demodulated sequence. Full-length decoding becomes impractical as the length of the message sequence increases. However, the most likely paths tend to have a common stem, and selecting

the trace value four or five times the constraint length prior to the present decoding depth results in near-optimum performance. Since the number of paths examined during each interval increases exponentially with the constraint length, the Viterbi algorithm becomes impractical for codes with large constraint length.

Maximum *A Posteriori* (MAP) Decoding

In contrast to ML decoding, **maximum *a posteriori* (MAP) decoding** selects the source sequence $\boldsymbol{m}_i$ that maximizes the *a posteriori* probability $P(\boldsymbol{m}_i|\boldsymbol{x})$, $i \in 0, 1, \ldots, 2^L - 1$, where $\boldsymbol{x}$ and L are as defined above. Using Bayes rule, this probability can be written

$$P(\boldsymbol{m}_i|\boldsymbol{x}) = \frac{P(\boldsymbol{x}|\boldsymbol{m}_i)P(\boldsymbol{m}_i)}{P(\boldsymbol{x})}, \quad i \in 0, 1, \ldots, 2^L - 1 \tag{13.9}$$

Since the denominator is constant for all $\boldsymbol{m}_i$ it does not affect selection of the source sequence. The difference between MAP and ML decoding lies in the fact that through the $P(\boldsymbol{m}_i)$ term in the numerator, MAP decoding takes into account the different probabilities of occurrence of different source sequences. When all source sequences are equally likely, ML and MAP decoding techniques are equivalent.

Several MAP decoding algorithms have been developed. One of the most efficient for non-recursive convolutional codes is the BCJR algorithm [Bahl et al., 1974]. This algorithm has been modified to accommodate recursive codes. The BCJR algorithm requires both a forward and reverse recursion through the trellis, evaluating and using metrics in a fashion similar to the Viterbi algorithm. This algorithm is practical only for codes with relatively small constraint lengths, and because of the dual pass nature of this algorithm, its complexity exceeds twice that of the Viterbi algorithm. Further, under most practical conditions, the performance of the Viterbi algorithm is almost identical to the performance of the BCJR algorithm. However, a strength of the BCJR algorithm is its ability to easily provide soft, rather than hard, decision information for each source symbol. This characteristic is necessary for decoding of turbo codes.

Sub-Optimal Decoding Techniques

The inherent complexity of the Viterbi and BCJR algorithms make them unsuitable for decoding convolutional codes with constraint lengths larger than about 12. Other decoding techniques, such as sequential and threshold decoding, can be used with larger constraint lengths. Sequential decoding was proposed by Wozencraft, and the most widely used algorithm was developed by Fano. Rather than tracking multiple paths through the trellis, the sequential decoder operates on a single path while searching the code tree for a path with high probability. It makes tentative decisions regarding the transmitted sequence, computes a metric between its proposed path and the demodulated sequence, and moves forward through the tree as long as the metric indicates that the path is likely. If the likelihood of the path becomes low, the decoder moves backward, searching other paths until it finds one with high probability. The number of computations involved in this procedure is almost independent of the constraint length and is typically quite small, but can be highly variable, depending on the channel. Buffers must be provided to store incoming sequences as the decoder searches the tree. Their overflow is a significant limiting factor in the performance of these decoders.

Figure 13.6 compares the performance of the Viterbi and sequential decoding algorithms for several convolutional codes operating on coherently-demodulated BPSK signals corrupted by AWGN. Other decoding algorithms have also been developed, including syndrome decoding methods such as table look-up feedback decoding and threshold decoding [Clark and Cain, 1981]. These algorithms are easily implemented but suffer from a reduction in performance.

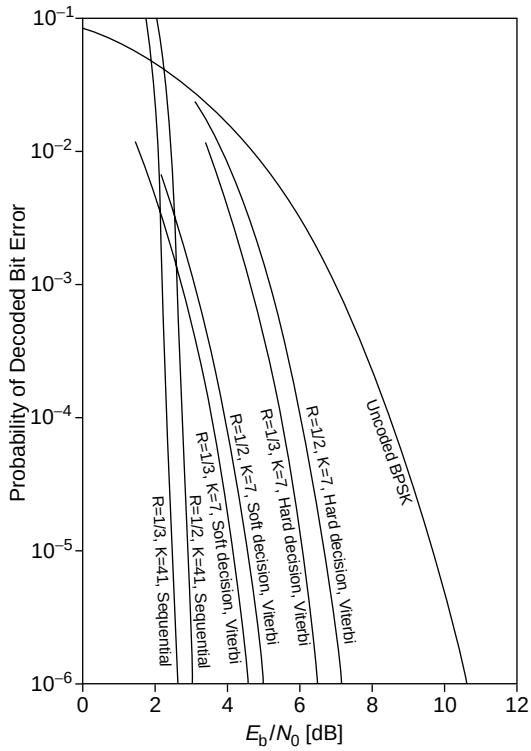


FIGURE 13.6 Convolutional code performance. (From Omura, J.K. and Levitt, B.K., 1982, Coded error probability evaluation for antijam communication systems, *IEEE Trans. Commun.*, COM-30(5), 896–903. With permission.)

13.7 Trellis-Coded Modulation

Trellis-coded modulation (TCM) has received considerable attention since its development by Ungerboeck in the late 1970s [Ungerboeck, 1987]. Unlike block and convolutional codes, TCM schemes achieve coding gain by increasing the size of the signal alphabet and using multi-level/phase signaling. Like convolutional codes, sequences of coded symbols are restricted to certain valid patterns. In TCM, these patterns are chosen to have large Euclidean distance from one another so that a large number of corrupted sequences can be corrected. The Viterbi algorithm is often used to decode these sequences. Since the symbol transmission rate does not increase, coded and uncoded signals require the same transmission bandwidth. If transmission power is held constant, the signal constellation of the coded signal is denser. However, the loss in symbol separation is more than overcome by the error correction capability of the code.

Ungerboeck investigated the increase in channel capacity which can be obtained by increasing the size of the signal set and restricting the pattern of transmitted symbols, and he concluded that almost all of the additional capacity can be gained by doubling the number of points in the signal constellation. This is accomplished by encoding the binary data with a rate $R = k/(k + 1)$ code and mapping sequences of $k + 1$ coded bits to points in a constellation of 2^{k+1} symbols. For example, the rate 2/3 encoder of Fig. 13.7(a) encodes pairs of source bits to three coded bits. Figure 13.7(b) depicts one stage in the trellis of the coded output where, as with the convolutional code, the state of the encoder is defined by the contents of the shift register. Note that unlike the trellis for the convolutional code, this trellis contains parallel paths between nodes.

The key to improving performance with TCM is to map the coded bits to points in the signal space such that the Euclidean distance between transmitted sequences is maximized. A method that ensures

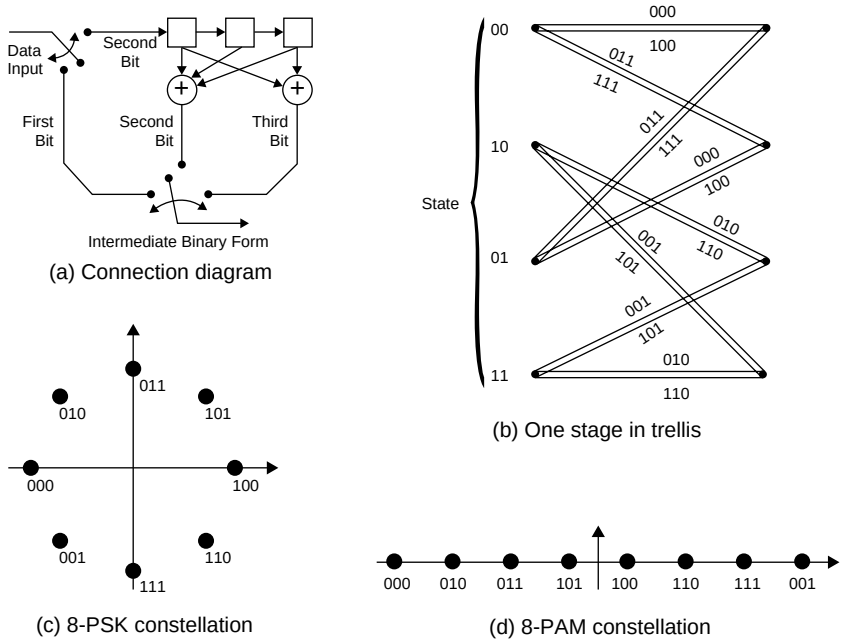


FIGURE 13.7 A rate 2/3 TCM code.

improved Euclidean distance is the method of **set partitioning**. This involves separating all parallel paths on the trellis with maximum distance and assigning the next greatest distance to paths that diverge from or merge onto the same node. Figure 13.7(c) and (d) gives examples of mappings for the rate 2/3 code with 8-PSK and 8-PAM signal constellations, respectively.

As with convolution codes, the free distance of a TCM code is defined as the minimum distance between paths through the trellis, where the distance of concern is now Euclidean distance rather than Hamming distance. The free distance of an uncoded signal is defined as the distance between the closest signal points. When coded and uncoded signals have the same average power, the coding gain of the TCM system is defined as

$$\text{coding gain} = 20\log_{10}\left(\frac{d_{\text{free, coded}}}{d_{\text{free, uncoded}}}\right). \tag{13.10}$$

It can be shown that the simple rate 2/3 8-PSK and 8-PAM TCM systems provide gains of 3 dB and 3.3 dB, respectively [Clark and Cain, 1981]. More complex TCM systems yield gains up to 6 dB. Tables of good codes are given in [Ungerboeck, 1987].

13.8 Additional Measures

When the demodulated sequence contains bursts of errors, the performance of codes designed to correct independent errors improves if coded sequences are **interleaved** prior to transmission and deinterleaved prior to decoding. Deinterleaving separates the burst errors, making them appear more random and increasing the likelihood of accurate decoding. It is generally sufficient to interleave several block lengths of a block coded signal or several constraint lengths of a convolutionally encoded signal. Block interleaving is the most straightforward approach, but delay and memory requirements are halved with convolutional and helical interleaving techniques. Periodicity in the way sequences are combined is avoided with pseudo-random interleaving.

Serially **concatenated codes**, first investigated by Forney, use two levels of coding to achieve a level of performance with less complexity than a single coding stage would require. The inner code interfaces with the modulator and demodulator and corrects the majority of the errors; the outer code corrects errors which appear at the output of the inner-code decoder. A convolutional code with Viterbi decoding is usually chosen as the inner code, and an RS code is often chosen as the outer code due to its ability to correct the bursts of bit errors that can result with incorrect decoding of trellis-coded sequences. Interleaving and deinterleaving outer-code symbols between coding stages offers further protection against the burst error output of the inner code.

Product codes effectively place the data in a two (or more)-dimensional array and use FEC techniques over the rows and columns of this array. Not only do these codes result in error protection in multiple dimensions, but the manner in which the array is constructed can offer advantages similar to those achieved through interleaving.

13.9 Turbo Codes

The most recent significant advancement in FEC coding is the development of **turbo codes** [Berrou, Glavieux, and Thitimajshima, 1993]. The principle of this coding technique is to encode the data with two or more simple **constituent codes** concatenated through an interleaver. The received sequence is decoded in an iterative, serial approach using soft-input, soft-output decoders. The iterative decoder involves feedback of information in a manner similar to processes within the turbo engine, giving this coding technique its name.

Turbo codes effectively result in the construction of relatively long codewords with few codewords being close in terms of Hamming distance, while at the same time constraining the implementation complexity of the decoder to practical limits. The first turbo codes that were developed used recursive systematic convolutional codes as the constituent codes and concatenated them in parallel through a pseudo-random interleaver, as depicted in Fig. 13.8(a). The use of other types of constituent codes, and the serial concatenation of constituent codes, has since been considered.

As in other multi-stage coding techniques, the complexity of the turbo decoder is limited through use of separate decoding stages for each constituent code. The input to the first stage is the soft output of the demodulator for a finite-length received symbol sequence. As shown in Fig. 13.8(b), subsequent stages use both the demodulator output and the **extrinsic information** from the previous decoding stage. Extrinsic information is the new, soft information evaluated in each stage of decoding that can be used as *a priori* information for each source symbol during the next stage of decoding.

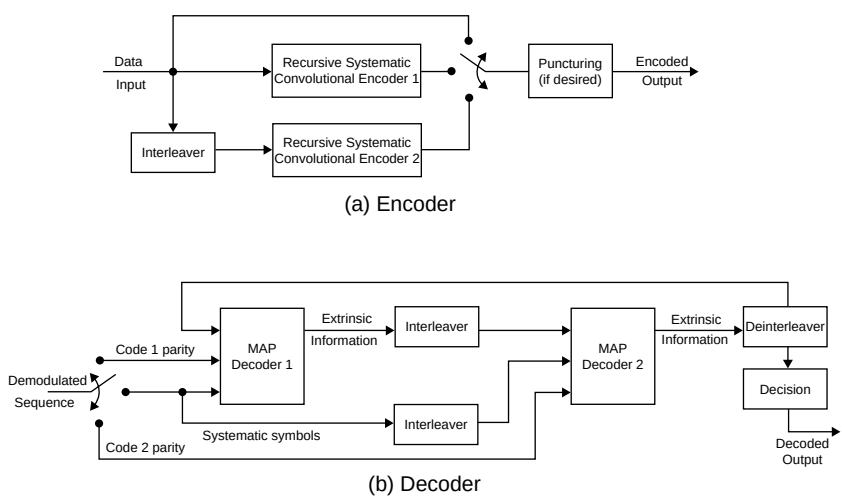


FIGURE 13.8 Turbo code with parallel concatenated constituent codes.

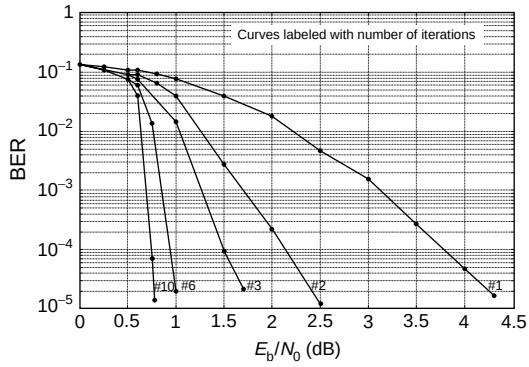


FIGURE 13.9 Performance in AWGN of a turbo code with parallel concatenated constituent codes of rate 1/2 and constraint length 4. Pseudo-random interleaver of size 65536.

Decoding proceeds by iterating through constituent decoders, each forwarding updated extrinsic information to the next decoder until a predefined number of iterations has been completed or the extrinsic information indicates that high reliability has been achieved. Typical simulation results are shown in Fig. 13.9. As indicated in this figure, performance tends to improve with the number of iterations, where the largest performance gains are made during the initial iterations. Excellent performance can be obtained at signal-to-noise ratios appreciably less than 1 dB, however at higher values of E_b/N_0, the performance curves tend to exhibit flattening similar to error floors. Flattening of the BER curves is a result of the relatively small overall Hamming distance typical of most turbo codes.

Since the discovery of turbo codes, a significant amount of effort has been directed towards explaining and analyzing their performance. Recent work demonstrating the relationship between turbo codes and belief propagation in Bayesian networks shows promise in furthering these goals. There has also been significant effort in determining optimal code parameters. With convolutional constituent codes, it has been found that coding gain through turbo decoding is only possible with recursive codes and that there is little advantage to using codes with constraint lengths larger than about five. These short constraint lengths allow the soft-in, soft-out BCJR algorithm to be used. It has also been recognized that decoding performance increases with decreased correlation between the extrinsic information and the demodulated data at the input of the constituent decoders, implying that highly random interleavers should be used. It has also been established that it is important for interleavers to ensure that sequences highly susceptible to error not occur simultaneously in both constituent codes.

Even though great strides have been made over the last few years in understanding the structure of these codes and relating them to serially concatenated and product codes, work regarding their analysis, optimization, and efficient implementation will interest coding theorists and practitioners for many years to come.

13.10 Applications

FEC coding remained of theoretical interest until advances in digital technology and improvements in decoding algorithms made their implementation possible. It has since become an attractive alternative to improving other system components or boosting transmission power. FEC codes are commonly used in digital storage systems, deep-space and satellite communication systems, terrestrial wireless and band-limited wireline systems, and have also been proposed for fiber optic transmission. Accordingly, the theory and practice of error correcting codes now occupies a prominent position in the field of communications engineering.

Deep-space systems began using forward error correction in the early 1970s to reduce transmission power requirements, and used multiple error correcting RS codes for the first time in 1977 to protect against corruption of compressed image data in the Voyager missions [Wicker and Bhargava, 1994]. The Consultative

Committee for Space Data Systems (CCSDS) has recommended use of a concatenated coding system which uses a rate 1/2, constraint length 7 convolutional inner code and a (255, 223) RS outer code.

Coding is now commonly used in satellite systems to reduce power requirements and overall hardware costs and allow closer orbital spacing of geosynchronous satellites [Berlekamp, 1987]. FEC codes play integral roles in the VSAT, MSAT, INTELSAT, and INMARSAT systems [Wu et al., 1987]. Further, a (31, 15) RS code is used in the joint tactical information distribution system (JTIDS), a (7, 2) RS code is used in the air force satellite communication system (AFSATCOM), and a (204, 192) RS code has been designed specifically for satellite TDMA systems. Another code designed for military applications involves concatenation of a Golay and RS code with interleaving to ensure an imbalance of 1s and 0s in the transmitted symbol sequence and enhance signal recovery under severe noise and interference [Berlekamp, 1987].

TCM has become commonplace in transmission of data over voiceband telephone channels. Modems developed since 1984 use trellis coded QAM modulation to provide robust communication at rates above 9.6 kb/s.

Various coding techniques are used in the new digital cellular and personal communication standards [Rappaport, 1996]. Convolutional codes are employed in the second generation GSM, TDM, and CDMA systems. More recent standards use turbo coding to provide error protection for data signals.

FEC codes have also been widely used in digital recording systems, most prominently in the compact disc digital audio system. This system uses two levels of coding and interleaving in the cross-interleaved RS coding (CIRC) system to correct errors which result from disc imperfections and dirt and scratches which accumulate during use. Steps are also taken to mute uncorrectable sequences [Wicker and Bhargava, 1994].

Defining Terms

Binary symmetric channel: A memoryless discrete data channel with binary signaling, hard-decision demodulation, and channel impairments that do not depend on the value of the symbol transmitted.

Bounded distance decoding: Limiting the error patterns which are corrected in an imperfect code to those with t or fewer errors.

Catastrophic code: A convolutional code in which a finite number of code symbol errors can cause an unlimited number of decoded bit errors.

Code rate: The ratio of source word length to codeword length, indicative of the amount of information transmitted per encoded symbol.

Coding gain: The reduction in signal-to-noise ratio required for specified error performance in a block or convolutional coded system over an uncoded system with the same information rate, channel impairments, and modulation and demodulation techniques. In TCM, the ratio of the squared free distance in the coded system to that of the uncoded system.

Column distance: The minimum Hamming distance between convolutionally encoded sequences of a specified length with different leading n-tuples.

Constituent codes: Two or more FEC codes that are combined in concatenated coding techniques.

Cyclic code: A block code in which cyclic shifts of code vectors are also code vectors.

Cyclic redundancy check: When the syndrome of a cyclic block code is used to detect errors.

Designed distance: The guaranteed minimum distance of a BCH code designed to correct up to t errors.

Discrete data channel: The concatenation of all system elements between FEC encoder output and decoder input.

Distance profile: The minimum Hamming distance after each encoding interval of convolutionally encoded sequences which differ in the first interval.

Erasure: A position in the demodulated sequence where the symbol value is unknown.

Extrinsic information: The output of a constituent soft decision decoder that is forwarded as *a priori* information for the source symbols to the next decoding stage in iterative decoding of turbo codes.

Finite field: A finite set of elements and operations of addition and multiplication which satisfy specific properties. Often called Galois fields and denoted GF(q), where q is the number of elements in the field. Finite fields exist for all q which are prime or the power of a prime.

Free distance: The minimum Hamming weight of convolutionally encoded sequences which diverge and remerge in the trellis. Equals the maximum column distance and the limiting value of the distance profile.

Generator matrix: A matrix used to describe a linear code. Code vectors equal the information vectors multiplied by this matrix.

Generator polynomial: The polynomial which is a divisor of all codeword polynomials in a cyclic block code; a polynomial which describes circuit connections in a convolutional encoder.

Hamming distance: The number of symbols in which codewords differ.

Hard decision: Demodulation which outputs only a symbol value for each received symbol.

Interleaving: Shuffling the coded bit sequence prior to modulation and reversing this operation following demodulation. Used to separate and redistribute burst errors over several codewords (block codes) or constraint lengths (trellis codes) for higher probability of correct decoding by codes designed to correct random errors.

Linear code: A code whose code vectors form a vector space. Equivalently, a code where the addition of any two code vectors forms another code vector.

Maximum *a posteriori* decoding: A decoding algorithm that selects the source sequence that maximizes the probability of the source sequence being transmitted, given reception of the received sequence.

Maximum distance separable: A code with the largest possible minimum distance given the block length and code rate. These codes meet the Singleton bound of $d_{\min} \leq n - k + 1$.

Maximum likelihood decoding: A decoding algorithm that selects the source sequence that maximizes the probability of the received sequence occurring given transmission of this source sequence.

Metric: A measure of goodness against which items are judged. In the Viterbi algorithm, an indication of the probability of a path being taken given the demodulated symbol sequence.

Minimum distance: In a block code, the smallest Hamming distance between any two codewords. In a convolutional code, the column distance after K intervals.

Parity check matrix: A matrix whose rows are orthogonal to the rows in the generator matrix of a linear code. Errors can be detected by multiplying the received vector by this matrix.

Perfect code: A t error correcting (n, k) block code in which $q^{n-k} - 1 = \sum_{i=1}^{t} \binom{n}{i}$.

Puncturing: Periodic deletion of code symbols from the sequence generated by a convolutional encoder for purposes of constructing a higher rate code. Also, deletion of parity bits in a block code.

Set partitioning: Rules for mapping coded sequences to points in the signal constellation which always result in a larger Euclidean distance for a TCM system than an uncoded system, given appropriate construction of the trellis.

Shannon limit: The ratio of energy per data bit E_b to noise power spectral density N_0 in an AWGN channel above which errorless transmission is possible when no bandwidth limitations are placed on the signal and transmission is at channel capacity. This limit has the value $\ln 2 = 0.693 = -1.6$ dB.

Soft decision: Demodulation which outputs an estimate of the received symbol value along with an indication of the reliability of this value. Usually implemented by quantizing the received signal to more levels than there are symbol values.

Standard array decoding: Association of an error pattern with each syndrome by way of a look-up table.

Syndrome: An indication of whether or not errors are present in the demodulated symbol sequence.

Systematic code: A code in which the values of the message symbols can be identified by inspection of the code vector.

Vector space: An algebraic structure comprised of a set of elements in which operations of vector addition and scalar multiplication are defined. For our purposes, a set of n-tuples consisting of symbols from GF(q) with addition and multiplication defined in terms of element-wise operations from this finite field.

Viterbi algorithm: A maximum-likelihood decoding algorithm for trellis codes which discards low probability paths at each stage of the trellis, thereby reducing the total number of paths which must be considered.

References

Bahl, L.R., Cocke, J., Jelinek, F., and Raviv, J., Optimal decoding of linear codes for minimizing symbol error rate, *IEEE Trans. Inf. Theory,* 20, 248, 1974.

Berlekamp, E.R., Peile, R.E., and Pope, S.P., The application of error control to communications, *IEEE Commn. Mag.,* 25(4), 44, 1987.

Berrou, C., Glavieux, A., and Thitimajshima, P., Near Shannon limit error-correcting coding and decoding: turbo codes, *Proceedings of ICC'93,* Geneva, Switzerland, 1064, 1993.

Bhargava, V.K., Forward error correction schemes for digital communications, *IEEE Commn. Mag.,* 21(1), 11, 1983.

Blahut, R.E., *Theory and Practice of Error Control Codes,* Addison-Wesley, Massachusetts, 1983.

Clark, G.C., Jr. and Cain, J.B., *Error Correction Coding for Digital Communications,* Plenum Press, New York, 1981.

Lin, S. and Costello, D.J., Jr., *Error Control Coding: Fundamentals and Applications,* Prentice-Hall, Englewood Cliffs, NJ, 1983.

Rappaport, T.S., *Wireless Communications, Principles and Practice.* Prentice-Hall and IEEE Press, Englewood Cliffs, NJ, 1996.

Shannon, C.E., A mathematical theory of communication, *Bell Syst. Tech. J.,* 27(3), 379, 1948.

Sklar, B., *Digital Communications: Fundamentals and Applications,* Prentice Hall, Englewood Cliffs, NJ, 1988.

Ungerboeck, G., Trellis-coded modulation with redundant signal sets, *IEEE Commun. Mag.,* 25(2), 5, 1987.

Wicker, S.B. and Bhargava, V.K., *Reed–Solomon Codes and Their Applications,* IEEE Press, New Jersey, 1994.

Wu, W.W., Haccoun, D., Peile, R., and Hirata, Y., Coding for satellite communication, *IEEE J. Selected Areas Commn.,* SAC-5(4), 724, 1987.

For Further Information

There is now a large amount of literature on the subject of FEC coding. An introduction to the philosophy and limitations of these codes can be found in the second chapter of Lucky's book *Silicon Dreams: Information, Man, and Machine*, St. Martin's Press, New York, 1989. More practical introductions can be found in overview chapters of many communications texts. The number of texts devoted entirely to this subject also continues to grow. Although these texts summarize the algebra underlying block codes, more in-depth treatments can be found in mathematical texts. Survey papers appear occasionally in the literature, but the interested reader is directed to the seminal papers by Shannon, Hamming, Reed and Solomon, Bose and Chaudhuri, Hocquenghem, Wozencraft, Fano, Forney, Berlekamp, Massey, Viterbi, Ungerboeck, and Berrou and Glavieux, among others. The most recent advances in the theory and implementation of error control codes are published in *IEEE Transactions on Information Theory, IEEE Transactions on Communications,* and special issues of *IEEE Journal on Selected Areas in Communications.*

14

Automatic Repeat Request

David Haccoun
École Polytechnique de Montréal

Samuel Pierre
Université du Québec

14.1 Introduction

In most digital communication systems, whenever error events occur in the transmitted messages, some action must be taken to correct these events. This action may take the form of an **error correction** procedure. In some applications where a two-way communication link exists between the sender and the receiver, the receiver may inform the sender that a message has been received in error and, hence, request a repeat of that message. In principle, the procedure may be repeated as many times as necessary until that message is received error free. An error control system in which the erroneously received messages are simply retransmitted is called *automatic-repeat-request* (ARQ).

In ARQ systems, the receiver must perform only an **error detection** procedure on the received messages without attempting to correct the errors. Hence, an error detecting code, in the form of specific redundant or parity-check symbols, must be added to the information-bearing sequence. In general, as the error detecting capability of the code increases, the number of added redundant symbols must also be increased. Clearly, with such a system, an erroneously received message is delivered to the user only if the receiver fails to detect the presence of errors. Since error detection coding is simple, powerful, and quite robust, ARQ systems constitute a simple and efficient method for providing highly reliable transfer of messages from the source to the user over a variety of transmission channels. ARQ systems are therefore widely used in data communication systems that are highly sensitive to errors, such as in computer-to-computer communications.

This chapter presents and discusses principles, performances, limitations, and variants of basic ARQ strategies. The fundamentals of the basic ARQ schemes are presented in Section 14.2, whereas the

performance analysis and the limitations are carried out in Section 14.3. Section 14.4 presents some of the most common variants of the basic ARQ schemes, and hybrid forward error control (FEC)/ARQ techniques are outlined in Section 14.5. Finally, an application problem is provided in Section 14.6.

14.2 Fundamentals and Basic Automatic Repeat Request Schemes

Messages transmitted on a communication channel are subjected to errors, which must be detected and/or corrected at the receiver. This section presents the basic principles underlying the concept of error control using retransmission schemes.

Basic Principles

In the seven-layer open system interconnection (OSI) model, error control refers to some basic functions that may be performed at several layers, especially at the *transport layer* (level 4) and at the *data link layer* (level 2). The transport layer is responsible for the end-to-end transport of information processed as *protocol data unit* (PDU) through a network infrastructure. An intermediate layer, called the *network layer* (level 3), ensures that messages, which are broken down into smaller blocks called *packets*, can successfully be switched from one node to another of the network.

The error control on each individual link of the network is performed by the data link layer, which transforms packets received from the network layer into *frames* by adding address, control, and error check fields prior to their transmission into the channel. Figure 14.1 shows an example of the frame structure, whose length, which depends on the particular data link protocol used, usually varies approximately between 50 and 200 bytes. The length of each field also depends on the particular data link control protocol employed. The data link layer must ensure a correct and orderly delivery of packets between switching nodes in the network. The objective of the error control is to maintain the integrity of the data in the frames as they transit through the links, overcoming the loss, duplication, desequencing, and damage of data and control information bits at these various levels. Most ARQ error control techniques integrate error detection using standard *cyclic redundancy check* (CRC), as well as requests for retransmission using PDU.

Error control coding consists essentially of adding to the *information* bits to be transmitted some extra or redundant bits, called *parity-check* bits, in a regular and controlled manner. These parity-check bits, which are removed at the receiver, do not convey information but allow the receiver to detect and possibly correct errors in the received frames.

The single parity-check bit used in ASCII is an example of a redundant bit for error detection purposes. It is an eighth bit added to the seven information bits in the frame representing the alphanumeric character to be transmitted. This bit takes the value 1 if the number of 1s in the first seven bits is odd (odd parity), and 0 otherwise. At the receiver, the parity-check bit of the received frame is computed again in order to verify whether it agrees with the received parity check bit. A mismatch of these parity-check bits indicates that an odd number of errors has occurred during transmission, leading the receiver to request a retransmission of the erroneous frame.

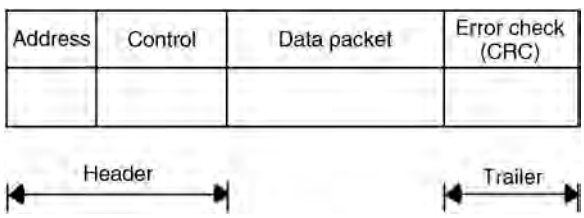


FIGURE 14.1 Structure of a frame.

ARQ refers to retransmission techniques, which basically operate as follows. Erroneously received frames are retransmitted until they are received or detected as being error-free; errors are detected using a simple detection code. Positive (ACK) or negative (NAK) acknowledgments are sent back by the receiver to the sender over a reliable **feedback channel** in order to report whether a previously transmitted frame has been received error-free or with errors. In principle, a positive acknowledgment signals the transmitter to send the next data frame, whereas a negative acknowledgment is a request for frame retransmission. *Stop-and-wait, go-back-N*, and *selective-repeat* are the three basic versions of ARQ, the last two schemes being usually referred to as *sliding window protocols*.

Stop-and-Wait Automatic Repeat Request

Stop-and-wait ARQ is a very simple acknowledgment scheme, which works as follows. Sender and receiver communicate through a **half-duplex** point-to-point **link**. Following transmission of a single frame to the receiver, the sender waits for an acknowledgment from the receiver before sending the next data frame or repeating the same frame. Upon receiving a frame, the receiver sends back to the sender a positive acknowledgment (ACK) if the frame is correct; otherwise it transmits a negative acknowledgment (NAK).

Figure 14.2 illustrates the frame flow for stop-and-wait ARQ. Frames $F1$, $F2$, $F3$,..are in queue for transmission. At time t_0, the sender transmits frame $F1$ to the receiver and stays idle until either an ACK or NAK is received for that frame. At time t_1, $F1$ is correctly received, hence an ACK is sent back. At time t_2, the sender receives this ACK and transmits the next frame $F2$, which is received at time t_3 and detected in error. Frame $F2$ is then discarded at the receiver and a NAK is therefore returned to the sender. At time t_4, the sender receiving a NAK for $F2$ retransmits a copy of that frame and waits again, until it receives at time t_6 either an ACK or NAK for $F2$. Assuming it is an ACK, as illustrated in Fig. 14.2, the sender then proceeds to transmit the next frame $F3$, and so on. Clearly, such a technique requires the sender to store in a buffer a copy of a transmitted frame until a positive acknowledgment has been received for that frame.

Obviously, this technique does not take into account all contingencies. Indeed, it is possible that some transmitted frames damaged by transmission impairments do not even reach the receiver, in which case the receiver cannot even acknowledge them. To overcome this situation, a timer is integrated in the sender for implementing a *time-out* mechanism. A time-out is some pre-established time interval beyond which a transmitted frame is considered lost if its acknowledgment is not received; a timed-out frame is usually simply retransmitted.

Another undesirable situation can be described as follows. The sender transmits a copy of some frame F, which is received correctly by the receiver. A positive acknowledgment (ACK) is then returned to the sender. Unfortunately, this ACK may be corrupted during its transmission in the feedback channel so that it becomes unrecognizable by the sender which therefore transmits another copy of the same frame. As a result, there is a possibility of frame duplication at the receiver leading to an ambiguity between a

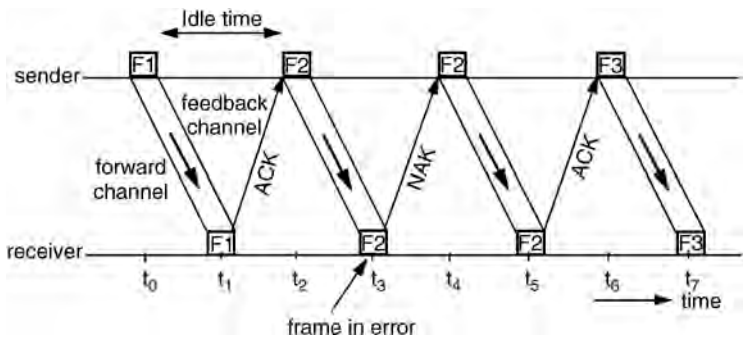


FIGURE 14.2 Stop-and-wait ARQ.

frame and its immediate predecessor or successor. This ambiguity can be easily resolved by attaching a sequence number in each transmitted frame. In practice, a single-bit (0 or 1) inserted in the header of each frame is sufficient to distinguish a frame from its duplicate at the receiver.

Stop-and-wait ARQ schemes offer the advantage of great simplicity. However, they are not very suitable for high-speed modern communication systems since data is transmitted in one direction only. Protocols in which the ACK for a frame must be received by the sender prior to the transmission of the next frame are referred to as *positive acknowledgment with retransmission* (PAR) protocols. The principal weakness of PAR schemes is an inefficient link utilization related to the time lost waiting for acknowledgments before sending another frame, especially if the propagation delay is significantly longer than the packet transmission time, such as in satellite links. To overcome this shortcoming, PAR schemes have been adapted to more practical situations, which require transmitting frames in both directions, using **full-duplex links** between sender and receiver. Clearly then, each direction of transmission acts as the **forward channel** for the frames and as the return channel for the acknowledgments. Now in order to improve further the efficiency of each channel, ACK/NAK control packets and information frames need not be sent separately. Instead, the acknowledgment about each previously received frame is simply appended to the next transmitted frame. In this technique, called *piggybacking*, transmission of the acknowledgment of a received frame is clearly delayed by the transmission of the next data frame. A fixed maximum waiting time for the arrival of a new frame is thus set. If no new frame has arrived by the end of that waiting time, then the acknowledgment frame is sent separately. Even with piggybacking, PAR schemes still require the sender to wait for an acknowledgment before transmitting another frame. An obvious improvement of the efficiency may be achieved by relaxing this condition and allowing the sender to transmit up to W, $W > 1$ frames without waiting for an acknowledgment. A further improvement still is to allow a continuous transmission of the frames without any waiting for acknowledgments. These protocols are called *Continuous ARQ* and are associated with the concept of sliding window protocols.

Sliding Window Protocols

To overcome the inefficiency of PAR schemes, an obvious solution is not to stay idle between frame transmissions. One approach is to use a sliding window protocol where a window refers to a subset of consecutive frames.

Consider a window that contains N frames numbered ω, $\omega + 1, .., \omega + N - 1$, as shown in Fig. 14.3 Every frame using a number smaller than ω has been sent and acknowledged, as identified in Fig. 14.3 as *frames before the window*, whereas no frame with a number larger than or equal to $\omega + N$, identified as *frames beyond the window*, has yet been sent. Frames in the window that have been sent but that have not yet been acknowledged are said to be *outstanding frames*. As outstanding frames are being acknowledged, the window shifts to the right to exclude acknowledged frames and, subsequently, to include the

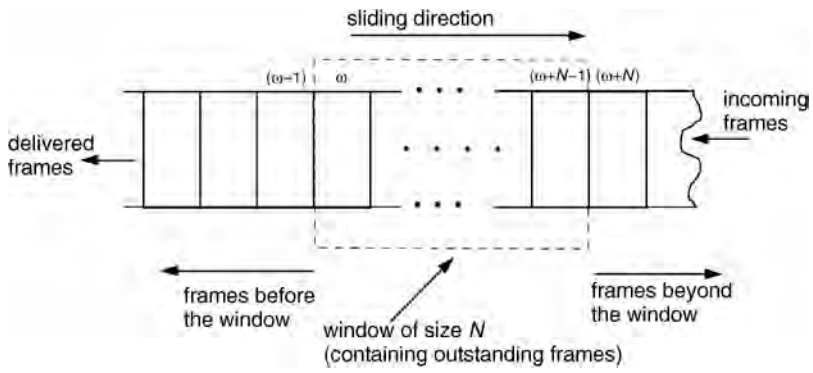


FIGURE 14.3 Frame sequencing in a sliding window protocol (sender side).

new frames that have to be sent. Since the window must always contain frames that are numbered consecutively, the frames are excluded from the window in the same order in which they were included.

To limit the number of outstanding frames, a limit on the window size may be determined. When this limit is reached, the sender accepts no more frames. A window size $N = 1$ corresponds to the stop-and-wait protocol, whereas a window size greater than the total number of frames that can be represented in the header field of the frame refers to an *unrestricted protocol*. Clearly, the window size has some influence on the network traffic and on buffering requirements. Go-back-*N* and selective-repeat may be considered as two common implementations of a sliding window protocol, as will be described next.

Go-Back-*N* Automatic Repeat Request

Go-back-*N* ARQ schemes use continuous transmission without waiting for ACK between frames. Clearly, a full-duplex link between the sender and the receiver is required, allowing a number of consecutive frames to be sent without receiving an acknowledgment. In fact, ACKs may not even be transmitted. Upon detecting a frame in error, the receiver sends a NAK for that frame and discards that frame and all succeeding frames until that erroneous frame has been correctly received. Following reception of a NAK, the sender completes the transmission of the current frame and retransmits the erroneous frame and all subsequent ones.

In accordance with the sliding window protocol, each frame is numbered by an integer $\ell = 0, 1, 2, \ldots, 2^k - 1$, where k denotes the number of bits in the number field of the frame. In general, in order to facilitate their interpretation, frames are all numbered consecutively modulo-2^k. For example, for $k = 3$, frames are numbered 0–7 repeatedly.

It is important to fix the maximum window size, that is, the maximum number of outstanding frames waiting to be transmitted at any one time. To avoid two different outstanding frames having the same number, as well as to prevent pathological protocol failures, the maximum window size is $N = 2^k - 1$. It follows that the number of discarded frames is at most equal to N; however, the actual exact number of discarded frames depends also on the propagation delay. For most usual applications, a window of size $N = 7$ (i.e., $k = 3$) is adequate, whereas for some satellite links the window size N is usually equal to 127.

Figure 14.4 illustrates the frame flow for a go-back-*N* ARQ scheme, with $k = 3$ and, hence, $N = 7$. Thus, the sender can transmit sequences of frames numbered $F0, F1, F2, \ldots, F7$. As shown in Fig. 14.4, since the receiver detects an error on frame $F3$, it returns NAK3 to the sender and discards both $F3$ and the succeeding frames $F4$, $F5$, and $F6$, which because of propagation delays have already been transmitted by the sender before it received the negative acknowledgment NAK3. Then, the sender retransmits $F3$ and the sequence of frames $F4$, $F5$, and $F6$, and proceeds to transmit $F7$ (using the current sequencing)

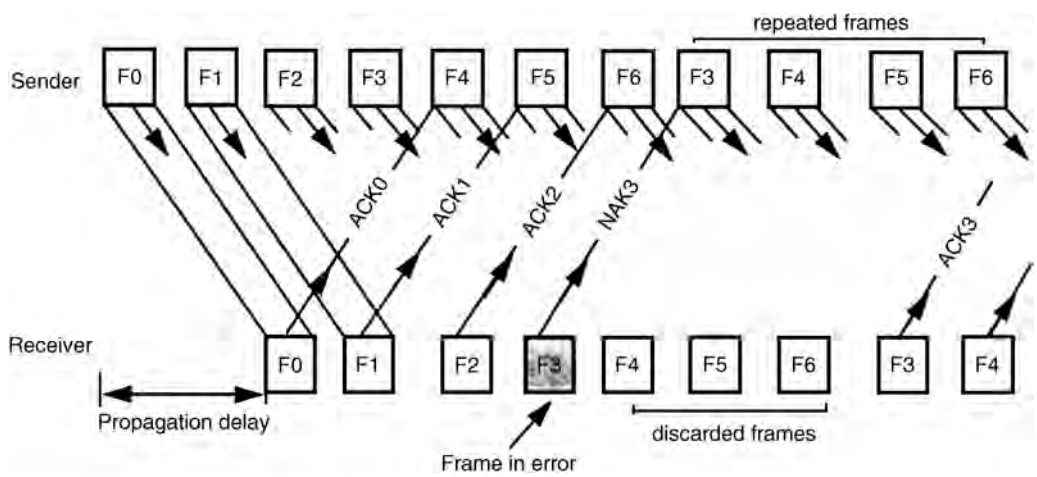


FIGURE 14.4 Go-back-*N* ARQ.

as well as frames $F0$, $F1$,.., using the next sequencing. Obviously, should any of the preceding frames be received in error, then the same repeat procedure is implemented, starting from that frame in error.

Go-back-N procedures have been implemented in various data link protocols such as high-level data link control (HDLC), which refers to layer 2 of the seven-layer OSI model, and transmission control protocol/internet protocol (TCP/IP), which supports the Internet, the world's largest interconnection of computer networks. Indeed, to ensure reliable end-to-end connections and improve the performance, TCP uses a sliding window protocol with timers. In accordance with this protocol, the sender can transmit several TCP messages or frames before an acknowledgment is received for any of the frames. The number of frames allowed in transit is negotiated dynamically using the window field in the TCP header.

As an implementation of the sliding window protocol, the go-back-N scheme is more efficient and more general than the stop-and-wait scheme, which is clearly equivalent to go-back-N, with $N = 1$. However, go-back-N presents the disadvantage of requiring a prerequisite numbering of frames as well as the buffering in the window of N frames awaiting positive acknowledgments. Furthermore, in the absence of a controlled source, all incoming frames not yet accepted by the window must be buffered while they wait to be inserted in the window. Inefficiency is also due to the discarding of all frames that follow the frame received in error, even though these frames may be error free. On the other hand, such a procedure presents the advantage of maintaining the proper sequencing of the frames accepted at the receiver.

Selective-Repeat Automatic Repeat Request

To overcome the inefficiency resulting from unnecessarily retransmitting the error-free frames, retransmission can be restricted to only those frames that have been detected in error. Repeated frames correspond to erroneous or damaged frames that are negatively acknowledged as well as frames for which the time-out has expired. Such a procedure is referred to as selective-repeat ARQ, which obviously should improve the performance over both stop-and-wait and go-back-N schemes. Naturally, a full-duplex link is again assumed.

Figure 14.5 illustrates the frame flow for selective-repeat ARQ. The frames appear to be disjointed but there is no idle time between consecutive frames. The sender transmits a sequence of frames $F1$, $F2$, $F3$,…. As illustrated in Fig. 14.5, the receiver detects an error on the third frame $F3$ and thus returns a negative acknowledgment denoted NAK3 to the sender. However, due to the continuous nature of the protocol, the succeeding frames $F4$, $F5$, and $F6$ have already been sent and are in the pipeline by the time the sender receives NAK3. Upon receiving NAK3, the sender completes the transmission of the current frame ($F6$) and then retransmits $F3$ before sending the next frame $F7$ in the sequence. The correctly

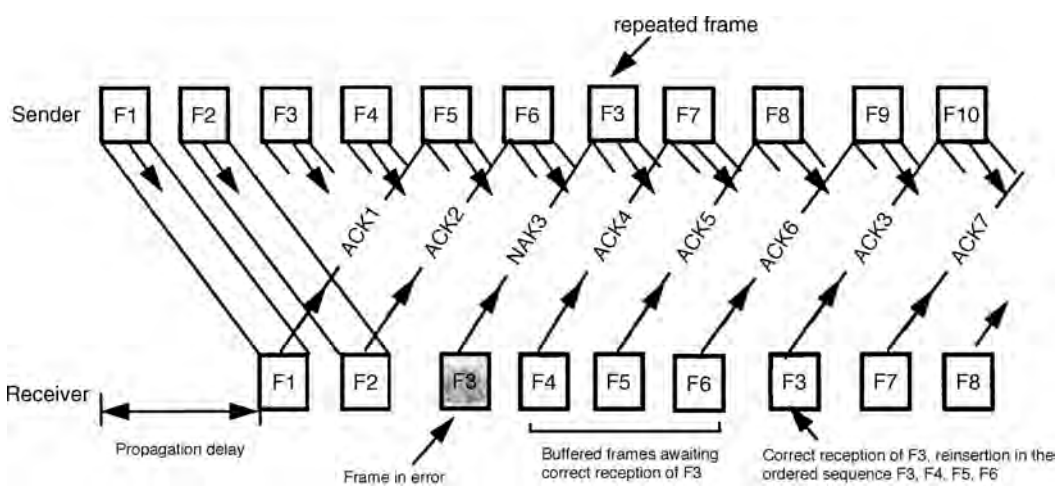


FIGURE 14.5 Selective-repeat ARQ.

received but incomplete sequence $F4$, $F5$, $F6$ must be buffered at the receiver, until $F3$ has been correctly received and inserted in its proper place to complete the sequence $F3$, $F4$, $F5$, $F6$,…, which is then delivered.

If $F3$ is correctly received at the first retransmission, the buffered sequence is $F4$, $F5$, and $F6$. However, if $F3$ is received in error during its retransmission and hence must be retransmitted again, then clearly the buffered sequence is $F4$, $F5$, $F6$, $F7$, $F8$, $F9$. Therefore, multiple retransmissions of a given frame lead to an ever larger buffering, which may be theoretically unbounded. To circumvent such a possibility and the ensuing buffer overflow, a time-out mechanism is usually incorporated in all practical selective-repeat ARQ systems.

In practice, each correctly received frame prior to the appearance of an erroneous frame denoted F is delivered to the user (i.e., the physical layer). Thereafter, all error-free successive frames are buffered at the receiver, until frame F is correctly received. It is then inserted at the beginning of the buffered sequence, which can thus be safely delivered to the user up to the next frame received in error. However, the reordering becomes somewhat more complex if other frames are received in error before frame F is finally received error free. As a consequence, proper buffering must be provided by the sender, which must also integrate the appropriate logic in order to send frames that are out of sequence. The receiver must also be provided with adequate buffering to store the out-of-order frames within the window until the frame in error is correctly received, in addition to being provided with the appropriate logic for inserting the accepted frames in their proper place in the sequence. The required buffers may be rather large for satellite transmission with long propagation delays and long pipelines corresponding to the many frames in transit between sender and receiver. As a result, the selective-repeat ARQ procedure tends to be less commonly used than the go-back-N ARQ scheme, even though it may be more efficient from a throughput point of view.

14.3 Performance Analysis and Limitations

In ARQ systems, the performance is usually measured using two parameters: throughput efficiency and undetected error probability on the data bits. The throughput efficiency is defined as the ratio of the average number of information bits per second delivered to the user to the average number of bits per second that have been transmitted in the system. This throughput efficiency is obviously smaller than 100%. For example, using an error-detecting scheme with a coding rate $R = 0.98$, an error-free transmission would then correspond to a throughput efficiency of 98%, and clearly, any frame retransmission yields a further decrease of the throughput efficiency.

The reliability of an ARQ system is measured by its frame error probability, which may take one of the following forms:

- P = Probability of error detection. It is the probability of an erroneously received frame detected in error by the error detecting code.
- P_u = Probability of undetected error. It is the probability of an erroneously received frame *not* detected in error by the error-detecting code.

Let P_c denote the probability of receiving an error-free frame. We then have $P_c = 1 - P - P_u$. All of these probabilities depend on both the channel error statistics and the error detection code implemented by the CRC. By a proper selection of the CRC, these error probabilities can be made very small. Assuming white noise on a channel having a bit error probability p, the probabilities of correctly receiving a frame of length L is $P_c = (1 - p)^L$. As for the undetectable error probability P_u, it is usually vanishingly small, typically $P_u < 1 \times 10^{-10}$.

Stop-and-Wait Automatic Repeat Request

For error-free operation using the stop-and-wait ARQ technique, the *utilization rate* U_R of a link is defined as

$$U_R = \frac{T_S}{T_T} \tag{14.1}$$

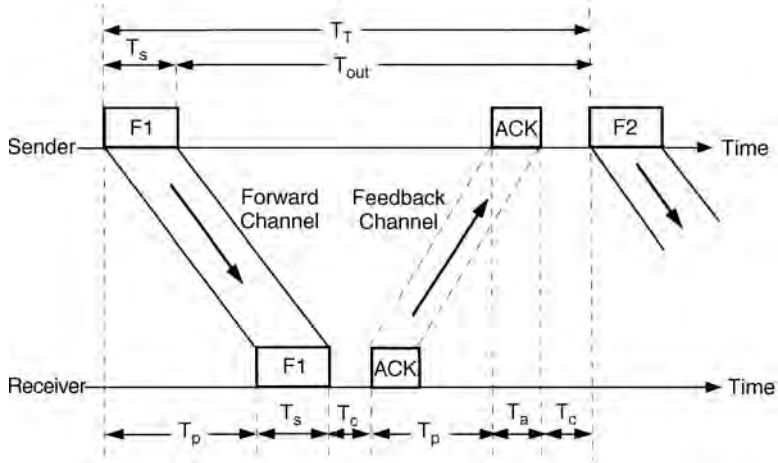


FIGURE 14.6 Timing of successive frames in stop-and-wait ARQ scheme.

where T_s denotes the time required to transmit a single frame, and T_T is the overall time between transmission of two consecutive frames, including processing and ACK/NAK transmissions. As shown in Fig. 14.6, the total time T_T can be expressed as

$$T_T = T_p + T_S + T_c + T_p + T_a + T_c \tag{14.2}$$

In this expression, T_p denotes the propagation delay, that is, the time needed for a transmitted bit to reach the receiver; T_c is the processing delay, that is, the time required for either the sender or the receiver to perform the necessary processing and error checking, whereas T_s and T_a denote the transmission duration of a data frame and of an ACK/NAK frame, respectively.

Assuming the processing time T_c is negligible with respect to T_p and that the sizes of the ACK/NAK frames are very small, leading to negligible value for T_a, then Eq. (14.2) becomes

$$T_T \;{}^{a}\; T_s + 2T_p \tag{14.3}$$

Defining the propagation delay ratio $\alpha = T_p/T_s$, then Eq. (14.1) can be written as

$$U_R = \frac{1}{1+2\alpha} \tag{14.4}$$

Clearly Eq. (14.4) may be seen as a lower bound on the utilization rate or throughput efficiency that the communication link can achieve over all of the frames.

The utilization rate expressed by Eq. (14.4) may be used to evaluate the throughput efficiency on a per frame basis. In the remainder of the chapter, unless specified otherwise, the throughput efficiency is considered on a per frame basis. Because of repetitions caused by transmission errors, using Eq. (14.1), the *average utilization rate*, or *throughput efficiency*, is defined as

$$\eta = \frac{U_R}{N_t} = \frac{T_s}{N_t T_T} \tag{14.5}$$

where N_t is the expected number of transmissions per frame.

In this definition of the throughput efficiency, the coding rate of the error detecting code, as well as the other overhead bits, are not taken into account. That is, definition (14.5) represents the throughput efficiency on a per frame performance basis. Some other definitions of the throughput efficiency may be related to the number of information bits actually delivered to the destination. In such a case, the new

throughput efficiency η^* is simply equal to

$$\eta^* = \eta(1-\rho)$$

where ρ represents the fraction of all redundant and overhead bits in the frame.

Assuming error-free transmissions of ACK and NAK frames, and assuming independence for each frame transmission, the probability of requiring exactly k attempts to successfully receive a given frame is $P_k = P^{k-1}(1-P)$, $k = 1, 2, 3, \ldots$, where P is the frame error probability. The average number of transmissions N_t that are required before a frame is accepted by the receiver is then

$$N_t = \sum_{k=1}^{\infty} kP_k = \sum_{k=1}^{\infty} kP^{k-1}(1-P) = \frac{1}{1-P} \tag{14.6}$$

Using Eqs. (14.4–14.6) the throughput efficiency can thus be written as

$$\eta = \frac{1-P}{1+2\alpha} \tag{14.7}$$

The average overall transmission time T_v, for an accepted frame, is then given by

$$T_v = \sum_{k=1}^{\infty} kP_k T_T = \frac{T_T}{1-P} \tag{14.8}$$

As expected, Eq. (14.8) indicates that if the channel is error-free and thus $P = 0$, then $T_v = T_T$, whereas if the channel is very noisy with $P \to 1$, then the average transmission time T_v may become theoretically unbounded.

Continuous Automatic Repeat Request: Sliding Window Protocols

For continuous ARQ schemes, the basic expression of the throughput efficiency, Eq. (14.5), must be used with the following assumptions: the transmission duration of a data frame is normalized to $T_s = 1$; the transmission time of an ACK/NAK frame, T_a, as well as the processing delay of any frame, T_c, are negligible. Since $\alpha = T_p/T_s$ and $T_s = 1$, then the propagation delay T_p is equal to α.

Let the sender start transmitting the first frame $F1$ at time t_0. Since $T_p = \alpha$, the beginning of that frame will reach the receiver at time $t_0 + \alpha$. Therefore, given that $T_s = 1$, frame $F1$ will be entirely received at time $(t_0 + \alpha + 1)$. The processing delay being negligible, the receiver can immediately acknowledge frame $F1$ by returning ACK1. As a result, since the transmission duration T_a of an ACK/NAK frame is negligible, ACK1 will be delivered to the sender at time $(t_0 + 2\alpha + 1)$. Let N be the window size, then two cases may be considered: $N \geq 2\alpha + 1$ and $N < 2\alpha + 1$.

If $N \geq 2\alpha + 1$, the acknowledgment ACK1 for frame $F1$ will reach the sender before it has exhausted (e.g., emptied) its own window, and hence the sender can transmit continuously without stopping. In this case, the utilization rate of the transmission link is 100%, that is, $U_R = 1$.

If $N < 2\alpha + 1$, the sender will have exhausted its window at time $t_0 + N$, and thus cannot transmit additional frames until time $(t_0 + 2\alpha + 1)$. In this case, the utilization rate is obviously smaller than 100% and can be expressed as $U_R = N/(2\alpha + 1)$. Furthermore, the protocol is no longer a continuous protocol since there is a break in the sender transmission.

Assuming there are no transmission errors, it follows that the utilization rate of the transmission link for a continuous ARQ using a sliding window protocol of size N is given by

$$U_R = \begin{cases} 1, & N \geq 2\alpha + 1 \\ \dfrac{N}{1+2\alpha}, & N < 2\alpha + 1 \end{cases} \tag{14.9}$$

Selective-Repeat Automatic Repeat Request

For selective-repeat ARQ scheme, the expression for the throughput efficiency can be obtained by dividing Eq. (14.9) by $N_t = 1/(1 - P)$ yielding

$$\eta_{SR} = \begin{cases} 1 - P, & N \geq 2\alpha + 1 \\ \dfrac{N(1-P)}{1+2\alpha}, & N < 2\alpha + 1 \end{cases} \tag{14.10}$$

Go-Back-*N* Automatic Repeat Request

For the go-back-*N* ARQ scheme, each frame in error necessitates the retransmission of M frames ($M \leq N$) rather than just one frame, where M depends on the roundtrip propagation delay and the frame size. Let $g(k)$ denote the total number of transmitted frames corresponding to a particular frame being transmitted k times. Since each repetition involves M frames, we can write [Stallings, 1994]

$$g(k) = 1 + (k-1)M = (1-M) + kM$$

Using the same approach as in Eq. (14.6), the average number of transmitted frames N_t to successfully transmit one frame can be expressed as

$$N_t = \sum_{k=1}^{\infty} g(k)P^{k-1}(1-P) = \frac{1-P+PM}{1-P} \tag{14.11}$$

If $N \geq 2\alpha + 1$, the sender transmits continuously and, consequently, M is approximately equal to $2\alpha + 1$. In this case, in accordance with Eq. (14.11), $N_t = (1 + 2\alpha P)/(1 - P)$. Dividing Eq. (14.9) by N_t, the throughput efficiency becomes $\eta = (1 - P)/(1 + 2\alpha P)$.

If $N < 2\alpha + 1$, then $M = N$, and hence $N_t = (1 - P + PN)/(1 - P)$. Dividing Eq. (14.9) again by N_t, we obtain $U = N(1 - P)/(1 + 2\alpha)\ (1 - P + PN)$.

Summarizing, the throughput efficiency of go-back-*N* ARQ is given by

$$\eta_{GB} = \begin{cases} \dfrac{1-P}{1+2\alpha p}, & N \geq 2\alpha + 1 \\ \dfrac{N(1-P)}{(1+2\alpha)(1-P+PN)}, & N < 2\alpha + 1 \end{cases} \tag{14.12}$$

Assuming negligible processing times, the average transmission time of a frame as seen by the sender is given by

$$T_v = T_s + N_t T_T = T_s + \sum_{k=1}^{\infty} kP^k(1-P)T_T \tag{14.13}$$

where T_s is the actual transmission duration of a frame and where again $T_T = T_s + 2T_p$. Since $\alpha = T_p/T_s$, then the total time T_T can be written as: $T_T = T_s + 2\alpha\, T_s = T_s(1 + 2\alpha)$. As a result, using Eq. (14.13) the average transmission time of a frame becomes

$$T_v = T_s + \frac{PT_T}{1-P} = T_s\left(\frac{1+2\alpha P}{1-P}\right) \tag{14.14}$$

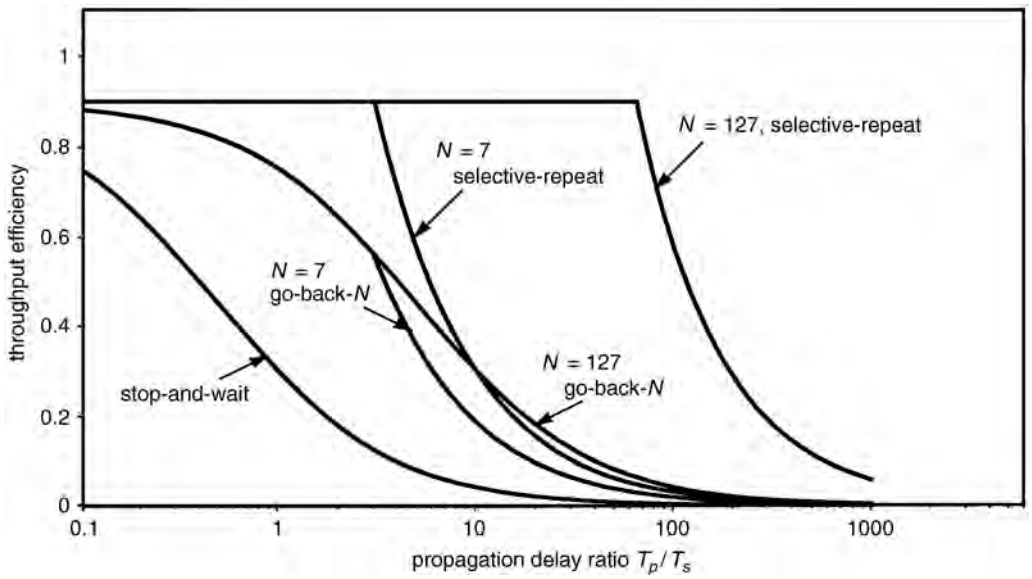


FIGURE 14.7 Throughput efficiency for various ARQ protocols ($P = 0.1$).

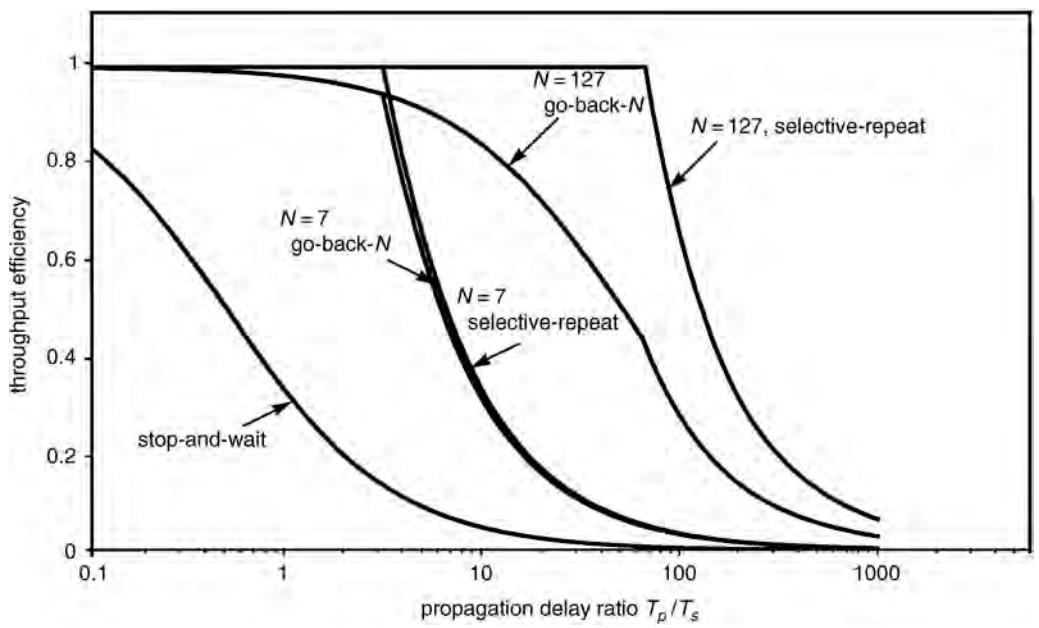


FIGURE 14.8 Throughput efficiency for various ARQ protocols ($P = 0.01$).

Figures 14.7–14.10 show the throughput efficiency η of the three basic ARQ schemes as a function of the propagation delay ratio α, for frame error probabilities $P = 1 \times 10^{-1}$, 1×10^{-2}, 1×10^{-3}, and 1×10^{-4}, respectively.

Basic ARQ schemes present some limitations. The stop-and-wait scheme suffers from inefficiency due to the fact that the channel is idle between the transmission of the frame and the reception of the acknowledgment from the receiver. As shown in Figs. 14.7–14.10, this inefficiency is especially severe for

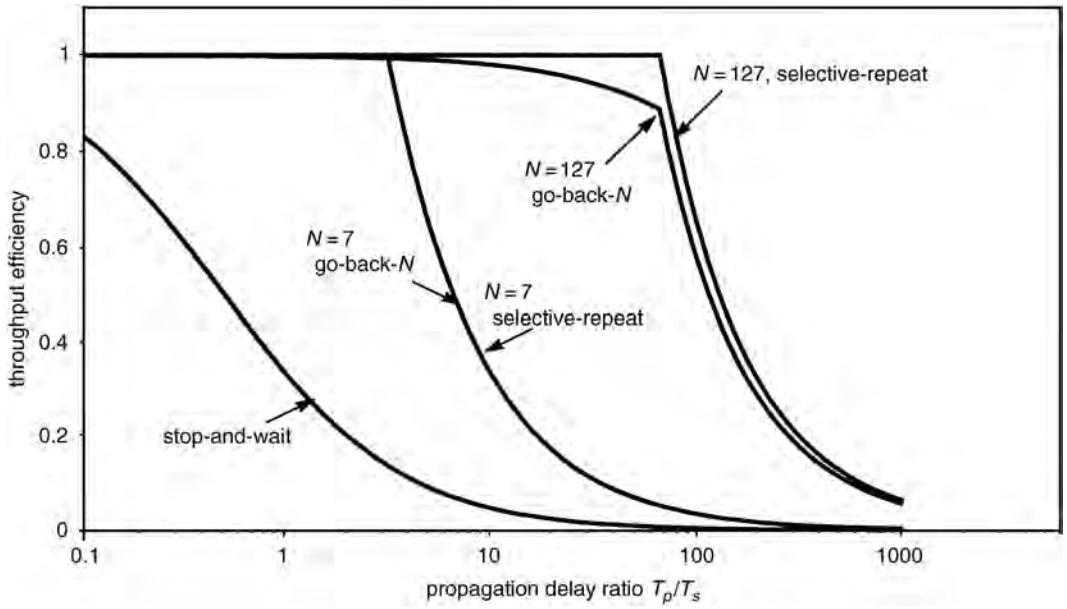


FIGURE 14.9 Throughput efficiency for various ARQ protocols ($P = 0.001$).

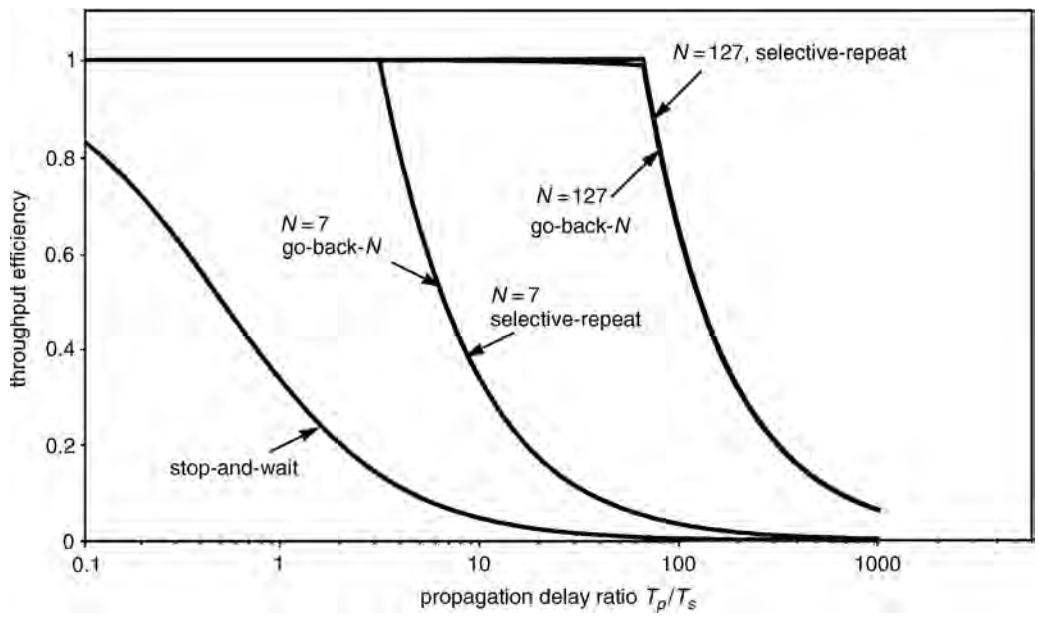


FIGURE 14.10 Throughput efficiency for various ARQ protocols ($P = 0.0001$).

large values of α, that is, when the roundtrip delay between the sender and the receiver is long compared to the transmission duration of a frame. In systems where the channel roundtrip delay is large and/or when the frame error probability P is relatively high, such as in satellite broadcast channels, the throughput efficiency of a stop-and-wait ARQ scheme becomes too poor to be acceptable. On the other hand, the selective-repeat ARQ scheme offers the best performance in terms of throughput efficiency, and, as shown in Figs. 14.7–14.10, it is insensitive to the propagation delay ratio. To overcome some of these limitations, hybrid FEC/ARQ schemes have been proposed.

14.4 Variants of the Basic Automatic Repeat Request Schemes

Several variants of the basic ARQ schemes with the objective of increasing the throughput efficiency and decreasing both buffer size and average transmission time have been proposed in the literature [Sastry, 1995; Despins and Haccoun, 1993]. Some variants have been proposed with the objective of alleviating particular drawbacks of specific ARQ schemes or satisfying special constraints for specific applications. However, these objectives may not be all satisfied, they may even be conflicting.

The main idea of these variants is to send multiple copies of a retransmitted frame rather than just sending a single copy. Variants of both stop-and-wait and continuous ARQ schemes have been proposed by Sastry [1995] for improving the throughput for channels having high error rates and long transmission delays such as satellite links. The modification consists of sending consecutively n copies, $n > 1$, of an erroneous frame before waiting for an ACK. This procedure substantially improves the probability of correctly receiving at least one error-free copy of that frame, which, hence, decreases multiple requests for retransmission of a given frame. For channels with long transmission delays, the increased transmission time for the n copies is more than compensated for by the fewer retransmission requests, leading to improved throughput effficiency over the usual stop-and-wait scheme. For continuous ARQ, whenever a NAK for a frame is received by the sender, that same frame is repeatedly transmitted until reception by the sender of its corresponding ACK. Here again the throughput may be substantially improved, especially for high error rate channels having a transmission delay much larger than the transmission duration of a frame.

A variation of the usual selective-repeat ARQ protocol has been recently proposed and analyzed by Despins and Haccoun [1993] for full-duplex channels suffering from a high error rate in both the forward and reverse directions. It consists essentially of sending a frame composed entirely of NAKs instead of piggybacking a single NAK on a data frame over the feedback channel. The idea is to provide a virtual error-free feedback channel for the NAKs, thus eliminating the need for implementing a time-out and buffer control procedure. For very severely degraded channels, the procedure is further refined by sending a number $m(m > 1)$ of NAK frames for each incorrectly received data frame. The number of NAK frames m is calculated to ensure a practically error-free feedback channel. It is shown that the elimination of needless retransmissions outweighs the additional overhead of the NAK frames, leading to higher throughput efficiency and smaller buffer requirements over conventional selective-repeat ARQ, especially over fading channels.

For point-to-multipoint communication, Wang and Silvester [1993] have proposed adaptive multireceiver ARQ schemes. These variants can be applied to all three basic ARQ schemes. They differ from the classic schemes in the way the sender utilizes the outcome of a previous transmission to alter its next transmission. Basically, upon receiving a NAK for an erroneous frame from some given receivers, instead of repeating the transmission of that frame to all of the receivers in the system, the sender transmits a number of copies of that frame to only those receivers that previously received that frame in error. It has been shown that the optimal number of copies that the sender should transmit depends on the transmission error probability, the roundtrip propagation delay, and the number of receivers that have incorrectly received the data frame.

14.5 Hybrid Forward Error Control/Automatic Repeat Request Schemes

In systems where the channel roundtrip delay is large and/or where there is a large number of receivers, such as in satellite broadcast channels, the throughput efficiencies of the basic ARQ schemes, especially stop-and-wait, become unacceptable. Thus, hybrid ARQ schemes, consisting of an FEC subsystem integrated into an ARQ system, can overcome the drawbacks of ARQ and FEC schemes considered separately. Such combined schemes are commonly named *hybrid FEC/ARQ schemes*.

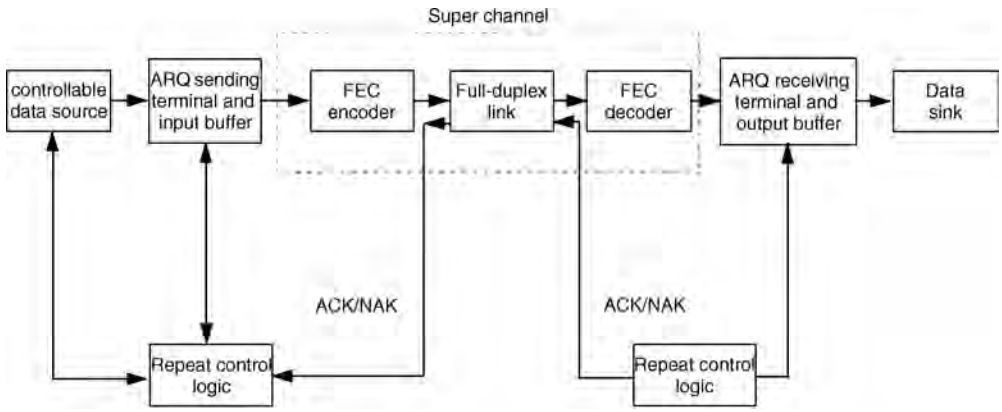


FIGURE 14.11 Principle of hybrid FEC/ARQ schemes.

The principle of hybrid FEC/ARQ schemes is to reduce the number of retransmissions by improving the channel using an FEC procedure. As shown in Fig. 14.11, a hybrid FEC/ARQ system is a concatenated system consisting of an inner FEC system and an outer ARQ system. The function of the inner FEC system is to improve the quality of the channel as seen by the ARQ system by correcting as many frames in error as possible, within its error correcting capability. Since only those erroneous frames that have not been corrected by the FEC will be retransmitted, clearly, under a poor channel condition where frame error rates are high and retransmission requests numerous, hybrid FEC/ARQ schemes can be especially useful for improving the efficiency of conventional ARQ schemes.

Hybrid FEC/ARQ schemes in data point-to-point communications over noisy channels are retransmission techniques that employ both error correction and error detection coding in order to achieve high throughput and low undetected error probabilities. There are two basic types of hybrid FEC/ARQ schemes: *type I hybrid* ARQ scheme, which includes parity bits for both error detection and error correction in each transmitted frame, and *type II hybrid* ARQ scheme, where parity bits for error correction only are sent on the first transmission.

In a type I hybrid FEC/ARQ scheme, when a frame is detected in error, the receiver first attempts to correct these errors. If the error-correcting capability allows the receiver to correct all of the errors detected in the frame, those errors are then corrected and the decoded frame is considered ready to be delivered to the user. In the case where a frame detected in error is revealed uncorrectable, the receiver rejects that frame and requests its retransmission, until that frame is successfully received or decoded. Clearly, a code designed to perform simultaneous error detection and correction is required. As a result, type I hybrid FEC/ARQ schemes require more parity-check bits than a code used only for error detection in a basic ARQ scheme. These extra redundant bits obviously increase the overhead of each transmission.

For low channel error probabilities, type I hybrid FEC/ARQ schemes provide lower throughput efficiencies than any basic ARQ scheme. However, for high channel error probabilities, since the number of retransmissions is reduced by the error-correcting capability of the system, type I hybrid FEC/ARQ schemes provide higher throughput efficiencies than any basic ARQ scheme. These schemes require that a fixed number of parity-check bits are sent with every frame for error correction purposes. Consequently, they are not adaptive to changing channel conditions and are best suited for channels with fairly constant noise or interference levels. For varying channel conditions, other adaptive error control techniques called type II hybrid FEC/ARQ schemes have been proposed [Rice and Wicker, 1994].

Type II hybrid FEC/ARQ schemes basically operate as follows. A frame F to be sent is coded, for its first transmission, with parity-check bits which only allow error detection. If that frame is detected in error in the form F^*, the receiver stores F^* in its buffer and then requests a retransmission. Instead of

retransmitting F in its original form, the sender retransmits a parity-check frame F_1 consisting of the original frame F and an error-correcting code. The receiver, which receives this parity-check frame F_1, uses it to correct the erroneous frame F^* currently stored in its buffer. In the case of error correction failure, a second retransmission of the original frame F is requested by the receiver. Depending on the retransmission strategy and the type of error-correcting codes that are used, the second retransmission may be either a repetition of the original frame F or the transmission of another parity-check frame denoted F_2 [Lin, Costello, and Miller, 1984].

Both type I and type II hybrid FEC/ARQ schemes may use block or convolutional error-correcting codes. It is shown, by means of analysis as well as computer simulations, that both hybrid schemes are capable of providing high throughput efficiencies over a wide range of signal-to-noise ratios. More precisely, performance analysis of type II hybrid FEC/ARQ scheme shows that it is fairly robust against nonstationary channel noise due to the use of parity retransmission. In summary, hybrid FEC/ARQ schemes are suitable for large file transfers in applications where high throughput and high reliability are required over noisy channels, such as in satellite communications.

14.6 Application Problem

In this section, some numerical calculations on performance are presented for illustration purposes.

A satellite channel with a data rate of 2 Mb/s transmits 1000-bit frames. The roundtrip propagation time is 250 ms and the frame error probability is $P = 1 \times 10^{-3}$. Assuming that ACK and NAK frames are error-free, and assuming negligible processing and acknowledgment times:

1. Calculate the throughput efficiency for stop-and-wait ARQ.
2. Determine the average time for correct reception of a frame.
3. Find the average transmission time and the throughput efficiency for error-free operation using go-back-N ARQ, with a window size $N = 127$.
4. Compare the throughput efficiency for go-back-N ARQ and for selective-repeat ARQ, with the same window size $N = 127$.
5. Repeat step 4 using a window size $N = 2047$. Discuss.

Solution

The propagation delay $T_p = 1/2 \times 250 = 125$ ms. For a channel transmitting 1000-bit frames, at 2 Mb/s, the transmission duration of a single frame is

$$T_s = 1000/(2 \times 10^6) = 500 \times 10^{-6}\text{s} = 0.5 \text{ ms}$$

The propagation delay ratio is then

$$\alpha = T_p/T_s = 125/(0.5) = 250$$

and, hence,

$$(2\alpha + 1) = 501$$

1. Given that ACK and NAK frames are error-free, and assuming the processing and acknowledgment times are negligible, the throughput efficiency for stop-and-wait ARQ can be calculated using Eq. (14.7)

$$\eta_{\text{SW}} = (1 - P)/(1 + 2\alpha) = 0.999/(501) = 2 \times 10^{-3}$$

The throughput efficiency of stop-and-wait ARQ schemes over such a channel is vanishingly low as also observed in Fig. 14.9. Clearly, stop-and-wait ARQ on satellite channels is not suitable.

2. Using Eq. (14.8), the average time for correct reception of a frame is $T_v = T_T/(1 - \mathrm{P})$, where T_T can be computed using Eq. (14.3)

$$T_T = T_s + 2T_p = 0.5 + 2 \times 125 = 250.5 \text{ ms}$$

It follows that

$$T_v = 250.5/0.999 = 250.75 \text{ ms}$$

Hence, due to the low value of P, the average time for the correct retransmission of a frame is nearly equal to the roundtrip propagation time.

3. The average transmission time for go-back-N ARQ is given by Eq. (14.14)

$$T_v = 0.5[(1 + 500 \times 10^{-3})/0.999] = 751 \text{ ms}$$

The throughput efficiency for go-back-N ARQ can be computed using Eq. (14.12). Since the window size is $N = 127$ and $(2\alpha + 1) = 501$, it follows that

$$\eta_{\mathrm{GB}} = 127(0.999)/[(501)(0.999 + 0.127)] = 0.225$$

This value for the throughput efficiency may also be read directly from Fig. 14.9. The improvement in throughput efficiency of the go-back-N ARQ over the stop-and-wait scheme is then substantial.

4. The throughput efficiency for selective-repeat ARQ is given by Eq. (14.10). With a window size $N = 127$ and $(2\alpha + 1) = 501$, then

$$\eta_{\mathrm{SR}} = 127(0.999)/501 = 0.253$$

Again, this value can be read directly from Fig. 14.9, which shows that for these values of N and α, the throughput efficiencies of go-back-N and selective-repeat ARQ are approximately equal.

5. Since $N = 2047$ and $(2\alpha + 1) = 501$, using Eqs. (14.12) and (14.10), respectively, it follows that for go-back-N the throughput efficiency is

$$\eta_{\mathrm{GB}} = (0.999)/[1 + (2 \times 250 \times 0.001)] = 0.667$$

For selective-repeat ARQ we obtain

$$\eta_{\mathrm{SR}} = (1 - 0.001) = 0.999$$

Therefore, the selective-repeat ARQ scheme provides nearly the maximum possible throughput efficiency and is substantially more efficient than go-back-N.

14.7 Conclusion

Fundamentals, performances, limitations, and variants of the basic ARQ schemes have been presented. These schemes offer reliability and robustness at buffering cost and reduced throughput. In systems such as packet radio and satellite networks, which are characterized by relatively large frame lengths and high noise or interference levels, the use of error correction coding may not provide the desired reliability. A combination of error correction and error detection coding may be therefore especially

attractive to overcome this shortcoming. Furthermore, over real channel systems characterized by great variations in the quality of the channel, ARQ systems tend to provide the robustness of the required error performance. New trends in digital communications systems using ARQ tend toward increasing the throughput efficiency without increasing buffering. These tendencies are being integrated into many applications involving point-to-multipoint communications over broadcast links used for file distribution, videotex systems, and teleconferencing. Such applications are becoming more and more popular, especially with the deployment of *Integrated Services Digital Networks* (ISDN) linked worldwide by satellite communication systems. Similarly, for wireless communication channels, which are corrupted by noise, fading, interference and shadowing, ARQ schemes can be used in order to provide adequate robustness and performance.

Defining Terms

Error correction: Procedure by which additional redundant symbols are added to the messages in order to allow the correction of errors in those messages.

Error detection: Procedure by which additional redundant symbols are added to the messages in order to allow the detection of errors in those messages.

Feedback channel: Return channel used by the receiver in order to inform the sender that erroneous frames have been received.

Forward channel: One-way channel used by the sender to transmit data frames to the receiver.

Full-duplex (or duplex) link: Link used for exchanging data or ACK/NAK frames between two connected devices in both directions simultaneously.

Half-duplex link: Link used for exchanging data or ACK/NAK frames between two connected devices in both directions but alternatively. Thus, the two devices must be able to switch between send and receive modes after each transmission.

References

Deng, R.H. 1994. Hybrid ARQ schemes employing coded modulation and sequence combining. *IEEE Trans. Commun.*, 42(June):2239–2245.

Despins, C. and Haccoun, D. 1993. A new selective-repeat ARQ protocol and its application to high rate indoor wireless cellular data links. Proceedings of the 4th International Symposium on Personal, Indoor and Mobile Radio Communications, Yokohama, Japan, Sept.

Lin, S., Costello, D.J., and Miller, M.J. 1984. Automatic repeat request error-control schemes. *IEEE Commun. Mag.*, 22(12):5–16.

Rice, M. and Wicker, S.B. 1994. Adaptive error control for slowly varying channels. *IEEE Trans. Commun.*, 42(Feb./March/April):917–926.

Sastry, A.R.K. 1975. Improving ARQ performance on satellite channels under high error rate conditions. *IEEE Trans. Commun.*, COM-23(April):436–439.

Stallings, W. 1994. *Data and Computer Communications*, 4th ed., Macmillan, New York.

Wang, J.L. and Silvester, J.A. 1993. Optimal adaptive multireceiver ARQ protocols. *IEEE Trans. Commun.*, COM-41(Dec.):1816–1829.

Further Information

Halsall, F. 1995. *Data Communications, Computer Networks, and Open Systems*. 4th ed., Chap. 4, pp.168–214. Addison–Wesley, Reading, MA.

Bertsekas, D. and Gallager, R.G. 1992. *Data Networks*, 2nd ed. Chap. 2, pp. 37–127. Prentice–Hall, Englewood Cliffs, NJ.

Tanenbaum, A.S. *Computer Networks*, 2nd ed., Chap. 4, pp. 196–264. Prentice–Hall, Englewood liffs, NJ.

Stallings, W. 1994. *Data and Computer Communications,* 4th ed., Chaps. 4 and 5, pp. 133–197. Macmillan, New York.

15

Spread Spectrum Communications

Laurence B. Milstein
University of California

Marvin K. Simon
Jet Propulsion Laboratory

15.1 A Brief History

Spread spectrum (SS) has its origin in the military arena where the friendly communicator is (1) susceptible to detection/interception by the enemy and (2) vulnerable to intentionally introduced unfriendly interference (jamming). Communication systems that employ spread spectrum to reduce the communicator's detectability and combat the enemy-introduced interference are respectively referred to as **low probability of intercept (LPI)** and **antijam** (AJ) **communication systems**. With the change in the current world political situation wherein the U.S. Department of Defense (DOD) has reduced its emphasis on the development and acquisition of new communication systems for the original purposes, a host of new commercial applications for SS has evolved, particularly in the area of cellular mobile communications. This shift from military to commercial applications of SS has demonstrated that the basic concepts that make SS techniques so useful in the military can also be put to practical peacetime use. In the next section, we give a simple description of these basic concepts using the original military application as the basis of explanation. The extension of these concepts to the mentioned commercial applications will be treated later on in the chapter.

15.2 Why Spread Spectrum?

Spread spectrum is a communication technique wherein the transmitted modulation is *spread* (increased) in bandwidth prior to transmission over the channel and then *despread* (decreased) in bandwidth by the same amount at the receiver. If it were not for the fact that the communication channel introduces some form of narrowband (relative to the spread bandwidth) interference, the receiver performance would be transparent to the spreading and despreading operations (assuming that they are identical inverses of each other). That is, after **despreading** the received signal would be identical to the transmitted signal prior to **spreading**. In the presence of narrowband interference, however, there is a significant advantage to employing the spreading/despreading procedure described. The reason for this is as follows.

Since the interference is introduced after the transmitted signal is spread, then, whereas the despreading operation at the receiver shrinks the desired signal back to its original bandwidth, at the same time it spreads the undesired signal (interference) in bandwidth by the same amount, thus reducing its power spectral density. This, in turn, serves to diminish the effect of the interference on the receiver performance, which depends on the amount of interference power in the despread bandwidth. It is indeed this very simple explanation which is at the heart of all spread spectrum techniques.

15.3 Basic Concepts and Terminology

To describe this process analytically and at the same time introduce some terminology that is common in spread spectrum parlance, we proceed as follows. Consider a communicator that desires to send a message using a transmitted power S Watts (W) at an information rate R_b bits/s (bps). By introducing an SS modulation, the bandwidth of the transmitted signal is increased from R_b Hz to W_{SS} Hz, where $W_{SS} \gg R_b$ denotes the **spread spectrum bandwidth**. Assume that the channel introduces, in addition to the usual thermal noise (assumed to have a single-sided power spectral density (PSD) equal to N_0 W/Hz), an additive interference (jamming) having power J distributed over some bandwidth W_J. After despreading, the desired signal bandwidth is once again now equal to R_b Hz and the interference PSD is now $N_J = J/W_{SS}$. Note that since the thermal noise is assumed to be white, i.e., it is uniformly distributed over all frequencies, its PSD is unchanged by the despreading operation and, thus, remains equal to N_0. Regardless of the signal and interferer waveforms, the equivalent bit energy-to-total noise spectral density ratio is, in terms of the given parameters,

$$\frac{E_b}{N_t} = \frac{E_b}{N_0 + N_J} = \frac{S/R_b}{N_0 + J/W_{SS}} \tag{15.1}$$

For most practical scenarios, the jammer limits performance, and, thus, the effects of receiver noise in the channel can be ignored. Thus, assuming $N_J \gg N_0$, we can rewrite Eq. (15.1) as

$$\frac{E_b}{N_t} \cong \frac{E_b}{N_J} = \frac{S/R_b}{J/W_{SS}} = \frac{S}{J}\frac{W_{SS}}{R_b} \tag{15.2}$$

where the ratio J/S is the *jammer-to-signal power ratio* and the ratio W_{SS}/R_b is the **spreading ratio** and is defined as the **processing gain** of the system. Since the ultimate error probability performance of the communication receiver depends on the ratio E_b/N_J, we see that from the communicator's viewpoint his goal should be to minimize J/S (by choice of S) and maximize the processing gain (by choice of W_{SS} for a given desired information rate). The possible strategies for the jammer will be discussed in the section on military applications dealing with AJ communications.

15.4 Spread Spectrum Techniques

By far the two most popular spreading techniques are **direct sequence (DS) modulation** and **frequency hopping (FH) modulation**. In the following subsections, we present a brief description of each.

Direct Sequence Modulation

A direct sequence modulation $c(t)$ is formed by linearly modulating the output sequence $\{c_n\}$ of a pseudorandom number generator onto a train of pulses, each having a duration T_c called the **chip time**. In mathematical form,

$$c(t) = \sum_{n=-\infty} c_n p(t - nT_c) \tag{15.3}$$

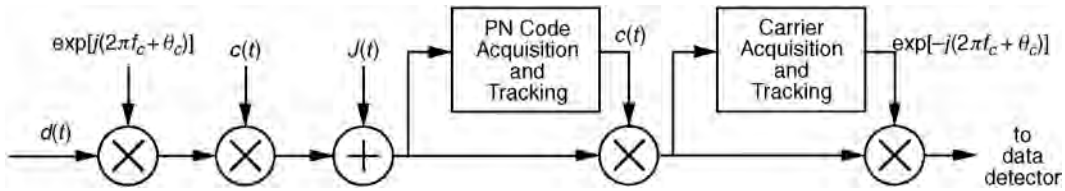


FIGURE 15.1 A DS-BPSK system (complex form).

where $p(t)$ is the basic pulse shape and is assumed to be of rectangular form. This type of modulation is usually used with binary phase-shift-keyed (BPSK) information signals, which have the complex form $d(t)\exp\{j(2\pi f_c t + \theta_c)\}$, where $d(t)$ is a binary-valued data waveform of rate $1/T_b$ bits/s and f_c and θ_c are the frequency and phase of the data-modulated carrier, respectively. As such, a DS/BPSK signal is formed by multiplying the BPSK signal by $c(t)$ (see Fig. 15.1), resulting in the real transmitted signal

$$x(t) = \mathrm{Re}\{c(t)d(t)\exp[j(2\pi f_c t + \theta_c)]\} \tag{15.4}$$

Since T_c is chosen so that $T_b \gg T_c$, then relative to the bandwidth of the BPSK information signal, the bandwidth of the DS/BPSK signal[1] is effectively increased by the ratio $T_b/T_c = W_{SS}/2R_b$, which is one-half the spreading factor or processing gain of the system. At the receiver, the sum of the transmitted DS/BPSK signal and the channel interference $I(t)$ (as discussed before, we ignore the presence of the additive thermal noise) are ideally multiplied by the identical DS modulation (this operation is known as despreading), which returns the DS/BPSK signal to its original BPSK form whereas the real interference signal is now the real wideband signal $\mathrm{Re}\{I(t)c(t)\}$. In the previous sentence, we used the word ideally, which implies that the PN waveform used for despreading at the receiver is identical to that used for spreading at the transmitter. This simple implication covers up a multitude of tasks that a practical DS receiver must perform. In particular, the receiver must first acquire the PN waveform. That is, the local PN random generator that generates the PN waveform at the receiver used for despreading must be aligned (synchronized) to within one chip of the PN waveform of the received DS/BPSK signal. This is accomplished by employing some sort of **search algorithm** which typically steps the local PN waveform sequentially in time by a fraction of a chip (e.g., half a chip) and at each position searches for a high degree of correlation between the received and local PN reference waveforms. The search terminates when the correlation exceeds a given threshold, which is an indication that the alignment has been achieved. After bringing the two PN waveforms into **coarse alignment**, a **tracking algorithm** is employed to maintain **fine alignment**. The most popular forms of tracking loops are the continuous time **delay-locked loop** and its time-multiplexed version the **tau–dither loop**. It is the difficulty in synchronizing the receiver PN generator to subnanosecond accuracy that limits PN chip rates to values on the order of hundreds of Mchips/s, which implies the same limitation on the DS spread spectrum bandwidth W_{SS}.

Frequency Hopping Modulation

A **frequency hopping (FH) modulation** $c(t)$ is formed by nonlinearly modulating a train of pulses with a sequence of pseudorandomly generated frequency shifts $\{f_n\}$. In mathematical terms, $c(t)$ has the complex form

$$c(t) = \sum_{n=-\infty}^{\infty} \exp\{j(2\pi f_n + \phi_n)\}p(t - nT_h) \tag{15.5}$$

[1] For the usual case of a rectangular spreading pulse $p(t)$, the PSD of the DS/BPSK modulation will have $(\sin x/x)^2$ form with first zero crossing at $1/T_c$, which is nominally taken as one-half the spread spectrum bandwidth W_{SS}.

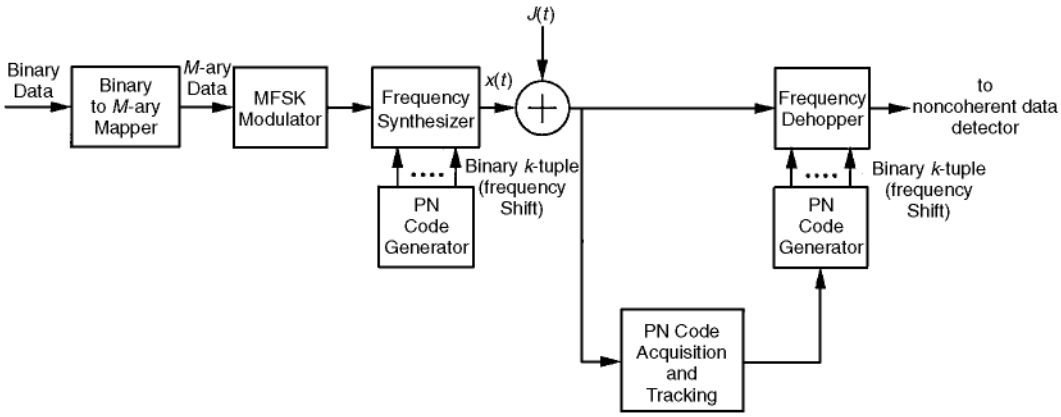


FIGURE 15.2 An FH-MFSK system.

where $p(t)$ is again the basic pulse shape having a duration T_h, called the **hop time**, and $\{\phi_n\}$ is a sequence of random phases associated with the generation of the hops. FH modulation is traditionally used with multiple-frequency-shift-keyed (MFSK) information signals, which have the complex form $\exp\{j[2\pi(f_c + d(t))t]\}$, where $d(t)$ is an M-level digital waveform (M denotes the symbol alphabet size) representing the information frequency modulation at a rate $1/T_s$ symbols/s (sps). As such, an FH/MFSK signal is formed by complex multiplying the MFSK signal by $c(t)$ resulting in the real transmitted signal

$$x(t) = \mathrm{Re}\{c(t)\exp\{j[2\pi(f_c + d(t))t]\}\} \tag{15.6}$$

In reality, $c(t)$ is never generated in the transmitter. Rather, $x(t)$ is obtained by applying the sequence of pseudorandom frequency shifts $\{f_n\}$ directly to the frequency synthesizer that generates the carrier frequency f_c (see Fig. 15.2). In terms of the actual implementation, successive (not necessarily disjoint) k-chip segments of a PN sequence drive a frequency synthesizer, which hops the carrier over 2^k frequencies. In view of the large bandwidths over which the frequency synthesizer must operate, it is difficult to maintain phase coherence from hop to hop, which explains the inclusion of the sequence $\{\phi_n\}$ in the Eq. (15.5) model for $c(t)$. On a short term basis, e.g., within a given hop, the signal bandwidth is identical to that of the MFSK information modulation, which is typically much smaller than W_{SS}. On the other hand, when averaged over many hops, the signal bandwidth is equal to W_{SS}, which can be on the order of several GHz, i.e., an order of magnitude larger than that of implementable DS bandwidths. The exact relation between W_{SS}, T_h, T_s and the number of frequency shifts in the set $\{f_n\}$ will be discussed shortly.

At the receiver, the sum of the transmitted FH/MFSK signal and the channel interference $I(t)$ is ideally complex multiplied by the identical FH modulation (this operation is known as **dehopping**), which returns the FH/MFSK signal to its original MFSK form, whereas the real interference signal is now the wideband (in the average sense) signal $\mathrm{Re}\{I(t)c(t)\}$. Analogous to the DS case, the receiver must acquire and track the FH signal so that the dehopping waveform is as close to the hopping waveform $c(t)$ as possible.

FH systems are traditionally classified in accordance with the relationship between T_h and T_s. **Fast frequency-hopped (FFH)** systems are ones in which there exist one or more hops per data symbol, that is, $T_s = NT_h$ (N an integer), whereas **slow frequency-hopped (SFH)** systems are ones in which there exist more than one symbol per hop, that is, $T_h = NT_s$. It is customary in SS parlance to refer to the FH/MFSK tone of shortest duration as a "chip," despite the same usage for the PN chips associated with the code generator that drives the frequency synthesizer. Keeping this distinction in mind, in an FFH system where, as already stated, there are multiple hops per data symbol, a chip is equal to a hop. For SFH, where there are multiple data symbols per hop, a chip is equal to an MFSK symbol. Combining these two statements,

the chip rate R_c in an FH system is given by the larger of $R_h = 1/T_h$ and $R_s = 1/T_s$ and, as such, is the highest system clock rate.

The frequency spacing between the FH/MFSK tones is governed by the chip rate R_c and is, thus, dependent on whether the FH modulation is FFH or SFH. In particular, for SFH where $R_c = R_s$, the spacing between FH/MFSK tones is equal to the spacing between the MFSK tones themselves. For noncoherent detection (the most commonly encountered in FH/MFSK systems), the separation of the MFSK symbols necessary to provide orthogonality[2] is an integer multiple of R_s. Assuming the minimum spacing, i.e., R_s, the entire spread spectrum band is then partitioned into a total of $N_t = W_{SS}/R_s = W_{SS}/R_c$ equally spaced FH tones. One arrangement, which is by far the most common, is to group these N_t tones into $N_b = N_t/M$ contiguous, nonoverlapping bands, each with bandwidth $MR_s = MR_c$; see Fig. 15.3a. Assuming symmetric MFSK modulation around the carrier frequency, then the center frequencies of the $N_b = 2^k$ bands represent the set of hop carriers, each of which is assigned to a given k-tuple of the PN code generator. In this fixed arrangement, each of the N_t FH/MFSK tones corresponds to the combination of a unique hop carrier (PN code k-tuple) and a unique MFSK symbol. Another arrangement, which provides more protection against the sophisticated interferer (jammer), is to overlap adjacent M-ary bands by an amount equal to R_c; see Fig. 15.3b. Assuming again that the center frequency of each band corresponds to a possible hop carrier, then since all but $M - 1$ of the N_t tones are available as center frequencies, the number of hop carriers has been increased from N_t/M to $N_t - (M - 1)$, which for $N_t \gg M$ is approximately an increase in randomness by a factor of M.

For FFH, where $R_c = R_h$, the spacing between FH/MFSK tones is equal to the hop rate. Thus, the entire spread spectrum band is partitioned into a total of $N_t = W_{SS}/R_h = W_{SS}/R_c$ equally spaced FH tones, each of which is assigned to a unique k-tuple of the PN code generator that drives the frequency synthesizer. Since for FFH there are R_h/R_s hops per symbol, then the metric used to make a noncoherent decision on a particular symbol is obtained by summing up R_h/R_s detected chip (hop) energies, resulting in a so-called *noncoherent combining loss*.

Time Hopping Modulation

Time hopping (TH) is to spread spectrum modulation what pulse position modulation (PPM) is to information modulation. In particular, consider segmenting time into intervals of T_f seconds and further segment each T_f interval into M_T increments of width T_f/M_T. Assuming a pulse of maximum duration equal to T_f/M_T, then a **time hopping spread spectrum** modulation would take the form

$$c(t) = \sum_{n=-\infty}^{\infty} p\left[t - \left(n + \frac{a_n}{M_T}\right) T_f\right] \tag{15.7}$$

where a_n denotes the pseudorandom position (one of M_T uniformly spaced locations) of the pulse within the T_f-second interval.

For DS and FH, we saw that *multiplicative* modulation, that is, the transmitted signal is the product of the SS and information signals, was the natural choice. For TH, *delay* modulation is the natural choice. In particular, a TH-SS modulation takes the form

$$x(t) = \mathrm{Re}\{c(t - d(t)) \exp[j(2\pi f_c + \phi_T)]\} \tag{15.8}$$

[2]An optimum noncoherent MFSK detector consists of a bank of energy detectors each matched to one of the M frequencies in the MFSK set. In terms of this structure, the notion of *orthogonality* implies that for a given transmitted frequency there will be no crosstalk (energy spillover) in any of the other $M - 1$ energy detectors.

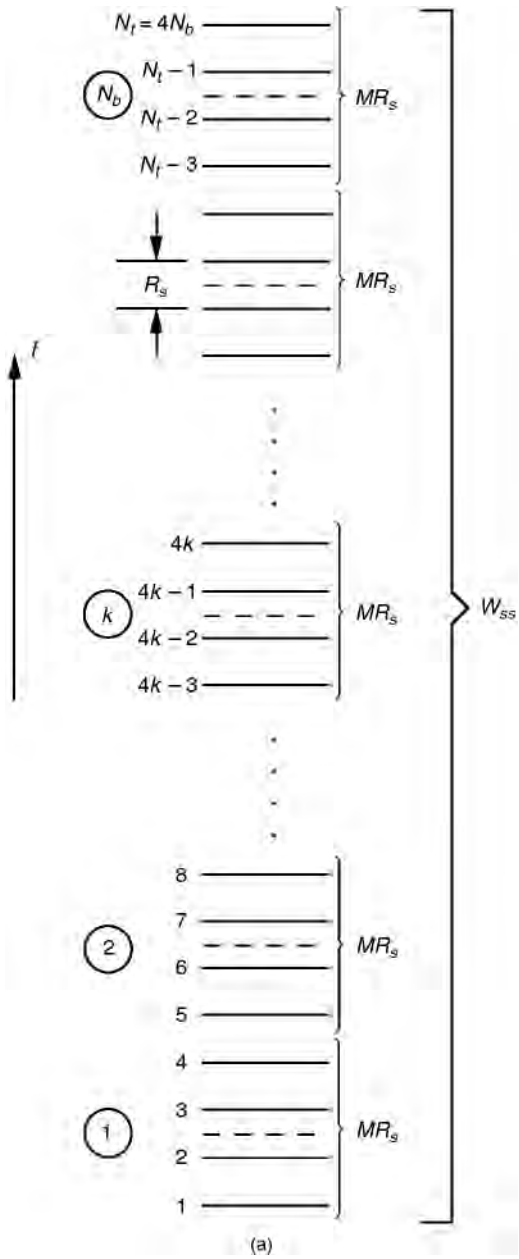


FIGURE 15.3(a) Frequency distribution for FH-4FSK—nonoverlapping bands. Dashed lines indicate location of hop frequencies.

where $d(t)$ is a digital information modulation at a rate $1/T_s$ sps. Finally, the dehopping procedure at the receiver consists of removing the sequence of delays introduced by $c(t)$, which restores the information signal back to its original form and spreads the interferer.

Hybrid Modulations

By blending together several of the previous types of SS modulation, one can form **hybrid** modulations that, depending on the system design objectives, can achieve a better performance against the interferer than can any of the SS modulations acting alone. One possibility is to multiply several of the $c(t)$ wideband

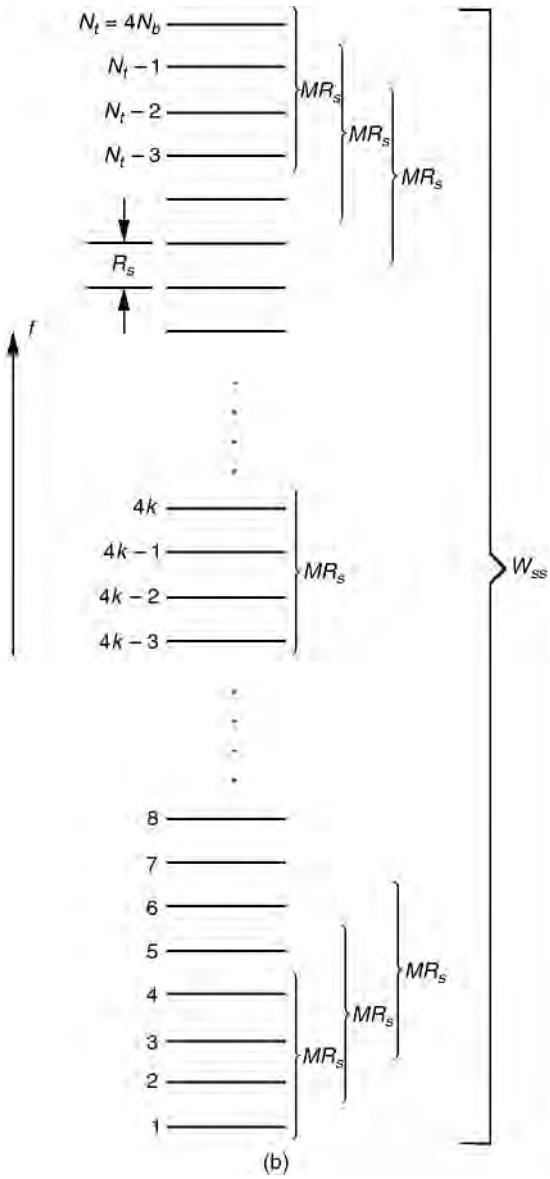


FIGURE 15.3(b) Frequency distribution for FH-4FSK—overlapping bands.

waveforms [now denoted by $c^{(i)}(t)$ to distinguish them from one another] resulting in a SS modulation of the form

$$c(t) = \prod_i c^{(i)}(t) \tag{15.9}$$

Such a modulation may embrace the advantages of the various $c^{(i)}(t)$, while at the same time mitigating their individual disadvantages.

15.5 Applications of Spread Spectrum

Military

Antijam (AJ) Communications

As already noted, one of the key applications of spread spectrum is for antijam communications in a hostile environment. The basic mechanism by which a **direct sequence spread spectrum** receiver attenuates a noise jammer was illustrated in Section 15.3. Therefore, in this section, we will concentrate on tone jamming.

Assume the received signal, denoted $r(t)$, is given by

$$r(t) = Ax(t) + I(t) + n_w(t) \tag{15.10}$$

where $x(t)$ is given in Eq. (15.4), A is a constant amplitude,

$$I(t) = \alpha\cos(2\pi f_c t + \theta) \tag{15.11}$$

and $n_w(t)$ is additive white Gaussian noise (AWGN) having two-sided spectral density $N_0/2$. In Eq. (15.11), α is the amplitude of the tone jammer and θ is a random phase uniformly distributed in $[0, 2\pi]$.

If we employ the standard correlation receiver of Fig. 15.4, it is straightforward to show that the final test statistic out of the receiver is given by

$$g(T_b) = AT_b + \alpha\cos\theta\int_0^{T_b} c(t)\mathrm{d}t + N(T_b) \tag{15.12}$$

where $N(T_b)$ is the contribution to the test statistic due to the AWGN. Noting that, for rectangular chips, we can express

$$\int_0^{T_b} c(t)\mathrm{d}t = T_c\sum_{i=1}^{M} c_i \tag{15.13}$$

where

$$M \triangleq \frac{T_b}{T_c} \tag{15.14}$$

is one-half of the processing gain, it is straightforward to show that, for a given value of θ, the signal-to-noise-plus-interference ratio, denoted by S/N_{total}, is given by

$$\frac{S}{N_{\text{total}}} = \frac{1}{\dfrac{N_0}{2E_b} + \left(\dfrac{J}{MS}\right)\cos^2\theta} \tag{15.15}$$

In Eq.(15.15), the jammer power is

$$J \triangleq \frac{\alpha^2}{2} \tag{15.16}$$

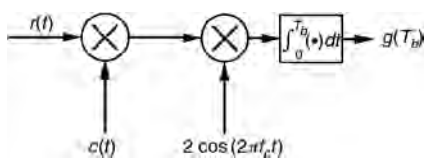


FIGURE 15.4 Standard correlation receiver.

and the signal power is

$$S \triangleq \frac{A^2}{2} \tag{15.17}$$

If we look at the second term in the denominator of Eq. (15.15), we see that the ratio J/S is divided by M. Realizing that J/S is the ratio of the jammer power to the signal power before despreading, and J/MS is the ratio of the same quantity after despreading, we see that, as was the case for noise jamming, the benefit of employing direct sequence spread spectrum signalling in the presence of tone jamming is to reduce the effect of the jammer by an amount on the order of the processing gain.

Finally, one can show that an estimate of the average probability of error of a system of this type is given by

$$P_e = \frac{1}{2\pi}\int_0^{2\pi} \phi\left(-\sqrt{\frac{S}{N_{\text{total}}}}\right) \mathrm{d}\theta \tag{15.18}$$

where

$$\phi(x) \triangleq \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-y^2/2}\mathrm{d}y \tag{15.19}$$

If Eq. (15.18) is evaluated numerically and plotted, the results are as shown in Fig. 15.5. It is clear from this figure that a large initial power advantage of the jammer can be overcome by a sufficiently large value of the processing gain.

Low-Probability of Intercept (LPI)

The opposite side of the AJ problem is that of LPI, that is, the desire to hide your signal from detection by an intelligent adversary so that your transmissions will remain unnoticed and, thus, neither jammed nor exploited in any manner. This idea of designing an LPI system is achieved in a variety of ways, including transmitting at the smallest possible power level, and limiting the transmission time to as short an interval in time as is possible. The choice of signal design is also important, however, and it is here that spread spectrum techniques become relevant.

The basic mechanism is reasonably straightforward; if we start with a conventional narrowband signal, say a BPSK waveform having a spectrum as shown in Fig. 15.6a, and then spread it so that its new spectrum is as shown in Fig. 15.6b, the peak amplitude of the spectrum after spreading has been reduced by an amount on the order of the processing gain relative to what it was before spreading. Indeed, a sufficiently large processing gain will result in the spectrum of the signal after spreading falling below the ambient thermal noise level. Thus, there is no easy way for an unintended listener to determine that a transmission is taking place.

That is not to say the spread signal cannot be detected, however, merely that it is more difficult for an adversary to learn of the transmission. Indeed, there are many forms of so-called intercept receivers that are specifically designed to accomplish this very task. By way of example, probably the best known and simplest to implement is a **radiometer**, which is just a device that measures the total power present in the received signal. In the case of our intercept problem, even though we have lowered the power spectral density of the transmitted signal so that it falls below the noise floor, we have not lowered its power (i.e., we have merely spread its power over a wider frequency range). Thus, if the radiometer integrates over a sufficiently long period of time, it will eventually determine the presence of the transmitted signal buried in the noise. The key point, of course, is that the use of the spreading makes the interceptor's task much more difficult, since he has no knowledge of the spreading code and, thus, cannot despread the signal.

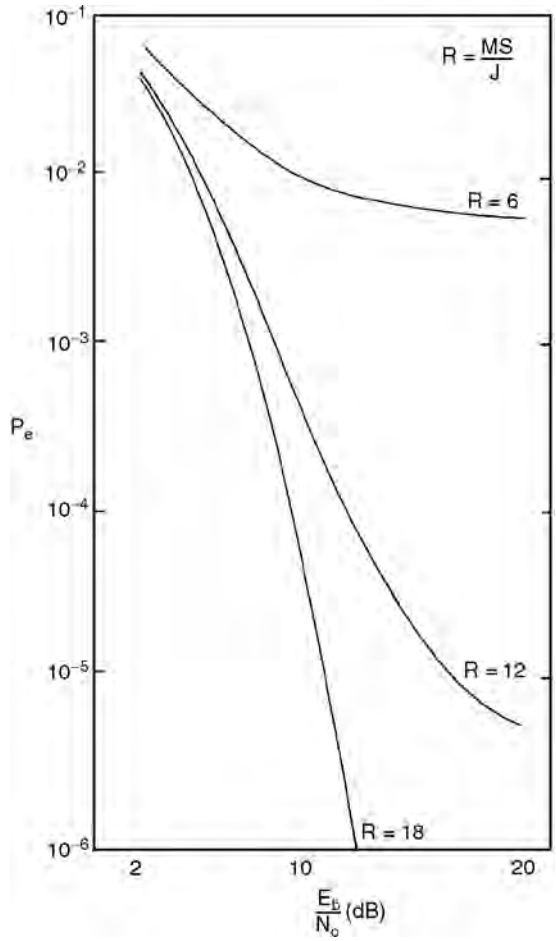


FIGURE 15.5 Plotted results of Eq. (11.18).

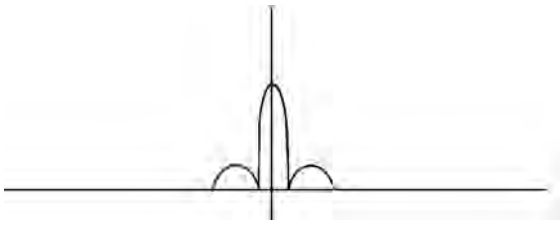


FIGURE 15.6a

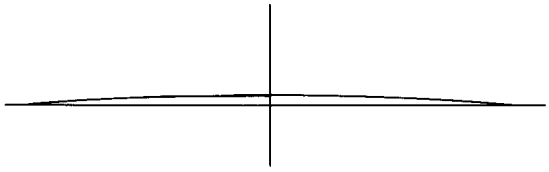


FIGURE 15.6b

Commercial

Multiple Access Communications

From the perspective of commercial applications, probably the most important use of spread spectrum communications is as a multiple accessing technique. When used in this manner, it becomes an alternative to either frequency division multiple access (FDMA) or time division multiple access (TDMA) and is typically referred to as either code division multiple access (CDMA) or spread spectrum multiple access (SSMA). When using CDMA, each signal in the set is given its own spreading sequence. As opposed to either FDMA, wherein all users occupy disjoint frequency bands but are transmitted simultaneously in time, or TDMA, whereby all users occupy the same bandwidth but transmit in disjoint intervals of time, in CDMA, all signals occupy the same bandwidth and are transmitted simultaneously in time; the different waveforms in CDMA are distinguished from one another at the receiver by the specific spreading codes they employ.

Since most CDMA detectors are correlation receivers, it is important when deploying such a system to have a set of spreading sequences that have relatively low-pairwise cross-correlation between any two sequences in the set. Further, there are two fundamental types of operation in CDMA, synchronous and asynchronous. In the former case, the symbol transition times of all of the users are aligned; this allows for orthogonal sequences to be used as the spreading sequences and, thus, eliminates interference from one user to another. Alternately, if no effort is made to align the sequences, the system operates asychronously; in this latter mode, multiple access interference limits the ultimate channel capacity, but the system design exhibits much more flexibility.

CDMA has been of particular interest recently for applications in wireless communications. These applications include cellular communications, personal communications services (PCS), and wireless local area networks. The reason for this popularity is primarily due to the performance that spread spectrum waveforms display when transmitted over a multipath fading channel.

To illustrate this idea, consider DS signalling. As long as the duration of a single chip of the spreading sequence is less than the multipath delay spread, the use of DS waveforms provides the system designer with one of two options. First, the multipath can be treated as a form of interference, which means the receiver should attempt to attenuate it as much as possible. Indeed, under this condition, all of the multipath returns that arrive at the receiver with a time delay greater than a chip duration from the multipath return to which the receiver is synchronized (usually the first return) will be attenuated because of the processing gain of the system.

Alternately, the multipath returns that are separated by more than a chip duration from the main path represent independent "looks" at the received signal and can be used constructively to enhance the overall performance of the receiver. That is, because all of the multipath returns contain information regarding the data that is being sent, that information can be extracted by an appropriately designed receiver. Such a receiver, typically referred to as a RAKE receiver, attempts to resolve as many individual multipath returns as possible and then to sum them coherently. This results in an *implicit* diversity gain, comparable to the use of *explicit* diversity, such as receiving the signal with multiple antennas.

The condition under which the two options are available can be stated in an alternate manner. If one envisions what is taking place in the frequency domain, it is straightforward to show that the condition of the chip duration being smaller than the multipath delay spread is equivalent to requiring that the spread bandwidth of the transmitted waveform exceed what is called the coherence bandwidth of the channel. This latter quantity is simply the inverse of the multipath delay spread and is a measure of the range of frequencies that fade in a highly correlated manner. Indeed, anytime the coherence bandwidth of the channel is less than the spread bandwidth of the signal, the channel is said to be *frequency selective* with respect to the signal. Thus, we see that to take advantage of DS signalling when used over a multipath fading channel, that signal should be designed such that it makes the channel appear frequency selective.

In addition to the desirable properties that spread spectrum signals display over multipath channels, there are two other reasons why such signals are of interest in cellular-type applications. The first has to do with a concept known as the reuse factor. In conventional cellular systems, either analog or digital, in order to avoid excessive interference from one cell to its neighbor cells, the frequencies used by a given

cell are not used by its immediate neighbors (i.e., the system is designed so that there is a certain spatial separation between cells that use the same carrier frequencies). For CDMA, however, such spatial isolation is typically not needed, so that so-called *universal reuse* is possible.

Further, because CDMA systems tend to be interference limited, for those applications involving voice transmission, an additional gain in the capacity of the system can be achieved by the use of *voice activity detection*. That is, in any given two-way telephone conversation, each user is typically talking only about 50% of the time. During the time when a user is quiet, he is not contributing to the instantaneous interference. Thus, if a sufficiently large number of users can be supported by the system, statistically only about one-half of them will be active simultaneously, and the effective capacity can be doubled.

Interference Rejection

In addition to providing multiple accessing capability, spread spectrum techniques are of interest in the commercial sector for basically the same reasons they are used in the military community, namely their AJ and LPI characteristics. However, the motivations for such interest differ. For example, whereas the military is interested in ensuring that the systems they deploy are robust to interference generated by an intelligent adversary (i.e., exhibit jamming resistance), the interference of concern in commercial applications is unintentional. It is sometimes referred to as cochannel interference (CCI) and arises naturally as the result of many services using the same frequency band at the same time. And while such scenarios almost always allow for some type of spatial isolation between the interfering waveforms, such as the use of narrow-beam antenna patterns, at times the use of the inherent interference suppression property of a spread spectrum signal is also desired. Similarly, whereas the military is very much interested in the LPI property of a spread spectrum waveform, as indicated in Section 15.3, there are applications in the commercial segment where the same characteristic can be used to advantage.

To illustrate these two ideas, consider a scenario whereby a given band of frequencies is somewhat sparsely occupied by a set of conventional (i.e., nonspread) signals. To increase the overall spectral efficiency of the band, a set of spread spectrum waveforms can be overlaid on the same frequency band, thus forcing the two sets of users to share a common spectrum. Clearly, this scheme is feasible only if the mutual interference that one set of users imposes on the other is within tolerable limits. Because of the interference suppression properties of spread spectrum waveforms, the despreading process at each spread spectrum receiver will attenuate the components of the final test statistic due to the overlaid narrowband signals. Similarly, because of the LPI characteristics of spread spectrum waveforms, the increase in the overall noise level as seen by any of the conventional signals, due to the overlay, can be kept relatively small.

Defining Terms

Antijam communication system: A communication system designed to resist intentional jamming by the enemy.

Chip time (interval): The duration of a single pulse in a direct sequence modulation; typically much smaller than the information symbol interval.

Coarse alignment: The process whereby the received signal and the despreading signal are aligned to within a single chip interval.

Dehopping: Despreading using a frequency-hopping modulation.

Delay-locked loop: A particular implementation of a closed-loop technique for maintaining fine alignment.

Despreading: The notion of decreasing the bandwidth of the received (spread) signal back to its information bandwidth.

Direct sequence modulation: A signal formed by linearly modulating the output sequence of a pseudorandom number generator onto a train of pulses.

Direct sequence spread spectrum: A spreading technique achieved by multiplying the information signal by a direct sequence modulation.

Fast frequency-hopping: A spread spectrum technique wherein the hop time is less than or equal to the information symbol interval, i.e., there exist one or more hops per data symbol.

Fine alignment: The state of the system wherein the received signal and the despreading signal are aligned to within a small fraction of a single chip interval.

Frequency-hopping modulation: A signal formed by nonlinearly modulating a train of pulses with a sequence of pseudorandomly generated frequency shifts.

Hop time (interval): The duration of a single pulse in a frequency-hopping modulation.

Hybrid spread spectrum: A spreading technique formed by blending together several spread spectrum techniques, e.g., direct sequence, frequency-hopping, etc.

Low-probability-of-intercept communication system: A communication system designed to operate in a hostile environment wherein the enemy tries to detect the presence and perhaps characteristics of the friendly communicator's transmission.

Processing gain (spreading ratio): The ratio of the spread spectrum bandwidth to the information data rate.

Radiometer: A device used to measure the total energy in the received signal.

Search algorithm: A means for coarse aligning (synchronizing) the despreading signal with the received spread spectrum signal.

Slow frequency-hopping: A spread spectrum technique wherein the hop time is greater than the information symbol interval, i.e., there exists more than one data symbol per hop.

Spread spectrum bandwidth: The bandwidth of the transmitted signal after spreading.

Spreading: The notion of increasing the bandwidth of the transmitted signal by a factor far in excess of its information bandwidth.

Tau–dither loop: A particular implementation of a closed-loop technique for maintaining fine alignment.

Time-hopping spread spectrum: A spreading technique that is analogous to pulse position modulation.

Tracking algorithm: An algorithm (typically closed loop) for maintaining fine alignment.

References

1. Cook, C.F., Ellersick, F.W., Milstein, L.B., and Schilling, D.L., *Spread Spectrum Communications,* IEEE Press, 1983.
2. Dixon, R.C., *Spread Spectrum Systems,* 3rd ed., John Wiley & Sons, New York, 1994.
3. Holmes, J.K., *Coherent Spread Spectrum Systems,* John Wiley & Sons, New York, 1982.
4. Simon, M.K., Omura, J.K., Scholtz, R.A., and Levitt, B.K., *Spread Spectrum Communications Handbook,* McGraw-Hill, 1994 (previously published as *Spread Spectrum Communications,* Computer Science Press, 1985).
5. Ziemer, R.E. and Peterson, R.L., *Digital Communications and Spread Spectrum Techniques,* Macmillan, New York, 1985.

16

Diversity

Arogyaswami J. Paulraj
Stanford University

16.1 Introduction

Diversity is a commonly used technique in mobile radio systems to combat signal **fading**. The basic principle of diversity is as follows. If several replicas of the same information-carrying signal are received over multiple channels with comparable strengths, which exhibit independent fading, then there is a good likelihood that at least one or more of these received signals will not be in a fade at any given instant in time, thus making it possible to deliver an adequate signal level to the receiver. Without diversity techniques, in noise limited conditions, the transmitter would have to deliver a much higher power level to protect the link during the short intervals when the channel is severely faded. In mobile radio, the power available on the reverse link is severely limited by the battery capacity of hand-held subscriber units. Diversity methods play a crucial role in reducing transmit power needs. Also, cellular communication networks are mostly interference limited and, once again, mitigation of channel fading through use of diversity can translate into reduced variability of carrier-to-interference ratio (C/I), which in turn means lower C/I margin and hence better reuse factors and higher system capacity.

The basic principles of diversity have been known since 1927, when the first experiments in space diversity were reported. There are many techniques for obtaining independently fading branches, and these can be subdivided into two main classes. The first are explicit techniques where explicit redundant signal transmission is used to exploit diversity channels. Use of dual polarized signal transmission and reception in many point-to-point radios is an example of explicit diversity. Clearly, such redundant signal transmission involves a penalty in frequency spectrum or additional power. In the second class are implicit diversity techniques: the signal is transmitted only once, but the decorrelating effects in the propagation medium such as multipaths are exploited to receive signals over multiple diversity channels. A good example of implicit diversity is the **RAKE receiver** in code division multiple access (CDMA) systems, which uses independent fading of resolvable multipaths to achieve diversity gain. Figure 16.1 illustrates the principle of diversity where two independently fading signals are shown along with the selection diversity output signal which selects the stronger signal. The fades in the resulting signal have been substantially smoothed out while also yielding higher average power.

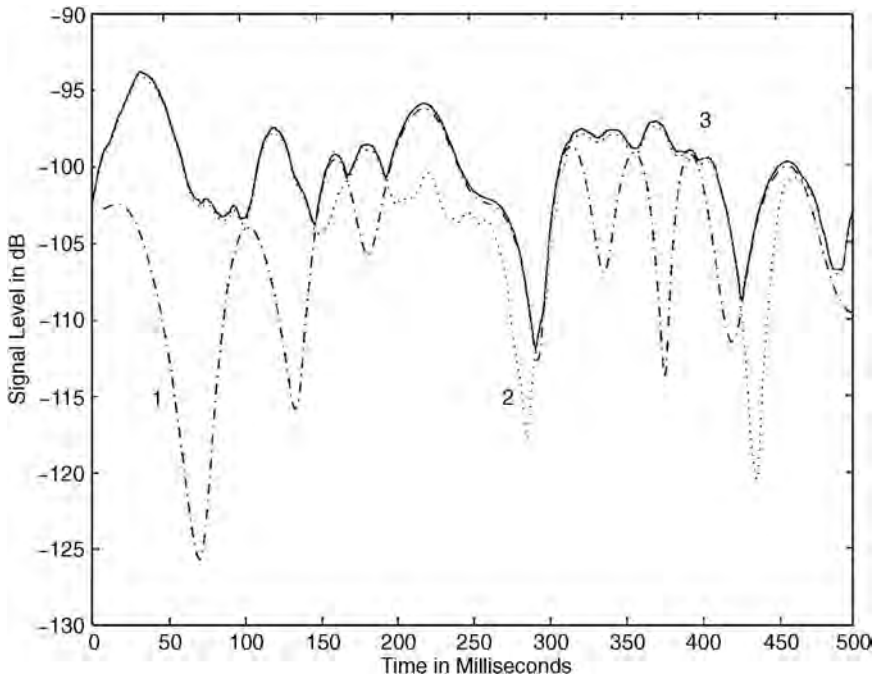


FIGURE 16.1 Example of diversity combining. Two independently fading signals 1 and 2. The signal 3 is the result of selecting the strongest signal.

If antennas are used in transmit, they can be exploited for diversity. If the transmit channel is known, the antennas can be driven with complex conjugate channel weighting to co-phase the signals at the receive antenna. If the forward channel is not known, we have several methods to convert space selective fading at the transmit antennas to other forms of diversity exploitable in the receiver.

Exploiting diversity needs careful design of the communication link. In explicit diversity, multiple copies of the same signal are transmitted in channels using either a frequency, time, or polarization dimension. At the receiver end we need arrangements to receive the different diversity branches (this is true for both explicit and implicit diversity). The different diversity branches are then combined to reduce signal **outage probability** or bit error rate.

In practice, the signals in the diversity branches may not show completely independent fading. The envelope cross correlation ρ between these signals is a measure of their independence.

$$r = \frac{E[[r_1 - \bar{r}_1][r_2 - \bar{r}_2]]}{\sqrt{E|r_1 - \bar{r}_1|^2 E|r_2 - \bar{r}_2|^2}}$$

where r_1 and r_2 represent the instantaneous envelope levels of the normalized signals at the two receivers and $\bar{r}_1$ and $\bar{r}_2$ are their respective means. It has been shown that a cross correlation of 0.7 [3] between signal envelopes is sufficient to provide a reasonable degree of diversity gain. Depending on the type of diversity employed, these diversity channels must be sufficiently *separated* along the appropriate diversity dimension. For spatial diversity, the antennas should be separated by more than the *coherence distance* to ensure a cross correlation of less than 0.7. Likewise in frequency diversity, the frequency separation must be larger than the *coherence bandwidth,* and in time diversity the separation between channel reuse in time should be longer than the *coherence time.* These coherence factors in turn depend on the channel characteristics. The coherence distance, coherence bandwidth, and coherence time vary inversely as the angle spread, delay spread, and Doppler spread, respectively.

If the receiver has a number of diversity branches, it has to combine these branches to maximize the signal level. Several techniques have been studied for diversity combining. We will describe three main techniques: selection combining, equal gain combining, and maximal ratio combining.

Finally, we should note that diversity is primarily used to combat fading, and if the signal does not show significant fading in the first place, for example, when there is a direct path component, diversity combining may not provide significant diversity gain. In the case of antenna diversity, array gain proportional to the number of antennas will still be available.

16.2 Diversity Schemes

There are several techniques for obtaining diversity branches, sometimes also known as diversity dimensions. The most important of these are discussed in the following sections.

Space Diversity

This has historically been the most common form of diversity in mobile radio base stations. It is easy to implement and does not require additional frequency spectrum resources. Space diversity is exploited on the reverse link at the base station receiver by spacing antennas apart so as to obtain sufficient decorrelation. The key for obtaining minimum uncorrelated fading of antenna outputs is adequate spacing of the antennas. The required spacing depends on the degree of multipath angle spread. For example, if the multipath signals arrive from all directions in the azimuth, as is usually the case at the mobile, antenna spacing (coherence distance) of the order of 0.5λ to 0.8λ is quite adequate [5]. On the other hand, if the multipath angle spread is small, as in the case of base stations, the coherence distance is much larger. Also, empirical measurements show a strong coupling between antenna height and spatial correlation. Larger antenna heights imply larger coherence distances. Typically, 10λ to 20λ separation is adequate to achieve $\rho = 0.7$ at base stations in suburban settings when the signals arrive from the broadside direction. The coherence distance can be 3 to 4 times larger for endfire arrivals. The endfire problem is averted in base stations with trisectored antennas as each sector needs to handle only signals arriving $\pm 60\infty$ off the broadside. The coherence distance depends strongly on the terrain. Large multipath angle spread means smaller coherence distance. Base stations normally use space diversity in the horizontal plane only. Separation in the vertical plane can also be used, and the necessary spacing depends upon vertical multipath angle spread. This can be small for distant mobiles making vertical plane diversity less attractive in most applications.

Space diversity is also exploitable at the transmitter. If the forward channel is known, it works much like receive space diversity. If it is not known, then space diversity can be transformed to another form of diversity exploitable at the receiver. (See the section on Transformed Diversity).

If antennas are used at transmit and receive, the M transmit and N receive antennas both contribute to diversity. It can be shown that if simple weighting is used without additional bandwidth or time/memory processing, then maximum diversity gain is obtained if the transmitter and receiver use the left and right singular vectors of the $M \times N$ channel matrix, respectively. However, to approach the maximum $M \times N$ order diversity will require the use of additional bandwidth or time/memory-based methods.

Polarization Diversity

In mobile radio environments, signals transmitted on orthogonal polarizations exhibit low fade correlation and, therefore, offer potential for diversity combining. Polarization diversity can be obtained either by explicit or implicit techniques. Note that with polarization, only two diversity branches are available as compared to space diversity, where several branches can be obtained using multiple antennas. In explicit polarization diversity, the signal is transmitted and received in two orthogonal polarizations. For a fixed total transmit power, the power in each branch will be 3 dB lower than if single polarization is used. In the implicit polarization technique, the signal is launched in a single polarization but is received with cross-polarized antennas. The propagation medium couples some energy into the cross-polarization plane. The observed cross-polarization coupling factor lies between 8 to 12 dB in mobile radio[8,1].

The cross-polarization envelope decorrelation has been found to be adequate. However, the large branch imbalance reduces the available diversity gain.

With hand-held phones, the handset can be held at random orientations during a call. This results in energy being launched with varying polarization angles ranging from vertical to horizontal. This further increases the advantage of cross-polarized antennas at the base station since the two antennas can be combined to match the received signal polarization. This makes polarization diversity even more attractive. Recent work [4] has shown that with variable launch polarization, a cross-polarized antenna can give comparable overall (matching plus diversity) performance to a vertically polarized space diversity antenna.

Finally, we should note that cross-polarized antennas can be deployed in a compact antenna assembly and do not need large physical separation as needed in space diversity antennas. This is an important advantage in the **PCS** base stations where low profile antennas are needed.

Angle Diversity

In situations where the angle spread is very high, such as indoors or at the mobile unit in urban locations, signals collected from multiple nonoverlapping beams offer low fade correlation with balanced power in the diversity branches. Clearly, since directional beams imply use of antenna aperture, angle diversity is closely related to space diversity. Angle diversity has been utilized in indoor wireless LANs, where its use allows substantial increase in LAN throughputs [2].

Frequency Diversity

Another technique to obtain decorrelated diversity branches is to transmit the same signal over different frequencies. The frequency separation between carriers should be larger than the coherence bandwidth. The coherence bandwidth, of course, depends on the multipath delay spread of the channel. The larger the delay spread, the smaller the coherence bandwidth and the more closely we can space the frequency diversity channels. Clearly, frequency diversity is an explicit diversity technique and needs additional frequency spectrum.

A common form of frequency diversity is multicarrier (also known as multitone) modulation. This technique involves sending redundant data over a number of closely spaced carriers to benefit from frequency diversity, which is then exploited by applying **interleaving** and **channel coding/forward error correction** across the carriers. Another technique is to use **frequency hopping** wherein the interleaved and channel coded data stream is transmitted with widely separated frequencies from burst to burst. The wide frequency separation is chosen to guarantee independent fading from burst to burst.

Path Diversity

This implicit diversity is available if the signal bandwidth is much larger than the channel coherence bandwidth. The basis for this method is that when the multipath arrivals can be resolved in the receiver and since the paths fade independently, diversity gain can be obtained. In CDMA systems, the multipath arrivals must be separated by more than one *chip* period and the RAKE receiver provides the diversity [9]. In TDMA systems, the multipath arrivals must be separated by more than one *symbol* period and the MLSE receiver provides the diversity.

Time Diversity

In mobile communications channels, the mobile motion together with scattering in the vicinity of the mobile causes time selective fading of the signal with Rayleigh fading statistics for the signal envelope. Signal fade levels separated by the *coherence time* show low correlation and can be used as diversity branches if the same signal can be transmitted at multiple instants separated by the coherence time. The coherence time depends on the Doppler spread of the signal, which in turn is a function of the mobile speed and the carrier frequency.

Time diversity is usually exploited via interleaving, forward-error correction (FEC) coding, and **automatic request for repeat** (ARQ). These are sophisticated techniques to exploit channel coding and time

diversity. One fundamental drawback with time diversity approaches is the delay needed to collect the repeated or interleaved transmissions. If the coherence time is large, as, for example, when the vehicle is slow moving, the required delay becomes too large to be acceptable for interactive voice conversation.

The statistical properties of fading signals depend on the field component used by the antenna, the vehicular speed, and the carrier frequency. For an idealized case of a mobile surrounded by scatterers in all directions, the autocorrelation function of the received signal $x(t)$ (note this is not the envelope $r(t)$) can be shown to be

$$E[x(t)x(t+t)] = J_0(2\pi\tau v/\lambda)$$

where J_0 is a Bessel function of the 0th order and v is the mobile velocity.

Transformed Diversity

In transformed diversity, the space diversity branches at the transmitter are transformed into other forms of diversity branches exploitable at the receiver. This is used when the forward channel is not known and shifts the responsibility of diversity combining to the receiver which has the necessary channel knowledge.

Space to Frequency

- *Antenna-delay.* Here the signal is transmitted from two or more antennas with delays of the order of a chip or symbol period in CDMA or TDMA, respectively. The different transmissions simulate resolved path arrivals that can be used as diversity branches by the RAKE or MLSE equalizer.
- *Multicarrier modulation.* The data stream after interleaving and coding is modulated as a multicarrier output using an inverse DFT. The carriers are then mapped to the different antennas. The space selective fading at the antennas is now transformed to frequency selective fading and diversity is obtained during decoding.

Space to Time

- *Antenna hopping/phase rolling.* In this method the data stream after coding and interleaving is switched randomly from antenna to antenna. The space selective fading at the transmitter is converted into a time selective fading at the receiver. This is a form of "active" fading.
- *Space-time coding.* The approach in space-time coding is to split the encoded data into multiple data streams, each of which is modulated and simultaneously transmitted from different antennas. The received signal is a superposition of the multiple transmitted signals. Channel decoding can be used to recover the data sequence. Since the encoded data arrive over uncorrelated fade branches, diversity gain can be realized.

16.3 Diversity Combining Techniques

Several diversity combining methods are known. We describe three main techniques: selection, maximal ratio, and equal gain. They can be used with each of the diversity schemes discussed above.

Selection Combining

This is the simplest and perhaps the most frequently used form of diversity combining. In this technique, one of the two diversity branches with the highest carrier-to-noise ratio (C/N) is connected to the output. See Fig. 16.2(a).

The performance improvement due to selection diversity can be seen as follows. Let the signal in each branch exhibit Rayleigh fading with mean power σ^2. The density function of the envelope is given by

$$p(r_i) = \frac{r_i}{\sigma^2} e^{\frac{-r_i^2}{2s^2}} \tag{16.1}$$

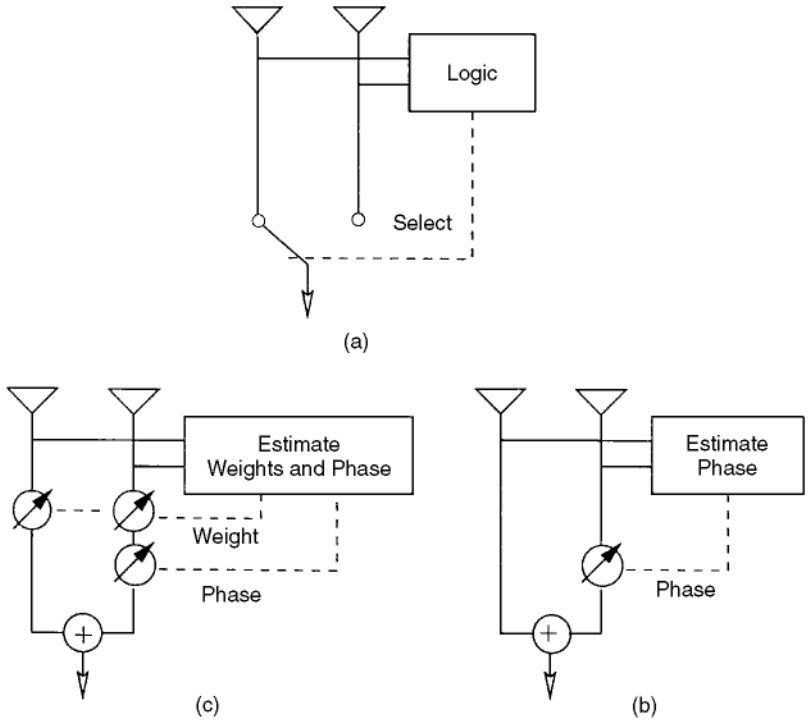


FIGURE 16.2 Diversity combining methods for two diversity branches.

where r_i is the signal envelope in each branch. If we define two new variables

$$g_i = \frac{\text{Instantaneous signal power in each branch}}{\text{Mean noise power}}$$

$$G = \frac{\text{Mean signal power in each branch}}{\text{Mean noise power}}$$

then the probability that the C/N is less than or equal to some specified value γ_s is

$$\text{Prob}[\gamma_i \le \gamma_s] = 1 - e^{-\gamma_s/\Gamma} \tag{16.2}$$

The probability that γ_i in all branches with independent fading will be simultaneously less than or equal to γ_s is then

$$\text{Prob}[\gamma_1, \gamma_2, \ldots \gamma_M \le \gamma_s] = (1 - e^{-\gamma_s/\Gamma})^M \tag{16.3}$$

This is the distribution of the best signal envelope from the two diversity branches. Figure 16.3 shows the distribution of the combiner output C/N for $M = 1,2,3$, and 4 branches. The improvement in signal quality is significant. For example, at 99% reliability level, the improvement in C/N is 10 dB for two branches and 16 dB for four branches.

Selection combining also increases the mean C/N of the combiner output and can be shown to be [3]

$$\text{Mean}(\gamma_s) = \Gamma \sum_{k=1}^{M} \frac{1}{k} \tag{16.4}$$

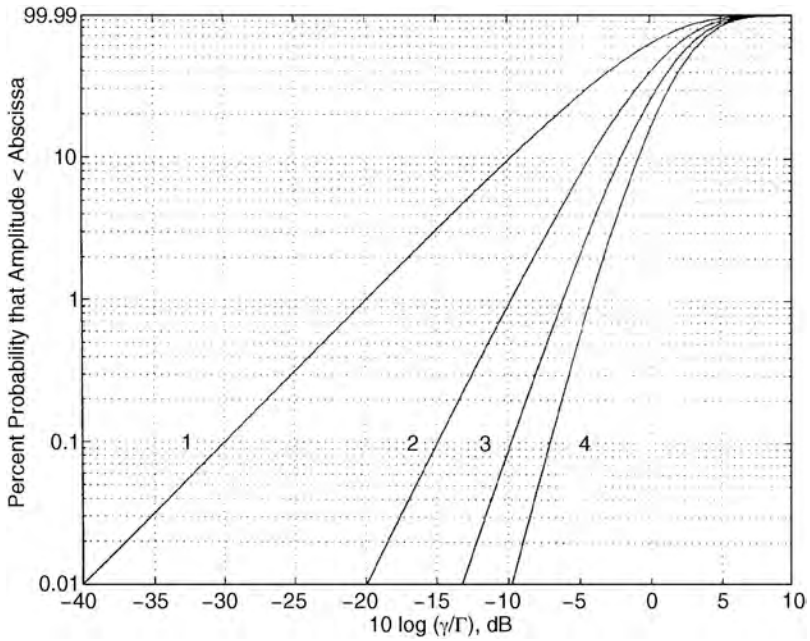


FIGURE 16.3 Probability distribution of signal envelope for selection combining.

This indicates that with 4 branches, for example, the mean C/N of the selected branch is 2.08 better than the mean C/N in any one branch.

Maximal Ratio Combining

In this technique the M diversity branches are first co-phased and then weighted proportionally to their signal level before summing. See Fig. 16.2(b). The distribution of the maximal ratio combiner has been shown to be [5]

$$\text{Prob}[\gamma \le \gamma_m] = 1 - e^{(-\gamma_m/\Gamma)} \sum_{k=1}^{M} \frac{(\gamma_m/\Gamma)^{k-1}}{(k-1)!} \tag{16.5}$$

The distribution of output of a maximal ratio combiner is shown in Fig. 16.4. Maximal ratio combining is known to be optimal in the sense that it yields the best statistical reduction of fading of any linear diversity combiner. In comparison to the selection combiner, at 99% reliability level, the maximal ratio combiner provides a 11.5 dB gain for two branches and a 19 dB gain for four branches, an improvement of 1.5 and 3 dB, respectively, over the selection diversity combiner.

The mean C/N of the combined signal may be easily shown to be

$$\text{Mean}(\gamma_m) = M\Gamma \tag{16.6}$$

Therefore, combiner output mean varies linearly with M. This confirms the intuitive result that the output C/N averaged over fades should provide gain proportional to the number of diversity branches. This is a situation similar to conventional beamforming.

Equal Gain Combining

In some applications, it may be difficult to estimate the amplitude accurately, the combining gains may all be set to unity, and the diversity branches merely summed after co-phasing. [See Fig. 16.2(c)].

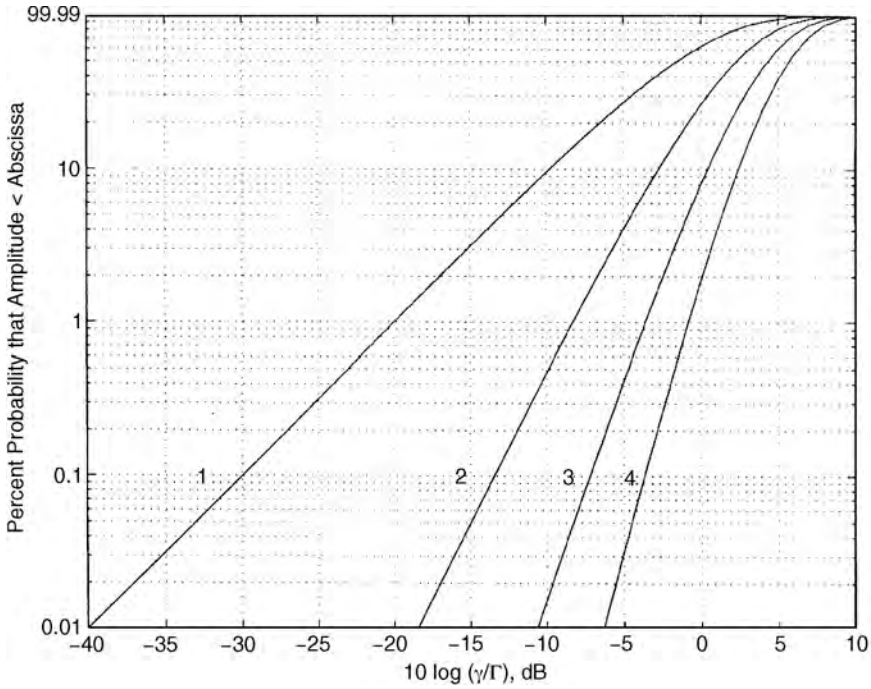


FIGURE 16.4 Probability distribution for signal envelope for maximal ratio combining.

The distribution of equal gain combiner does not have a neat expression and has been computed by numerical evaluation. Its performance has been shown to be very close to within a decibel to maximal ratio combining. The mean C/N can be shown to be [3]

$$\text{Mean}(\gamma_e) = \Gamma\left[1 + \frac{\pi}{4}(M-1)\right] \tag{16.7}$$

Like maximal ratio combining, the mean C/N for equal gain combining grows almost linearly with M and is approximately only one decibel poorer than maximal ratio combining even with an infinite number of branches.

Loss of Diversity Gain Due to Branch Correlation and Unequal Branch Powers

The above analysis assumed that the fading signals in the diversity branches were all uncorrelated and of equal power. In practice, this may be difficult to achieve and, as we saw earlier, the branch cross-correlation coefficient $\rho = 0.7$ is considered to be acceptable. Also, equal mean powers in diversity branches are rarely available. In such cases we can expect a certain loss of diversity gain. However, since most of the damage in fading is due to deep fades, and also since the chance of coincidental deep fades is small even for moderate branch correlation, one can expect a reasonable tolerance to branch correlation.

The distribution of the output signal envelope of maximal ratio combiner has been shown to be [6]:

$$\text{Prob}[\gamma_m] = \sum_{n=1}^{M} \frac{A_n}{2\lambda_n} e^{-\gamma_m/2\lambda_n} \tag{16.8}$$

where λ_n are the eigenvalues of the M ¥ M branch envelope covariance matrix whose elements are defined by

$$\mathbf{R}_{ij} = E[r_i r_j^*] \tag{16.9}$$

and A_n is defined by

$$A_n = \prod_{\substack{k=1 \\ k \neq n}}^{M} \frac{1}{1 - \lambda_k/\lambda_n} \tag{16.10}$$

16.4 Effect of Diversity Combining on Bit Error Rate

So far we have studied the distribution of the instantaneous envelope or C/N after diversity combining. We will now briefly survey how diversity combining affects BER performance in digital radio links; we assume maximal ratio combining.

To begin let us first examine the effect of Rayleigh fading on the BER performance of digital transmission links. This has been studied by several authors and is summarized in [7]. Table 16.1 gives the BER expressions in the large E_b/N_0 case for coherent binary PSK and coherent binary orthogonal FSK for unfaded and Rayleigh faded AWGN (additive white Gaussian noise channels) channels. $\bar{E}_b/N_0$ represents the average E_b/N_0 for the fading channel.

Observe that error rates decrease only inversely with SNR as compared to exponential decreases for the unfaded channel. Also note that for fading channels, coherent binary PSK is 3 dB better than coherent binary FSK, exactly the same advantage as in the unfaded case. Even for the modest target BER of 10^{-2} that is usually needed in mobile communications, the loss due to fading can be very high—17.2 dB.

To obtain the BER with maximal ratio diversity combining, we have to average the BER expression for the unfaded BER with the distribution obtained for the maximal ratio combiner given in Eq. (16.5). Analytical expressions have been derived for these in [7]. For a branch SNR greater than 10 dB, the BER after maximal ratio diversity combining is given in Table 16.2.

We observe that the probability of error varies as $1/\bar{E}_b/N_0$ raised to the Lth power. Thus, diversity reduces the error rate exponentially as the number of independent branches increases.

TABLE 16.1 Comparison of BER Performance for Unfaded and Rayleigh Faded Signals

Modulaton	Unfaded BER	Faded BER
Coh BPSK	$\frac{1}{2}\text{erfc}(\sqrt{E_b/N_0})$	$\frac{1}{4(\bar{E}_b/N_0)}$
Coh FSK	$\frac{1}{2}\text{erfc}\left(\sqrt{\frac{1}{2}E_b/N_0}\right)$	$\frac{1}{2(\bar{E}_b/N_0)}$

TABLE 16.2 BER Performance for Coherent BPSK and FSK with Diversity

Modulaton	Post Diversity BER
Coherent BPSK	$\left(\frac{1}{4\bar{E}_b/N_0}\right)^L \binom{2L-1}{L}$
Coherent FSK	$\left(\frac{1}{2\bar{E}_b/N_0}\right)^L \binom{2L-1}{L}$

16.5 Concluding Remarks

Diversity provides a powerful technique for combating fading in mobile communication systems. Diversity techniques seek to generate and exploit multiple branches over which the signal shows low fade correlation. To obtain the best diversity performance, the multiple access, modulation, coding, and antenna design of the wireless link must all be carefully chosen so as to provide a rich and reliable level of well-balanced, low-correlation diversity branches in the target propagation environment. Successful diversity exploitation can impact a mobile network in several ways. Reduced power requirements can result in increased coverage or improved battery life. Low signal outage improves voice quality and handoff performance. Finally, reduced fade margins directly translate to better reuse factors and, hence, increased system capacity.

Defining Terms

Automatic request for repeat: An error control mechanism in which received packets that cannot be corrected are retransmitted.

Channel coding/forward error correction: A technique that inserts redundant bits during transmission to help detect and correct bit errors during reception.

Fading: Fluctuation in the signal level due to shadowing and multipath effects.

Frequency hopping: A technique where the signal bursts are transmitted at different frequencies separated by random spacing that are multiples of signal bandwidth.

Interleaving: A form of data scrambling that spreads burst of bit errors evenly over the received data allowing efficient forward error correction.

Outage probability: The probability that the signal level falls below a specified minimum level.

PCS: Personal communications services.

RAKE receiver: A receiver used in direct sequence spread spectrum signals. The receiver extracts energy in each path and then adds them together with appropriate weighting and delay.

References

1. Adachi, F., Feeney, M.T., Williason, A.G., and Parsons, J.D., Crosscorrelation between the envelopes of 900 MHz signals received at a mobile radio base station site. *Proc. IEE,* 133(6), 506–512, 1986.
2. Freeburg, T.A., Enabling technologies for in-building network communications—four technical challenges and four solutions. *IEEE Trans. Veh. Tech.,* 29(4), 58–64, 1991.
3. Jakes, W.C., *Microwave Mobile Communications,* John Wiley & Sons, New York, 1974.
4. Jefford, P.A., Turkmani, A.M.D., Arowojulu, A.A., and Kellet, C.J., An experimental evaluation of the performance of the two branch space and polarization schemes at 1800 MHz. *IEEE Trans. Veh. Tech.,* VT-44(2), 318–326, 1995.
5. Lee, W.C.Y., *Mobile Communications Engineering,* McGraw-Hill, New York, 1982.
6. Pahlavan, K. and Levesque, A.H., *Wireless Information Networks,* John Wiley & Sons, New York, 1995.
7. Proakis, J.G., *Digital Communications,* McGraw-Hill, New York, 1989.
8. Vaughan, R.G., Polarization diversity system in mobile communications. *IEEE Trans. Veh. Tech.,* VT-39(3), 177–186, 1990.
9. Viterbi, A.J., *CDMA: Principle of Spread Spectrum Communications,* Addison-Wesley, Reading, MA, 1995.

17

Information Theory

Bixio Rimoldi
Swiss Federal Institute of Technology, Lausanne, Switzerland

Rüdiger Urbanke
Swiss Federal Institute of Technology, Lausanne, Switzerland

17.1 Introduction

The field of information theory has its origin in Claude Shannon's 1948 paper, "A mathematical theory of communication." Shannon's motivation was to study "[The problem] of reproducing at one point either exactly or approximately a message selected at another point." Whereas in this section we will be concerned only with Shannon's original problem, one should keep in mind that information theory is a growing field of research whose profound impact has reached various areas such as statistical physics, computer science, statistical inference, and probability theory. For an excellent treatment of information theory that extends beyond the area of communication we recommend Cover and Thomas [1991]. For the reader who is strictly interested in communication problems we also recommend Gallager [1968], Blahut [1987], and McEliece [1977].

17.2 The Communication Problem

A rather general diagram for a (point-to-point) communication system is shown in Fig. 17.1.

The source might be digital (e.g., a data file) or analog (e.g., a video signal). Since an analog source can be sampled without loss of information (see chapter on sampling, this volume), without loss of generality we consider only discrete-time sources and model them as discrete-time stochastic processes.

The channel could be a pair of wires, an optical fiber, a radio link, etc. The channel model specifies the set of possible channel inputs and, for each input, specifies the output process. Most real-world channels are waveform channels, meaning that the input and output sets are sets of waveforms. It is often the case that the communication engineer is given a waveform channel with a modulator and a demodulator. In a satellite communication system, the modulator might be a phase-shift keying modulator whose input alphabet V is the set $\{0, 1\}$; the channel might be modeled by the additive white Gaussian noise channel, and the demodulator attempts to output a guess (perhaps incorrect) of the modulator input. In this case, the modulator output alphabet W equals V. In more sophisticated cases, where $|W| > |V|$ and $|\cdot|$ denotes the number of elements in the enclosed set, the digital data demodulator can also furnish information about the reliability of the decision. Either way, one would consider the modulator and the demodulator as part of the channel. If the statistics of the channel output at a given time depend

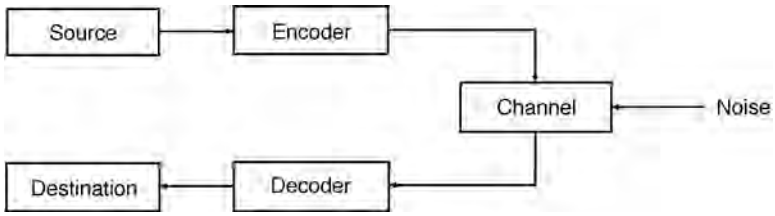


FIGURE 17.1 Block diagram of a communications system.

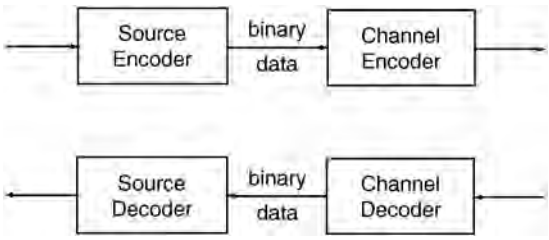


FIGURE 17.2 Block diagram of an encoder and a decoder, each split into two parts.

only on the value of the corresponding position of the input sequence, then the channel is a discrete memoryless channel. This is an important class of channel models that will receive particular attention in this chapter. Even if the modulator and the demodulator are not given, assuming that the bandwidth of the waveform channel is limited, one can always use the sampling theorem to convert a waveform channel into a channel with discrete-time input and output.

The destination is merely a place holder to remind us that the user has some expectation concerning the quality of the reproduced signal. If the source output symbols are elements of a finite set, then the destination may specify a maximum value for the probability of error. Otherwise, it may specify a maximum distortion computed according to some specified criterion (e.g., mean square error).

It is generally assumed that the source, the channel, and the destination are given and fixed. On the other hand, the communication engineer usually is completely free to design the encoder and the decoder to meet the desired performance specifications.

An important result of information theory is that, without loss of optimality, the encoder in Fig. 17.1 can be decomposed into two parts: a source encoder that produces a sequence of binary data and a channel encoder, as shown in Fig. 17.2. Similarly, the decoder may be split into a channel decoder and a source decoder.

The objective of the source encoder is to compress the source, that is, to minimize the average number of bits necessary to represent a source symbol. The most important questions concerning the source encoder are the following: What is the minimum number of bits required (on average) to represent a source output symbol? How do we design a source encoder that achieves this minimum? These questions will be considered here. It turns out that the output of an ideal source encoder is a sequence of independent and uniformly distributed binary symbols. The purpose of the channel encoder/decoder pair is to create a reliable bit pipe for the binary sequence produced by the source encoder. Here the most important question is whether or not it is possible to design such an encoder/decoder pair. This question will also be considered in some detail. The fact that we can split the encoder into two parts as described is important. First of all, it allows us to study the source and the channel separately and to assume that they are connected by a binary interface. Second, it tells us that for a given channel we may design the channel encoder and the channel decoder to maximize the rate of the resulting (virtually error-free) bit pipe without having to know the nature of the source that will use it.

17.3 Source Coding for Discrete-Alphabet Sources

We start by considering the simplest possible source, namely, a discrete-time discrete-alphabet memoryless information source modeled by a random variable X taking values on a finite alphabet X. A (binary source) **code C** for X is a mapping from X into the set of finite length binary sequences called codewords. Such a code can be represented by a binary tree, as shown in the following example.

Example 1

Let X take on values in {1, 2, 3, 4, 5, 6, 7}. A possible binary code for X is given by the (ordered) set of binary sequences {1, 010, 011, 0000, 0001, 0010, 0011}. The corresponding tree is shown in Fig. 17.3. Notice that the codeword corresponding to a given source output symbol is the label sequence from the root to the node corresponding to that symbol.

In Example 1, source output symbols correspond to leaves in the tree. Such a code is called *prefix free* since no codeword is the prefix of another codeword. Given a concatenation of codewords of a prefix-free code, we can parse it in an unique way into codewords. Codes with this property are called *uniquely decodable*. Although there are uniquely decodable codes which are *not* prefix free, we will restrict our attention to prefix-free codes, as it can be shown that the performance of general uniquely decodable codes is no better than that of prefix-free codes (see Cover and Thomas [1991]).

For each i in X, let $C(i)$ be the codeword associated to the symbol i, let l_i be its length, $L \int \max_i l_i$, and let p_i be the probability that $X = i$. A complete binary tree of depth L has 2^L leaves. To each leaf at depth l, l £ L, of the binary code tree (which is not necessarily complete) there correspond 2^{L-l} leaves at depth L of the corresponding complete tree (obtained by extending the binary code tree). Further, any two distinct leaves in the code tree have distinct associated leaves at depth L in the complete binary tree. Hence, summing up over all leaves of the code tree (all codewords) we get

$$2^L \geq \sum_i 2^{L-l_i} \quad \sum_i 2^{-l_i} \leq 1$$

The right-hand side is called the *Kraft inequality* and is a necessary condition on the codeword lengths of any prefix-free code. Conversely, for any set of codeword lengths satisfying the displayed inequality, we can construct a prefix-free code having codeword lengths l_i (start with a complete binary tree of sufficient length; to each l_i associate a node in this tree at level l_i and make this node a leaf by deleting all of its descendants). The problem of finding the best source code then reduces to that of finding a set of lengths that satisfies the Kraft inequality and minimizes the average codeword lengths $L^* = \hat{A}_i l_i p_i$. Such a code is called *optimal*.

To any source X with probabilities p_i, we associate the quantity $H(X)$, which we call **entropy** and which is defined as

$$H(X) = -\sum_i p_i \log p_i$$

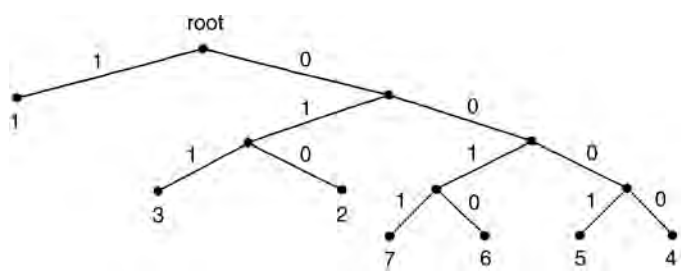


FIGURE 17.3 Example of a code tree.

Usually we take the base of the logarithm to be 2, in which case the units of entropy are called *bits*.

It is straightforward to show by means of Lagrange multipliers that if we neglect the integer constraint on l_i, then the minimization of L^* subject to the Kraft inequality yields $l_i^* = -\log p_i$. This noninteger choice of codeword lengths yields

$$L^* = \sum_i p_i l_i^* = -\sum_i p_i \log p_i = H(X)$$

On the other hand, the (not necessarily optimal) integer choice $l_i = \lceil -\log p_i \rceil$ satisfies the Kraft inequality and yields an expected codeword length

$$L(C) = \sum_i p_i \lceil -\log p_i \rceil \le -\sum_i p_i \log p_i + \sum_i p_i = H(X) + 1$$

Hence, we have the following theorem.

Theorem 1. The average length L^* of the optimal prefix-free code for the random variable X satisfies $H(X) \le L^* \le H(X) + 1$.

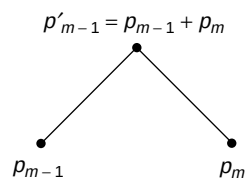
Example 2

Let X be a random variable that takes on the values 1 and 0 with probabilities $\theta (0 \le \theta \le 1)$ and $1 - \theta$, respectively. Then $H(X) = h(\theta)$, where $h(\theta) = -\theta \log \theta - (1 - \theta)\log(1 - \theta)$. The function $h(\theta)$ is called the *binary entropy function*. In particular, it can be shown that $h(\theta) \le 1$ with equality if and only if $q = \frac{1}{2}$. More generally, if X takes on $|X|$ values, then $H(X) \le \log |X|$ with equality if and only if X is uniformly distributed.

A simple algorithm that generates optimal prefix-free codes is the Huffman algorithm, which can be described as follows:

Step 1. Re-index to arrange the probabilities in decreasing order, $p_1 \ge p_2 \ge \cdots \ge p_m$.

Step 2. Form a subtree by combining the last two probabilities p_{m-1} and p_m into a single node of weight $p'_m = p_{m-1} + p_m$.



Step 3. Recursively execute steps 1 and 2, decreasing the number of nodes each time, until a single node is obtained.

Stpe 4. Use the tree constructed above to assign codewords.

Example 3

As in Example 1, let X take on the values {1, 2, 3, 4, 5, 6, 7}. Assume their corresponding probabilities are {0.4, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1}. Figure 17.4 shows the steps in the construction of the **Huffman code**. For simplicity, probabilities are not shown. The final tree is given in Example 1.

Given a block of n random variables $X_1,\ldots, X_n$, with joint distribution $p(X_1,\ldots, X_n)$, define their entropy as

$$H(X_1,\ldots,X_n) = -\sum_{x_1}\sum_{x_2}\cdots\sum_{x_n} p(x_1,\ldots,x_n)\log p(x_1,\ldots,x_n)$$

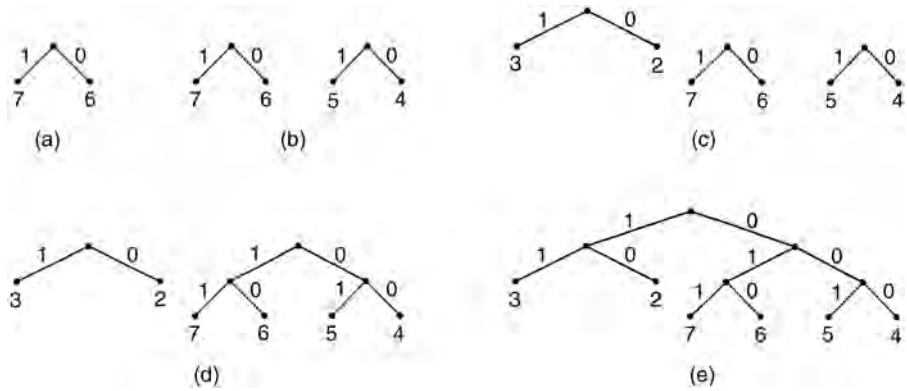


FIGURE 17.4 Construction of a Huffman code of X as given in Example 1.

Note that for the special case of n independent random variables, we have $H(X_1,\ldots,X_n) = nH(X)$. This agrees with the intuitive notion that if it takes $H(X)$ bits to encode one source output symbol, then it should take $nH(X)$ bits to encode n independent output symbols. For a sequence of random variables $X_1, X_2,\ldots, X_n$, we define $H_n = \frac{1}{n}H(X_1,\ldots,X_n)$.If $X_1, X_2,\ldots,X_n$ is a stationary random process, as n goes to infinity H_n converges to a limit $H_\bullet$ called the *entropy rate* of the process.

Applying Theorem 1 to blocks of random variables, we obtain

$$H_n \le \frac{L}{n} < H_n + \frac{1}{n}$$

In the special case that the process $X_1, X_2,\ldots, X_n$ is stationary, the left and the right side will both tend to $H_\bullet$ as $n \to \infty$. This shows that by encoding a sufficiently large number of source output symbols at a time, one can get arbitrarily close to the fundamental limit of H_∞ bits per symbol.

17.4 Universal Source Coding

The Huffman algorithm produces optimal codes, but it requires knowing the statistics of the source. It is a surprising fact that there are source coding algorithms that are asymptotically optimal without requiring prior knowledge of the source statistics. Such algorithms are called *universal* source coding algorithms. The best known universal source coding algorithm is the **Lempel–Ziv** algorithm, which has been implemented in various forms on computers (as PKZIP command for Windows and Linux-based systems). There are two versions of the Lempel–Ziv algorithm: the LZ77, see Ziv and Lempel [1977], and the LZ78, see Lempel and Ziv [1978]. We will describe the basic idea of the LZ77.

Consider the sequence $X_1, X_2,\ldots, X_k$ taking values in $\{a, b, c\}$, shown in Fig. 17.5, and assume that the first w letters ($w = 6$ in our example) are passed to the decoder with no attempt at compression. At this point the encoder identifies the longest string $X_{w+1}, X_{w+2},\ldots,X_{w+L}$ such that a copy, $X_{w-p+1},\ldots,X_{w-p+L}$, begins (but not necessarily ends) in the portion of data already available to the decoder. In our example this string is *aabaa*, $L = 5$, and $p = 3$; see Fig. 17.5. Next, the encoder transmits a binary representation of the pair (L, p) as this is sufficient for the decoder to reconstruct $X_{w+1},\ldots,X_{w+L}$. At this point, the procedure can be repeated with w replaced by $w + L$. There are two exceptions: the first occurs when $X_{w+\ell}$ is a new letter that does not appear in $X_l,\ldots, X_w$; the second exception occurs when it is more efficient to encode $X_{w+1},\ldots,X_{w+L}$ directly than to encode (L, p). In order to handle such special cases, the algorithm has the option of encoding a substring directly, without attempt at compression. It is surprising that the Lempel–Ziv algorithm is asymptotically optimal for all finite alphabet stationary ergodic sources. This has been shown in Wyner and Ziv [1994].

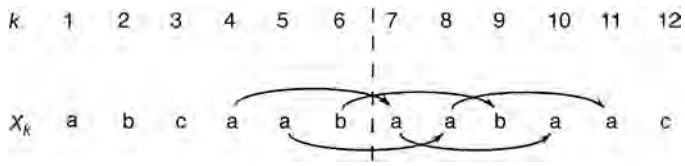


FIGURE 17.5 Encoding procedure of the Lempel–Ziv algorithm.

17.5 Rate Distortion Theory

Thus far we have assumed that X is a discrete alphabet and that the decoder must be able to perfectly reconstruct the source output. Now we drop both assumptions.

Let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables. Let $\mathcal{X}$ denote the source alphabet and $\mathcal{Y}$ be a suitably chosen representation alphabet (where $\mathcal{Y}$ is not necessarily equal to $\mathcal{X}$). A *distortion measure* is a function $d: \mathcal{X} \times \mathcal{Y} \to \mathrm{R}^+$. The most common distortion measures are the *Hamming distortion* d_H and the *squared error* distortion d_E defined as follows:

$$d_H(x, y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{if } x \neq y \end{cases}$$

$$d_E(x, y) = (x - y)^2$$

Such a single-letter distortion measure can be extended to a distortion measure on n-tuples by defining

$$d(x^n, y^n) = \frac{1}{n} \sum_{i=1}^{n} d(x_i, y_i)$$

The *encoder* maps an n-length source output X^n to an index $U \in \{1, 2, \ldots, 2^{\lfloor nR \rfloor}\}$, where R denotes the number of bits per source symbol that we are allowed or willing to use. The *decoder* maps U into an n-tuple Y^n, called the representation of X^n. We will call U together with the associated encoding and decoding function a *rate distortion code*. Let f be the function that maps X^n to its representation Y^n, that is, f is the mapping describing the concatenation of the encoding and the decoding function. The expected disortion D is then given by

$$D = \sum_{x^n} p(x^n) d(x^n, f(x^n))$$

The objective is to minimize the average number of bits per symbol, denoted by R, for a given average distortion D. What is the minimum R?

Definition. The rate distortion pair (R, D) is said to be *achievable* if there exists a rate distortion code of rate R with expected distortion D. The **rate distortion function** $R(D)$ is the infimum of rates R such that (R, D) is achievable for a given D.

Entropy played a key role in describing the limits of lossless coding. When distortion is allowed, a similar role is played by a quantity called mutual information.

Definition. Given the random variables X and Y with their respective distributions $p(x)$, $p(y)$, and $p(x, y)$, the **mutual information** $I(X; Y)$ between X and Y is defined as

$$I(X; Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)}$$

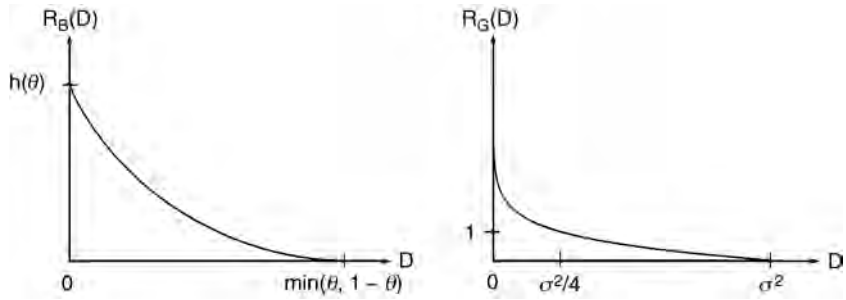


FIGURE 17.6 Examples of rate distortion functions.

This definition can be extended to blocks of random variables X^n and Y^n by replacing $p(x)$, $p(y)$, and $p(x, y)$ with their higher-dimensional counterparts. We can now state the fundamental theorem of rate distortion theory. For its proof see Cover and Thomas [1991].

Theorem 2. The rate distortion function for an independent identically distributed source X with distribution $p(x)$ and distortion function $d(x, y)$ is

$$R(D) = \min_{p(y|x):\Sigma_{x,y} p(x)p(y|x)d(x,y) \le D} I(X; Y)$$

The rate distortion functions for a *binary source* that outputs 1 with probability θ and Hamming distortion and for a *Gaussian source* of variance σ^2 and squared error distortion are as follows:

$$R_B(D) = \begin{cases} h(\theta) - h(D), & 0 \le D \le \min\{\theta, 1-\theta\} \\ 0, & D > \min\{\theta, 1-\theta\} \end{cases}$$

$$R_G(D) = \begin{cases} \frac{1}{2}\log\frac{\sigma^2}{D}, & 0 \le D \le \sigma^2 \\ 0, & D > \sigma^2 \end{cases}$$

These two functions are plotted in Fig. 17.6.

Some insight can be gained by considering the extreme values of both $R_B(D)$ and $R_G(D)$. Assume that $X_1, X_2, \ldots$ is a sequence of independent and identically distributed binary random variables and let θ be the probability that $X_i = 1$. Without loss of generality, we may assume that $\theta \ge 1 - \theta$. If we let $Y^n = (0, \ldots, 0)$ regardless of X^n, then the expected Hamming distortion equals θ. Hence, it must be true that $R_B(D) = 0$ for $D \ge \theta$. On the other hand, $D = 0$ means that the reproduction Y_n must be a perfect copy of X^n. From Theorem 1 and Example 2 we know that this means $R_B(0) = H(X) = h(q)$. Similar considerations hold for $R_G(D)$. It is interesting to observe that $R_G(D)$ tells us that if we are allowed to use R bits per symbol to describe a sequence of independent Guassian random variables with variance σ^2, then we need to accept a mean squared error of $\sigma^2 2^{-2R}$.

17.6 Channel Coding

Now we consider the fundamental limit for the amount of information that can be transmitted through a noisy channel and discuss practical methods to approach it. For simplicity, we assume that the channel can be modeled as a discrete memoryless channel, defined as follows:

A **discrete memoryless channel (DMC)** is a system consisting of an input alphabet $\mathcal{X}$, an output alphabet $\mathcal{Y}$, and a collection of probability mass functions $p(y|x)$, one for each $x \in \mathcal{X}$. A specific DMC will be denoted by $(\mathcal{X}, p(y|x), \mathcal{Y})$.

An **(M, n) channel code** for the DMC $(\mathrm{X}, p(y|x), \mathrm{Y})$ consists of the following:

1. an index set $\{1, 2, \ldots, M\}$
2. an encoding function X^n: $\{1, 2, \ldots, M\} \to \mathrm{X}^n$, yielding codewords $X^n(1), X^n(2), \ldots, X^n(M)$ (the set of codewords is called the *code book*)
3. a decoding function g: $\mathrm{Y}^n \to \{1, 2, \ldots, M\}$, which is a deterministic rule that assigns a guess to each possible received vector

The most important figures of merit for a channel code for a DMC are the *maximal probability of error*, defined by

$$\lambda^{(n)} \equiv \max_{i \in \{1, 2, \ldots, M\}} Pr\{g(Y^n) \neq i | X^n = X^n(i)\}$$

and the *transmission rate*,

$$R = \frac{\log M}{n} \quad \text{bits/channel use}$$

It makes sense to define the rate in this way since the M codewords can be labeled with $\log M$ bits. Hence, every time we transmit a codeword, we actually transmit $\log M$ bits of information. Since it takes n uses of the channel to transmit one codeword, the resulting information rate is $(\log M)/n$ bits per channel use.

For a given DMC, a rate R is said to be *achievable* if, for any desired probability of error P_e and sufficiently large block length n, there exists a $(\lceil 2^{nR} \rceil, n)$ code for that DMC with maximal probability of error $\lambda^{(n)} \leq P_e$.

The *operational capacity* of a discrete memoryless channel is the supremum of all achievable rates. One would expect that determining the operational capacity would be a formidable task. One of the most remarkable results of information theory is a relatively easy-to-compute way to determine the operational capacity.

We define the **information channel capacity** of a discrete memoryless channel $(\mathcal{X}, p(y|x), \mathcal{Y})$ as

$$C \equiv \max_{p(x)} I(X; Y)$$

where the maximum is taken over all possible input distributions $p(x)$. As we will show in the next example, it is straightforward to compute the information channel capacity for many discrete memoryless channels of interest. Shannon's channel coding theorem establishes that the information channel capacity equals the operational channel capacity.

Theorem (Shannon's channel coding theorem). For a DMC, the *information* channel capacity equals the operational channel capacity, that is, for every rate $R < C$, there exists a sequence of $(\lceil 2^{nR} \rceil, n)$ codes for the DMC with maximum probability of error $\lambda^{(n)} \to 0$. Conversely, for any sequence of $(\lceil 2^{nR} \rceil, n)$ for codes this DMC with the property that $\lambda^{(n)} \to 0$, the rate R must satisfy $R \leq C$.

This is perhaps the most important result of information theory. In order to shed some light on the interpretation of mutual information and to provide an efficient way to determine C, it is convenient to rewrite $I(X; Y)$ in terms of conditional entropy. Given two random variables X and Y with joint probability mass function $p(x, y)$, marginal probability mass function $p(x) = \sum_y p(x, y)$, and conditional probability mass function $p(x|y) = p(x, y)/p(y)$, we define the conditional entropy of X given that $Y = y$ as

$$H(X \mid Y = y) = -\sum_x p(x|y) \log p(x|y)$$

Its average is the conditional entropy of X given Y, denoted by $H(X|Y)$ and defined as

$$\begin{aligned} H(X\,|\,Y) &= \sum_y p(y)H(X|Y=y) \\ &= -\sum_{x,y} p(x,y)\log p(x|y) \\ &= E[-\log p(X|Y)] \end{aligned}$$

Recall that for the discrete variable X with probability mass function $p(x)$, $H(X)$ is the average number of binary symbols necessary to describe X. Now let X be the random variable at the input of a DMC and let Y be the output. The knowledge that $Y = y$ changes the probability mass function of the channel input from $p(x)$ to $p(x|y)$ and its entropy from $H(X)$ to $H(X|Y=y)$. According to our previous result, $H(X\,|\,Y=y)$ is the average number of bits necessary to describe X after the observation that $Y = y$. Since $H(X)$ is the average number of bits needed to describe X (without knowledge of Y), $H(X) - H(X|Y=y)$ is the average amount of information about X acquired by the observation that $Y = y$, and

$$H(X) - H(X\,|\,Y)$$

is the average amount of information about X acquired by the observation of Y. One can easily verify that the latter expression is another way to write $I(X; Y)$. Moreover,

$$\begin{aligned} I(X;\,Y) &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X) \\ &= I(Y;\,X) \end{aligned}$$

Hence, the amount of information that Y gives about X is the same as the amount of information that X gives about Y. For this reason, $I(X; Y)$ is called the mutual information between X and Y.

Example (Binary Symmetric Channel). Let $\mathbb{X} = \{0, 1\}$, $\mathbb{Y} = \{0, 1\}$, and $p(y|x) = p$ when $x \neq y$ and $1 - p$ otherwise. This DMC can be conveniently depicted as in Fig. 17.7. For obvious reasons it is called the *binary symmetric channel* (BSC).

For the BSC, we bound the mutual information by

$$\begin{aligned} I(X;\,Y) &= H(Y) - H(Y\,|\,X) \\ &= H(Y) - \sum_x p(x)H(Y\,|\,X=x) \\ &= H(Y) - h(p) \\ &\leq 1 - h(p) \end{aligned}$$

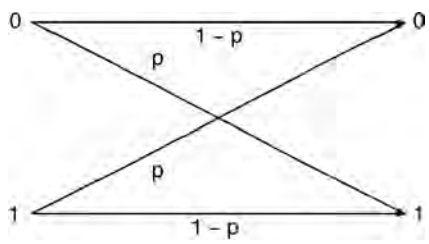


FIGURE 17.7 Discrete memoryless channel with crossover probability p.

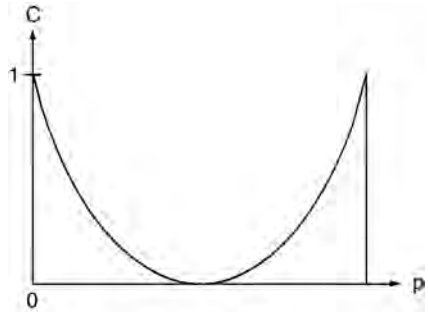


FIGURE 17.8 Capacity of the BSC as a function of p.

where the last inequality follows because Y is a binary random variable (see Example 2). Equality is achieved when the input distribution is uniform since in this case the output distribution is also uniform. Hence, the information capacity of a binary symmetric channel with parameter p is

$$C = 1 - h(p) \qquad \text{[bits/channel use]}$$

This function is plotted in Fig. 17.8.

17.7 Simple Binary Codes

Shannon's channel coding theorem promises the existence of block codes that allow us to transmit reliably through an unreliable channel at any rate below the channel capacity. Unfortunately, the codes constructed in all known proofs of this fundamental result are impractical in that they do not exhibit any structure which can be employed for an efficient implementation of the encoder and the decoder. This is an important issue, as good codes must have long block lengths and, hence, a large number of codewords. For these codes, simple table lookup encoding and decoding procedures are not feasible. The search for good and practical codes has led to the development of coding theory. Although powerful block codes are, in general, described by sophisticated techniques, the main idea is simple. We collect a number k of information symbols which we wish to transmit, we append r check symbols, and transmit the entire block of $n = k + r$ channel symbols. Assuming that the channel changes a sufficiently small number of symbols within an n-length block, the r check symbols may provide the receiver with sufficient information to detect and/or correct the errors.

Traditionally, good codes have been constructed by sophisticated algebraic schemes. More recent code construction techniques are inspired by Shannon's original work. The idea is to construct the code by making random choices on a suitably defined structure. A key property of the structure is that it leads to codes that may be decoded using powerful iterative techniques. We will demonstrate these two quite distinct concepts by simple examples.

Hamming Codes

For any positive integer m, there is a Hamming code with parameters $k = 2^m - m - 1$ and $n = 2^m - 1$. To illustrate, we consider the $k = 4$, $n = 7$ Hamming code. All operations will be done modulo 2. Consider the set of all nonzero binary vectors of length $m = 3$. Arrange them in columns to form the matrix

$$H = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

TABLE 17.1 Codewords of the (7, 4) Hamming Code

0	0	0	0	0	0	0	1	0	0	0	1	0	1
0	0	0	1	0	1	1	1	0	0	1	1	1	0
0	0	1	0	1	1	0	1	0	1	0	0	1	1
0	0	1	1	1	0	1	1	0	1	1	0	0	0
0	1	0	0	1	1	1	1	1	0	0	0	1	0
0	1	0	1	1	0	0	1	1	0	1	0	0	1
0	1	1	0	0	0	1	1	1	1	0	1	0	0
0	1	1	1	0	1	0	1	1	1	1	1	1	1

The row space of H is a vector space of binary 7-tuples of dimension $m = 3$. The $k = 4$, $n = 7$ Hamming code is the null space of H, that is, the vector space of dimension $n - m = 4$ consisting of all binary 7-tuples c such that $cH^T = 0$. Hence, it contains the $M = 2^k = 16$ codewords shown in Table 17.1 and its rate is $R = (\log M)/n = 4/7$. If c is transmitted and u is received, then the error is the unique binary n-tuple e such that $u = c + e$, where addition is component-wise modulo 2. The decoder performs the operation

$$uH^T = (c+e)H^T = cH^T + eH^T = eH^T$$

If there is a single error at position i, that is, if e is zero except for the ith component (which is one), then uH^T is the ith row of H^T. This tells us the error location. If e contains 2 ones (channel errors), then uH^T is the sum of two rows of H^T. Since this cannot be zero, two channel errors are detectable. They are not correctable, however, since, as one can easily verify when two channel errors occur, there is always a codeword that agrees with the received word in $n - 1$ positions. This is the codeword that would be selected by the decoder if it were forced to make a decision since the transmitted codeword agrees with the received word in only $n - 2$ positions. A simple encoding procedure can be specified as follows: let G be a $k \times n$ matrix whose rows span the null space of H. Such a matrix is

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Then an information k-tuple a can be mapped into a codeword $c = aG$ by a simple matrix multiplication.

Low-Density Parity-Check Codes

Low-density parity-check (LDPC) codes constitute a family of a new breed of coding techniques variably described as "codes on graphs" or "iterative coding techniques." They were originally introduced by Gallager in his thesis in 1962 (see Gallager [1963]), then long forgotten, and later revived in the wake of the discovery of turbo codes by Berrou et. al. (see Berrou et al. [1993] and MacKay [1999]).

LDPC codes are codes which possess a sparse parity check matrix. Because of this sparseness, they can be decoded very efficiently in an iterative fashion. Although this iterative decoding procedure is, in general, suboptimal, it is sufficiently strong to allow transmission seemingly arbitrarily close to channel capacity.

Consider the code represented by the following parity check matrix.

$$H = \begin{pmatrix} 1110111000 \\ 0011111100 \\ 0110100111 \\ 1001010111 \\ 1101001011 \end{pmatrix}$$

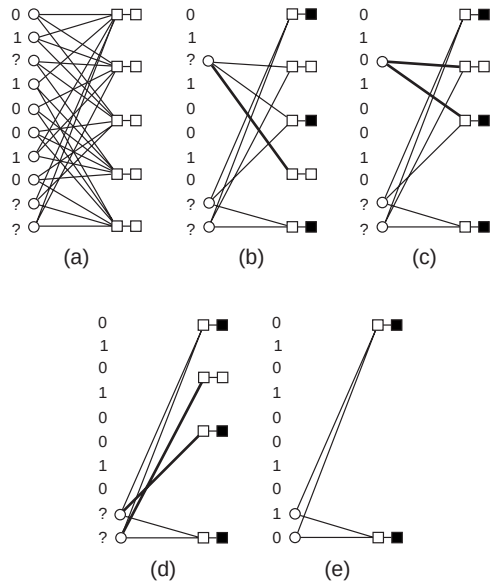


FIGURE 17.9 Iterative decoding.

Note that each row contains exactly 6 ones and that each column contains exactly 3 ones. Such a code is referred to as a (3, 6)-regular LDPC code. A convenient graphical representation of this code is given by means of the bipartite graph shown in Fig. 17.9(a). This graph is usually referred to as Tanner graph. In this graph, each left node (variable node) corresponds to one column of the parity check matrix and each right node (check node) corresponds to one row of H. A left node is connected to a right node if and only if the corresponding position in the parity check matrix contains a 1.

Let $(x = x_0, \ldots, x_9)$ be the codeword (0101001010). Assume that we transmit x over a binary erasure channel (BEC) and that the received word is $y = (??01001?10)$ where ? indicates an erasure. Rather than employing a maximum likelihood decoder, we will use the iterative decoding algorithm described in Figs. 17.9(a)–(e). The received word y is written on the left (y_0 on the bottom and y_a on the top). We start by passing all known values from the left to the right along the corresponding edges. We next accumulate at each check the partial sum (mod 2) of all known values and delete all edges along which messages have already been passed. The result is shown in Fig. 17.9(b), where a black box indicates a partial sum equal to 1. Now note that check node one (the second from the bottom) has received already all values except the one from the erased bit seven. Since, by the code constraints, the sum of all values must be equal to zero, we see that we can deduce the value of variable node seven to be equal to zero. This value can now again be propagated to the right and all involved edges be deleted. The result is shown in Fig. 17.9 (d). After one more iteration all erased values are known. (Fig. 17.9(e)).

As we can see, in the above example the given iterative decoding strategy successfully recovered the erased bits. This might seem like a lucky coincidence. Nevertheless, it has been shown that long and suitably defined LDPC codes can achieve capacity on the BEC under iterative decoding (see Luby et. al. [1997]).

Even more, the above iterative decoding strategy, suitably extended, can be used to transmit reliably close to capacity over a wide array of channels (see Richardson et. al. [2001]).

Defining Terms

Binary source code: A mapping from a set of messages into binary strings.
Channel capacity: The highest rate at which information can be transmitted reliably across a channel.
Discrete memoryless channel: A channel model characterized by discrete input and output alphabets and a probability mass function on the output conditioned on the input.

Entropy: A measure of the average uncertainty of a random variable. For a random variable with distribution $p(x)$, the entropy $H(X)$ is defined as $-\sum_x p(x)\log p(x)$.

Huffman coding: A procedure that constructs a code of minimum average length for a random variable.

Lempel-Ziv coding: A procedure for coding that does not use the probability distribution of the source but nevertheless is asymptotically optimal.

Mutual information: A measure for the amount of information that a random variable gives about another.

Rate distortion function: The minimum rate at which a source can be described to the given average distortion.

References

Berrou, C., Glavieux, A., and Thitimajshima, P. 1993. Near Shannon limit error-correcting coding and decoding, in *Proceedings of ICC'93,* Geneva, Switzerland, 1064–1070.

Blahut, R. 1987, *Principles and Practice of information Theory,* Addison–Wesley, Reading, MA.

Csiszár, I. and Körner, J. 1981. *Information Theory: Coding Theorems for Discrete Memoryless Systems,* Academic Press, New York.

Cover, T.M. and Thomas, J.A. 1991. *Elements of Information Theory,* Wiley, New York.

Gallager, R.G. 1963. *Low-Density Parity-Check Codes,* Monograph, M.I.T. Press, Cambridge, MA.

Gallager, R.G. 1968. *Information Theory and Reliable Communication,* Wiley, New York.

Luby, M., Mitzenmacher, M., Shokrollahi, A., Spielman, D., and Stemann, V. 1997. Practical loss-resilient codes, in *Proceedings of the 29th Annual ACM Symposium on the Theory of Computing,* 150–159.

McEliece, R.J. 1977. *The Theory of Information and Coding*, Addison–Wesley, Reading, MA.

MacKay, D. 1999. Good error correcting codes based on very sparse matrices, *IEEE Trans. Inf. Theory,* 45:399–431.

Richardson, T., Shokrollahi, A., and Urbanke, R. 2001. Design of probably good low-density parity check codes, *IEEE Trans. Inf. Theory,* 47:619–637.

Wyner, A.D. and Ziv, J. 1994. The sliding-window Lempel–Ziv algorithm is asymptotically optimal. *Communications and Cryptography*, Blahut, R.E., Costello, D.J., Jr., Maurer, V., and Mittclholzer, T., Eds., Kluwer Academic Publishers, Boston, MA.

Ziv, J. and Lempel, A. 1977. A universal algorithm for sequential data compression. *IEEE Trans. Inf. Theory*, IT-23 (May):337–343.

Ziv, J. and Lempel, A. 1978. Compression of individual sequences by variable rate coding. *IEEE Trans. Inf. Theory,* IT-24 (Sept.):530–536.

Further Information

For a lucid and up-to-date treatment of information theory extending beyond the area of communication, a most recommended reading is Cover and Thomas [1991]. For readers interested in continuous-time channels and sources we recommend Gallager [1968]. For mathematically inclined readers who do not require much physical motivation, strong results for discrete memoryless channels and sources may be found in Csiszár and Körner [1981]. Other excellent readings are Blahut [1987] and McEliece [1977]. Most results in information theory are published in *IEEE Transactions on Information Theory.*

18

Digital Communication System Performance*[1]

Bernard Sklar
Communications Engineering Services

18.1 Introduction

In this section we examine some fundamental tradeoffs among bandwidth, power, and error performance of digital communication systems. The criteria for choosing modulation and coding schemes, based on whether a system is bandwidth limited or power limited, are reviewed for several system examples. Emphasis is placed on the subtle but straightforward relationships we encounter when transforming from data-bits to channel-bits to symbols to chips.

The design or definition of any digital communication system begins with a description of the communication link. The *link* is the name given to the communication transmission path from the modulator and transmitter, through the channel, and up to and including the receiver and demodulator. The *channel* is the name given to the propagating medium between the transmitter and receiver. A link description quantifies the average signal power that is received, the available bandwidth, the noise statistics, and other

[1*] A version of this chapter has appeared as a paper in the *IEEE Communications Magazine*, November 1993, under the title "Defining, Designing, and Evaluating Digital Communication Systems."

impairments, such as fading. Also needed to define the system are basic requirements, such as the data rate to be supported and the error performance.

The Channel

For radio communications, the concept of *free space* assumes a channel region free of all objects that might affect radio frequency (RF) propagation by absorption, reflection, or refraction. It further assumes that the atmosphere in the channel is perfectly uniform and nonabsorbing and that the earth is infinitely far away or its reflection coefficient is negligible. The RF energy arriving at the receiver is assumed to be a function of distance from the transmitter (simply following the inverse-square law as used in optics). In practice, of course, propagation in the atmosphere and near the ground results in refraction, reflection, and absorption, which modify the free space transmission.

The Link

A radio transmitter is characterized by its average output signal power P_t and the gain of its transmitting antenna G_t. The name given to the product P_tG_t, with reference to an isotropic antenna, is *effective radiated power* (*EIRP*) in watts (or dBW). The average signal power S arriving at the output of the receiver antenna can be described as a function of the *EIRP*, the gain of the receiving antenna G_r, the path loss (or space loss) L_s, and other losses, L_o, as follows [1,2]:

$$S = \frac{EIRP\ G_r}{L_s L_o} \tag{18.1}$$

The path loss L_s can be written as follows [2]:

$$L_s = \left(\frac{4\pi d}{\lambda}\right)^2 \tag{18.2}$$

where d is the distance between the transmitter and receiver and λ is the wavelength.

We restrict our discussion to those links distorted by the mechanism of additive white Gaussian noise (AWGN) only. Such a noise assumption is a very useful model for a large class of communication systems. A valid approximation for average received noise power N that this model introduces is written as follows [3,4]:

$$N \cong kT^\circ W \tag{18.3}$$

where k is Boltzmann's constant (1.38×10^{-23} joule/K), T° is system effective temperature in kelvin, and W is bandwidth in hertz. Dividing Eq. (18.3) by bandwidth enables us to write the received noise-power spectral density N_0 as follows:

$$N_0 = \frac{N}{W} = kT^\circ \tag{18.4}$$

Dividing Eq. (18.1) by N_0 yields the received average signal-power to noise-power spectral density S/N_0 as

$$\frac{S}{N_0} = \frac{EIRP\ G_r/T^\circ}{kL_s L_o} \tag{18.5}$$

where G_r/T° is often referred to as the receiver figure of merit. A link budget analysis is a compilation of the power gains and losses throughout the link; it is generally computed in decibels and thus takes on the bookkeeping appearance of a business enterprise, highlighting the assets and liabilities of the link. Once the value of S/N_0 is specified or calculated from the link parameters, we then shift our attention to optimizing the choice of signalling types for meeting system bandwidth and error performance requirements.

In Eq. (18.4), N_0 characterizes all the noise in the receiving system since T^o is the system effective temperature; thus, S/N_0 is referenced to the predetection point in the receiver [2]. Then we can write the ratio of bit energy-to-$N_0(E_b/N_0)$ at this point, for any data rate R, as

$$\frac{E_b}{N_0} = \frac{ST_b}{N_0} = \frac{S}{N_0}\left(\frac{1}{R}\right) \tag{18.6}$$

Equation(18.6) follows from the basic definitions that received bit energy is equal to received average signal power times the bit duration and that bit rate is the reciprocal of bit duration. Received E_b/N_0 is a key parameter in defining a digital communication system. Its value indicates the apportionment of the received waveform energy among the bits that the waveform represents. At first glance, one might think that a system specification should entail the symbol-energy to noise-power spectral density E_s/N_0 associated with the arriving waveforms. We will show, however, that for a given S/N_0, the value of E_s/N_0 is a function of the modulation and coding. The reason for defining systems in terms of E_b/N_0 stems from the fact that E_b/N_0 depends only on S/N_0 and R and is unaffected by any system design choices, such as modulation and coding.

18.2 Bandwidth and Power Considerations

Two primary communications resources are the received power and the available transmission bandwidth. In many communication systems, one of these resources may be more precious than the other and, hence, most systems can be classified as either bandwidth limited or power limited. In bandwidth-limited systems, spectrally efficient modulation techniques can be used to save bandwidth at the expense of power; in power-limited systems, power efficient modulation techniques can be used to save power at the expense of bandwidth. In both bandwidth- and power-limited systems, error-correction coding (often called channel coding) can be used to save power or to improve error performance at the expense of bandwidth. Recently, trellis-coded modulation (TCM) schemes have been used to improve the error performance of bandwidth-limited channels without any increase in bandwidth [5], but these methods are beyond the scope of this chapter.

The Bandwidth Efficiency Plane

Figure 18.1 shows the abscissa as the ratio of bit-energy to noise-power spectral density E_b/N_0 (in decibels) and the ordinate as the ratio of throughput, R (in bits per second), that can be transmitted per hertz in a given bandwidth W. The ratio R/W is called bandwidth efficiency since it reflects how efficiently the bandwidth resource is utilized. The plot stems from the Shannon–Hartley capacity theorem [2,6,7], which can be stated as

$$C = W\log_2\left(1 + \frac{S}{N}\right) \tag{18.7}$$

where S/N is the ratio of received average signal power to noise power. When the logarithm is taken to the base 2, the capacity C is given in bits per second. The capacity of a channel defines the maximum number of bits that can be reliably sent per second over the channel. For the case where the data (information) rate R is equal to C, the curve separates a region of practical communication systems from a region where such communication systems cannot operate reliably [2,6].

M-ary Signalling

Each symbol in an M-ary alphabet can be related to a unique sequence of m bits, expressed as

$$M = 2^m \qquad \text{or} \qquad m = \log_2 M \tag{18.8}$$

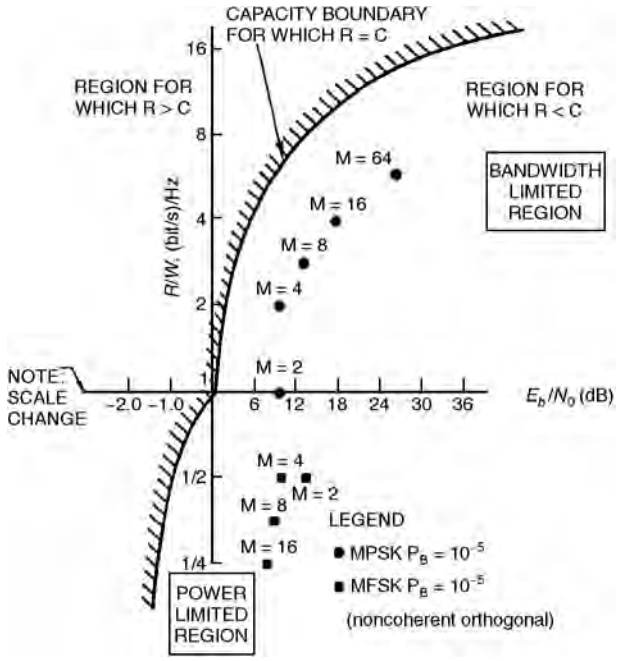


FIGURE 18.1 Bandwidth-efficiency plane.

where M is the size of the alphabet. In the case of digital transmission, the term symbol refers to the member of the M-ary alphabet that is transmitted during each symbol duration T_s. To transmit the symbol, it must be mapped onto an electrical voltage or current waveform. Because the waveform represents the symbol, the terms symbol and waveform are sometimes used interchangeably. Since one of M symbols or waveforms is transmitted during each symbol duration T_s, the data rate R in bits per second can be expressed as

$$R = \frac{m}{T_s} = \frac{\log_2 M}{T_s} \tag{18.9}$$

Data-bit-time duration is the reciprocal of data rate. Similarly, symbol-time duration is the reciprocal of symbol rate. Therefore, from Eq. (18.9), we write that the effective time duration T_b of each bit in terms of the symbol duration T_s or the symbol rate R_s is

$$T_b = \frac{1}{R} = \frac{T_s}{m} = \frac{1}{m R_s} \tag{18.10}$$

Then, using Eqs. (18.8) and (18.10) we can express the symbol rate R_s in terms of the bit rate R as follows:

$$R_s = \frac{R}{\log_2 M} \tag{18.11}$$

From Eqs. (18.9) and (18.10), any digital scheme that transmits $m = \log_2 M$ bits in T_s seconds, using a bandwidth of W hertz, operates at a bandwidth efficiency of

$$\frac{R}{W} = \frac{\log_2 M}{W T_s} = \frac{1}{W T_b} \quad \text{(b/s)/Hz} \tag{18.12}$$

where T_b is the effective time duration of each data bit.

Bandwidth-Limited Systems

From Eq. (18.12), the smaller the WT_b product, the more bandwidth efficient will be any digital communication system. Thus, signals with small WT_b products are often used with bandwidth-limited systems. For example, the European digital mobile telephone system known as Global System for Mobile Communications (GSM) uses Gaussian minimum shift keying (GMSK) modulation having a WT_b product equal to 0.3 Hz/(b/s), where W is the 3-dB bandwidth of a Gaussian filter [8].

For uncoded bandwidth-limited systems, the objective is to maximize the transmitted information rate within the allowable bandwidth, at the expense of E_b/N_0 (while maintaining a specified value of bit-error probability P_B). The operating points for coherent M-ary phase-shift keying (MPSK) at $P_B = 10^{-5}$ are plotted on the bandwidth-efficiency plane of Fig. 18.1. We assume Nyquist (ideal rectangular) filtering at baseband [9]. Thus, for MPSK, the required double-sideband (DSB) bandwidth at an intermediate frequency (IF) is related to the symbol rate as follows:

$$W = \frac{1}{T_s} = R_s \tag{18.13}$$

where T_s is the symbol duration and R_s is the symbol rate. The use of Nyquist filtering results in the minimum required transmission bandwidth that yields zero intersymbol interference; such ideal filtering gives rise to the name Nyquist minimum bandwidth.

From Eqs. (18.12) and (18.13), the bandwidth efficiency of MPSK modulated signals using Nyquist filtering can be expressed as

$$R/W = \log_2 M \quad \text{(b/s)/Hz} \tag{18.14}$$

The MPSK points in Fig. 18.1 confirm the relationship shown in Eq. (18.14). Note that MPSK modulation is a bandwidth-efficient scheme. As M increases in value, R/W also increases. MPSK modulation can be used for realizing an improvement in bandwidth efficiency at the cost of increased E_b/N_0. Although beyond the scope of this chapter, many highly bandwidth-efficient modulation schemes have been investigated [10].

Power-Limited Systems

Operating points for noncoherent orthogonal M-ary FSK (MFSK) modulation at $P_B = 10^{-5}$ are also plotted in Fig. 18.1. For MFSK, the IF minimum bandwidth is as follows [2]:

$$W = \frac{M}{T_s} = MR_s \tag{18.15}$$

where T_s is the symbol duration and R_s is the symbol rate. With MFSK, the required transmission bandwidth is expanded M-fold over binary FSK since there are M different orthogonal waveforms, each requiring a bandwidth of $1/T_s$. Thus, from Eqs. (18.12) and (18.15), the bandwidth efficiency of noncoherent orthogonal MFSK signals can be expressed as

$$\frac{R}{W} = \frac{\log_2 M}{M} \quad \text{(b/s)/Hz} \tag{18.16}$$

The MFSK points plotted in Fig. 18.1 confirm the relationship shown in Eq. (18.16). Note that MFSK modulation is a bandwidth-expansive scheme. As M increases, R/W decreases. MFSK modulation can be used for realizing a reduction in required E_b/N_0 at the cost of increased bandwidth.

In Eqs. (18.13) and (18.14) for MPSK and Eqs. (18.15) and (18.16) for MFSK, and for all the points plotted in Fig. 18.1, ideal filtering has been assumed. Such filters are not realizable! For realistic channels

TABLE 18.1 Symbol Rate, Minimum Bandwidth, Bandwidth Efficiency, and Required E_b/N_0 for MPSK and Noncoherent Orthogonal MFSK Signalling at 9600 bit/s

M	m	R (b/s)	R_s (symb/s)	MPSK Minimum Bandwidth (Hz)	MPSK R/W	MPSK E_b/N_0 (dB) $P_B = 10^{-5}$	Noncoherent Orthog MFSK Min Bandwidth (Hz)	MFSK R/W	MFSK E_b/N_0 (dB) $P_B = 10^{-5}$
2	1	9600	9600	9600	1	9.6	19,200	1/2	13.4
4	2	9600	4800	4800	2	9.6	19,200	1/2	10.6
8	3	9600	3200	3200	3	13.0	25,600	3/8	9.1
16	4	9600	2400	2400	4	17.5	38,400	1/4	8.1
32	5	9600	1920	1920	5	22.4	61,440	5/32	7.4

and waveforms, the required transmission bandwidth must be increased in order to account for realizable filters.

In the examples that follow, we will consider radio channels that are disturbed only by additive white Gaussian noise (AWGN) and have no other impairments, and for simplicity, we will limit the modulation choice to constant-envelope types, i.e., either MPSK or noncoherent orthogonal MFSK. For an uncoded system, MPSK is selected if the channel is bandwidth limited, and MFSK is selected if the channel is power limited. When error-correction coding is considered, modulation selection is not as simple, because coding techniques can provide power-bandwidth tradeoffs more effectively than would be possible through the use of any M-ary modulation scheme considered in this chapter [11].

In the most general sense, M-ary signalling can be regarded as a waveform-coding procedure, i.e., when we select an M-ary modulation technique instead of a binary one, we in effect have replaced the binary waveforms with better waveforms—either better for bandwidth performance (MPSK) or better for power performance (MFSK). Even though orthogonal MFSK signalling can be thought of as being a coded system, i.e., a first-order Reed-Muller code [12], we restrict our use of the term coded system to those traditional error-correction codes using redundancies, e.g., block codes or convolutional codes.

Minimum Bandwidth Requirements for MPSK and MFSK Signalling

The basic relationship between the symbol (or waveform) transmission rate R_s and the data rate R was shown in Eq. (18.11). Using this relationship together with Eqs. (18.13–18.16) and R = 9600 b/s, a summary of symbol rate, minimum bandwidth, and bandwidth efficiency for MPSK and noncoherent orthogonal MFSK was compiled for M = 2, 4, 8, 16, and 32 (Table 18.1). Values of E_b/N_0 required to achieve a bit-error probability of 10^{-5} for MPSK and MFSK are also given for each value of M. These entries (which were computed using relationships that are presented later in this chapter) corroborate the tradeoffs shown in Fig. 18.1. As M increases, MPSK signalling provides more bandwidth efficiency at the cost of increased E_b/N_0, whereas MFSK signalling allows for a reduction in E_b/N_0 at the cost of increased bandwidth.

18.3 Example 1: Bandwidth-Limited Uncoded System

Suppose we are given a bandwidth-limited AWGN radio channel with an available bandwidth of W = 4000 Hz. Also, suppose that the link constraints (transmitter power, antenna gains, path loss, etc.) result in the ratio of received average signal-power to noise-power spectral density S/N_0 being equal to 53 dB-Hz. Let the required data rate R be equal to 9600 b/s, and let the required bit-error performance P_B be at most 10^{-5}. The goal is to choose a modulation scheme that meets the required performance. In general, an error-correction coding scheme may be needed if none of the allowable modulation schemes can meet the requirements. In this example, however, we shall find that the use of error-correction coding is not necessary.

Solution to Example 1

For any digital communication system, the relationship between received S/N_0 and received bit-energy to noise-power spectral density, E_b/N_0 was given in Eq. (18.6) and is briefly rewritten as

$$\frac{S}{N_0} = \frac{E_b}{N_0} R \tag{18.17}$$

Solving for E_b/N_0 in decibels, we obtain

$$\begin{aligned} \frac{E_b}{N_0}\,(\text{dB}) &= \frac{S}{N_0}\,(\text{dB-Hz}) - R\ (\text{dB-b/s}) \\ &= 53\ \text{dB-Hz} - (10 \times \log_{10} 9600)\ \text{dB-b/s} \\ &= 13.2\ \text{dB (or 20.89)} \end{aligned} \tag{18.18}$$

Since the required data rate of 9600 b/s is much larger than the available bandwidth of 4000Hz, the channel is bandwidth limited. We therefore select MPSK as our modulation scheme. We have confined the possible modulation choices to be constant-envelope types; without such a restriction, we would be able to select a modulation type with greater bandwidth efficiency. To conserve power, we compute the *smallest possible* value of M such that the MPSK minimum bandwidth does not exceed the available bandwidth of 4000 Hz. Table 18.1 shows that the smallest value of M meeting this requirement is $M = 8$. Next we determine whether the required bit-error performance of $P_B \leq 10^{-5}$ can be met by using 8-PSK modulation alone or whether it is necessary to use an error-correction coding scheme. Table 18.1 shows that 8-PSK alone will meet the requirements, since the required E_b/N_0 listed for 8-PSK is less than the received E_b/N_0 derived in Eq. (18.18). Let us imagine that we do not have Table 18.1, however, and evaluate whether or not error-correction coding is necessary.

Figure 18.2 shows the basic modulator/demodulator (MODEM) block diagram summarizing the functional details of this design. At the modulator, the transformation from data bits to symbols yields an output symbol rate R_s, that is, a factor $\log_2 M$ smaller than the input data-bit rate R, as is seen in Eq. (18.11). Similarly, at the input to the demodulator, the symbol-energy to noise-power spectral density E_S/N_0 is a factor $\log_2 M$ larger than E_b/N_0, since each symbol is made up of $\log_2 M$ bits. Because E_S/N_0 is larger than E_b/N_0 by the same factor that R_s is smaller than R, we can expand Eq. (18.17), as

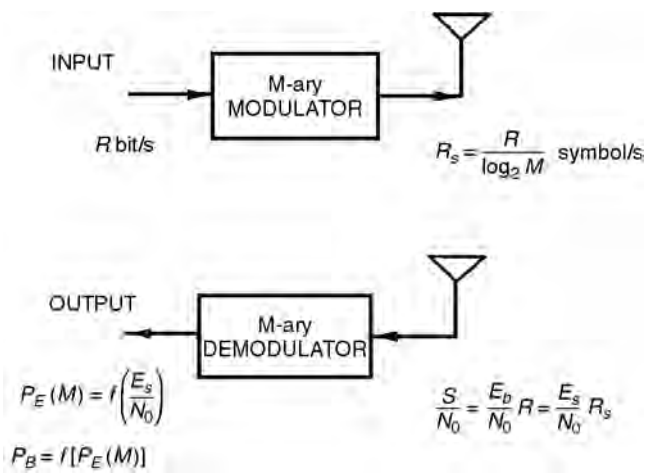


FIGURE 18.2 Basic modulator/demodulator (MODEM) without channel coding.

follows:

$$\frac{S}{N_0} = \frac{E_b}{N_0}R = \frac{E_s}{N_0}R_s \tag{18.19}$$

The demodulator receives a waveform (in this example, one of $M = 8$ possible phase shifts) during each time interval T_s. The probability that the demodulator makes a symbol error $P_E(M)$ is well approximated by the following equation for $M > 2$ [13]:

$$P_E(M) \cong 2Q\left[\sqrt{\frac{2E_s}{N_0}}\sin\left(\frac{\pi}{M}\right)\right] \tag{18.20}$$

where $Q(x)$, sometimes called the complementary error function, represents the probability under the tail of a zero-mean unit-variance Gaussian density function. It is defined as follows [14]:

$$Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty \exp\left(-\frac{u^2}{2}\right)\mathrm{d}u \tag{18.21}$$

A good approximation for $Q(x)$, valid for $x > 3$, is given by the following equation [15]:

$$Q(x) \cong \frac{1}{x\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right) \tag{18.22}$$

In Fig. 18.2 and all of the figures that follow, rather than show explicit probability relationships, the generalized notation $f(x)$ has been used to indicate some functional dependence on x.

A traditional way of characterizing communication efficiency in digital systems is in terms of the received E_b/N_0 in decibels. This E_b/N_0 description has become standard practice, but recall that there are no bits at the input to the demodulator; there are only waveforms that have been assigned bit meanings. The received E_b/N_0 represents a bit-apportionment of the arriving waveform energy.

To solve for $P_E(M)$ in Eq. (18.20), we first need to compute the ratio of received symbol-energy to noise-power spectral density E_s/N_0. Since, from Eq. (18.18),

$$\frac{E_b}{N_0} = 13.2 \text{ dB (or } 20.89)$$

and because each symbol is made up of $\log_2 M$ bits, we compute the following using $M = 8$:

$$\frac{E_s}{N_0} = (\log_2 M)\frac{E_b}{N_0} = 3 \times 20.89 = 62.67 \tag{18.23}$$

Using the results of Eq. (18.23) in Eq. (18.20) yields the symbol-error probability $P_E = 2.2 \times 10^{-5}$. To transform this to bit-error probability, we use the relationship between bit-error probability P_B and symbol-error probability P_E, for multiple-phase signalling [12] $P_E \ll 1$ as follows:

$$P_B \cong \frac{P_E}{\log_2 M} = \frac{P_E}{m} \tag{18.24}$$

which is a good approximation when Gray coding is used for the bit-to-symbol assignment [13]. This last computation yields $P_B = 7.3 \times 10^{-6}$, which meets the required bit-error performance. No error-correction coding is necessary, and 8-PSK modulation represents the design choice to meet the requirements of the bandwidth-limited channel, which we had predicted by examining the required E_b/N_0 values in Table 18.1.

18.4 Example 2: Power-Limited Uncoded System

Now, suppose that we have exactly the same data rate and bit-error probability requirements as in Example 1, but let the available bandwidth W be equal to 45 kHz and the available S/N_0 be equal to 48 dB-Hz. The goal is to choose a modulation or modulation/coding scheme that yields the required performance. We shall again find that error-correction coding is not required.

Solution to Example 2

The channel is clearly not bandwidth limited since the available bandwidth of 45kHz is more than adequate for supporting the required data rate of 9600 bit/s. We find the received E_b/N_0 from Eq. (18.18), as follows:

$$\frac{E_b}{N_0}\,(\text{dB}) = 48\ \text{dB-Hz} - (10 \times \log_{10} 9600)\ \text{dB-b/s} = 8.2\ \text{dB (or 6.61)} \tag{18.25}$$

Since there is abundant bandwidth but a relatively small E_b/N_0 for the required bit-error probability, we consider that this channel is power limited and choose MFSK as the modulation scheme. To conserve power, we search for the *largest possible M* such that the MFSK minimum bandwidth is not expanded beyond our available bandwidth of 45 kHz. A search results in the choice of $M = 16$ (Table 18.1). Next, we determine whether the required error performance of $P_B \le 10^{-5}$ can be met by using 16-FSK alone, i.e., without error-correction coding. Table 18.1 shows that 16-FSK alone meets the requirements, since the required E_b/N_0 listed for 16-FSK is less than the received E_b/N_0 derived in Eq. (18.25). Let us imagine again that we do not have Table 18.1 and evaluate whether or not error-correction coding is necessary.

The block diagram in Fig. 18.2 summarizes the relationships between symbol rate R_s and bit rate R, and between E_s/N_0 and E_b/N_0, which is identical to each of the respective relationships in Example 1. The 16-FSK demodulator receives a waveform (one of 16 possible frequencies) during each symbol time interval T_s. For noncoherent orthogonal MFSK, the probability that the demodulator makes a symbol error $P_E(M)$ is approximated by the following upper bound [16]:

$$P_E(M) \le \frac{M-1}{2}\exp\left(-\frac{E_s}{2N_0}\right) \tag{18.26}$$

To solve for $P_E(M)$ in Eq. (18.26), we compute E_S/N_0, as in Example 1. Using the results of Eq. (18.25) in Eq. (18.23), with $M = 16$, we get

$$\frac{E_s}{N_0} = (\log_2 M)\frac{E_b}{N_0} = 4 \times 6.61 = 26.44 \tag{18.27}$$

Next, using the results of Eq. (18.27) in Eq. (18.26) yields the symbol-error probability $P_E = 1.4 \times 10^{-5}$. To transform this to bit-error probability, P_B, we use the relationship between P_B and P_E for orthogonal signalling [16], given by

$$P_B = \frac{2^{m-1}}{(2^m - 1)}P_E \tag{18.28}$$

This last computation yields $P_B = 7.3 \times 10^{-6}$, which meets the required bit-error performance. Thus, we can meet the given specifications for this power-limited channel by using 16-FSK modulation, without any need for error-correction coding, as we had predicted by examining the required E_b/N_0 values in Table 18.1.

18.5 Example 3: Bandwidth-Limited and Power-Limited Coded System

We start with the same channel parameters as in Example 1 ($W = 4000$ Hz, $S/N_0 = 53$ dB-Hz, and $R = 9600$ b/s), with one exception.

In this example, we specify that P_B must be at most 10^{-9}. Table 18.1 shows that the system is both bandwidth limited and power limited, based on the available bandwidth of 4000 Hz and the available E_b/N_0 of 20.2 dB, from Eq. (18.18); 8-PSK is the only possible choice to meet the bandwidth constraint; however, the available E_b/N_0 of 20.2 dB is certainly insufficient to meet the required P_B of 10^{-9}. For this small value of P_B, we need to consider the performance improvement that error-correction coding can provide within the available bandwidth. In general, one can use convolutional codes or block codes.

The Bose–Chaudhuri–Hocquenghem (BCH) codes form a large class of powerful error-correcting cyclic (block) codes [17]. To simplify the explanation, we shall choose a block code from the BCH family. Table 18.2 presents a partial catalog of the available BCH codes in terms of n, k, and t, where k represents the number of information (or data) bits that the code transforms into a longer block of n coded bits (or channel bits), and t represents the largest number of incorrect channel bits that the code can correct within each n-sized block. The rate of a code is defined as the ratio k/n; its inverse represents a measure of the code's redundancy [17].

Solution to Example 3

Since this example has the same bandwidth-limited parameters given in Example 1, we start with the same 8-PSK modulation used to meet the stated bandwidth constraint. We now employ error-correction coding, however, so that the bit-error probability can be lowered to $P_B \leq 10^{-9}$.

To make the optimum code selection from Table 18.2, we are guided by the following goals.

1. The output bit-error probability of the combined modulation/coding system must meet the system error requirement.

TABLE 18.2 BCH Codes (Partial Catalog)

n	k	t
7	4	1
15	11	1
	7	2
	5	3
31	26	1
	21	2
	16	3
	11	5
63	57	1
	51	2
	45	3
	39	4
	36	5
	30	6
127	120	1
	113	2
	106	3
	99	4
	92	5
	85	6
	78	7
	71	9
	64	10

TABLE 18.3 Bandwidth-Compatible BCH Codes

n	k	t	Coding Gain, G (dB) MPSK, $P_B = 10^{-9}$
31	26	1	2.0
63	57	1	2.2
	51	2	3.1
127	120	1	2.2
	113	2	3.3
	106	3	3.9

2. The rate of the code must not expand the required transmission bandwidth beyond the available channel bandwidth.
3. The code should be as simple as possible. Generally, the shorter the code, the simpler will be its implementation.

The uncoded 8-PSK minimum bandwidth requirement is 3200 Hz (Table 18.1) and the allowable channel bandwidth is 4000 Hz, and so the uncoded signal bandwidth can be increased by no more than a factor of 1.25 (i.e., an expansion of 25%). The very first step in this (simplified) code selection example is to eliminate the candidates in Table 18.2 that would expand the bandwidth by more than 25%. The remaining entries form a much reduced set of bandwidth-compatible codes (Table 18.3).

In Table 18.3, a column designated Coding Gain G (for MPSK at $P_B = 10^{-9}$) has been added. Coding gain in decibels is defined as follows:

$$G = \left(\frac{E_b}{N_0}\right)_{\text{uncoded}} - \left(\frac{E_b}{N_0}\right)_{\text{coded}} \tag{18.29}$$

G can be described as the reduction in the required E_b/N_0 (in decibels) that is needed due to the error-performance properties of the channel coding. G is a function of the modulation type and bit-error probability, and it has been computed for MPSK at $P_B = 10^{-9}$ (Table 18.3). For MPSK modulation, G is relatively independent of the value of M. Thus, for a particular bit-error probability, a given code will provide about the same coding gain when used with any of the MPSK modulation schemes. Coding gains were calculated using a procedure outlined in the subsequent Calculating Coding Gain section.

A block diagram summarizes this system, which contains both modulation and coding (Fig. 18.3). The introduction of encoder/decoder blocks brings about additional transformations. The relationships that exist when transforming from R b/s to R_c channel-b/s to R_s symbol/s are shown at the encoder/modulator. Regarding the channel-bit rate R_c, some authors prefer to use the units of channel-symbol/s (or code-symbol/s). The benefit is that error-correction coding is often described more efficiently with nonbinary digits. We reserve the term symbol for that group of bits mapped onto an electrical waveform for transmission, and we designate the units of R_c to be channel-b/s (or coded-b/s).

We assume that our communication system cannot tolerate any message delay, so that the channel-bit rate R_c must exceed the data-bit rate R by the factor n/k. Further, each symbol is made up of $\log_2 M$ channel bits, and so the symbol rate R_s is less than R_c by the factor $\log_2 M$. For a system containing both modulation and coding, we summarize the rate transformations as follows:

$$R_c = \left(\frac{n}{k}\right)R \tag{18.30}$$

$$R_s = \frac{R_c}{\log_2 M} \tag{18.31}$$

At the demodulator/decoder in Fig. 18.3, the transformations among data-bit energy, channel-bit energy, and symbol energy are related (in a reciprocal fashion) by the same factors as shown among the rate

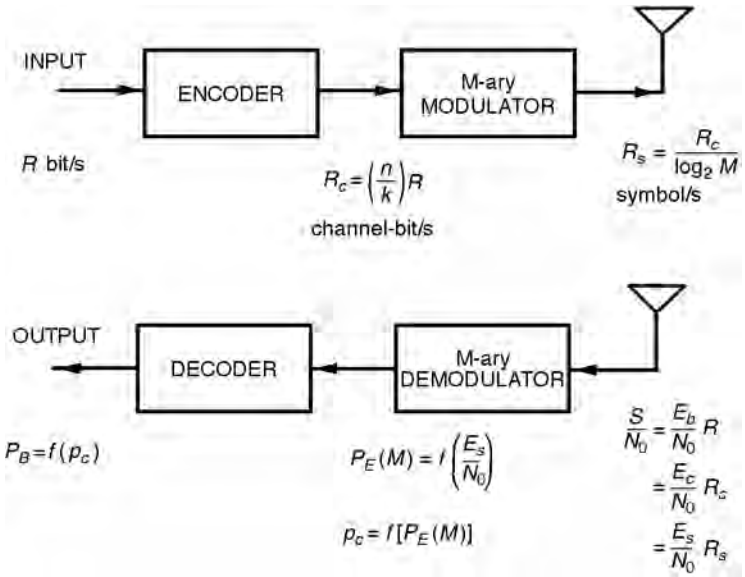


FIGURE 18.3 MODEM with channel coding.

transformations in Eqs. (18.30) and (18.31). Since the encoding transformation has replaced k data bits with n channel bits, then the ratio of channel-bit energy to noise-power spectral density E_c/N_0 is computed by decrementing the value of E_b/N_0 by the factor k/n. Also, since each transmission symbol is made up of $\log_2 M$ channel bits, then E_S/N_0, which is needed in Eq. (18.20) to solve for P_E, is computed by incrementing E_c/N_0 by the factor $\log_2 M$. For a system containing both modulation and coding, we summarize the energy to noise-power spectral density transformations as follows:

$$\frac{E_c}{N_0} = \left(\frac{k}{n}\right)\frac{E_b}{N_0} \tag{18.32}$$

$$\frac{E_s}{N_0} = (\log_2 M)\frac{E_c}{N_0} \tag{18.33}$$

Using Eqs. (18.30) and (18.31), we can now expand the expression for S/N_0 in Eq. (18.19) as follows (see Appendix):

$$\frac{S}{N_0} = \frac{E_b}{N_0}R = \frac{E_c}{N_0}R_c = \frac{E_s}{N_0}R_s \tag{18.34}$$

As before, a standard way of describing the link is in terms of the received E_b/N_0 in decibels. However, there are no data bits at the input to the demodulator, and there are no channel bits; there are only waveforms that have bit meanings and, thus, the waveforms can be described in terms of bit-energy apportionments.

Since S/N_0 and R were given as 53 dB-Hz and 9600 b/s, respectively, we find as before, from Eq. (18.18), that the received $E_b/N_0 = 13.2$ dB. The received E_b/N_0 is fixed and independent of n, k, and t (see Appendix). As we search in Table 18.3 for the ideal code to meet the specifications, we can iteratively repeat the computations suggested in Fig. 18.3. It might be useful to program on a personal computer (or calculator) the

following four steps as a function of n, k, and t. Step 1 starts by combining Eqs. (18.32) and (18.33), as follows.

Step 1:

$$\frac{E_s}{N_0} = (\log_2 M)\frac{E_c}{N_0} = (\log_2 M)\left(\frac{k}{n}\right)\frac{E_b}{N_0} \tag{18.35}$$

Step 2:

$$P_E(M) \cong 2Q\left[\sqrt{\frac{2E_s}{N_0}}\sin\left(\frac{\pi}{M}\right)\right] \tag{18.36}$$

which is the approximation for symbol-error probability P_E rewritten from Eq. (18.20). At each symbol-time interval, the demodulator makes a symbol decision, but it delivers a channel-bit sequence representing that symbol to the decoder. When the channel-bit output of the demodulator is quantized to two levels, 1 and 0, the demodulator is said to make hard decisions. When the output is quantized to more than two levels, the demodulator is said to make soft decisions [2]. Throughout this paper, we shall assume hard-decision demodulation.

Now that we have a decoder block in the system, we designate the channel-bit-error probability out of the demodulator and into the decoder as p_c, and we reserve the notation P_B for the bit-error probability out of the decoder. We rewrite Eq. (18.24) in terms of p_c for $P_E \ll 1$ as follows.

Step 3:

$$p_c \cong \frac{P_E}{\log_2 M} = \frac{P_E}{m} \tag{18.37}$$

relating the channel-bit-error probability to the symbol-error probability out of the demodulator, assuming Gray coding, as referenced in Eq. (18.24).

For traditional channel-coding schemes and a given value of received S/N_0, the value of E_s/N_0 with coding will always be less than the value of E_s/N_0 without coding. Since the demodulator with coding receives less E_s/N_0, it makes more errors! When coding is used, however, the system error-performance does not only depend on the performance of the demodulator, it also depends on the performance of the decoder. For error-performance improvement due to coding, the decoder must provide enough error correction to more than compensate for the poor performance of the demodulator.

The final output decoded bit-error probability P_B depends on the particular code, the decoder, and the channel-bit-error probability p_c. It can be expressed by the following approximation [18].

Step 4:

$$P_B \cong \frac{1}{n}\sum_{j=t+1}^{n} j\binom{n}{j}p_c^j(1-p_c)^{n-j} \tag{18.38}$$

where t is the largest number of channel bits that the code can correct within each block of n bits. Using Eqs. (18.35–18.38) in the four steps, we can compute the decoded bit-error probability P_B as a function of n, k, and t for each of the codes listed in Table 18.3. The entry that meets the stated error requirement with the largest possible code rate and the smallest value of n is the double-error correcting (63, 51) code. The computations are as follows.

Step 1:

$$\frac{E_s}{N_0} = 3\left(\frac{51}{63}\right)20.89 = 50.73$$

where $M = 8$, and the received $E_b/N_0 = 13.2$ dB (or 20.89).

Step 2:

$$P_E \cong 2Q\left[\sqrt{101.5} \times \sin\left(\frac{\pi}{8}\right)\right] = 2Q(3.86) = 1.2 \times 10^{-4}$$

Step 3:

$$p_c \cong \frac{1.2 \times 10^{-4}}{3} = 4 \times 10^{-5}$$

Step 4:

$$\begin{aligned} P_B &\cong \frac{3}{63}\binom{63}{3}(4 \times 10^{-5})^3(1 - 4 \times 10^{-5})^{60} \\ &\quad + \frac{4}{63}\binom{63}{4}(4 \times 10^{-5})^4(1 - 4 \times 10^{-5})^{59} + \cdots \\ &= 1.2 \times 10^{-10} \end{aligned}$$

where the bit-error-correcting capability of the code is $t = 2$. For the computation of P_B in step 4, we need only consider the first two terms in the summation of Eq. (18.38) since the other terms have a vanishingly small effect on the result. Now that we have selected the (63, 51) code, we can compute the values of channel-bit rate R_c and symbol rate R_s using Eqs. (18.30) and (18.31), with $M = 8$,

$$\begin{aligned} R_c &= \left(\frac{n}{k}\right)R = \left(\frac{63}{51}\right) 9600 \approx 11{,}859 \text{ channel-b/s} \\ R_s &= \frac{R_c}{\log_2 M} = \frac{11859}{3} = 3953 \text{ symbol/s} \end{aligned}$$

Calculating Coding Gain

Perhaps a more direct way of finding the simplest code that meets the specified error performance is to first compute how much coding gain G is required in order to yield $P_B = 10^{-9}$ when using 8-PSK modulation alone; then, from Table 18.3, we can simply choose the code that provides this performance improvement. First, we find the uncoded E_s/N_0 that yields an error probability of $P_B = 10^{-9}$, by writing from Eqs. (18.24) and (18.36), the following:

$$P_B \cong \frac{P_E}{\log_2 M} \cong \frac{2Q\left[\sqrt{\frac{2E_s}{N_0}}\sin\left(\frac{\pi}{M}\right)\right]}{\log_2 M} = 10^{-9} \tag{18.39}$$

At this low value of bit-error probability, it is valid to use Eq. (18.22) to approximate $Q(x)$ in Eq. (18.39) By trial and error (on a programmable calculator), we find that the uncoded $E_s/N_0 = 120.67 = 20.8$ dB, and since each symbol is made up of $\log_2 8 = 3$ bits, the required $(E_b/N_0)_{\text{uncoded}} = 120.67/3 = 40.22 = 16$ dB. From the given parameters and Eq. (18.18), we know that the received $(E_b/N_0)_{\text{coded}} = 13.2$ dB. Using Eq. (18.29), the required coding gain to meet the bit-error performance of $P_B = 10^{-9}$ in decibels is

$$G = \left(\frac{E_b}{N_0}\right)_{\text{uncoded}} - \left(\frac{E_b}{N_0}\right)_{\text{coded}} = 16 - 13.2 = 2.8$$

To be precise, each of the E_b/N_0 values in the preceding computation must correspond to exactly the same value of bit-error probability (which they do not). They correspond to $P_B = 10^{-9}$ and $P_B = 1.2 \times 10^{-10}$, respectively. At these low probability values, however, even with such a discrepancy, this computation still provides a good approximation of the required coding gain. In searching Table 18.3 for the simplest code that will yield a coding gain of at least 2.8 dB, we see that the choice is the (63, 51) code, which corresponds to the same code choice that we made earlier.

18.6 Example 4: Direct-Sequence (DS) Spread-Spectrum Coded System

Spread-spectrum systems are not usually classified as being bandwidth or power limited. They are generally perceived to be power-limited systems, however, because the bandwidth occupancy of the information is much larger than the bandwidth that is intrinsically needed for the information transmission. In a direct-sequence spread-spectrum (DS/SS) system, spreading the signal bandwidth by some factor permits lowering the signal-power spectral density by the same factor (the total average signal power is the same as before spreading). The bandwidth spreading is typically accomplished by multiplying a relatively narrowband data signal by a wideband spreading signal. The spreading signal or spreading code is often referred to as a pseudorandom code or PN code.

Processing Gain

A typical DS/SS radio system is often described as a two-step BPSK modulation process. In the first step, the carrier wave is modulated by a bipolar data waveform having a value +1 or −1 during each data-bit duration; in the second step, the output of the first step is multiplied (modulated) by a bipolar PN-code waveform having a value +1 or −1 during each PN-code-bit duration. In reality, DS/SS systems are usually implemented by first multiplying the data waveform by the PN-code waveform and then making a single pass through a BPSK modulator. For this example, however, it is useful to characterize the modulation process in two separate steps—the outer modulator/demodulator for the data, and the inner modulator/demodulator for the PN code (Fig. 18.4).

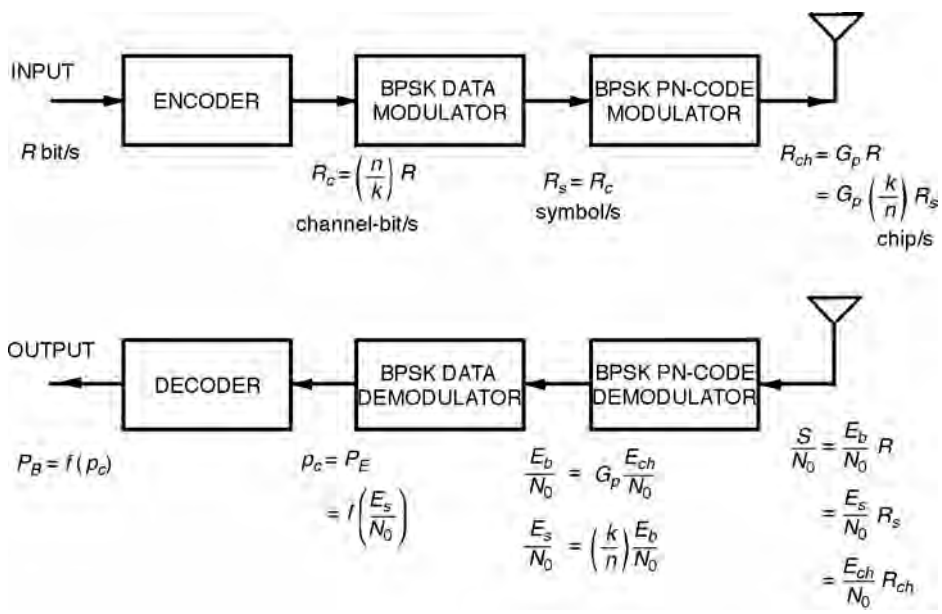


FIGURE 18.4 Direct-sequence spread-spectrum MODEM with channel coding.

A spread-spectrum system is characterized by a processing gain G_p that is defined in terms of the spread-spectrum bandwidth W_{ss} and the data rate R as follows [19]:

$$G_p = \frac{W_{ss}}{R} \tag{18.40}$$

For a DS/SS system, the PN-code bit has been given the name chip, and the spread-spectrum signal bandwidth can be shown to be about equal to the chip rate R_{ch} as follows:

$$G_p = \frac{R_{ch}}{R} \tag{18.41}$$

Some authors define processing gain to be the ratio of the spread-spectrum bandwidth to the symbol rate. This definition separates the system performance that is due to bandwidth spreading from the performance that is due to error-correction coding. Since we ultimately want to relate all of the coding mechanisms relative to the information source, we shall conform to the most usually accepted definition for processing gain, as expressed in Eqs. (18.40) and (18.41).

A spread-spectrum system can be used for interference rejection and for multiple access (allowing multiple users to access a communications resource simultaneously). The benefits of DS/SS signals are best achieved when the processing gain is very large; in other words, the chip rate of the spreading (or PN) code is much larger than the data rate. In such systems, the large value of G_p allows the signalling chips to be transmitted at a power level well below that of the thermal noise. We will use a value of $G_p = 1000$. At the receiver, the despreading operation correlates the incoming signal with a synchronized copy of the PN code and, thus, accumulates the energy from multiple (G_p) chips to yield the energy per data bit. The value of G_p has a major influence on the performance of the spread-spectrum system application. We shall see, however, that the value of G_p has no effect on the received E_b/N_0. In other words, spread spectrum techniques offer no error-performance advantage over thermal noise. For DS/SS systems, there is no disadvantage either! Sometimes such spread-spectrum radio systems are employed only to enable the transmission of very small power-spectral densities and thus avoid the need for FCC licensing [20].

Channel Parameters for Example 4

Consider a DS/SS radio system that uses the same (63, 51) code as in the previous example. Instead of using MPSK for the data modulation, we shall use BPSK. Also, we shall use BPSK for modulating the PN-code chips. Let the received $S/N_0 = 48$ dB-Hz, the data rate $R = 9600$ b/s, and the required $P_B \leq 10^{-6}$. For simplicity, assume that there are no bandwidth constraints. Our task is simply to determine whether or not the required error performance can be achieved using the given system architecture and design parameters. In evaluating the system, we will use the same type of transformations used in the previous examples.

Solution to Example 4

A typical DS/SS system can be implemented more simply than the one shown in Fig. 18.4. The data and the PN code would be combined at baseband, followed by a single pass through a BPSK modulator. We will, however, assume the existence of the individual blocks in Fig. 18.4 because they enhance our understanding of the transformation process. The relationships in transforming from data bits, to channel bits, to symbols, and to chips, shown in Fig. 18.4, have the same pattern of subtle but straightforward transformations in rates and energies as previous relationships (Figs. 18.2 and 18.3). The values of R_c, R_s, and R_{ch} can now be calculated immediately since the (63, 51) BCH code has already been selected. From Eq. (18.30) we write

$$R_c = \left(\frac{n}{k}\right)R = \left(\frac{63}{51}\right)9600 \approx 11{,}859 \text{ channel-b/s}$$

Since the data modulation considered here is BPSK, then from Eq. (18.31) we write

$$R_s = R_c \approx 11{,}859 \text{ symbol/s}$$

and from Eq. (18.41), with an assumed value of $G_p = 1000$

$$R_{\text{ch}} = G_p R = 1000 \times 9600 = 9.6 \times 10^6 \text{chip/s}$$

Since we have been given the same S/N_0 and the same data rate as in Example 2, we find the value of received E_b/N_0 from Eq. (18.25) to be 8.2 dB (or 6.61). At the demodulator, we can now expand the expression for S/N_0 in Eq. (18.34) and the Appendix as follows:

$$\frac{S}{N_0} = \frac{E_b}{N_0}R = \frac{E_c}{N_0}R_c = \frac{E_s}{N_0}R_s = \frac{E_{\text{ch}}}{N_0}R_{\text{ch}} \tag{18.42}$$

Corresponding to each transformed entity (data bit, channel bit, symbol, or chip) there is a change in rate and, similarly, a reciprocal change in energy-to-noise spectral density for that received entity. Equation (18.42) is valid for any such transformation when the rate and energy are modified in a reciprocal way. There is a kind of *conservation of power* (or energy) phenomenon that exists in the transformations. The total received average power (or total received energy per symbol duration) is fixed regardless of how it is computed, on the basis of data bits, channel bits, symbols, or chips.

The ratio E_{ch}/N_0 is much lower in value than E_b/N_0. This can be seen from Eqs. (18.42) and (18.41), as follows:

$$\frac{E_{\text{ch}}}{N_0} = \frac{S}{N_0}\left(\frac{1}{R_{\text{ch}}}\right) = \frac{S}{N_0}\left(\frac{1}{G_p R}\right) = \left(\frac{1}{G_p}\right)\frac{E_b}{N_0} \tag{18.43}$$

But, even so, the despreading function (when properly synchronized) accumulates the energy contained in a quantity G_p of the chips, yielding the same value $E_b/N_0 = 8.2$ dB, as was computed earlier from Eq. (18.25). Thus, the DS spreading transformation has no effect on the error performance of an AWGN channel [2], and the value of G_p has no bearing on the value of P_B in this example.

From Eq. (18.43), we can compute, in decibels,

$$\begin{aligned} \frac{E_{\text{ch}}}{N_0} &= E_b/N_0 - G_p \\ &= 8.2 - (10 \times \log_{10} 1000) \\ &= -21.8 \end{aligned} \tag{18.44}$$

The chosen value of processing gain ($G_p = 1000$) enables the DS/SS system to operate at a value of chip energy well below the thermal noise, with the same error performance as without spreading.

Since BPSK is the data modulation selected in this example, each message symbol therefore corresponds to a single channel bit, and we can write

$$\frac{E_s}{N_0} = \frac{E_c}{N_0} = \left(\frac{k}{n}\right)\frac{E_b}{N_0} = \left(\frac{51}{63}\right) \times 6.61 = 5.35 \tag{18.45}$$

where the received $E_b/N_0 = 8.2$ dB (or 6.61). Out of the BPSK data demodulator, the symbol-error probability P_E (and the channel-bit error probability p_c) is computed as follows [2]:

$$p_c = P_E = Q\left(\sqrt{\frac{2E_c}{N_0}}\right) \tag{18.46}$$

Using the results of Eq. (18.45) in Eq. (18.46) yields

$$p_c = Q(3.27) = 5.8 \times 10^{-4}$$

Finally, using this value of p_c in Eq. (18.38) for the (63, 51) double-error correcting code yields the output bit-error probability of $P_B = 3.6 \times 10^{-7}$. We can, therefore, verify that for the given architecture and design parameters of this example, the system does, in fact, achieve the required error performance.

18.7 Conclusion

The goal of this chapter has been to review fundamental relationships used in evaluating the performance of digital communication systems. First, we described the concept of a link and a channel and examined a radio system from its transmitting segment up through the output of the receiving antenna. We then examined the concept of bandwidth-limited and power-limited systems and how such conditions influence the system design when the choices are confined to MPSK and MFSK modulation. Most important, we focused on the definitions and computations involved in transforming from data bits to channel bits to symbols to chips. In general, most digital communication systems share these concepts; thus, understanding them should enable one to evaluate other such systems in a similar way.

Appendix: Received E_b/N_0 Is Independent of the Code Parameters

Starting with the basic concept that the received average signal power S is equal to the received symbol or waveform energy, E_s, divided by the symbol-time duration, T_s (or multiplied by the symbol rate, R_s), we write

$$\frac{S}{N_0} = \frac{E_s/T_s}{N_0} = \frac{E_s}{N_0}R_s \tag{A18.1}$$

where N_0 is noise-power spectral density characterizing all the noise in the receiving system.

Using Eqs. (18.27) and (18.25), rewritten as

$$\frac{E_s}{N_0} = (\log_2 M)\frac{E_c}{N_0} \qquad \text{and} \qquad R_s = \frac{R_c}{\log_2 M}$$

let us make substitutions into Eq. (A18.1), which yields

$$\frac{S}{N_0} = \frac{E_c}{N_0}R_c \tag{A18.2}$$

Next, using Eqs. (18.26) and (18.24), rewritten as

$$\frac{E_c}{N_0} = \left(\frac{k}{n}\right)\frac{E_b}{N_0} \qquad \text{and} \qquad R_c = \left(\frac{n}{k}\right)R$$

let us now make substitutions into Eq. (A18.2), which yields the relationship expressed in Eq. (18.11)

$$\frac{S}{N_0} = \frac{E_b}{N_0}R \tag{A18.3}$$

Hence, the received E_b/N_0 is only a function of the received S/N_0 and the data rate R. It is independent of the code parameters n, k, and t. These results are summarized in Fig. 18.3.

References

1. Sklar, B., What the system link budget tells the system engineer or how I learned to count in decibels, *Proc. of the Intl. Telemetering Conf.,* San Diego, CA, Nov. 1979.
2. Sklar, B., *Digital Communications: Fundamentals and Applications,* 2nd ed., Prentice-Hall, Englewood Cliffs, NJ, 2001.
3. Johnson, J.B., Thermal agitation of electricity in conductors, *Phys. Rev.,* 32, 97–109, Jul. 1928.
4. Nyquist, H., Thermal agitation of electric charge in conductors, *Phys. Rev.,* 32, 110–113, Jul. 1928.
5. Ungerboeck, G., Trellis-coded modulation with redundant signal sets, Pt. I and II, *IEEE Comm. Mag.,* 25, 5–21. Feb. 1987.
6. Shannon, C.E., A mathematical theory of communication, *BSTJ.* 27, 379–423, 623–657, 1948.
7. Shannon, C.E., Communication in the presence of noise, *Proc. IRE.* 37(1), 10–21, 1949.
8. Hodges, M.R.L., The GSM radio interface, *British Telecom Technol. J.,* 8(1), 31–43, 1990.
9. Nyquist, H., Certain topics on telegraph transmission theory, *Trans. AIEE,* 47, 617–644, Apr. 1928.
10. Anderson, J.B. and Sundberg, C. E., Advances in constant envelope coded modulation, *IEEE Commun. Mag.,* 29(12), 36–45, 1991.
11. Clark Jr., G.C. and Cain, J.B., *Error-Correction Coding for Digital Communications,* Plenum Press, New York, 1981.
12. Lindsey, W.C. and Simon, M.K., *Telecommunication Systems Engineering,* Prentice-Hall, Englewood Cliffs, NJ, 1973.
13. Korn, I., *Digital Communications,* Van Nostrand Reinhold Co., New York, 1985.
14. Van Trees, H.L., *Detection, Estimation, and Modulation Theory,* Pt. I, John Wiley & Sons, New York, 1968.
15. Borjesson, P.O. and Sundberg, C.E., Simple approximations of the error function $Q(x)$ for communications applications, *IEEE Trans. Comm.,* COM-27, 639–642, Mar. 1979.
16. Viterbi, A.J., *Principles of Coherent Communication,* McGraw-Hill, New York, 1966.
17. Lin, S. and Costello Jr., D.J., *Error Control Coding: Fundamentals and Applications,* Prentice-Hall, Englewood Cliffs, NJ, 1983.
18. Odenwalder, J.P., *Error Control Coding Handbook,* Linkabit Corp., San Diego, CA, Jul. 15, 1976.
19. Viterbi, A.J., Spread spectrum communications—myths and realities, *IEEE Comm. Mag.,* 11–18, May, 1979.
20. Title 47, *Code of Federal Regulations,* Part 15 Radio Frequency Devices.

Further Information

A useful compilation of selected papers can be found in *Cellular Radio & Personal Communications–A Book of Selected Readings,* edited by Theodore S. Rappaport, Institute of Electrical and Electronics Engineers, Inc., Piscataway, New Jersey, 1995. Fundamental design issues, such as propagation, modulation, channel coding, speech coding, multiple-accessing, and networking, are well represented in that volume.

Another useful sourcebook that covers the fundamentals of mobile communications in great detail is *Mobile Radio Communications*, edited by Raymond Steele, Pentech Press, London, 1992. That volume is also available through the Institute of Electrical and Electronics Engineers, Inc., Piscataway, New Jersey.

For spread spectrum systems, an excellent reference is *Spread Spectrum Communications Handbook,* by Marvin K. Simon, Jim K. Omura, Robert A. Scholtz, and Barry K. Levitt, McGraw-Hill Inc., New York, 1994.

19

Synchronization

Costas N. Georghiades
Texas A&M University

Erchin Serpedin
Texas A&M University

19.1 Introduction

Etymologically, the word synchronization refers to the process of making two or more events occur at the same time, or, by duality, at the same frequency. In a digital communication context, various levels of synchronization must be established before data decoding can take place, including *carrier synchronization, symbol synchronization*, and *frame synchronization*.

In radio frequency communications, carrier synchronization refers to the process of generating a sinusoidal signal that closely tracks the *phase* and *frequency* of a received noisy carrier, transmitted by a possibly distant transmitter. Thus, carrier synchronization in general refers to both frequency and phase acquisition and tracking. Of the two, in many cases the more difficult problem is extracting carrier phase, which is often a much faster varying process compared to frequency offset. This is especially so since the advent of highly stable crystal oscillators operating in the UHF or lower frequency bands, although the problem is still prevalent at the higher, microwave frequencies where crystal oscillators are not available. Frequency acquisition is a also a problem in mobile radio applications where, due to the Doppler effect, there is an offset in the frequency of the received carrier.

Communication systems that have available, or somehow extract and make use of, good (theoretically perfect) carrier frequency and phase information are known as *coherent systems*, in contrast to *incoherent systems* that neglect the carrier phase. Systems that attempt to acquire phase information but do not do a perfect job are known as *partially coherent systems*. Coherent systems are known to perform better than incoherent ones, at the price, however, of more complexity required for carrier synchronization. This chapter will look at some of the classical techniques for carrier acquisition as well as some of the modern techniques, which often involve operating on sampled instead of analog data.

Symbol synchronization is the process of deriving at the receiver timing signals that indicate where in time the transmitted symbols are located. The decision part of the receiver subsequently uses this information in order to decide what the symbols are. As with carrier synchronization, the data available to the receiver for making timing estimates is noisy. Thus, perfect timing information cannot be obtained in practice, although practical systems come close.

Once symbol synchronization is achieved, the next highest synchronization level is frame synchronization. Frame synchronization is necessary in systems for which the unit of information is not a symbol, but rather a sequence of symbols. Such systems are, for example, coded systems where the unit of information is a codeword which consists of a number of symbols. In this case, it is clear that knowing where the symbols are is not enough, and a further knowledge of where the codewords are is needed. It is easily seen that the existence of frame synchronization automatically implies symbol synchronization, but the converse is not true. Thus, one might be tempted to attempt frame synchronization before symbol synchronization is achieved, thereby achieving both at once. Such an approach, although in theory resulting in better performance, has the disadvantage of requiring more complex processing than the approach of first achieving lower level synchronization before higher levels are attempted. In practice, almost invariably the latter approach is followed. A rather standard approach to achieving frame synchronization is for the transmitter to insert at the start of every frame a special synchronization pattern, whose detection at the receiver locates frame boundaries. We will look at the optimal frame synchronization processing as well as some of the desirable characteristics of synchronization sequences later.

There are two general methodologies for achieving the various synchronization levels needed by the receiver. One way is to provide at the receiver side an unmodulated carrier, or a carrier modulated by a known sequence, which can be used solely for the purpose of synchronization. This approach has the advantage of decoupling the problem of data detection and synchronization and makes the synchronization system design easier. On the other hand, the overall communication efficiency suffers since signal energy and time is used but no information is sent. A second approach, which is often preferred, is to derive synchronization from the data-modulated carrier, the same signal used for symbol decisions. In this way, no efficiency is sacrificed, but the processing becomes somewhat more involved. In the sequel, we will only consider algorithms that utilize a modulated received signal to derive synchronization. We first start with a look at carrier synchronization algorithms. Excellent general treatments of this and the other synchronization problems studied later can be found in the classic texts of Stiffler [1] and Lindsey and Simon [2], and more recently by Meyr [3, 16] and Mengali and D'Andrea [17].

19.2 Carrier Synchronization

Carrier Frequency Synchronization

The objective of a carrier frequency synchronization system consists of estimating and compensating the carrier frequency offset that may be induced at the receiver by oscillator instabilities and/or Doppler shifts. According to the degree of knowledge of the transmitted symbols, carrier frequency synchronizers are classified into three main categories: data-aided (DA), decision-directed (DD) and non-data aided (NDA) or blind methods. DA methods assume perfect knowledge of the transmitted symbols, while NDA methods do not require such knowledge. Being more spectrally efficient, the NDA methods are well suited for burst applications. As an intermediate category between the DA and NDA methods, the DD methods rely on knowledge of the symbols obtained at the output of a symbol-by-symbol decoder. According to the magnitude of the carrier frequency offset that they can cope with, carrier frequency synchronizers may be classified in two classes:

1. Carrier frequency synchronizers that can compensate frequency offsets much smaller than the symbol rate (1/T), in general less than 10% of the symbol rate.
2. Carrier frequency synchronizers that can compensate large frequency offsets on the order of the symbol rate (1/T).

During initial carrier frequency acquisition, the second class of carrier synchronizers is employed in order to reduce large carrier frequency offsets to a small percentage of the symbol rate and to facilitate other carrier or symbol synchronization operations. Compensation of large frequency offsets with magnitudes in the range of 100% of the symbol rate can be performed using either closed-loop (feedback) or open-loop (feedforward) frequency synchronizers. Within the class of closed-loop frequency synchronizers, the most

used carrier recovery systems are the approximate maximum likelihood (ML) frequency error detectors [23, 25], quadricorrelators [24, 26, 33, 34], and dual filter detectors [21, 33], which were shown to be equivalent in [22]. Open-loop frequency recovery schemes [27, 28], referred to also as delay-and-multiply methods, present shorter acquisition time and are simpler to implement than the closed-loop frequency recovery methods. They also exhibit comparable performance to the approximate ML frequency error detectors [23, 25]. Due to these features, open-loop carrier recovery schemes are used in spontaneous packet transmissions, where the frequency synchronization must be performed within a fixed time interval.

Once that the compensation of the large carrier frequency offsets has been accomplished and the receiver operates under steady-state conditions, carrier frequency recovery systems that can track and compensate carrier frequency offsets of magnitude much less than the symbol rate are usually employed. Compensation of frequency offsets with magnitudes in the range of 10% of the symbol rate can be performed by employing DA methods, such as those proposed by Fitz [30, 31], Luise and Reggiannini [32], and the approximate ML estimator [17, 28]. These methods appear to be the best methods available in the literature in terms of implementation complexity and performance. All these methods practically achieve the Cramer–Rao bound at a signal-to-noise ratio (SNR) of zero dB. For burst mode applications, open-loop NDA frequency recovery schemes have been proposed for arbitrary QAM and M-ary PSK input symbol constellations [29]. Open-loop NDA frequency recovery schemes estimate the frequency offset based on certain higher-order statistics computed from the received samples and present limited performance due to the self-noise that is induced by computing the associated higher-order statistics.

Carrier Phase Synchronization

This section will formulate the carrier synchronization problem in mathematical terms as an estimation problem and will then see how the optimal equations derived can be practically implemented through appropriate approximations. Our optimality criterion is in the sense of maximum-likelihood, under which estimates maximize with respect to the parameter to be estimated, the conditional probability density of the data given the parameter (see, for example, [4]). As pointed out in the introduction, carrier synchronization may be achieved rather easily by tracking the phase of an unmodulated carrier that is frequency multiplexed with the modulated carrier. In this case, we need not worry about the noise (uncertainty) introduced by the random modulation. On the other hand, the resulting system is inefficient since part of the transmitter power carries no information and is used solely for carrier phase estimation. Although we will look at such synchronizers, our emphasis will be on the more efficient carrier synchronizers that derive their carrier phase estimates from suppressed carrier signals.

Unmodulated Carrier

The received unmodulated signal $r(t)$ is

$$r(t) = s(t;\phi) + n(t), \quad t \in T_0 \tag{19.1}$$

where $s(t, \phi) = A\cos(2\pi f_c t - \phi)$ is the transmitted (unmodulated) signal, ϕ is the carrier phase, and $n(t)$ is a zero-mean, white Gaussian noise process with spectral density $N_0/2$. The maximum-likelihood (ML) estimate of the received noisy carrier is the value of ϕ that maximizes the likelihood function, given by (see, for example, [4])

$$L(\phi) = \exp\left[\frac{2}{N_0}\int_{T_0} r(t)s(t;\phi)dt - \frac{1}{N_0}\int_{T_0} s^2(t;\phi)dt\right] \tag{19.2}$$

Since the second integral above is not a function of ϕ, it can be dropped. Taking the logarithm of the resulting expression, we can equivalently maximize the log-likelihood function given by

$$\ell(\phi) = \frac{2}{N_0}\int_{T_0} r(t)s(t;\phi)\,dt \tag{19.3}$$

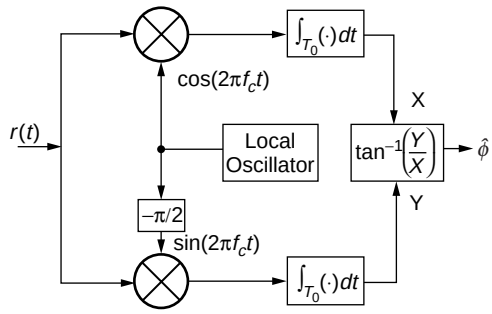


FIGURE 19.1 ML synchronizer: open-loop.

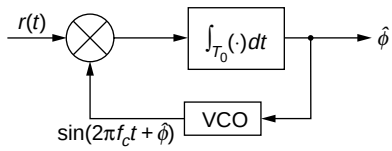


FIGURE 19.2 Phase-locked loop estimator.

Differentiating with respect to ϕ and setting to zero, we obtain the following necessary condition for the maximum-likelihood estimate $\hat{\phi}$:

$$\int_{T_0} r(t)\sin(2\pi f_c t - \hat{\phi})dt = 0 \tag{19.4}$$

Solving for $\hat{\phi}$ we obtain

$$\hat{\phi} = \tan^{-1}\left[\frac{\int_{T_0} r(t)\sin(2\pi f_c t)dt}{\int_{T_0} r(t)\cos(2\pi f_c t)dt}\right]$$

Figure 19.1 shows how the above estimator can be implemented in block-diagram form in what is referred to as an *open-loop* realization.

A *closed-loop* or tracking synchronizer that uses the optimality condition in Eq. (19.4) in a tracking loop referred to as a *phase-locked loop* (PLL) is shown in Fig. 19.2. In this figure, VCO stands for Voltage Controlled Oscillator and is a device that produces a sinusoid at the carrier frequency f_c and has an instantaneous frequency which is proportional to its input (or, equivalently, a phase that is the integral of its input). The integrator in Fig. 19.2 over the interval T_0 is a linear filter that, in general, can be modeled by an impulse response $g(t)$ or a transfer function $G(s)$ and is referred to as the *loop filter*.

Let us now investigate the performance of the PLL synchronizer. We have the following equation describing the PLL in Fig. 19.2:

$$\frac{d\phi(t)}{dt} = \left\{r(t)\sin[2\pi f_c t + \hat{\phi}(t)]\right\} * g(t) \tag{19.5}$$

where $*$ represents convolution. Substituting for $r(t)$ from Eq. (19.1) the product term becomes

$$r(t)\sin[2\pi f_c t + \hat{\phi}(t)] = \frac{A}{2}\left\{\sin[4\pi f_c t + \hat{\phi}(t) + \phi] + \sin[\hat{\phi}(t) - \phi]\right\} + n'(t) \tag{19.6}$$

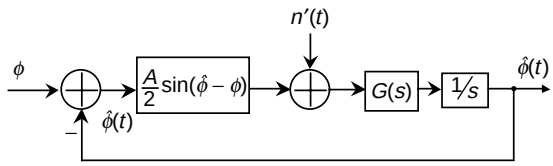


FIGURE 19.3 Equivalent PLL model for performance analysis.

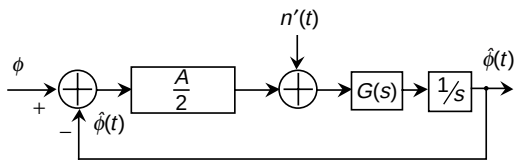


FIGURE 19.4 Linearized PLL model for performance computation.

where $n'(t)$ can be argued to be white and Gaussian with zero mean and spectral density $N_0/4$. Dropping the double-frequency term in Eq. (19.6), since it will be filtered out by the much lower bandwidth of the loop, we obtain the following model for the PLL

$$\frac{d\phi(t)}{dt} = \left[\frac{A}{2}\sin[\hat{\phi}(t) - \phi] + n'(t)\right] * g(t) \tag{19.7}$$

Equation (19.7) is a nonlinear stochastic differential equation that describes the evolution of the phase estimate and is modeled in Fig. 19.3. If we let $e(t) = [\hat{\phi}(t) - \phi]$ be the phase error, then it is easily seen (since ϕ is not a function of time) that the phase-error is described by

$$\frac{de(t)}{dt} = \left[\frac{A}{2}\sin[e(t)] + n'(t)\right] * g(t) \tag{19.8}$$

An analytical solution for the density of the error $e(t)$ in steady-state has been derived (see Viterbi [5]) for the special case when $G(s) = 1$ and is given by the *Tichonov density function*

$$p(e) = \frac{\exp[\alpha\cos(e)]}{2\pi I_0(\alpha)}, \quad -\pi \le e \le \pi \tag{19.9}$$

where $\alpha = 4A/N_0$ and $I_0(\cdot)$ is the zero-order modified Bessel function. For large α, i.e., large signal-to-noise ratios (SNRs), the variance of the phase-error computed from Eq. (19.9) can be approximated by $\sigma_e^2 \cong 1/\alpha$. Further, for large SNRs, the error will be small on the average, and so, in Eq. (19.8), $\sin(e) \cong e$. Under this approximation, Eq. (19.8) (and, consequently, Eq. (19.7)) become linear. Figure 19.4 shows the linearized model for a PLL.

For the linear model in Fig. 19.4, an expression for the variance of the estimation error in steady-state can be derived easily by computing the variance of the output ϕ when the input is just the noise $n'(t)$. The following expression for the variance of the estimation error can be derived (which is a good approximation to the actual variance at high SNRs when the linearized model is valid)

$$\sigma_e^2 = \frac{2N_0B_L}{A^2}, \tag{19.10}$$

where the parameter B_L is known as the (one-sided) *loop noise equivalent bandwidth* and is given by

$$B_L = \frac{1}{\max|H(f)|^2}\int_0^\infty |H(f)|^2\, df. \tag{19.11}$$

In Eq. (19.11), $H(s)$ is the closed-loop transfer function of the (linearized) loop and is given by

$$H(s) = \frac{AG(s)}{2s + AG(s)}. \tag{19.12}$$

A simple computation of the variance given by Eq. (19.9) for the special case when $G(s) = 1$ yields $1/\alpha$, as expected. The advantage of Eq. (19.10) is that it can be used for general $G(s)$, provided the signal-to-noise ratio is large enough for the linear approximation to hold. Note also that the estimation error becomes smaller with decreasing bandwidth B_L. Thus, one might be tempted to reduce B_L to zero. The problem in this case is that the transient performance of the loop degrades to the extent that it may take a longer period of time to achieve a small error. Also, in practice, the input phase ϕ is time-varying, in which case the loop bandwidth should be large enough so that the loop can track the changes in the input. We now turn our attention to finding the performance of the ML carrier synchronizer in Fig. 19.1.

We make use of the Cramer–Rao lower bound [4] that states that under some conditions, the variance of the estimation error for *any* unbiased estimate of a parameter ϕ is lower-bounded by

$$\sigma_e^2 = \frac{-1}{E\left[\dfrac{\partial^2 \ln p(r|\phi)}{\partial \phi^2}\right]}, \tag{19.13}$$

where $\ln p(r|\phi)$ is the log-likelihood function given by Eq. (19.3) for the phase estimation problem. Using Eq. (19.13), we obtain

$$\sigma_e^2 \geq \left\{\frac{2A}{N_0}\int_{T_0} E[r(t)]\cos(2\pi f_c t + \phi)\,dt\right\} = \frac{2N_0}{A^2}\frac{1}{2T_0}. \tag{19.14}$$

A brief computation shows that $1/T_0$ corresponds to the equivalent noise bandwidth of the integrator. Thus, if the equivalent noise bandwidths in Eqs. (19.10) and (19.14) are the same, the performances of the ML and the PLL synchronizers are the same (assuming that the bound in Eq. (19.14) is achieved closely). Next, we address carrier extraction from a modulated carrier, which, as pointed out earlier, is, in practice, the preferred approach for efficiency reasons.

Synchronization from a Modulated Carrier

A suboptimal approach, often used in practice, to extract carrier synchronization from a modulated signal is to first nonlinearly preprocess the signal to "wipe off" the modulation, and then follow that by, for example, a PLL, as described above. For M-ary phase-shift keying (PSK), the nonlinear preprocessing involves taking the Mth power of the signal. This has the effect of multiplying the carrier phase by M, and thus a subsequent division by M is needed to match the original carrier phase. For PSK signaling, a side effect of the power-law nonlinearity is to introduce a phase ambiguity, which must be resolved either by using pilot symbols or through the use of *differential encoding* of the data (information is conveyed by the change in the phase relative to the previous baud interval, rather than absolute phase), which is the case in practice.

Besides the suboptimal power-law technique, ML estimation can be used to suggest the optimal processing and possible approximations, which we study next. Let the modulated data be

$$r(t) = s(t;d,\phi) + n(t) \tag{19.15}$$

where d is a sequence of N modulation symbols. For simplicity, we will assume binary antipodal signals in which case each component, d_k, of d is either 1 or −1. If we let the baud rate be $1/T$ symbols/s, then the signal part in Eq. (19.15) can be expressed as

$$s(t;\, d,\phi) \;=\; A\sum_{k=0}^{N-1} d_k \cos(2\pi f_c t + \phi)\, p(t-kT), \tag{19.16}$$

where $p(t)$ is a baseband pulse which determines to a large extent the spectral content of the transmitted signal, and we are assuming a data window of length N. For simplicity, we will assume a unit height rectangular pulse next, but the results can be generalized to any arbitrary pulse shape. Assuming for the moment that the modulation sequence d is known, we have the following conditional likelihood function:

$$L(\phi) \;=\; \exp\left[\frac{2}{N_0}\int_0^{NT} r(t)s(t;d,\phi)dt\right] \;=\; \exp\left[\frac{2A}{N_0}\sum_{k=0}^{N-1} d_k \int_{kT}^{(k+1)T} r(t)\cos(2\pi f_c t + \phi)dt\right]. \tag{19.17}$$

All we need to do now is take the expectation of the conditional likelihood function in Eq. (19.17) with respect to the random modulation sequence. Assuming that each binary symbol occurs with probability 1/2 and that symbols are independent, we obtain the following log-likelihood function after dropping terms that are not functions of ϕ and taking the logarithm of the resulting expression:

$$\Lambda(\phi) \;=\; \sum_{k=0}^{N-1} \ln\cosh\left[\frac{2A}{N_0}\int_{kT}^{(k+1)T} r(t)\cos(2\pi f_c t + \phi)dt\right]. \tag{19.18}$$

A maximum likelihood estimator maximizes the above expression with respect to the phase ϕ. In practice, to reduce complexity we use the following substitutions, $\ln\cosh(x) \propto x^2$ and $\ln\cosh(x) \propto |x|$, valid for small and large signal-to-noise ratios, respectively. In the case of small SNRs, the log-likelihood simplifies to

$$\ell(\phi) \;=\; \sum\left[\int_{kT}^{(k+1)T} r(t)\cos(2\pi f_c t + \phi)dt\right]^2. \tag{19.19}$$

Taking the derivative of Eq. (19.19) with respect to ϕ, we obtain the following necessary condition for the ML estimate $\hat{\phi}$ of ϕ

$$\sum_{k=0}^{N-1}\int_{kT}^{(k+1)T} r(t)\cos(2\pi f_c t + \hat{\phi})dt \times \int_{kT}^{(k+1)T} r(t)\sin(2\pi f_c t + \hat{\phi})dt \;=\; 0. \tag{19.20}$$

A tracking loop that dynamically forces the condition in Eq. (19.20) is shown in Fig. 19.5. Note that the product of the two integrals effectively removes the modulation. In practice, the summation operator may be replaced by a digital filter that applies different weights to the past and present data in order to improve response.

To Explore Further: Carrier Acquisition for QAM Constellations

The need for high throughputs required by several high-speed applications (such as digital TV) has pushed system designers towards more throughput-efficient modulation schemes. Because of their relatively good performance, large QAM constellations are being used in many of these applications. One of the problems associated with their use is that of carrier acquisition, which, for efficiency reasons, must

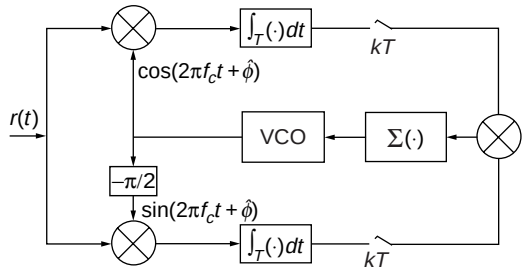


FIGURE 19.5 Carrier synchronization from modulated data.

often be done without the use of a preamble. The problem is further complicated for cross QAM constellations, for which the high SNR corner points used by some simple carrier phase estimators are not available. Clearly, due to the phase symmetry of QAM constellations, only phase offsets modulo $\pi/2$ are detectable, and differential encoding is used to resolve the ambiguity.

The phase synchronization problem is invariably divided into an acquisition and a tracking part. In many practical systems, tracking is done simply and efficiently in a decision-directed (DD) mode after acquisition has been established, and it is the acquisition problem that is the most troublesome, especially in applications where no preamble is allowed. For square QAM constellations, a simple technique for phase acquisition is based on detecting the signals at the four corners and using them to produce an estimate of the phase-offset which can be averaged in time to converge to a reliable estimate. The problem is more complicated for cross constellations which do not have the corner points.

We first look at the ML carrier phase estimator. Let

$$r_k = d_k e^{j\theta} + n_k$$

be the baud-rate samples of the output of a matched filter, where d_k is a complex number denoting the transmitted QAM symbol at time kT ($1/T$ is the signaling rate), and θ denotes the unknown phase-offset to be estimated; the n_k are complex, independent identically distributed (i.i.d.), zero-mean, Gaussian random variables with independent real and imaginary parts of variance σ^2, modeling the effects of noise. Without loss of generality, we assume that $E[d_k^2] = 1$(i.e., a unit average energy constellation), in which case the signal-to-noise ratio per symbol is SNR = $1/2\sigma^2$. Then the ML phase estimate of θ from data over a window of length N is easily obtained as the value of ϕ that maximizes the log-likelihood function

$$L(\phi) = \sum_{k=0}^{N} \ln\left[\sum_{d} \exp\left(-\frac{1}{2\sigma^2}\left|r_k - d \cdot e^{j\theta}\right|^2\right)\right],$$

where the inner summation is performed with respect to the data d present in the constellation. The complexity of the ML algorithm is due in part to this inner summation, which, even for small constellations, will require more computations than possible, especially for high-speed systems. Another complication in implementing the ML estimator is the need to solve a nonlinear maximization problem in order to find the ML estimate of ϕ. Some simplifications of the ML estimator can be obtained for square constellations, but they are not sufficient to bring the ML estimator into a practical form.

For square constellations, a simple algorithm can be used to extract the carrier phase by detecting the presence of one of the four corner points in the constellation. These points can be detected by setting an amplitude threshold between the peak amplitude of the corner points and the second largest amplitude. The angles of these four points can be expressed as

$$\phi^i = \frac{\pi}{4} + i \cdot \frac{\pi}{2}, \quad i = 0, 1, 2, 3.$$

Thus,

$$\phi^i = \frac{\pi}{4} \mod \left(\frac{\pi}{2}\right).$$

Since only phase rotations modulo $\pi/2$ are required, a simple estimator looks at the angle of the received sample r_k modulo $\pi/2$ and subtracts it from $\pi/4$. The result is the required estimate, which can be refined in time as more data are observed.

Another often used algorithm, which has been shown in [6] to be asymptotically the ML estimator in the limit of small SNRs, is the Mth power-law estimator, where $M = 4$ for QAM constellations, and it equals the size of the constellation for PSK signaling. For QAM signaling, the 4th power estimator extracts a phase estimate according to

$$\hat{\phi} = \frac{1}{4}\arg\left[E[d^{*4}] \cdot \sum_{k=1}^{N} r_k^4\right].$$

The approximate mean-square error performance of the above estimator was also obtained in [6]. More recently, another algorithm that seems to work well for both square and cross constellations has been reported in [7]. The algorithm referred to as the histogram algorithm (HA) does the following: (1) For each received sample r_k it finds the set of signals whose magnitude is closest to $|r_k|$. (2) Compute the angle of the subset of signals from step 1 belonging to the first quadrant. (3) Subtract each of the angles computed at step 2 from the angle of r_k. (4) Uniformly quantize the angle interval from 0 to 90 degrees into L bins, and associate a counter with each; increment the counters corresponding to the quantization intervals where the angles computed at step 3 fall in. (5) Repeat this process for new data r_k as they arrive. (6) When enough data is received, find the bin that has the largest counter value. The angle corresponding to this bin is produced as the phase estimate.

Figure 19.6 compares the performance of the ML and HA algorithms obtained through simulation to the Cramer–Rao bound for the 128-QAM (cross) constellation. Results for the 4th power estimator are shown separately in Fig. 19.7 since the performance of this estimator is about two orders of magnitude worse than the ML and HA algorithms. The reason behind the bad performance of the 4th power estimator is the existence of large self-noise, partly due to the absence of the corner points. These results indicate that the 4th power estimator is not an option for cross constellations, at least not for sizes greater than or equal to 128.

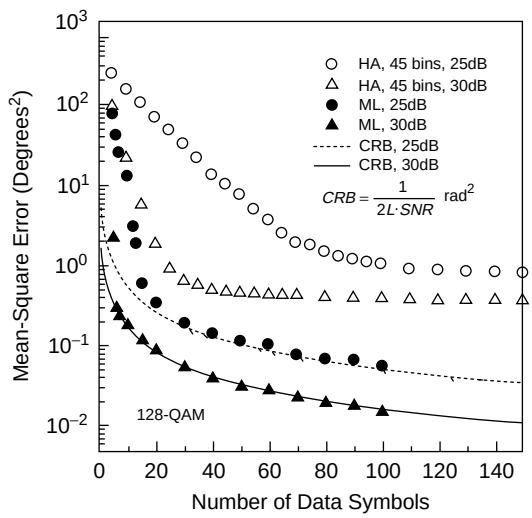


FIGURE 19.6 Mean square error for the various algorithms and 128-QAM.

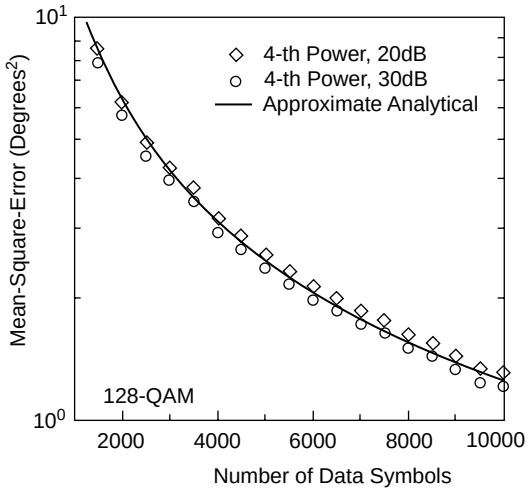


FIGURE 19.7 The performance of the 4th power estimator for 128-QAM.

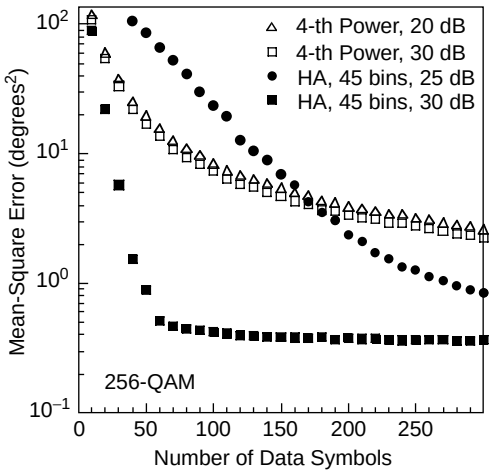


FIGURE 19.8 The HA vs. the 4th power algorithms for 256-QAM.

Figure 19.8 compares the HA and the 4th power estimators for the 256-QAM (square) constellation. As can be seen, the 4th power estimator performs much better with square constellations, and, in fact, it outperforms the HA for a range of data sequence lengths at 25dB SNR. As the SNR increases, however, the self-noise dominates and the performance of the 4th power estimator degrades.

In closing this section, we mention that an algorithm for joint estimation of carrier phase and frequency for 16-QAM input constellations has been proposed in [18] and shown to achieve the Cramer–Rao bound for SNRs greater than 15 dB.

19.3 Symbol Synchronization

We first investigate symbol synchronizers that are optimal in a maximum-likelihood sense. The ML symbol synchronizer can be used to suggest suboptimal but more easily implementable algorithms and provides a benchmark against which the performance of other synchronizers can be compared. Let T be the symbol duration and $r(t)$ be the received data. We assume an additive white Gaussian noise channel,

so $r(t)$ is modeled by

$$r(t) = s(t; d, \tau) + n(t), \tag{19.21}$$

where $n(t)$ is zero-mean white Gaussian noise having spectral density $N_0/2$, d is a sequence of modulation symbols d_k, $k = \cdots, -1, 0, 1, \cdots$, and τ is the timing-error. Assuming pulse-amplitude modulation (PAM), the signal can be described explicitly by

$$s(t; d, \tau) = \sum_k d_k p(t - kT - \tau), \tag{19.22}$$

where $p(t)$ is a baseband pulse. Our problem is to process the received signal $r(t)$ in order to obtain an estimate of the timing-error τ. To avoid loss in communication efficiency, we will do this in the presence of modulation symbols. For simplicity, we will assume a binary system with antipodal signals (one signal is just the negative of the other), in which case $d_k \in \{1, -1\}$.

As previously for carrier phase estimation, ML synchronization requires the probability density function of the data given the timing-error τ. Conditioned on knowing the data sequence d, the likelihood-function is given by Eq. (19.2) (with the signal part now given by Eq. (19.22)). Since the quadratic term does not depend on τ, the ML function can be reduced to

$$L(\tau, d) = \exp\left[\frac{2}{N_0}\sum_k d_k \int_{-\infty}^{\infty} r(t) p(t - kT - \tau)\, dt\right]. \tag{19.23}$$

All we need to do now is take the expectation of the above conditional likelihood function with respect to the data sequence to obtain the likelihood function. Performing the expectation and assuming independent and equiprobable data, we obtain (after taking the logarithm of the resulting expression) the required log-likelihood function

$$\ell(\tau) = \sum \ln \cosh\left[\frac{2 q_k(\tau)}{N_0}\right], \tag{19.24}$$

where

$$q_k(\tau) = \int_{-\infty}^{\infty} r(t) p(t - kT - \tau)\, dt. \tag{19.25}$$

A maximum-likelihood synchronizer finds the estimate $\hat{\tau}$, which maximizes the log-likelihood function in Eq. (19.24) based on the received data. Several problems arise as we try to implement this optimal synchronizer: (1) obtaining an ML estimate requires maximizing in real-time Eq. (19.24), an impossible task in most practical cases; (2) implementation of Eq. (19.24) requires knowledge of the signal-to-noise ratio, which is not readily available and must be estimated; and (3) there is no simple way to exploit Eq. (19.24) to extract timing estimates in real-time. The problem associated with (1) above can be partly alleviated by approximating the $\ln\cosh(\cdot)$ function for large and small SNRs, as was done before, which results in

$$\begin{aligned} \ell(\tau) &\cong \sum_k q_k^2(\tau), \quad \text{low SNR}, \\ \ell(\tau) &\cong \sum_k |q_k(\tau)|, \quad \text{high SNR}. \end{aligned} \tag{19.26}$$

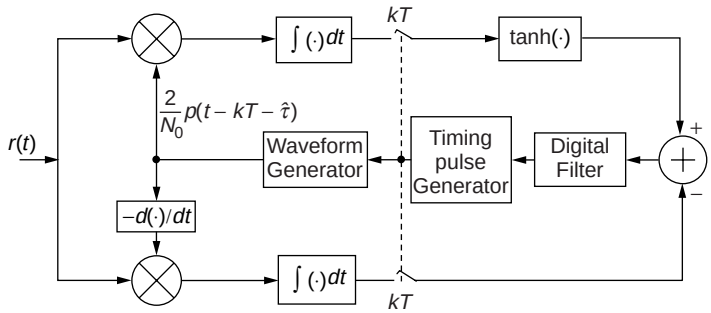


FIGURE 19.9 Closed-loop timing synchronizer.

In addition to simplifying the log-likelihood function, the approximations above also obviate the need for knowing the SNR. The need, however, to maximize a nonlinear function in real-time still exists.

As for the carrier synchronization problem, a number of open-loop realizations of the ML timing estimator exist. Perhaps the more interesting algorithms from a practical viewpoint are the closed-loop algorithms, two of which we motivate next. Taking the derivative of Eq. (19.24) with respect to τ and equating to zero results in an equation whose solution is the ML timing estimate:

$$\left.\frac{\partial \ell(\tau)}{\partial \tau}\right|_{\tau=\hat{\tau}} = \sum_k \left[\frac{2}{N_0}\int_{-\infty}^{\infty} r(t)\frac{\partial p(t-kT-\tau)}{\partial \tau}\,dt\right] \times \tanh\left[\frac{2}{N_0}\int_{-\infty}^{\infty} r(t)p(t-kT-\tau)\,dt\right] = 0, \quad (19.27)$$

where we have assumed that $p(-\infty) = p(\infty) = 0$. Note that if the timing τ is other than the ML estimate τ that makes the left-hand side of Eq. (19.27) equal to zero, the above derivative will be either positive or negative depending on the sign of the error $(\tau - \tau)$. Thus, the derivative can be used in a tracking loop to provide a correcting signal in a system that dynamically produces the ML timing estimate. Such a system is shown in Fig. 19.9, where the timing pulse generator adjusts the phase of the timing depending on the output of the accumulator once every T seconds. In practice, the accumulator may be replaced by a digital filter whose response is such that it puts more emphasis on recent data and less on past data. Clearly, when the timing-jitter is fast changing, better results may be obtained by having a short memory filter. On the other hand, if the timing-jitter is slowly varying, a filter with a long memory (low bandwidth) will yield better results.

Further simplifications to the above tracking synchronizer can be made under the assumptions of low or high signal-to-noise ratios, in which case Eq. (19.26) instead of Eq. (19.24) may be used. If the derivative of the likelihood function with respect to τ is approximated by the difference

$$\frac{\partial \ell(\tau)}{\partial \tau} \cong \frac{\ell(\tau+\delta/2) - \ell(\tau-\delta/2)}{\delta} \quad (19.28)$$

then, under the high SNR approximation (for example), Eq. (19.27) can be replaced by

$$\frac{\partial \ell(\tau)}{\partial \tau} \cong \sum_k \left[\left|\int_{-\infty}^{\infty} r(t)p(t-kT-\tau-\delta/2)dt\right| - \left|\int_{-\infty}^{\infty} r(t)p(t-kT-\tau+\delta/2)dt\right|\right] \quad (19.29)$$

A tracking-loop synchronizer, known as an *early-late gate symbol synchronizer*, implements Eq. (19.29) and is shown in Fig. 19.10. A similar synchronizer for a low SNR approximation can be implemented. The intuitive explanation of how early-late gate symbol synchronizers work is simple and can be easily illustrated for the case of non-return to zero (NRZ) pulses whose pulse shape $p(t)$ and autocorrelation function $a(t)$ are shown in Fig. 19.11.

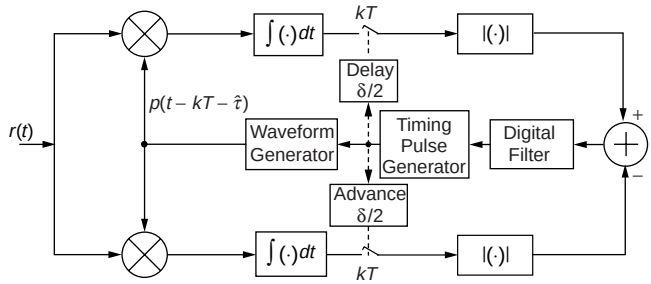


FIGURE 19.10 The early-late gate synchronizer.

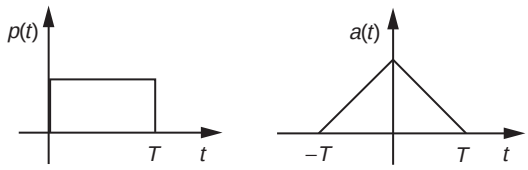


FIGURE 19.11 Example for NRZ pulses.

In the absence of timing-error, the receiver samples the output of the matched filter at the times corresponding to the peak of the autocorrelation function of the NRZ pulse (which results in the largest SNR). When a timing error exists, the samples occur at either side of the peak depending on whether the error is positive or negative. In either case, because of the symmetry of the autocorrelation function, the samples are of the same value (on the average). In an early-late gate synchronizer, two samples are taken, separated by δ seconds and centered around the current timing estimate. Depending on whether the error is positive or negative, the difference between the absolute values of these samples will be positive or negative, thus providing a control signal to increase or decrease τ in the desired direction to reduce the error. Notice that on the average, due to the even symmetry of the autocorrelation function, the difference between the absolute values of the two samples is zero at the actual symbol timing phase, i.e., when the timing error is zero. Thus, at least intuitively, the system in Fig. 19.10 is a stable loop for tracking the symbol timing phase. We end this section by noting that a number of practical implementations based on some further simplification of the synchronizers discussed above are used in practice.

To Browse Further

In modern receivers, more and more of the processing is done in the discrete domain, which allows for more complicated algorithms to be accurately implemented compared to analog implementations. For timing recovery, the preferred technique is to process samples taken at the output of a matched (or other suitable) filter at rates as low as the baud rate to a few samples per baud. The advantage of baud-rate sampling is that it uses the same samples that the detector uses to make symbol decisions, and it is the lowest rate possible, making the sampler less costly and the processing of samples faster. The disadvantage is that baud-rate sampling is below the Nyquist rate and, thus, acquisition performance tends to suffer somewhat. For bandlimited signaling, two or more samples per baud are at or above the Nyquist rate and, thus, all information contained in the original analog signal is preserved by the sampling process. This means that the sampler can be free-running without the need to adjust its sampling phase since that can be done in the discrete domain through interpolation.

Perhaps the most well known paper on timing-recovery from baud-rate samples is that by Mueller and Muller [8]. The timing algorithms studied therein are decision-directed (make use of tentative symbol decisions) and use the baud-rate samples in order to estimate the timing error. The timing error information is then used to adjust the phase of the sampler towards reducing the timing error. Other works that consider two or more samples per baud have been reported in [9–11].

19.4 Frame Synchronization

As noted in the introduction, frame synchronization is obtained in practice by locating at the receiver the position of a frame synchronization pattern (referred to also as a *marker*), periodically inserted in the data stream by the transmitter. In most systems, partly for simplicity and partly because the periodicity of the marker insertion makes it easily identifiable when enough frames are processed, the marker is not prevented from appearing in the random data stream. This means that it should be long enough compared to the frame length to make the probability of it appearing in the data small. If we let the synchronization pattern be of length L and the frame size (including the marker) be of length N, then the efficiency of such a system, measured by the number of data symbols per channel symbol, is

$$e = 1 - \frac{L}{N}. \tag{19.30}$$

The efficiency can be made arbitrarily close to one by increasing N for a fixed L or by decreasing L for a fixed N. In both cases, however, the probability of correctly detecting the position of the marker is reduced. In practice, good first pass acquisition probabilities can be achieved with efficiencies of about 97%. Figure 19.12 shows the contents of a frame.

As for the symbol synchronization case, we will first introduce the optimum (ML) frame synchronizer and then investigate some suboptimum synchronizers. For simplicity, we will only look at binary antipodal baseband signaling, although, in qualitative terms, similar results hold for non-binary systems. In the following, we assume that perfect symbol synchronization is present. As usual, we assume an additive white Gaussian noise channel, in which case the sufficient statistic (loosely, the simplest function of the data required by the optimum synchronizer) is the baud-rate samples of the output of a matched filter. Let $r = (r_1, r_2, \ldots, r_N)$ be the vector of the observed data obtained by sampling the output of a matched filter (matched to the baseband pulse) at the correct symbol rate and phase (no symbol timing-error), but not necessarily at the correct frame phase. Under a Gaussian noise assumption, the discrete matched filter samples can be modeled by

$$r_k = \sqrt{E} d_k + n_k, \tag{19.31}$$

where E is the signal energy, $d_k \in \{1, -1\}$ is the kth modulation symbol, and the n_ks constitute a sequence of independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and variance σ^2.

It is clear that since the frame length is N, there is exactly one frame marker within the observation window. Our problem is to locate the position $m \in (0, 1, 2, \ldots, N-1)$ of the marker from the observed data r. If m is the actual position of the marker, then the data vector corresponding to the observed vector r is

$$d = (d_1, d_2, \ldots, d_{m-1}, d_m, \ldots, d_{m+L-1}, \ldots, d_N),$$

where the L-symbol sequence starting at position m is the marker. If we denote the marker by S, this means

$$S = (s_1, s_2, \ldots, s_L) = (d_m, d_{m+1}, \ldots, d_{m+L-1}).$$

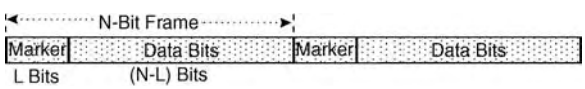


FIGURE 19.12 The composition of a frame.

For maximum-likelihood estimation of m, we need to maximize the following conditional density:

$$p(r|m) = \sum_{d'} p[r|m, d']\Pr(d') \tag{19.32}$$

where d' is the $(N - L)$-symbol data sequence that surrounds the marker. Assuming equiprobable symbols, Massey [12] derived the following log-likelihood function:

$$L(m) = \sum_{k=1}^{L} s_k r_{k+m} - \frac{\sigma^2}{\sqrt{E}} \sum \ln \cosh\left(\frac{\sqrt{E}}{\sigma^2} r_{k+m}\right). \tag{19.33}$$

To account for the periodicity of the marker, indices in Eq. (19.33) are interpreted modulo N. An optimal frame synchronizer computes the above expression for all values of m and chooses as its best estimate of the marker position the value that maximizes $L(m)$.

A few observations are now in order regarding the above likelihood function. First, we note that $L(m)$ is the sum of two terms: a linear term and a nonlinear term. The first term can be recognized as the correlation between the received data r and the known marker; the second term can be interpreted as an energy correction term that accounts for the random data surrounding the marker.

For practical implementation, some approximations of the optimal rule can be obtained easily by approximating $\ln \cosh(\cdot)$. For high SNRs, by replacing $\ln\cosh(x)$ by $|x|$, we obtain

$$L(m) \cong \sum_{k=1}^{L} s_k r_{k+m} - \sum_{k=1}^{L} |r_{k+m}|. \tag{19.34}$$

For low SNRs, replacing $\ln \cosh(x)$ by $x^2/2$, the optimal rule becomes

$$L(m) \cong \sum_{k=1}^{L} s_k r_{k+m} - \frac{\sqrt{E}}{2\sigma^2} \sum_{k=1}^{L} r_{k+m}^2. \tag{19.35}$$

A further approximation that is quite often used is to drop the second nonlinear term altogether. The resulting rule then becomes

$$L(m) \cong \sum_{k=1}^{L} s_k r_{k+m}, \tag{19.36}$$

and is known as the *simple correlation rule* for obvious reasons. The high SNR approximation and the simple correlation rule have the added advantage that no knowledge of the SNR is needed for implementation, compared to the optimum and the low SNR approximation rules.

Practical frame synchronizers use the periodicity of the marker in order to improve performance in time and usually include algorithms for detecting loss of synchronization in which re-acquisition is initiated. The above algorithms can be used as the basis for these practical synchronizers in estimating the marker position from a frame's-worth of data. Their performance in correctly identifying the marker position significantly affects the overall performance of the frame synchronizer as measured not only by the probability of correct acquisition, but also by the time it takes for the algorithm to acquire.

Other techniques for marker acquisition (besides those based on the maximum-likelihood principle) can be used as well. For example, in some practical implementations of the simple correlation rule, sequential detection of the marker is often implemented: the correlation of the marker with the data is

computed sequentially for each frame position and the result is compared to some threshold. When for some value of m the computed correlation exceeds the threshold, the frame synchronizer declares a marker presence. Otherwise the search continues. The value of the threshold is critical for performance and it is usually chosen to minimize the time to marker acquisition.

Another important aspect of frame synchronization design is the design of good marker sequences. Although the above algorithms work with any chosen sequence, the resulting performance of the synchronizer depends critically on the sequence used. In general, sequences that have good autocorrelation properties perform well as frame markers. These sequences have the property that their autocorrelation function is uniformly small for all shifts other than the zero shift. Examples of such sequences include the *Barker sequences* [13] and the *Neuman–Hofman sequences* [14]. Barker sequences are binary sequences whose largest side-lobe (nonzero shift correlation) is at most 1. Unfortunately, the largest known Barker sequence is of length 13, and there is proof that no Barker sequences of length between 14 and 6084 exist. In many cases, however, there is a need for larger sequences to improve performance. Neuman–Hofman sequences were specifically designed to maximize performance when a simple correlation rule is used. Thus, these sequences perform somewhat better than Barker sequences when a correlation rule is used. What is more important, though, is that Neuman–Hofman sequences of large length exist. Examples of Barker and Neuman–Hofman sequences of length seven and thirteen are

$$(1, -1, 1, 1, -1, -1, -1), \quad \text{Barker}, L = 7,$$

$$(1, 1, 1, 1, 1, -1, -1, 1, 1, -1, 1, -1, 1), \quad \text{Barker}, L = 13,$$

$$(-1, -1, -1, -1, -1, -1, 1, 1, -1, -1, 1, -1, 1), \quad \text{Neuman–Hofman}, L = 13.$$

Performance

We address briefly the performance of the ML and two suboptimal synchronizers given above, as measured by the probability of erroneous synchronization. First, we look at the question of how well *any* frame synchronizer can perform, as a function of the frame length N, marker length L, and SNR. Clearly, in the limit of infinite SNR we obtain the best performance (smallest probability of erroneous marker detection). In this case, an error can be made when the marker appears randomly in one or more positions in the random data part. For *bifix-free* sequences (i.e., sequences for which no prefix is also a suffix), Nielsen [15] has obtained the following expression for probability of erroneous synchronization in this case:

$$P_{LB} = \sum_{k=1}^{R} \frac{(-1)^{k+1}}{k+1} \cdot \binom{N-L-k(L-1)}{k} \cdot M^{-kL}, \tag{19.37}$$

where

$$R = \left\lfloor \frac{N-L}{L} \right\rfloor, \tag{19.38}$$

and M is the size of the modulation ($M = 2$ for binary signaling). The bifix-free condition guarantees that no partial overlap of the marker with itself results in a perfect match for the overlapped parts. Figure 19.13 shows simulation results for the performance of the ML, high SNR approximation, and simple-correlation rules of Eqs. (19.33), (19.34) and (19.36), respectively. Also shown is the lower-bound in Eq. (19.37), which is achieved by the ML rule and its high SNR approximation. On the other hand, the simple-correlation rule performs significantly worse.

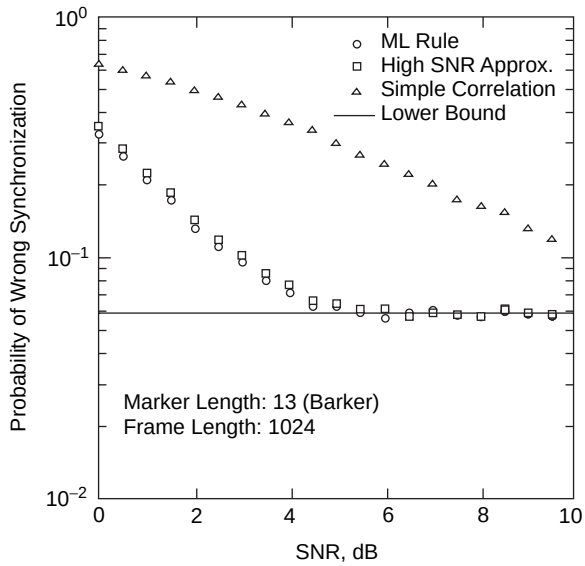


FIGURE 19.13 The performance of the ML and other rules.

References

1. J.J. Stiffler, *Theory of Synchronous Communications,* Prentice-Hall, Englewood Cliffs, NJ, 1971.
2. W. Lindsey and M. Simon, *Telecommunication Systems Engineering,* Prentice-Hall, Englewood Cliffs, NJ, 1973.
3. H. Meyr and G. Ascheid, *Synchronization in Digital Communications,* John Wiley & Sons, New York, 1990.
4. H.L. Van Trees, *Detection, Estimation, and Modulation Theory,* John Wiley & Sons, New York, 1968.
5. A. Viterbi, *Principles of Coherent Communication,* McGraw-Hill, New York, 1966.
6. M. Moeneclaey and G. de Jonghe, ML-oriented NDA carrier synchronization for general rotationally symmetric signal constellations, *IEEE Trans. Commun.,* 42, 2531–2533, 1994.
7. C.N. Georghiades, Blind carrier phase acquisition for QAM constellations, *IEEE Trans. Commun.,* 45, 1477–1486, 1997.
8. K.H. Mueller and M. Muller, Timing recovery in digital synchronous data receivers, *IEEE Trans. Commun.,* COM-24, 516–531, 1976.
9. O. Agazzi, C.P.J. Tzeng, D.G. Messerschmitt, and D.A. Hodges, Timing recovery in digital subscriber loops, *IEEE Trans. Commun.,* COM-33, 558–569, 1985.
10. F.M. Gardner, A BPSK/QPSK timing-error detector for sampled receivers, *IEEE Trans. Commun.,* COM-34, 423–429, 1986.
11. C.N. Georghiades and M. Moeneclaey, Sequence estimation and synchronization from nonsynchronized samples, *IEEE Trans. Inf. Theory,* 37, 1649–1657, 1991.
12. J.L. Massey, Optimum frame synchronization, *IEEE Trans. Commun.,* COM-20, 115–119, 1972.
13. R.H. Barker, Group synchronization of binary systems, in *Communication Theory,* W. Jackson, Ed., London, 273–287, 1953.
14. F. Neuman and L. Hofman, New pulse sequences with desirable correlation properties, Proceedings of the National Telemetry Conference, Washington, D.C., April 1971, 272–282.
15. P.T. Nielsen, On the expected duration of a search for a fixed pattern in random data, *IEEE Trans. Inf. Theory,* 702–704, 1973.
16. H. Meyr, M. Moeneclaey, and S.A. Fechtel, *Digital Communication Receivers,* John Wiley & Sons, New York, 1998.

17. U. Mengali and A.N. D'Andrea, *Synchronization Techniques for Digital Receivers,* Plenum Press, New York, 1997.
18. M. Morelli, A.N. D'Andrea, and U. Mengali, Feedforward estimation techniques for carrier recovery in 16-QAM modulation, in M. Luise and S. Pupolin, Eds., *Broadband Wireless Communications,* Springer-Verlag, London, 1998.
19. F.M. Gardner, Multipath Link and Modem Design for Satellite-Mobile Communications, Final Report, ESTEC Contract No. 5146/82/NL/GM(SC), Gardner Research Company, Palo Alto, CA, August 1983.
20. F.M. Gardner, On-board Processing for Mobile-Satellite Communications, Final Technical Report, ESTEC Contract No. 5889/84/NL/GM, Gardner Research Company, Palo Alto, CA, May 1985.
21. T. Alberty and V. Hespelt, A new pattern jitter free frequency error detector, *IEEE Trans. Commun.,* 37, 159–163, 1989.
22. M. Moeneclaey, Overview of Digital Algorithms for Carrier Frequency Synchronization, International ESA Workshop on DSP Techniques Applied to Space Communications, September 1992, 1.1–1.7.
23. F.M. Gardner, Frequency Detectors for Digital Demodulators via Maximum-Likelihood Derivation, Final Report: Part II, ESTEC Contract No. 8022/88/NL/DG, ESA, June 4, 1990.
24. A.N. D'Andrea and U. Mengali, Design of quadricorrelators for automatic frequency control systems, *IEEE Trans. Commun.,* 41, 988–997, 1993.
25. A.N. D'Andrea and U. Mengali, Noise performance of two frequency-error detectors derived from maximum likelihood estimation methods, *IEEE Trans. Commun.,* 42, 793–802, 1994.
26. A.N. D'Andrea and U. Mengali, Performance of a quadricorrelator driven by modulated signals, *IEEE Trans. Commun.,* 38, 1952–1957, 1990.
27. F. Classen and H. Meyr, Two frequency Estimation Schemes Operating Independently of Timing Information, Globecom'93 Conference, Houston, TX, November 29-December 2, 1993, p. 1996–2000.
28. F. Classen, H. Meyr, and P. Sehier, Maximum Likelihood Open Loop Carrier Synchronizer for Digital Radio, ICC'93, Geneva, May 1993, 493–497.
29. J.C.I. Chuang and N.R. Sollenberger, Burst coherent demodulation with combined symbol timing, frequency offset estimation, and diversity selection, *IEEE Trans. Commun.,* 39, 1157–1164, 1991.
30. M.P. Fitz, Planar Filtered Techniques for Burst Mode Carrier Synchronization, Globecom'91, Phoenix, AZ, December 1991.
31. M.P. Fitz, Further results in the fast estimation of a single frequency, *IEEE Trans. Commun.,* 42, 862–864, 1994.
32. M. Luise and R. Reggiannini, Carrier frequency recovery in all-digital modems for burst-mode transmissions, *IEEE Trans. Commun.,* 43, 1169–1178, 1995.
33. F.M. Gardner, Demodulator Reference Recovery Techniques Suited for Digital Implementation, Final Report, ESTEC Contract No. 6847/86/NL/DG, European Space Agency, August 1988.
34. F.M. Gardner, Properties of frequency difference detectors, *IEEE Trans. Commun.,* 33, 131–138, 1985.

Further Information

The proceedings of the International Communications Conference (ICC) and Global Telecommunications Conference (GLOBECOM) are good sources of current information on synchronization work. Other sources of archival value are the *IEEE Transactions on Communications* and the *IEEE Transactions on Information Theory.*

20

Digital Modulation Techniques

Ezio Biglieri
Politecnico di Torino

20.1 Introduction

The goal of a digital communication system is to deliver information represented by a sequence of binary symbols, through a physical channel, to a user. The mapping of these symbols into signals, selected to match the features of the physical channel, is called **digital modulation**.

The digital modulator is the device used to achieve this mapping. The simplest type of modulator has no memory, that is, the mapping of blocks of binary digits into signals is performed independently of the blocks transmitted before or after. If the modulator maps the binary digits 0 and 1 into a set of two different waveforms, then the modulation is called *binary*. Alternatively, the modulator may map symbols formed by h binary digits at a time onto $M = 2^h$ different waveforms. This modulation is called *multilevel*.

The general expression of a signal modulated by a modulator with memory is

$$v(t) = \sum_{n=-\infty}^{\infty} s(t - nT; \xi_n, \sigma_n)$$

where $\{s(t; i, j)\}$ is a set of waveforms, x_n is the symbol emitted by the source at time nT, and s_n is the state of the modulator at time n. If the modulator has no memory, then there is no dependence on s_n. A *linear* memoryless modulation scheme is one such that $s(t - nT; \mathsf{x}_n; \mathsf{s}_n) = \mathsf{x}_n s(t - nT)$.

Given a signal set $\{s_i(t)\}_{i=1}^{M}$ used for modulation, its compact characterization may be given in terms of a **geometric representation**. From the set of M waveforms we first construct a set of $N \leq M$ orthonormal

waveforms $\{\psi_i(t)\}_{i=1}^{N}$ (the Gram–Schmidt procedure is the appropriate algorithm to do this). Next we express each signal $s_k(t)$ as a linear combination of these waveforms,

$$s_k(t) = \sum_{i=1}^{N} s_{ki}\psi_i(t)$$

The coefficients of this expansion may be interpreted as the components of M vectors that geometrically represent the original signal set or, equivalently, as the coordinates in a Euclidean N-dimensional space of a set of points called the **signal constellation**.

20.2 The Challenge of Digital Modulation

The selection of a digital modulation scheme should be done by making the best possible use of the resources available for transmission, namely, bandwidth, power, and complexity, in order to achieve the reliability required.

Bandwidth

There is no unique definition of signal bandwidth. Actually, any signal $s(t)$ strictly limited to a time interval T would have an infinite bandwidth if the latter were defined as the support of the Fourier transform of $s(t)$. For example, consider the bandpass linearly modulated signal

$$v(t) = \Re\left[\sum_{n=-\infty}^{\infty} \xi_k s(t-nT)e^{j2\pi f_0 t}\right]$$

where $\Re$ denotes real part, f_0 is the carrier frequency, $s(t)$ is a rectangular pulse with duration T and amplitude 1, and (ξ_k) is a stationary sequence of complex uncorrelated random variables with $\mathbf{E}(\xi_n) = 0$ and $\mathbf{E}(|\xi_n|^2) = 1$. Then the power density spectrum of $v(t)$ is given by

$$\mathcal{G}(f) = \frac{1}{4}[G(-f-f_0) + G(f-f_0)]$$

where

$$G(f) = T\left[\frac{\sin \pi f T}{\pi f T}\right]^2 \tag{20.1}$$

The function $G(f)$ is shown in Fig. 20.1.

The following are possible definitions of the bandwidth:

- *Half-power bandwidth*. This is the interval between the two frequencies at which the power spectrum is 3 dB below its peak value.
- *Equivalent noise bandwidth*. This is given by

$$B_{\text{eq}} = \frac{1}{2}\frac{\int_{-\infty}^{\infty}\mathcal{G}(f)\mathrm{d}f}{\max_f \mathcal{G}(f)}$$

  This measures the basis of a rectangle whose height is $\max_f \mathcal{G}(f)$ and whose area is one-half of the power of the modulated signal.
- *Null-to-null bandwidth*. This represents the width of the main spectral lobe.

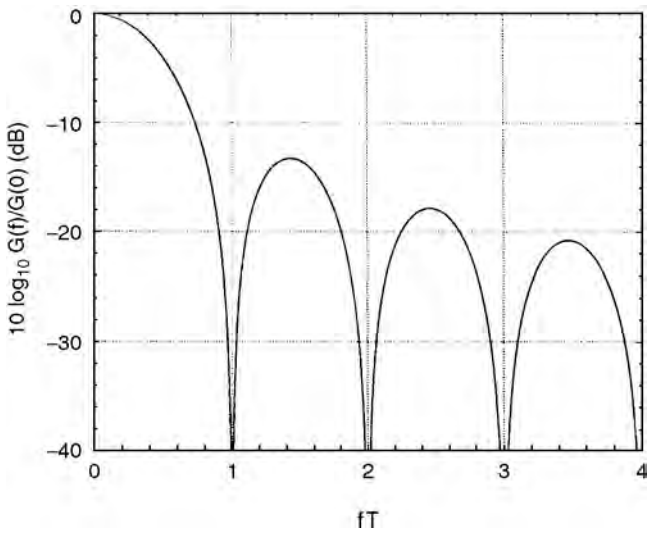


FIGURE 20.1 Power density spectrum of a linearly modulated signal with rectangular waveforms.

- *Fractional power containment bandwidth*. This bandwidth definition states that the occupied bandwidth is the band that contains $(1 - \varepsilon)$ of the total signal power.
- *Bounded-power spectral density bandwidth*. This states that everywhere outside this bandwidth the power spectral density must fall at least a certain level (e.g., 35 or 50 dB) below its maximum value.

Although the actual value of the signal bandwidth depends on the definition that has been accepted for the specific application, in general, we can say that

$$B = \frac{\alpha}{T}$$

where T is the duration of one of the waveforms used by the modulator, and α reflects the definition of bandwidth and the selection of waveforms. For example, for Eq. (20.1) the null-to-null bandwidth provides $B = 2/T$, that is, $\alpha = 2$. For 3-dB bandwidth, $\alpha = 0.88$. For equivalent-noise bandwidth, we have $\alpha = 1$.

Shannon Bandwidth

To make it possible to compare different modulation schemes in terms of their bandwidth efficiency, it is useful to consider the following definition of bandwidth. Consider a signal set and its geometric representation based on the orthonormal set of signals $\{\psi_i(t)\}_{i=1}^{N}$ defined over a time interval with duration T. The value of N is called the *dimensionality* of the signal set. We say that a real signal $x(t)$ with Fourier transform $X(f)$ is time limited to the interval $-T/2 < t < T/2$ at level ϵ if

$$\int_{|t|>T/2} x^2(t)\,\mathrm{d}t < \epsilon$$

and is bandlimited with bandwidth B at level ϵ if

$$\int_{|f|>B} |X(f)|^2\,\mathrm{d}f < \epsilon$$

Then for large BT the space of signals that are time limited and bandlimited at level ϵ has dimensionality $N = 2BT$ [Slepian, 1976]. Consequently, the Shannon bandwidth [Massey, 1995] of the signal set is defined as

$$B = \frac{N}{2T}$$

and is measured in dimensions per second.

Signal-to-Noise Ratio

Assume from now on that the information source emits independent, identically distributed binary digits with rate R_s digits per second, and that the transmission channel adds to the signal a realization of a white Gaussian noise process with power spectral density $N_0/2$.

The rate, in bits per second, that can be accepted by the modulator is

$$R_s = \frac{\log_2 M}{T}$$

where M is the number of signals of duration T available at the modulator, and $1/T$ is the signaling rate. The average signal power is

$$\mathcal{P} = \frac{\mathcal{E}}{T} = \mathcal{E}_b R_s$$

where $\mathcal{E}$ is the average signal energy and $\mathcal{E}_b = \mathcal{E}/\log_2 M$ is the energy required to transmit one binary digit. As a consequence, if B denotes the bandwidth of the modulated signal, the ratio between signal power and noise power is

$$\frac{\mathcal{P}}{N_0 B} = \frac{\mathcal{E}_b}{N_0}\frac{R_s}{B}$$

This shows that the **signal-to-noise ratio** is the product of two quantities, namely, the ratio $\mathcal{E}_b/N_0$, the energy per transmitted bit divided by twice the noise spectral density, and the ratio R_s/B, representing the *bandwidth efficiency* of the modulation scheme.

In some instances the *peak energy* $\mathcal{E}_p$ is of importance. This is the energy of the signal with the maximum amplitude level.

Error Probability

The performance of a modulation scheme is measured by its **symbol error probability** $P(e)$, which is the probability that a waveform is detected incorrectly, and by its *bit error probability,* or *bit error rate* (BER) $P_b(e)$, the probability that a bit sent is received incorrectly. A simple relationship between the two quantities can be obtained by observing that, since each symbol carries $\log_2 M$ bits, one symbol error causes at least one and at most $\log_2 M$ bits to be in error,

$$\frac{P(e)}{\log_2 M} \le P_b(e) \le P(e)$$

When the transmission takes place over a channel affected by additive white Gaussian noise, and the modulation scheme is memoryless, the symbol error probability is upper bounded as follows:

$$P(e) \le \frac{1}{2M}\sum_{i=1}^{M}\sum_{\substack{j=1\\ j\neq 1}}^{M} \operatorname{erfc}\left(\frac{d_{ij}}{2\sqrt{N_0}}\right)$$

where d_{ij} is the *Euclidean distance* between signals $s_i(t)$ and $s_j(t)$,

$$d_{ij}^2 = \int_0^T [s_i(t) - s_j(t)]^2 \mathrm{d}t$$

and erfc(·) denotes the Gaussian integral function

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-z^2} \mathrm{d}z.$$

Another function, denoted $Q(x)$, is often used in lieu of erfc(·). This is defined as

$$Q(x) = \frac{1}{2}\mathrm{erfc}\left(\frac{x}{\sqrt{2}}\right)$$

A simpler upper bound on error probability is given by

$$P(e) \le \frac{M-1}{2}\mathrm{erfc}\left(\frac{d_{\min}}{2\sqrt{N_0}}\right)$$

where $d_{\min} = \min_{i \ne j} d_{ij}$.

A simple lower bound on symbol error probability is given by

$$P(e) \ge \frac{1}{M}\mathrm{erfc}\left(\frac{d_{\min}}{2\sqrt{N_0}}\right)$$

By comparing the upper and the lower bound we can see that the symbol error probability depends exponentially on the term $d_{\min}$, the minimum Euclidean distance among signals of the constellation. In fact, upper and lower bounds coalesce asymptotically as the signal-to-noise ratio increases. For intermediate signal-to-noise ratios, a fair comparison among constellations should take into account the *error coefficient* as well as the minimum distance. This is the average number ν of nearest neighbors [i.e., the average number of signals at distance $d_{\min}$ from a signal in the constellation; for example, this is equal to 2 for M-ary **phase-shift keying (PSK)**, $M > 2$]. A good approximation to $P(e)$ is given by

$$P(e) \approx \frac{\nu}{2}\mathrm{erfc}\left(\frac{d_{\min}}{2\sqrt{N_0}}\right)$$

Roughly, at $P(e) = 10^{-6}$, doubling ν accounts for a loss of 0.2 dB in the signal-to-noise ratio.

20.3 One-Dimensional Modulation: Pulse-Amplitude Modulation (PAM)

Pulse-amplitude modulation (PAM) is a linear modulation scheme in which a signal $s(t)$ is modulated by random variables ξ_n taking on values in the set of amplitudes $\{a_i\}_{i=1}^M$, where

$$a_i = (2i - 1 - M)\frac{d_{\min}}{2}, \quad i = 1, 2, \ldots, M$$

The transmitter uses the set of waveforms $\{a_i s(t)\}_{i=1}^M$, where $s(t)$ is a unit-energy pulse. The geometric representation of this signal set is shown in Fig. 20.2 for $M = 4$.

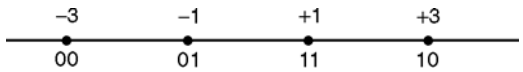


FIGURE 20.2 Quaternary PAM constellation.

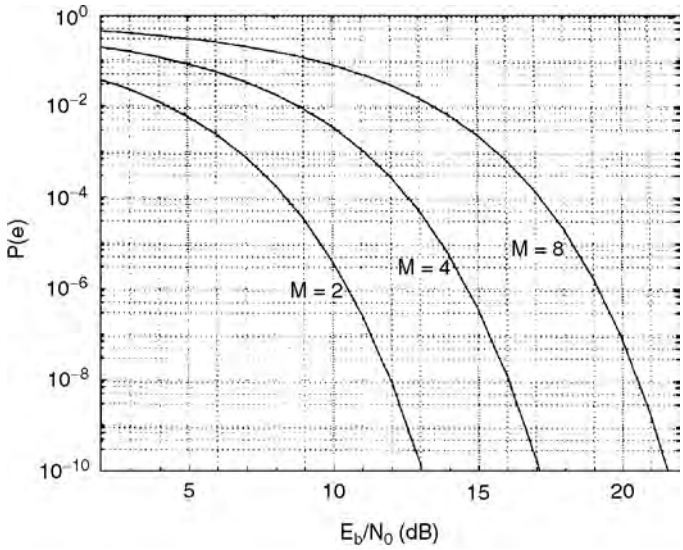


FIGURE 20.3 Symbol error probability of coherently demodulated PAM.

The symbol-error probability is given by

$$P(e) = \left(1 - \frac{1}{M}\right)\text{erfc}\left(\sqrt{\frac{3\log_2 M}{M^2 - 1}\frac{\mathcal{E}_b}{N_0}}\right)$$

The ratio between peak energy and average energy is

$$\frac{\mathcal{E}_p}{\mathcal{E}} = 3\frac{M-1}{M+1}$$

The bandwidth efficiency of this modulation is

$$\frac{R_s}{B} = \frac{\log_2 M}{T} \cdot 2T = 2\log_2 M$$

20.4 Two-Dimensional Modulations

Phase-Shift Keying (PSK)

This is a linear modulation scheme generating the signal

$$v(t) = \mathcal{R}\left\{\sum_{n=-\infty}^{\infty} \xi_n s(t - nT)e^{j2\pi f_0 t}\right\}$$

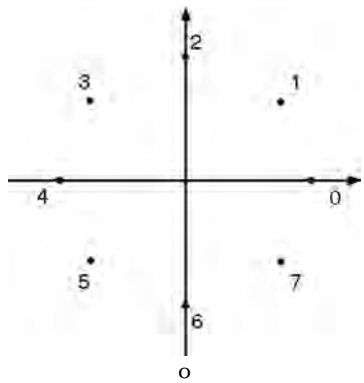


FIGURE 20.4 An 8-PSK constellation.

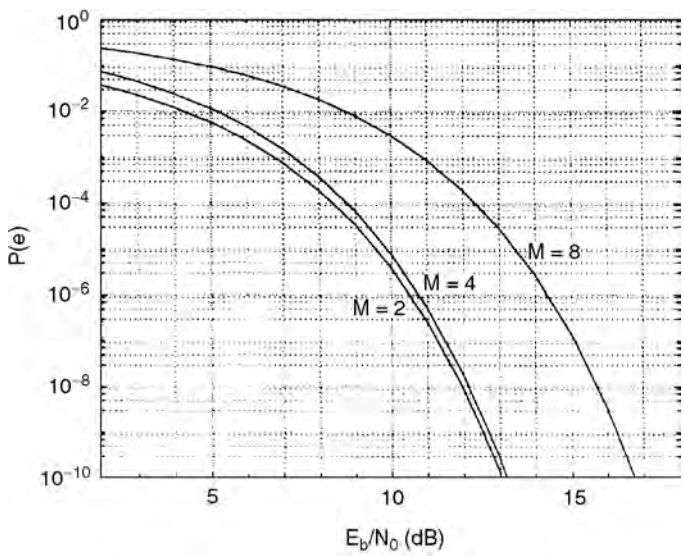


FIGURE 20.5 Symbol error probability of coherently demodulated PSK.

where $\xi_n = e^{j\phi_n}$ takes values in the set

$$\left\{\frac{2\pi}{M}(i-1)+\Phi\right\}_{i=1}^{M}$$

where Φ is an arbitrary phase. The signal constellation is shown, for $M = 8$, in Fig. 20.4.

The symbol-error probability of M-ary PSK is closely approximated by

$$P(e) \approx \operatorname{erfc}\left(\sqrt{\frac{\mathcal{E}_b}{N_0}\log_2 M}\sin\frac{\pi}{M}\right)$$

for high signal-to-noise ratios (see Fig. 20.5). The bandwidth efficiency of PSK is

$$\frac{R_s}{B} = \frac{\log_2 M}{T}\cdot T = \log_2 M$$

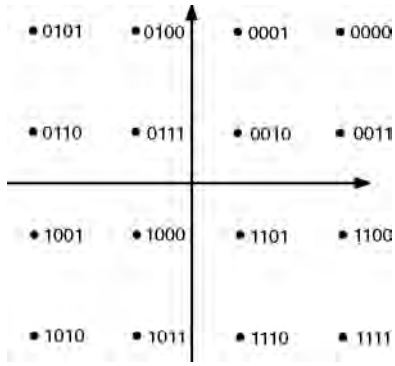


FIGURE 20.6 A 16-QAM constellation.

Quadrature Amplitude Modulation (QAM)

Quadrature amplitude modulation (QAM) is a linear modulation scheme for which the modulated signal takes the form

$$v(t) = \Re\left\{\sum_{n=-\infty}^{\infty} \xi_n s(t-nT)e^{j2\pi f_0 t}\right\}$$

where

$$\xi_n = \alpha_n + j\beta_n$$

and α_n, β_n take on equally spaced values. Figure 20.6 shows a QAM constellation with 16 points.

When $\log_2 M$ is an even integer, we have

$$P(e) = 1-(1-p)^2$$

with

$$p = \left(1-\frac{1}{\sqrt{M}}\right)\text{erfc}\left(\sqrt{\frac{3\log_2 M}{2(M-1)}\frac{\mathcal{E}_b}{N_0}}\right)$$

When $\log_2 M$ is odd, the following upper bound holds:

$$P(e) < 2\text{erfc}\left(\sqrt{\frac{3\log_2 M}{2(M-1)}\frac{\mathcal{E}_b}{N_0}}\right)$$

(See Fig. 20.7.)

20.5 Multidimensional Modulations: Frequency-Shift Keying (FSK)

In the modulation technique called **frequency-shift keying (FSK)** the transmitter uses the waveforms

$$s_i(t) = A\cos 2\pi f_i t, \quad 0 \le t < T$$

where $f_i = f_0 + (2i-1-M)f_d$, for $i = 1,\ldots, M$, and f_0 denotes the carrier frequency. The waveforms have

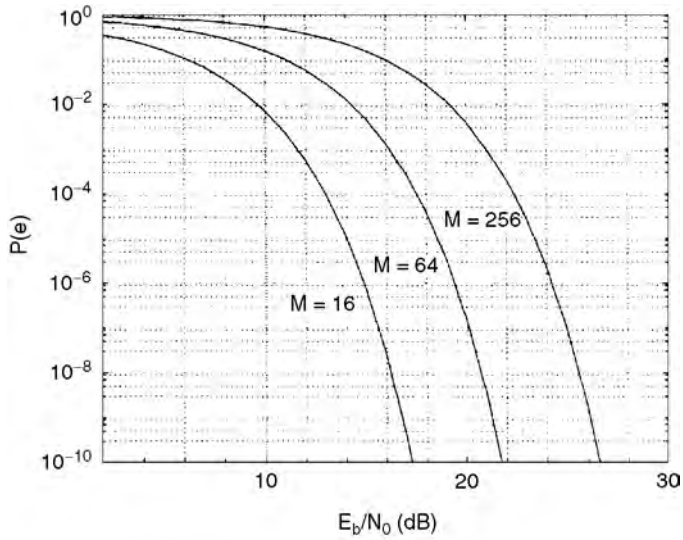


FIGURE 20.7 Symbol error probability of coherently demodulated QAM.

constant energy $\mathcal{E} = A^2\, T/2$, and the modulated signal has constant envelope A. The distance properties of these waveforms for a given energy depend on the value of their normalized frequency separation $2f_d\, T$.

Generation of these waveforms may be accomplished with a set of M separate oscillators, each tuned to the frequency f_i. In this case, the phase discontinuities at frequency-transition times cause large sidelobes in the power density spectrum of the modulated signal. A narrower spectrum can be obtained by using a single oscillator whose frequency is modulated by the source bits. This results in a modulation scheme with memory called continuous-phase FSK (discussed in Section 20.7).

The signals $s_i(t)$ are orthogonal, that is,

$$\int_0^T \cos(2\pi f_i t)\cos(2\pi f_j t)\mathrm{d}t = 0 \qquad i \neq j$$

if the frequency separation between adjacent signals is such that $2f_dT = m/2$, where m is any nonzero integer. Thus, the minimum frequency separation for orthogonality if $2f_d = 1/2T$.

The symbol error probability for M orthogonal signals with energy $\mathcal{E}$ is upper and lower bounded by

$$\frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{\varepsilon_b \log_2 M}{2N_0}}\right) \leq P(e) \leq \frac{M-1}{2}\mathrm{erfc}\left(\sqrt{\frac{\varepsilon_b \log_2 M}{2N_0}}\right).$$

For $M = 2$ the bounds merge into the exact value. The upper bound to $P(e)$ is plotted in Fig. 20.8.

It can be observed from Fig. 20.8 that the error probability for a given value of signal-to-noise ratio decreases as M increases, contrary to what is observed for PAM, PSK, and QAM. On the other hand, the bandwidth efficiency *decreases* as M increases, its value being given by

$$\frac{R_s}{B} = \frac{2\log_2 M}{M}$$

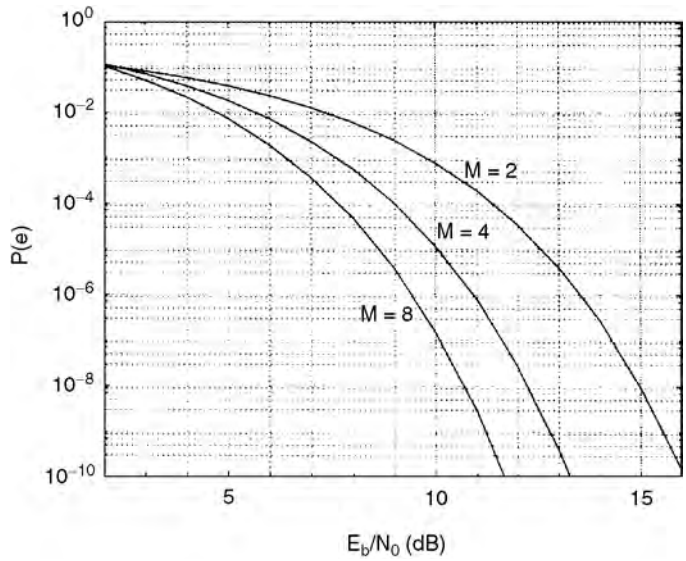


FIGURE 20.8 Symbol error probability of coherently demodulated orthogonal FSK.

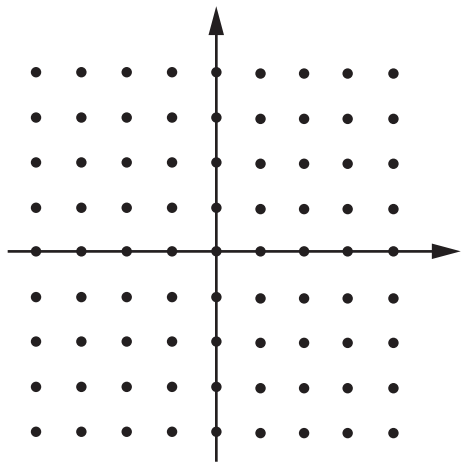


FIGURE 20.9 The lattice Z^2.

20.6 Multidimensional Modulations: Lattices

The QAM constellations may be considered as carved from an infinite constellation of regularly spaced points in the plane, called the **lattice** Z^2. This is represented in Fig. 20.9. This concept can be generalized to other lattices, both two-dimensional and multidimensional. In this framework, the design of a signal constellation consists of choosing a lattice, then carving a finite constellation of signal points from the lattice.

In general, a lattice Λ is defined as a set closed under ordinary addition and multiplication by integers. As such, it forms an additive group as well as a vector space. Specifically, an N-dimensional lattice Λ is an infinite set of N vectors.

If $d_{\min}$ is the minimum distance between any two points in the lattice, the number ν of adjacent lattice points located at distance $d_{\min}$, that is, the number of nearest neighbors of any lattice point, is called its *kissing number*.

Examples of Lattices

The Lattice Z^N

The set Z^N of all N-tuples with integer coordinates is called the *cubic lattice*, or *integer lattice*. The minimum distance is $d_{\min} = 1$, and the kissing number is $\nu = 2N$.

The N-Dimensional Lattices A_N and A_N^*

A_N is the set of all vectors with $(N+1)$ integer coordinates whose sum is zero. This lattice may be viewed as the intersection of Z^{N+1} and a hyperplane cutting the origin. The minimum distance is $d_{\min} = \sqrt{2}$, and the kissing number is $\nu = N(N+1)$.

The N-Dimensional Lattice D_N

D_N is the set of all N-dimensional points whose integer coordinates have an even sum. It may be viewed as a punctured version of Z^N, in which the points are colored alternately black and white with a checkerboard coloring, and the white points (those with odd sums) are removed. We have $d_{\min} = \sqrt{2}$, and $\nu = 2N(N-1)$. D_4 represents the densest lattice packing in R^4. This means that if unit-radius, four-dimensional spheres with centers in the lattice points are used to pack R^4, then D_4 is the lattice with the largest number of spheres per unit volume.

The Gosset Lattice E_8

In the *even coordinate system*, E_8 consists of the points

$$\left\{(x_1,\ldots,x_8):\quad \forall x_i \in Z \quad \text{or} \quad \forall x_i \in Z+\frac{1}{2},\quad \sum_{i=1}^{8} x_i \equiv 0 \bmod 2\right\}$$

In words, E_8 consists of the eight-vectors whose components are all integers, or all halves of odd integers, and whose sum is even.

The *odd* coordinate system is obtained by changing the sign of any coordinate: the points are

$$\left\{(x_1,\ldots,x_8):\quad \forall x_i \in Z \quad \text{or} \quad \forall x_i \in Z+\frac{1}{2},\quad \sum_{i=1}^{8} x_i \equiv 2x_8 \bmod 2\right\}$$

This lattice has $d_{\min} = \sqrt{2}$ and $\nu = 240$.

Other Lattices

The description and the properties of other important lattices, such as the 16-dimensional Barnes–Wall lattice Λ_{16} and the 24-dimensional Leech lattice Λ_{24}, can be found in Chapter 4 of Conway and Sloane [1988].

The Coding Gain of a Lattice

The *coding gain* $\gamma_c(\Lambda)$ of the lattice Λ is defined as

$$\gamma_c(\Lambda) = \frac{d_{\min}^2(\Lambda)}{V(\Lambda)^{2/N}} \tag{20.2}$$

where $V(\Lambda)$ is the *fundamental lattice volume*, that is, the reciprocal of the number of lattice points per unit volume [for example, $V(Z^N) = 1$]. Table 20.1 lists the coding gains of some of the most popular lattices.

TABLE 20.1 Coding Gains of Lattices

Name	Λ	N	$\gamma_c(\Lambda)$, (dB)
Integer lattice	Z	1	0.00
Hexagonal lattice	A_2	2	0.62
Schläfli	D_4	4	1.51
Gosset	E_8	8	3.01
Barnes-Wall	Λ_{16}	16	4.52
Leech	Λ_{24}	24	6.02

Carving a Signal Constellation out of a Lattice

We denote with $C(\Lambda, \mathcal{R})$ a constellation obtained from Λ (or from its translate $\Lambda + a$) by retaining only the points that fall in the region $\mathcal{R}$. The resulting constellation has a number of signals approximately equal to the ratio between the volume of $\mathcal{R}$ and the volume of Λ,

$$M \approx \frac{V(\mathcal{R})}{V(\Lambda)}$$

provided that $V(\mathcal{R}) \gg V(\Lambda)$, that is, that M is large enough.

The *figure of merit* (CFM) of $C = C(\Lambda, \mathcal{R})$ is the ratio between $d^2_{\min}$ and the average energy of the constellation *per dimension pair*,

$$\mathrm{CFM}(C) = \frac{d^2_{\min}}{\mathcal{E}/(N/2)}$$

To express the figure of merit of the constellation $C(\Lambda, \mathcal{R})$ in terms of parameters related to the lattice Λ and to the region $\mathcal{R}$, we introduce the definition of the *shape gain* $\gamma_s(\mathcal{R})$ of the region $\mathcal{R}$ [Forney and Wei, 1990]. This is the reduction in average energy (per dimension pair) required by a constellation bounded by $\mathcal{R}$ compared to that which would be required by a constellation bounded by an N-dimensional cube of the same volume $V(\mathcal{R})$. In formulas, the shape gain is the ratio between the *normalized second moment* of any N-dimensional cube (which is equal to 1/12) and the normalized second moment of $\mathcal{R}$,

$$\gamma_s(\mathcal{R}) = \frac{1/12}{G(\mathcal{R})} \tag{20.3}$$

where

$$G(\mathcal{R}) = \frac{\int_{\mathcal{R}} \|\mathbf{r}\|^2 \mathrm{d}\mathbf{r}}{N V(\mathcal{R})^{1+2/N}} \tag{20.4}$$

The following result holds [Forney and Wei, 1990]: The figure of merit of the constellation $C(\Lambda, \mathcal{R})$ is given by

$$\mathrm{CFM}(C) \approx \gamma_0 \cdot \gamma_c(\Lambda) \cdot \gamma_s(\mathcal{R}) \tag{20.5}$$

where γ_0 is the figure of merit of the one-dimensional PAM constellation with the same bit rate (chosen as the baseline), $\gamma_c(\Lambda)$ is the coding gain of the lattice Λ [see Eq. (20.2)], and $\gamma_s(\mathcal{R})$ is the shaping gain of the region $\mathcal{R}$. The approximation holds for large constellations.

This result shows that, at least for large constellations, the gain from shaping by the region $\mathcal{R}$ is almost completely decoupled from the coding gain due to Λ. Thus, for a good design it makes sense to optimize separately $\gamma_c(\Lambda)$ (i.e., the choice of the lattice) and $\gamma_s(\mathcal{R})$ (i.e., the choice of the region).

Spherical Constellations

The maximum shape gain achieved by an N-dimensional region $\mathcal{R}$ is that of a sphere, for which

$$\gamma_s = \frac{\pi(n+1)}{6(n!)^{1/n}}$$

where $n = N/2$ and N is even. As $N \to \infty$, the γ_s approaches $\pi e/6$, or 1.53 dB. The last figure is the maximum achievable shaping gain. A problem with spherical constellations is that the complexity of the encoding procedure (mapping input symbols to signals) may be too high. The main goal of N-dimensional lattice-constellation design is to obtain a shape gain as close to that of the N sphere as possible, while maintaining a reasonable implementation complexity and other desirable constellation characteristics.

20.7 Modulations with Memory

The modulation schemes considered so far are *memoryless* in the sense that the waveform transmitted in one symbol interval depends only on the symbol emitted by the source in that interval. In contrast, there are modulation schemes with memory.

One example of a modulation scheme with memory is given by *continuous-phase frequency-shift keying* (CPFSK), which in turn is a special case of **continuous-phase modulation (CPM)**. To describe CPFSK, consider standard FSK modulation, whose signals are generated by separate oscillators. As a consequence of this generation technique, at the end of each symbol interval the phase of the carrier changes abruptly whenever the frequency changes, because the oscillators are not phase synchronized. Since spectral occupancy increases by decreasing the smoothness of a signal, these sudden phase changes cause spectrum broadening. To reduce the spectral occupancy of a frequency-modulated signal, one option is CPFSK, in which the frequency is changed continuously.

To describe CPFSK, consider the PAM signal

$$v(t) = \sum_{n=-\infty}^{\infty} \xi_n q(t - nT)$$

where (ξ_n) is the sequence of source symbols, taking on values $\pm 1, \pm 3, \ldots, \pm(M-1)$, and $q(t)$ is a rectangular pulse of duration T and area 1/2. If the signal $v(t)$ is used to modulate the frequency of a sinusoidal carrier, we obtain the signal

$$u(t) = A\cos\left[2\pi f_0 t + 2\pi h \int_{-\infty}^{t} v(\tau)\,\mathrm{d}\tau + \varphi_0\right]$$

here f_0 is the unmodulated-carrier frequency, φ_0 is the initial phase, and h is a constant called the *modulation index*. The phase shift induced on the carrier, that is,

$$\theta(t) = 2\pi h \int_{-\infty}^{t} v(\tau)\,\mathrm{d}\tau$$

turns out to be a continuous function of time t, so that a continuous-phase signal is generated. The trajectories followed by the phase, as reduced mod 2π, form the *phase trellis* of the modulated signal. Figure 20.10 shows a segment (for 7 symbol intervals) of the phase trellis of binary CPFSK with $h = 1/2$ [this is called *minimum-shift keying* or (MSK)].

With MSK, the carrier-phase shift induced by the modulation in the time interval $nT \le t \le (n+1)\,T$ is given by

$$\theta(t) = \theta_n + \frac{\pi}{2}\left(\frac{t - nT}{T}\right)\xi_n \qquad nT \le t \le (n+1)T$$

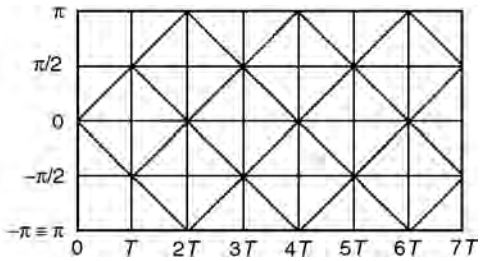


FIGURE 20.10 Phase trellis of binary CPFSK with $h = 1/2$ (MSK).

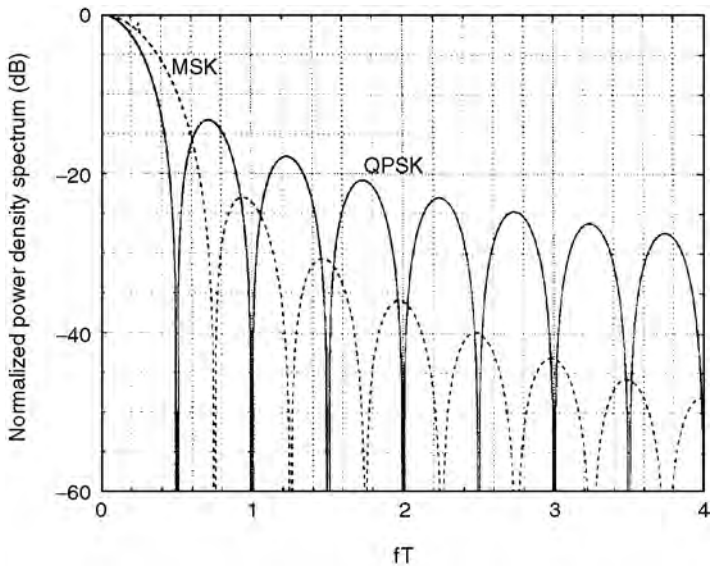


FIGURE 20.11 Power density spectrum of MSK and of quaternary PSK.

where

$$\theta_n = \frac{\pi}{2}\sum_{k=-\infty}^{n-1} \xi_k$$

The corresponding transmitted signal is

$$u(t) = A\cos\left[2\pi\left(f_0 + \frac{1}{4T}a_n\right)t - \frac{n\pi}{2}a_n + \theta_n\right]$$

which shows that MSK is an FSK using the two frequencies

$$f_1 = f_0 - \frac{1}{4T} \qquad f_2 = f_0 + \frac{1}{4T}$$

The frequency separation $f_2 - f_1 = 1/2T$ is the minimum separation for orthogonality of two sinusoids, which explains the name given to this modulation scheme.

The spectrum of CPFSK depends of the value of h. For $h < 1$, as the modulation index decreases, so does the spectral occupancy of the modulated signal. For MSK the power density spectrum is shown in Fig. 20.11.

CPM is a generalization of CPFSK. Here the carrier-phase shift in the interval $nT \le t \le (n+1)T$ is given by

$$\theta(t) = 2\pi \sum_{k=-\infty}^{n} a_k h_k q(t - nT)$$

where the waveform $q(t)$ is the integral of a frequency pulse $g(t)$ of arbitrary shape

$$q(t) = \int_{-\infty}^{t} g(\tau)\mathrm{d}\tau$$

subject to the constraints $g(t) = 0$ for $t < 0$ and

$$\int_{0}^{\infty} g(\tau)\mathrm{d}\tau = \frac{1}{2}$$

When the modulation index h_k varies (usually with a periodic law) from one symbol to another, the modulation is called *multi-h*. If $g(t) \neq 0$ for $t > T$, then the modulated signal is called *partial-response* CPM. Otherwise, it is called *full-response* CPM.

An example of partial-response CPM is given by Gaussian (GMSK), whose frequency pulse is obtained by passing a rectangular waveform into a low-pass filter whose impulse response $h(t)$ has a Gaussian shape,

$$h(t) = \sqrt{\frac{2\pi}{\ln 2}} B \exp\left\{-\frac{2\pi^2 B^2}{\ln 2} t^2\right\}$$

where B denotes the filter bandwidth. By decreasing B, the shape of the frequency pulse becomes smoother and the spectrum occupancy of the modulated signal is reduced. This modulation scheme is employed in the Global System for Mobile Communications (GSM) standard for cellular mobile radio with the choice $BT = 0.3$.

Defining Terms

Bandwidth: The frequency interval in which the density spectrum of a signal is significantly different from zero.

Continuous-phase modulation: A digital modulation scheme derived from phase-shift keying in which the carrier has no phase jumps, that is, the phase transitions occur continuously.

Digital modulation: The mapping of information-source symbols into signals, performed to carry information through the transmission channel.

Error probability: Probability that a symbol emitted by the information source will be received incorrectly by the end user.

Frequency-shift keying: A digital modulation scheme in which the source information is carried by the frequency of a sinusoidal waveform.

Geometric signal representation: Representation of a finite set of signals as a set of vector.

Lattice: An infinite signal constellation of points regularly located in space.

Phase-shift keying: A digital modulation scheme in which the source information is carried by the phase of a sinusoidal waveform, called the carrier.

Pulse-amplitude modulation: A digital modulation scheme in which the source information is carried by the amplitude of a waveform.

Quadrature amplitude modulation: A digital modulation scheme in which the source information is carried by the amplitude and by the phase of a sinusoidal waveform.

Signal constellation: A set of signals geometrically represented in the form of a set of vectors.
Signal-to-noise ratio: The ratio of the signal power and noise power. It is an index of channel quality.

References

Anderson, J.B., Aulin, T., and Sundberg, C.E.W. 1986. *Digital Phase Modulation,* Plenum, New York.

Benedetto, S., Biglieri, E., and Castellani, V. 1987. *Digital Transmission Theory,* Prentice–Hall, Englewood Cliffs, NJ.

Conway, J.H. and Sloane, N.J.A. 1988. *Sphere Packings, Lattices and Groups,* Springer–Verlag, New York.

Forney, G.D. Jr. and Wei, L.F. 1990. Multidimensional constellations—Part I: Introduction, figures of merit, and generalized cross constellations. *IEEE J. Selected Areas Comm.,* 7(6):877–892.

Massey, J.L. 1995. Towards an information theory of spread-spectrum systems. In: *Code Division Multiple Access Communications,* Eds. S.G. Glisic and P.A. Leppänen, Kluwer Academic, Boston.

Proakis, J.G. 1995. *Digital Communications,* 3rd ed., McGraw–Hill, New York.

Simon, M.K., Hinedi, S.M., and Lindsey, W.C. 1995. *Digital Communication Techniques. Signal Design and Detection,* Prentice–Hall, Englewood Cliffs, NJ.

Slepian, D. 1976. On bandwidth. *IEEE Proc.,* 64(2):292–300.

Further Information

The monthly journal *IEEE Transactions on Communications* reports advances in digital modulation techniques.

The books by Benedetto, Biglieri, and Castellani, 1987, Proakis, 1995, and Simon, Hinedi, and Lindsey, 1995 are good introductions to the theory of digital modulation.

II

Telephony

21

Plain Old Telephone Service (POTS)

A. Michael Noll
University of Southern California

21.1 Introduction

The acronym POTS stands for plain old telephone service. Over the years, POTS acquired an undeserved reputation for connoting *old-fashioned* and even *obsolete*. The telephone and the switched public network enable us to reach anyone anywhere on this planet at anytime and to speak to them in our natural voices. The minutes of telephone usage, in the number of access lines, and in the revenue of telephone companies, attest to the central importance of telephone service in today's information-age, global economy.

The word *old* in the acronym implies familiarity and ease of use, a major reason for the continued popularity of telephone service. The word *telephone* and the nuances expressed by natural human speech is what it is all about. *Service* used to mean responsiveness to the public, and it is sad that this dimension of the acronym has become so threatened by emphasis on short-term profits, particularly in the new world of competition that characterizes telephone service on all levels. The term *plain* is indeed obsolete, and the provision of today's telephone service utilizes some of the most sophisticated and advanced transmission and switching technology. In fact, telephone service today with all its intelligent and functional new features, advanced technology, and improved quality is truly fantastic. Allowing for a little misspelling, the acronym POTS can still be used, but with the "P" standing for phantastic! So, POTS it was, and POTS is still where the real action is.

Part of the excitement about the telephone network is that it can be used with a wide variety of devices to create exciting and useful services, such as facsimile for graphical communication and modems for access to the Internet. Cellular wireless telephony, pagers, and phones in airplanes extend telephony wherever we travel. Increased functionality in the network brings us voice mail, call forwarding, call waiting, caller ID, and a host of such services. The future of an ever-evolving POTS, indeed, looks bright and exciting.

Bell's Vision

The telegraph is older than the telephone, but special knowledge of Morse code was necessary to use the telegraph, thereby relegating it to use by specialists. The telephone uses normal human speech and, therefore, can be used by anyone and can convey all the nuances of the inflections of human speech.

Dr. Alexander Graham Bell was very prophetic in his vision of a world wired for two-way telecommunication using natural human speech. Bell and his assistant Thomas A. Watson demonstrated the first working model of a telephone on March 10, 1876, but Bell had applied for a patent a month earlier on February 14, 1876. Elisha Grey had submitted a disclosure of invention for a telephone on that same day in February, but ultimately the U.S. Supreme Court upheld Bell's invention, although in a split decision.

21.2 The Network

The telephone has come to signify a public switched network capable of reaching any other telephone on Earth. This switched network interconnects not only telephones but also facsimile machines, cellular telephones, and personal computers—anything that is connected to the network. As shown in Fig. 21.1, the telephones and other station apparatus in homes are all connected by pairs of copper wire to a single point, called the protector block, which offers simple protection to the network from overvoltages. A twisted pair of copper wires then connects the protector block all the way back to the central office. Many twisted pairs are all carried together in a cable that can be buried underground, placed in conduit, or strung between telephone poles. The twisted pair of copper wires connecting the station apparatus to the central office is called the local loop.

The very first stage of switching occurs at the central office. From there, telephone calls may be connected to other central offices over interoffice trunks. Calls may also be carried over much greater distances by connection to the long-distance networks operated by a number of **interexchange carriers (IXC)** such as AT&T, MCI, and Sprint. The point where the connection is made from the local service provider to the interexchange carrier is known as the **point of presence (POP)**.

The local portion of the telephone network is today known as a **local access and transport area (LATA)**. The local Bell telephone companies (commonly called the Baby Bells) were restricted by the Bell breakup of 1984 solely to the provision of intra-LATA service and were forbidden from providing inter-LATA service. Although the technology of telephony has progressed impressively over the last 100 years, policy and regulation have also had great impact on the telephone industry.

In the past, telephone service in the U.S. was mostly under the control of AT&T and its Bell system and a number of smaller independent telephone companies. Today, a number of competing companies

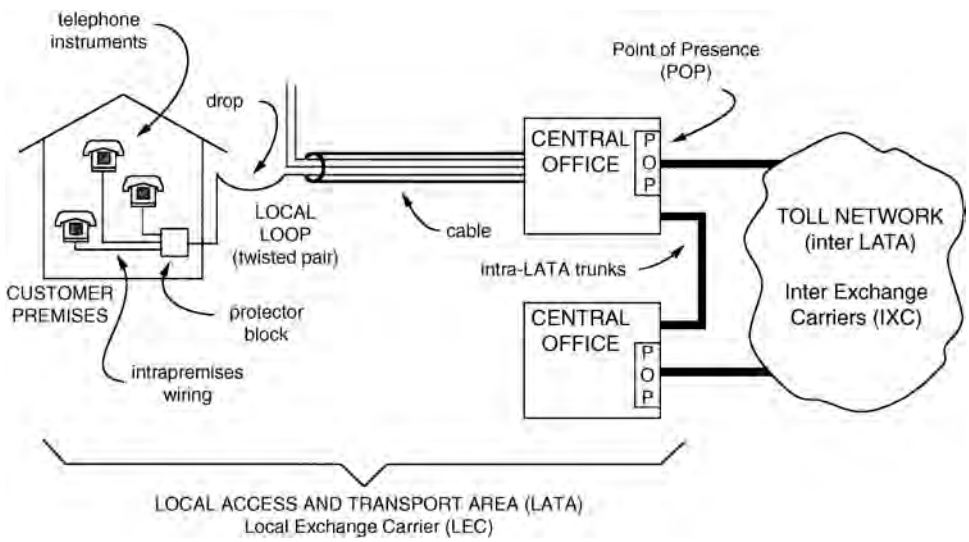


FIGURE 21.1 Network.

own and operate their own long-distance networks, and some futurists believe that local competition will occur soon.

The various long-distance and local networks all interconnect and offer what has been called *a network of networks*. Data communication for the Internet is carried over data networks. With voice being converted to digital bits, a technological convergence of voice and data (which is already encoded as bits) is occurring.

21.3 Station Apparatus

Telephones of the past were black with a rotary dial with limited functionality. Today's telephones come in many colors and sizes and offer push-button dialing along with a variety of intelligent features, such as repertory dialing and display of the dialed number. However, the basic functions of the telephone instrument have not changed and are shown in Fig. 21.2. The telephone needs to signal the central office when service is desired. This is accomplished by lifting the handset, which then closes contacts in the telephone—the switch hook—so that the telephone draws DC over the local loop from the central office. The switching machine at the central office senses this flow of DC and thus knows that the customer desires service. The common battery at the central office has an electromotive force (EMF) of 48 V, and the telephone draws at least about 20 mA of current over the loop. The maximum loop resistance can not exceed about 1300 Ω.

The user needs to specify the telephone number of the called party. This is accomplished by a process called dialing. Older telephones accomplished dialing with a rotary dial that interrupted the flow of DC with short pulses at a rate of about 10 dial pulses per second. Most telephones today use touch-tone dialing and push buttons, as shown in Fig. 21.3. When a specific digit is pushed, a unique combination of two sinusoidal tones is transmitted over the line to the central office. For example, an 8 is indicated by the combination of sine waves at 852 Hz and 1336 Hz. Filters are used at the switching machine at the central office to detect the frequencies of the tones and thus decode the dialed digits. Touch-tone dialing is also known as **dual-tone multifrequency** (**DTMF**) dialing.

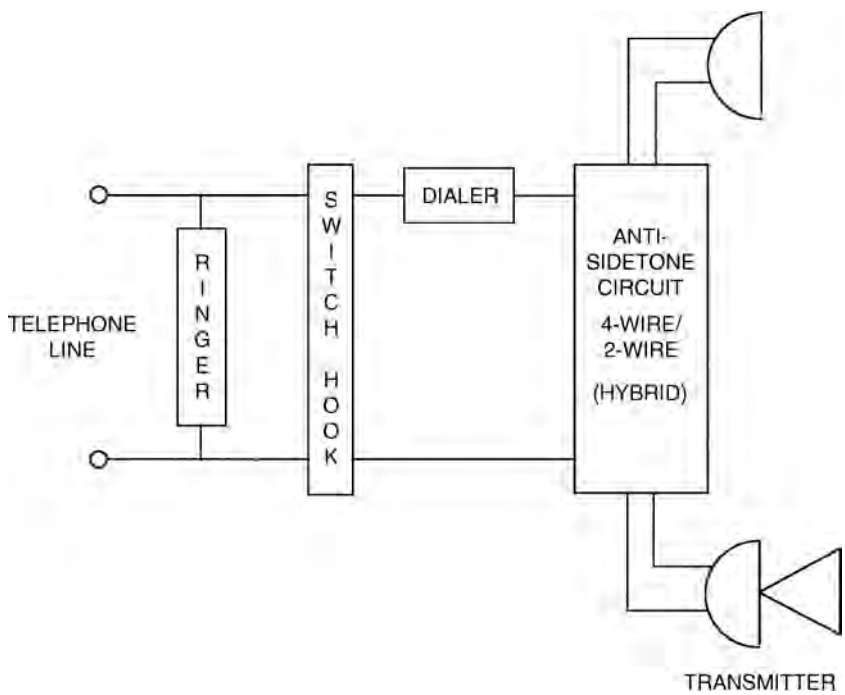


FIGURE 21.2 Telephone.

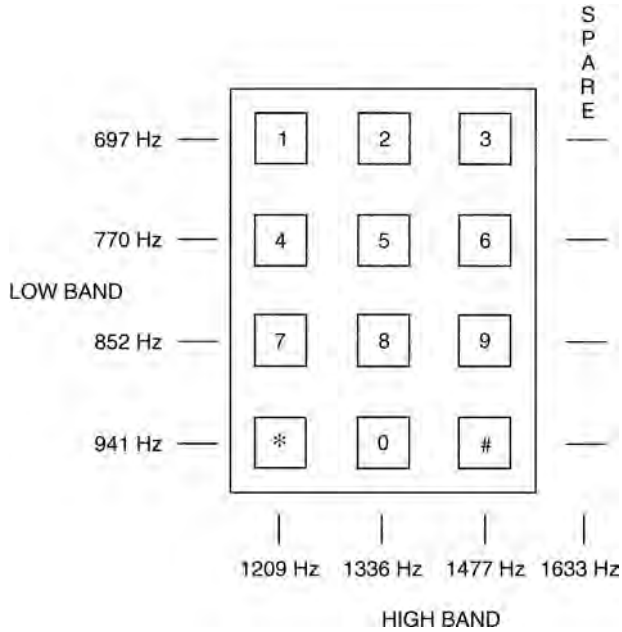


FIGURE 21.3 Touch-tone.

Bell's first telephone microphone, or transmitter, used the principle of variable resistance to create a large varying current. It consisted of a wire that moved in and out of a small cup of acid in response to the acoustic speech signal. Clearly, such a transmitter was not very practical, and it was soon replaced by the use of loosely packed carbon, invented in 1878 by Henry Hummings. In 1886, Thomas Alva Edison improved on the carbon transmitted by using roasted granules of anthracite coal. Today's telephone transmitters use high-quality, small, variable-capacitance, electret microphones. The telephone receiver is a small loudspeaker using a permanent magnet, coil of wire, and metal diaphragm. It was invented in 1876 by Thomas Watson, Bell's assistant, and the basic principles have changed little since then.

The pair of wires going to the telephone transmitter and receiver constitutes a four-wire circuit. The transmitter sends a speech signal down the telephone line, and the receiver receives the signal from the central office. However, a two-wire local loop connects the telephone instrument to the central office, and, hence, two-wire to four-wire conversion is needed within the telephone instrument. A center-tapped transformer, called a **hybrid**, accomplishes this conversion. The leakage current in the secondary receiver circuit depends on how well a balance network exactly matches the impedance of the telephone line. Since this balance network can never match the line perfectly, a small amount of the received transmitted signal leaks into the receiver circuit, and the user hears one's own speech, an effect known as sidetone. Actually, a small amount of sidetone is desirable because it makes the telephone seem live and natural, and, thus, the balance network is designed to allow an optimum amount of sidetone. Too much sidetone results in the user pulling the handset away from the head, which reduces the transmitted speech signal—an undesirable effect. The use of an induction coil to balance the electrical sidetone was patented in 1918 by G.A. Campbell, an AT&T research engineer. The induction coil has been replaced in modern telephones by a speech network that electronically cancels the sidetone leakage and performs the two-wire to four-wire conversion.

The telephone ringer is connected in parallel across the telephone line before the switch hook's contacts. Thomas Watson applied for the first ringer patent in 1878, and today's electromechanical ringers have changed little since then. A hammer attached to an armature with a magnetic field strengthened by a permanent magnet moves in response to the ringer current loudly striking two bells. The high-impedance ringer was invented in 1890 by John J. Carty, a Bell engineer who had invented the two-wire local loop

in 1881. A capacitor is placed in series with the ringer to prevent DC from flowing through it. The ringer signal consists of a short 2-s burst of a 75-V (rms), 20-Hz sine wave followed by 4 s of silence. Piezoelectric transducers and small loudspeakers have replaced electromechanical ringers in today's telephones.

Telephone instruments have progressed greatly in their functionality from the basic-black, rotary-dial phones of the past. Today's phones frequently include repertory dialers, speakerphones, and liquid crystal displays (LCDs). Tomorrow's phones will most likely build up this functionality and extend it to control central office features, to perform e-mail, and to integrate voice and data. Although some people still believe that the telephone of the future will also include a two-way video capability, the videophone, most consumers do not want to be seen while speaking on the phone, and, thus, the videophone will most probably remain an element of science fiction.

The public switched network can be used to transmit and switch any signal that remains within its baseband, namely, about 4 kHz. Thus, devices other than just a telephone can be used on the telephone network. The recent success of facsimile is one example; modems operating at speeds of 56 kb/s are another.

21.4 Transmission

A wide variety of transmission media have been and are used in providing telephone service. At the local level, twisted pairs of copper wire today connect most customers to the central office, although open copper wire was used in the distant past and in rural areas. Many pairs of wire are placed together in a cable, which is then either placed underground or strung between telephone poles. Coaxial cable carried telephone calls across the country. Microwave radio carried telephone calls terrestrially from microwave tower to tower across the country, with each tower located about 26 miles from the next. Microwave radio also carries telephone calls across oceans and continents by communication satellites located in geosynchronous orbits 22,300 miles above the surface of the earth. Today's transmission medium of choice for carrying telephone calls over long distances and between central offices is optical fiber.

Multiplexing is the means by which a number of communication signals are combined together to share a single communication medium. With analog multiplexing, signals are combined by frequency-division multiplexing; with digital multiplexing, signals are combined by time-division multiplexing.

Today analog multiplexing is obsolete in telephony. AT&T replaced all its analog multiplexing with digital multiplexing in the late 1980s; MCI followed suit in the early 1990s. Analog multiplexing was accomplished by A-type channel banks. Each baseband telephone channel was shifted in frequency to its own unique 4-kHz channel. The frequency-division multiplexing was accomplished in hierarchial stages. A hierarchy of multiplexing was created with 12 baseband channels forming a group, 5 groups forming a supergroup, 10 supergroups forming a mastergroup, and 6 mastergroups forming a jumbo group. A jumbo multiplex group contained 10,800 telephone channels and occupied a frequency range from 654 to 17,548 kHz.

With digital multiplexing, each baseband analog voice signal is converted to digital using a sampling rate of 8000 samples/s with 8-b nonlinear quantization for an overall bit rate of 64,000 b/s. A hierarchy of digital multiplexing has evolved with 24 digital telephone signals forming a DS1 signal requiring 1.544 Mb/s, 4 DS 1 signals forming a DS2 signal, 7 DS2 signals forming a DS3 signal, 6 DS3 signals forming a DS4 signal. A single digital telephone signal at 64 kb/s is called a DS0 signal. A DS4 signal multiplexes 4032 DS0 signals and requires an overall bit rate of about 274 Mb/s.

The transmission media and systems used for long-distance telephone service have progressed over the decades. The Ll-carrier system, first installed in 1946, utilized three working pairs of coax in a buried cable to carry 1800 telephone circuits across the country using analog, frequency-division multiplexing. The L5E-carrier system, installed in 1978, carried 132,000 telephone circuits in 10 coax pairs. Terrestrial microwave radio has been used to carry telephone signals from towers located about every 26 miles across the country. The first system, TD-2, became available in 1950 and carried 2400 voice circuits. The use of polarized radio waves to reduce channel interference, the horn antenna to allow simultaneous operation in both the 6-GHz and 4-GHz bands, solid-state technology, and single-sideband suppressed-carrier amplitude modulation resulted in a total system capacity in 1981 of 61,800 voice circuits.

Communication satellites located in a geosynchronous orbit 22,300 mi above the Earth's equator have been used to carry telephone signals. Here, too, the technology has progressed, offering ever increasing capacities. But geosynchronous communication satellites suffer a serious shortcoming. The time required for the radio signals to travel back and forth between satellite and Earth stations creates a round-trip delay of about 0.5 s, which is quite annoying to most people.

Today's transmission medium of choice is optical fiber utilizing digital, time-division multiplexing of the voice circuits. The basic idea of guiding light through thin glass fibers is quite old and was described by the British physicist Charles Vernon Boys in 1887. Today's optical fiber utilizes ultrapure silica. Optical fiber itself has progressed from multimode stepped index and graded index fibers to today's single-mode fiber. Solid-state lasers have also progressed in their use as light sources, and detector technology is also an area of much technological advancement. Today's fiber strands each carry a few gigabits per second. Technological advances include color mutliplexing, in which a number of light signals at different frequencies are carried on the same fiber strand and erbium doped fiber amplifiers that increase the strength of the light signal without the need to convert signals back to an electrical form for regeneration. Usually, a number of fiber strands are placed together in a single cable, but the capacity of each fiber strand is so great that many of the strands are not used and are called *dark fiber*. The theoretical capacity of a single strand is as much as 100 terabits per second. The synchronous optical network (SONET) standard facilitates the interconnection of optical networks operating at rates measured in gigabits per second.

Long-distance transmission systems and local carrier systems all utilize separate paths for each direction of transmission, thereby creating four-wire circuits. These four-wire circuits need to be connected to the two-wire local loop. The hybrids that accomplish this connection and conversion can not perfectly match the transmission characteristics of each and every local loop. The result is that a small portion of the signal leaks through the hybrid and is heard by the speaking party as a very annoying echo. The echo suppressor senses which party is speaking and then introduces loss in the return path to prevent the echo, but this solution also prevents simultaneous talking. Today, echo elimination is accomplished by an **echo canceler**, which uses an adaptive filter to create a synthetic echo, which is then subtracted from the return signal, thereby eliminating the echo entirely but also allowing simultaneous double talking. An echo canceler is required at each end of the transmission circuit.

21.5 Switching

In the old days, one telephone was connected to another telephone at switchboards operated by humans. The human operators used cords with a plugs at each end to make the connections. Each plug had a tip and a ring about the tip to create the electric circuit to carry the signals. A sleeve was used for signalling purposes to indicate whether a circuit was in use. Each human operator could reach as many as 10,000 jacks. The automation of the switchboard came early in the history of telephony with Almon B. Strowger's invention in 1892 of an electromechanical automatic switch and the creation of his Automatic Electric Company to manufacture and sell his switching systems, mostly to non-Bell telephone companies. The Strowger switching was ignored by the Bell System until 1919 when it was finally adopted. Now electromechanical switching is totally obsolete in the U.S., and today's telephone system utilizes electronic switching systems. However, the Strowger system, known as step-by-step in the Bell System, was a thing of great mechanical ingenuity that gave an intuitive grasp of switching with turning and stepping switch contacts that is not possible with today's computerized electronic systems. Electromechanical switching was subject to much wear and tear, however, and required considerable space and costly maintenance. Furthermore, electromechanical switching was inflexible and could not be reprogrammed.

In general, a switching system consists of two major functional parts, as shown in Fig. 21.4: (1) the switching network itself, where one telephone call is connected to another and (2) the means of control that determines the specific connections. Calls can be connected by physically connecting wires to create an electrical path, a technique called **space switching**. With space switching, individual telephone circuits

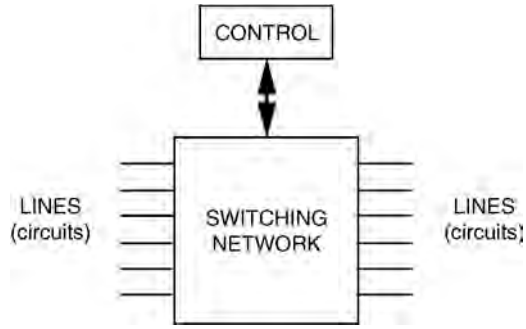


FIGURE 21.4 Switching system.

are connected physically to each other by some form of electromechanical or electronic switch. Calls can also be connected by reordering the time sequence of digitized samples, a technique called **time switching**. Modern digital switching systems frequently utilize both techniques in the switching network.

In the past, the switching network utilized electromechanical technology to accomplish space switching. This technology progressed over time from the automated Strowger switch to the Bell System's crossbar switch. The first crossbar switching system was installed in 1938, and crossbar switching systems were still in use in the U.S. in the early 1990s.

The switching network in today's switching systems is completely digital. Telephone signals either arrive in digital or are converted to digital. The digital signals are then switched, usually using a combination of electronic space switching along with time switching of the sequence of digitized samples. The space switches are shared by a number of digital calls connecting each of them for short durations while a small number of bits in each sample are transferred.

Yesterday's switching systems were controlled by hard-wired electromechanical relays. Today's switching systems are controlled by programmable digital computers, thereby offering great flexibility. The use of a digital computer to control the operation of a switching network is called electronic switching or stored-program control. The intelligence of stored-program control, coupled with the capabilities of modern signalling systems, enables a wide variety of functional services tailored to the needs of individual users.

21.6 Signalling

A variety of signals are sent over the telephone network to control its operation, an aspect of POTS known as signalling. The familiar dial tone, busy signal, and ring-back tone are signals presented to the calling party. Ringing is accomplished by a 20-Hz signal that is on for 2 s and off for 4 s. In addition to these more audible signals that tell us when to dial, whether lines are busy, and when to answer the telephone, other signals are sent over the telephone network itself to control its operation.

In the past, the telephone network needed to know whether a trunk is idle or not, and the presence or absence of DC indicated whether a local trunk was in use or idle. Long-distance trunks used in-band and out-of-band tones to indicate whether a circuit was idle or not. A single-frequency tone of 2600 Hz, which is within the voice band, was placed on an idle circuit to indicate its availability. The telephone number was sent as a sequence of two tones at a rate of 10 combinations per second, a technique known as multifrequency key pulsing (MFKP). Signalling today is accomplished by common channel interoffice signalling (CCIS).

With **common-channel signalling**, a separate dedicated data channel is used solely to carry signalling information in the form of short packets of data. Common-channel signalling is known as signalling system 7 (SS7) in the U.S. It offers advanced 800-service such as time of day routing, identification of the calling party, and various software defined network features. In addition to new features and services,

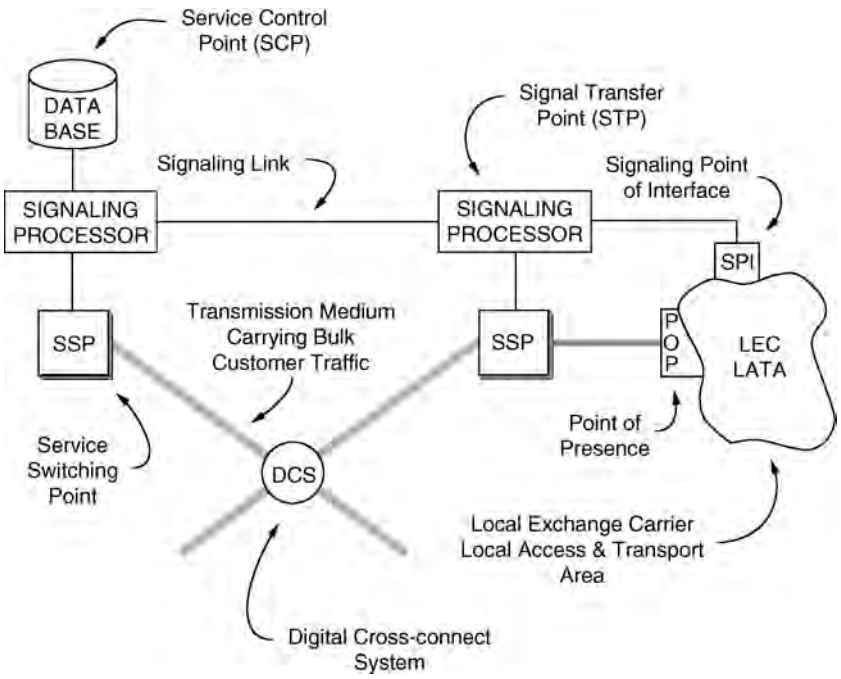


FIGURE 21.5 Intelligent network.

common-channel signalling offers more efficient assignment of telephone circuits and operation of the telephone network. Although first used for long-distance networks, common-channel signalling is also used at the local level.

Signalling has been integrated into a modern telecommunications network, depicted in Fig. 21.5, and when coupled with centralized data-bases, provides many of the features associated with today's intelligent network. The database, known as a service control point (SCP), contains the information needed to translate 800-numbers to the appropriate telephone location, among other items. The signalling information is sent over its own signalling links from one signalling processor to another, located at nodes called signal transfer points (STP). The signalling processors determine the actual switching of the customer circuits, performed by switching systems at service switching points. The bulk traffic carried over transmission media can be switched in times of service failures or to balance loads by digital cross-connect systems (DCS). The signalling links connect to the local network at a signalling point of interface (SPI), and the customer circuits connect at a point of presence. Today's signalling systems add much functionality to the network.

21.7 Functionality

A wide variety of intelligent features was easily available when humans operated the telephone network and its switchboards. The operator could announce the name of the calling party, hold calls if you were busy, transfer calls to other phones and locations, and interrupt a call if another more important one arrived. However, human operators were far too costly and were replaced by automated electromechanical switching systems. This made telephone service more affordable to more people, but lost the functionality of the human intelligence of the operators. Today's telephone switching systems are controlled by programmable computers, and, once again, intelligence has returned to the telephone network so that the functional services of the past can again be offered using today's computer-controlled technology.

Call waiting, call forwarding, and caller-ID are examples of some of these functional services. But not all of these services are wanted by all telephone users.

Caller-ID transmits the telephone number of the calling party over the local loop to the called party where the number is displayed on a small visual display. The number is encoded as digital data and is sent in a short burst, using phase shift keying, during the interval between the first and second ringing signal. Advanced caller-ID systems also transmit the name associated with the directory listing for the calling number. With caller-ID, it is possible to know who is calling before answering the telephone. However, some people consider their telephone number to be very private and personal and do not want it transmitted to others. This privacy issue delayed the availability of caller-ID in some states and is a good example of the importance of understanding the social impacts of the telephone.

21.8 The Future

Based on the false promises of new products and services such as picturephones, videotex, and high-definition television (HDTV), skepticism is warranted toward most new ideas, particularly when so many of them are really reincarnations of past failures. Accepting these words of warning, some last thoughts about the future of POTS will nevertheless be opined.

The telephone system invented by Alexander Graham Bell simply enabled people to convey their speech over distance. Bell's network has evolved to a system that enables people to stay in contact wherever they may be through the use of paging and radio-based cellular telephone services, and even telephones in commercial airplanes. The telephone network carries not only voice signals but also facsimile and data signals.

Bell's vision of a wired world communicating by human speech has been mostly achieved, although there are still many places on this planet for which a simple telephone call is a luxury. A global system of low earth orbit satellites could solve this problem, but it would be far too costly to offer telephone service to the poorer inhabitants of those places without conventional telephone service. Bell was very wise in emphasizing human speech over the Morse code of telegraphy. However, alphanumeric keyboards negate the need for knowledge of Morse code or American Standard Code for Information Interchange (ASCII) bits, and now everyone can communicate by text, or e-mail, as such communication is generally called.

Textual communication can have long-holding times but very low average data transmission. Today's packet-switched networks are most appropriate and efficient for this form of communication. Whether packet-switching will become dominant for voice telecommunication, evolving into a form of an integrated services digital network (ISDN), is promoted by many and yet unclear given today's bandwidth glut in backbone networks.

Much is said about the convergence of telephony and television, of telecommunication and entertainment, and of telephony and community antenna television (CATV). Yet the purpose of passive entertainment seems much different than the interactivity and two-way nature of telephony and most data telecommunication. The entertainment center is quite different than the communication centers in most homes. Yet the myth of convergence continues, although the past would tell us that convergence really is a blurring of boundaries between technologies and industry segments.

A trip to a central office will show that although much progress has been made in substituting electronic switching for electromechanical switching, there are still tens of thousands of physical wires for the provision of service over the local loops. The solution is the use of time-division multiplexing so that thousands of circuits are carried over a few physical optical fibers. However, there are still many engineering and cost challenges that must be solved before copper local loops are eliminated, but someday they clearly will be eliminated, resulting in lower costs and further increases in productivity.

Although the progress of the technology of telephony has been most impressive over the last century of the provision of POTS, many other factors are equally important in shaping the future of telecommunication. Policy, regulatory, and competitive factors caused the breakup of the Bell System in 1984, and this breakup has had tremendous impact on the provision of telecommunication in the U.S. The entry of long-distance companies into local service and the entry of local telephone companies into long

distance will likewise have considerable impact on the balance of power within the industry. Consumer reactions have halted videophones and delayed caller-ID. The financial risks of the provision of CATV and entertainment have had sobering impact on the plans of telephone companies to expand telephone service to these other businesses.

Defining Terms

Common channel signalling: Uses a separate dedicated path to carry signalling information in the form of short packets of data.

Dual tone multifrequency (DTMF) dialing: Touch-tone dialing where pairs of tones denote a dialed number.

Echo canceler: A device that uses an adaptive filter to create synthetic echo to subtract from the return signal to eliminate the echo.

Hybrid: A center tapped transformer that accomplishes two-wire to four-wire conversion.

Interexchange carriers (IXC): Companies that operate long-distance networks, such as AT&T, MCI, and Sprint.

Local access and transport area (LATA): The local portion of the telephone network.

Multiplexing: Method by which a number of communication channels are combined together to share a signal to eliminate the echo.

Point of presence (POP): The point where the connection is made from the local service provider to the interexchange carrier.

Space switching: Physically connecting wires to create an electrical path.

Time switching: Connecting calls by reordering the time sequence of digitized samples.

Further Information

Much of the material in this chapter is based on:

Noll, A.M. 1998. *Introduction to Telephones and Telephone Systems,* 3rd ed., Artech House, Norwood, MA.

Pierce, J.R. and Noll, A.M. 1990. *Signals: The Science of Telecommunications,* Scientific American Library, New York.

Other references are:

Elbert, B.R. 1987. *Introduction to Satellite Communication,* Artech House, Norwood, MA.

Hills, M.T. 1979. *Telecommunications Switching Principles,* MIT Press, Cambridge, MA.

Noll, A.M. 1996. *Highway of Dreams: A Critical View Along the Information Superhighway,* Lawrence Erlbaurn Associates, Mahwah, NJ.

Noll, A.M. 2001. *Principles of Modern Communications Technology,* Artech House, Norwood, MA.

Parker, S.P., Ed. 1987. *Communications Source Book,* McGraw-Hill, New York.

22

FDM Hierarchy

Pierre Catala
Texas A & M University

22.1 Introduction

Circuits linking telephone central offices nationwide carry from dozens to thousands of voice channels, all operating simultaneously. It would be very inefficient and prohibitively expensive to let each pair of copper wires carry a single voice communication. Therefore, very early in the history of the telephone network, telecommunications engineers searched for ways to combine telephone conversations so that several of them could be simultaneously transmitted over one circuit. At the receiver end, the combined channels would be separated back into individual channels.

Such a combining method is called *multiplexing*. There are two main methods of multiplexing:

- Time-division multiplexing (TDM), which can be used only with digital signals and is a more "recent" (1960s) technology.
- Frequency-division multiplexing (FDM), which was first implemented in 1918 by "Ma Bell" (AT&T) between Baltimore and Pittsburgh and could carry four simultaneous conversations per pair[1] of wires.

By the early 1970s, FDM microwave links commonly supported close to 2000 (and sometimes 2700) voice channels on a single radio carrier. In 1981, AT&T introduced its AR6A single sideband (SSB) microwave radio, which carried 6000 channels [Rey, 1987]. Coaxial cable carrier systems, such as the AT&T L5E, could carry in excess of 13,000 channels using frequency-division multiplexing.

This chapter reviews the principle of FDM then describes the details of how FDM channels are combined in a hierarchy sometimes referred to as the *analog hierarchy*.

[1]In reality, you need one pair for each direction since, in telephony, it is necessary to have full-duplex communications, i.e., to be able to talk and listen at the same time. These two pairs are usually labelled "GO" and "RETURN" and they create what is called a four-wire (4W) circuit.

22.2 Background Information

Voice-Channel Bandwidth

The voice-spectrum bandwidth in telephone circuits is internationally limited to frequencies in the range from 0.3 to 3.4 kHz. Trunk signaling for that voice channel is generally done *out of band* at 3825 Hz. Furthermore, since bandpass filters are not perfect (the flanks of the filters are not perfectly vertical), it is necessary to provide a guard band between channels. To take all of these factors into account, the overall voice channel is standardized as a 4-kHz channel.

Although most of the time the voice channels will, indeed, be carrying analog voice signals, they can also be used to transmit *analog data signals,* that is, signals generated by data modems. This consideration is important in the calculation of the loading created by FDM **baseband signals**.

22.3 Frequency-Division Multiplexing (FDM)

Frequency-division multiplexing combines the different voice channels by stacking them one above the other in the frequency domain, as shown in Fig. 22.1, before transmitting them. Therefore, each 4-kHz voice channel is shifted up to a frequency 4 kHz above the previous channel.

The FDM concept is not specific to telephony transmission. The TV channels in cable TV (or on-the-air broadcasting for that matter) are also stacked in frequency (channels 1–90, etc.) and are, therefore, frequency-division multiplexed. The TV receiver is the demultiplexer in that case. The difference here is that the TV receiver only needs to receive one channel at a time, whereas the telephone FDM demultiplexer must receive all channels simultaneously.

Implementation Considerations

Practical reasons preclude the assembly (*stacking*) of the different channels, one above the other, continuously from DC on. Problems such as AC power supply hum make the use of low frequencies unwise,

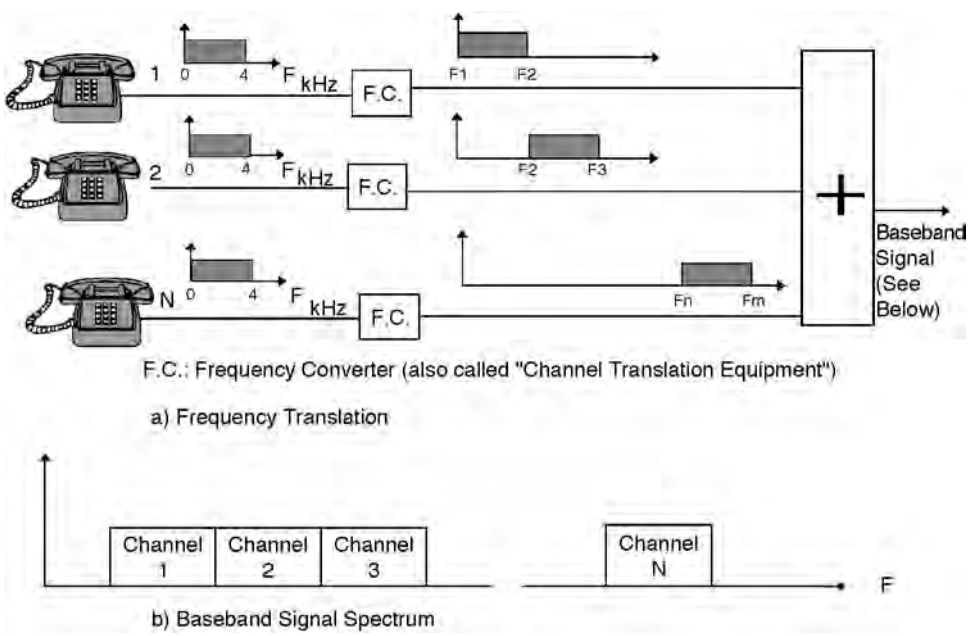


FIGURE 22.1 FDM principle.

thus the first voice channel is usually *transposed* to a frequency of 60 kHz (although for small capacity systems, the first channel can be as low as 12 kHz).

The necessity of having compatible equipment so that FDM links could be established between equipment of different vendors, and even between different countries, dictated the establishment of standards that specified how channels should be grouped for multiplexing. This also allowed the manufacturing of modular equipment, which can be combined to increase the capacity of the links.

All of these considerations led to the development of FDM baseband signals that follow a specific organization and occupy given sets of frequencies, as described in the following sections.

22.4 The Hierarchy

Group

The first step in FDM is to combine 12 voice channels together. This first level of the analog multiplexing hierarchy is called a *group* (also called *basic group* by the International Telecommunications Union and a *primary group* by several countries). The FDM group is, therefore, 48 kHz wide and composed of 12 voice channels stacked from 60 to 108 kHz, as shown in Fig. 22.2. For a *thin route*, that is, a small capacity system, the group can instead be translated to a range of 12–60 kHz so as to reduce the overall baseband bandwidth required.

The equipment performing the frequency translation of each channel uses a combination of mixers and single side band (SSB) modulation with appropriate bandpass filters to create this baseband group. That equipment is called *channel bank* in North America and *channel translation equipment* (CTE) in other English-speaking countries.

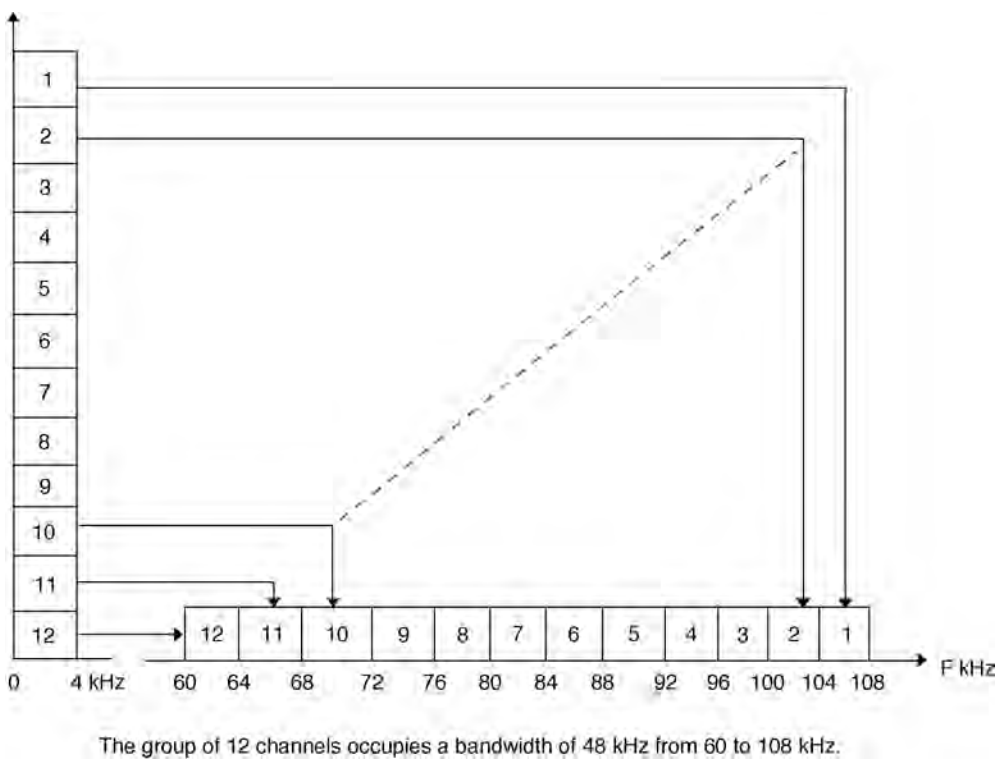


FIGURE 22.2 Basic FDM group.

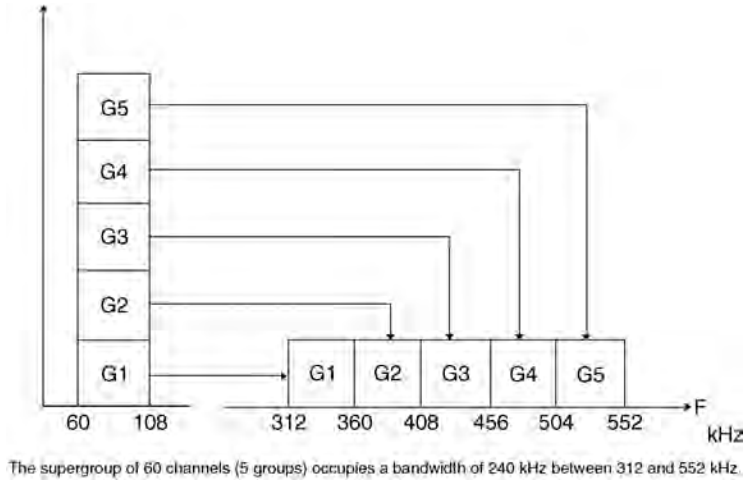


FIGURE 22.3 FDM supergroup.

Supergroup

If more than 12 channels are needed, the next level in the hierarchy is created by combining 5 groups together, thus providing a capacity of 60 voice channels. This second level is called a *basic supergroup*, usually abbreviated as supergroup, and has a bandwidth of 240 kHz going from 312 to 552 kHz, as shown in Fig. 22.3. Some countries call this 60-channel assembly a *secondary group*.

The related frequency translation equipment is called *group bank* in the U.S. and Canada but *group translation equipment* (GTE) in most other countries.

Mastergroup

The international agreements concerning FDM standards unfortunately stop at the supergroup level. Although both the ITU-T and the U.S./Canadian standards include a third level called *mastergroup*, the number of voice channels differ and the two are, therefore, not compatible! The most common mastergroup in North America was the old Western Electric U600 scheme, which combines 10 supergroups, thus creating a 600-channel system, as shown in Fig. 22.4.

There are several variations of the Western Electric U600 mastergroup such as the AT&T L1 coax-based carrier system, which also carries 600 channels but translated between 60–2788 kHz rather than 564–3084 kHz.

The frequency translation equipment creating mastergroups is called *supergroup bank* or *supergroup translation equipment* (STE). The ITU-T mastergroup on the other hand combines only five supergroups to create a 300-channel baseband, as shown in Fig. 22.5.

Higher Levels

There are many different multiplexing schemes for the higher density FDM systems. They become quite complex and will only be listed here. The reference section of this paper lists several sources where details of these higher level multiplexed signals can be found.

- *IYU-T supermastergroup*. This combines three CCITT mastergroups for a total of 900 channels occupying a 3.9 MHz baseband spectrum from 8.516 to 12.388 MHz.
- *ITU-T 15-Supergroup Assembly*. This also provides 900 channels but bypasses the mastergroup level by directly combining 15 supergroups ($15 \times 60 = 900$) occupying a 3.7-MHz spectrum from 312 to 4028 kHz. This assembly is sometimes called a *hypergroup*. There are variations of the 15-supergroup assembly such as a 16-supergroup assembly (960 channels).

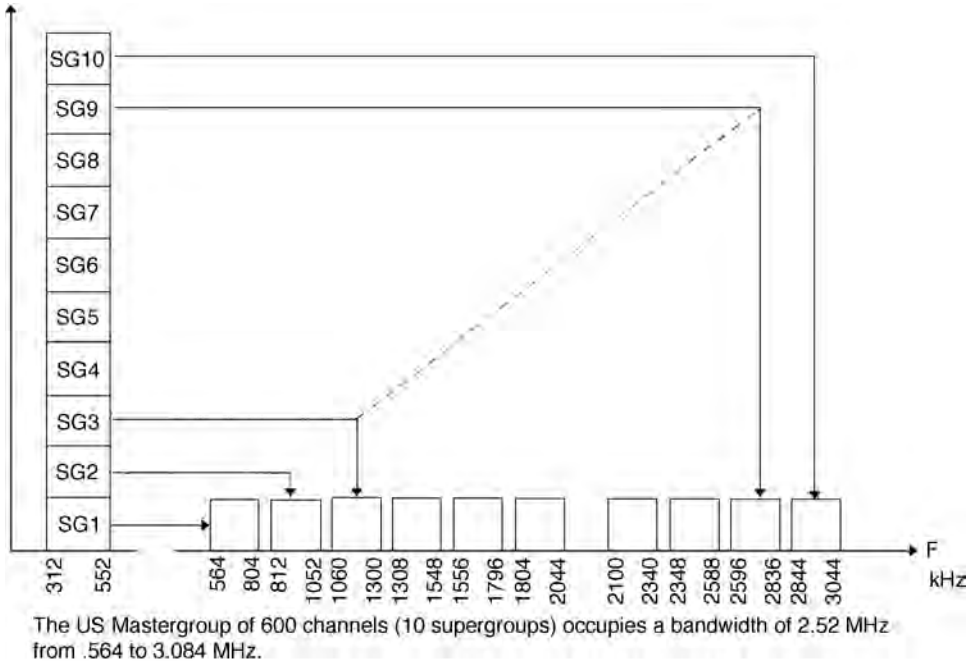


FIGURE 22.4 U.S. (AT&T) basic mastergroup.

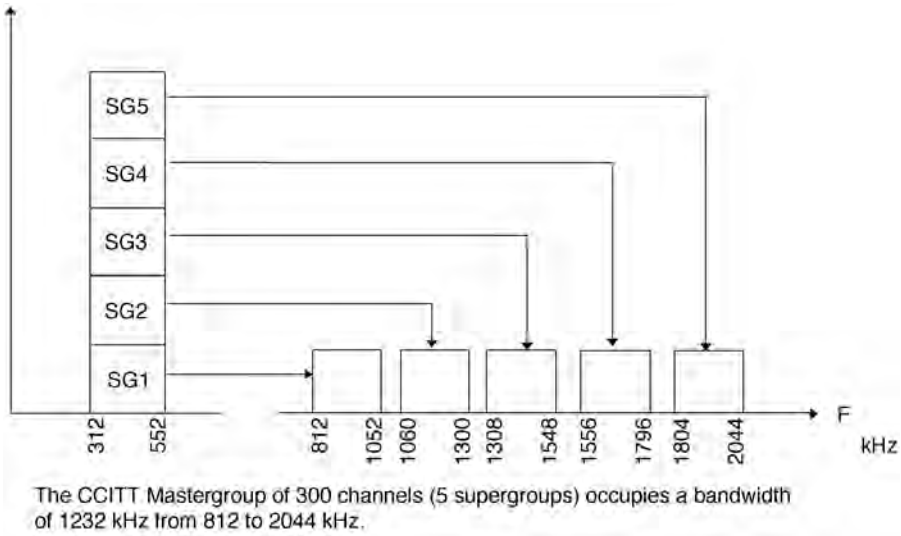


FIGURE 22.5 CCITT basic mastergroup.

For the very-high capacity FDM systems, grouping of supermastergroups or of 15-supergroup assemblies are used, with a preference for the latter method.

The AT&T high-level FDM hierarchy is, of course, based on the 600-channel mastergroup. The AT&T *mastergroup multiplex* (MMX) combines these mastergroups to a variety of levels, shown in Table 22.1, by multiplexing from two to eight mastergroups (except for the four-mastergroup level, which is not used).

Two of the 3000-channel intermediate levels are further combined to form the baseband signal for the 6000-channel AR6A microwave radio. At 6000 channels per radio, the AR6A radio system can transmit 42,000 voice channels on a single antenna using seven different RF carriers [Rey, 1987].

TABLE 22.1 AT&T Mastergroup Multiplex Hierarchy

No. of Multiplexed Mastergroups	Resulting No. of Channels	AT&T Application
2	1200	TD microwave radio
3	1800	TH microwave radio
5	3000	Intermediate level
6	3600	Jumbogroup
7	4200	Intermediate level
8	4800	Intermediate level

Two of the 4200-channel intermediate levels are combined with one 4800-channel intermediate level to form the line signal for the L5E, 13,200-channel, coaxial cable system, which occupies a frequency spectrum of 61.592 MHz starting at 3.252 MHz.

Jumbogroup

As was seen in Table 22.1, the basic jumbogroup is composed of six mastergroups (3600 channels). It was the basis for the L4 coaxial-cable system and has a spectrum ranging from 0.564 to 17.548 MHz.

When the L5 coaxial system was implemented in 1974, it used a line signal created by multiplexing three jumbogroups, thus providing a capacity of 10,800 channels occupying frequencies from 3.124 to 60.566 MHz.

22.5 Pilots

In addition to the voice channels (and related signalling tone), special tones called pilots are introduced to provide the receiver end (demultiplexer) with frequency and level references.

Each level of the FDM hierarchy (group, supergroup, mastergroup, etc.) will have its own *reference* pilot tone. Each of the respective demultiplexers monitors the appropriate tone to detect interruptions (faults) in which case it generates an alarm indicating which group, supergroup, etc., is defective. Alarms are also generated if the level of the tone is 4 dB below (or above) its normal value.

Other pilots, called *line regulating*, used to be found in gaps between the FDM building groups (e.g., supergroups). Their name came from the fact that they were used for automatic gain control (AGC) of the miscellaneous amplifiers/repeaters. As FDM equipment became better, the regulation function was transferred to the reference pilots just mentioned and line regulating pilots were no longer implemented. A regulation of 0.5 dB is a typical objective [Freeman, 1989].

Microwave radios use an additional type of pilot tone, the *continuity pilot*, which is generally inserted at the top of the baseband spectrum and controls the receiver AGC and also serves as an overall continuity pilot.

In older FDM equipment, there were also *frequency synchronization pilots* used to ensure that the demultiplexing carriers were within a few hertz of the multiplexing pilots. Otherwise, audio distortion would occur. Newer FDM equipment has sufficiently precise and stable oscillators so that synchronization pilots are no longer needed.

22.6 Direct to Line (DTL)

If interfacing with existing FDM equipment is not needed on a particular link, then a cheaper method of multiplexing can be done with special multiplexing equipment, at least for the low-capacity systems. The direct-to-line (DTL) method of forming the baseband signal moves the voice channels directly to their respective line frequencies, thus bypassing the group and supergroup building blocks. Such equipment is frequently used on thin-route microwave links carrying 60 channels or less but can also be found in multiplexers providing basebands of up to 600 channels.

The DTL method decreases the number of steps in the multiplexing process and, consequently, the number of circuit boards and filters. Therefore, it is cheaper, more reliable, and more flexible. DTL equipment, however, is not compatible with the standard FDM equipment.

22.7 Summary

Multiplexing in the frequency domain (FDM) was the mainstay of carrier systems, on twisted pairs, coaxial cable, and microwave links, until digital carrier systems became prevalent in the early 1980s. FDM is done by stacking voice channels, one above the other, in the frequency domain. In other words, each voice channel is translated from its original 0–4-kHz spectrum to a frequency above the previous channel. This is generally done in building blocks called groups (12 channels), supergroups (5 groups), mastergroups (5 or 10 supergroups), etc., which follow international ITU-T standards or AT&T standards. FDM equipment with fewer than 600 channels can also be implemented using a simpler and cheaper technique called direct to line, where the building blocks are bypassed and the voice channels are transposed directly to the appropriate line (baseband) frequency. However, DTL is not a standard, and, therefore, direct interconnection with the public telephone network is not possible.

Defining Term

Baseband signal: Modulating signal at the transmitter (i.e., signal found before modulation of the transmitter). At the receiver, it would be the demodulated signal, that is, the signal found after the demodulator. In the case of FDM, the baseband signal (also called line frequency) is, therefore, the combined signal of all the voice channels being transmitted.

References

CCITT (ITU-T). 1989. *International Analogue Carrier Systems.* Blue Book, Vol. III, Fascicle III. 2, Consultative Committee on International Telephony and Telegraphy/International Telecommunications Union, Geneva, Switzerland.

Freeman, R. 1989. *Telecommunication System Engineering,* Wiley-Interscience, New York.

Rey, R.F., Tech. Ed. 1987. Transmission systems, in *Engineering and Operations in the Bell System,* 2nd ed., AT&T Bell Labs., Murray Hill, NJ, 345–356.

Tomasi, W. 1992. *Advanced Electronics Communications Systems,* Prentice-Hall, Englewood Cliffs, NJ.

Further Information

General concepts of multiplexing can be found in Chapter 7 of this handbook. Principles of mixers, frequency converters (frequency translation), and single sideband modulation can be found in any electronic-communications or communication-system type of book.

23

Analog Telephone Channels and the Subscriber Loop

Whitham D. Reeve
Reeve Engineers

23.1 Telephone Band

A voice channel is considered to require a nominal bandwidth of 4000 Hz. For all practical purposes, however, the usable bandwidth of an end-to-end call in the **public switched telephone network (PSTN),** including the loop, is considered to fall between approximately 300 and 3400 Hz (this is called the *voice band*), giving a bandwidth of 3100 Hz. This bandwidth is entirely acceptable from a voice transmission point of view, giving subscriber satisfaction levels above 90% [IEEE, 1984].

Although it is not unduly difficult to provide a 3100-Hz bandwidth in the subscriber loop, the loop and the associated terminal equipment have evolved in such a way that this bandwidth is essentially fixed at this value and will continue to be for analog voice and voiceband data transmission. The bandwidth is somewhat restrictive for data transmission when using voiceband modems, and the speeds attained are below approximately 64 kb/s. However, the problem of bandwidth restriction has encouraged a number of innovative solutions in modem design, particularly adaptive equalizers and modulation methods.

Obviously, on a loop derived entirely from copper cable, the frequency response of the loop itself would extend down to DC (zero frequency). The lower response is lost, however, once the loop is switched or connected to other transmission and signaling equipment, all of which are AC coupled. Where DC continuity is not available or not practical, special tone signaling equipment is used to replace the DC signals. When voice signals or other signals with frequency content approaching zero frequency are placed on the loop, the transmission is considered to be in the baseband.

Similarly, the upper voiceband frequency limit is not exactly 3400 Hz. Depending on how it is specified or the type of cable, the limit may be much higher. In practice, the loop does not generally set the upper limit of a voice channel; the upper limit is mostly due to the design of filters in the equipment that interfaces with the loop.

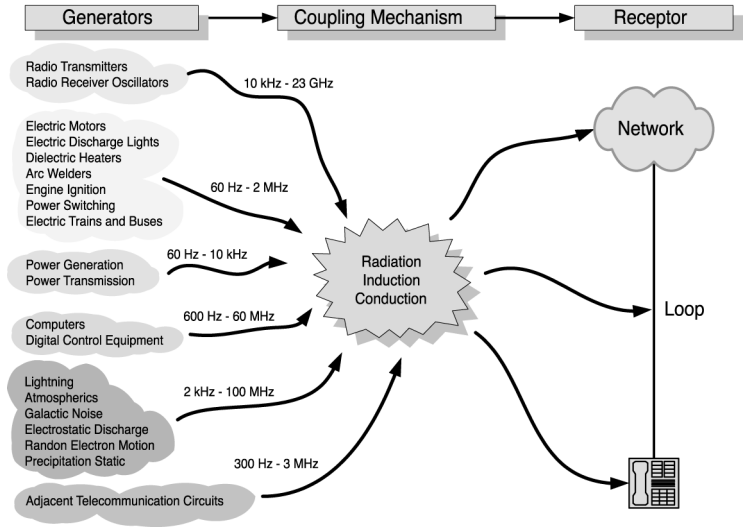


FIGURE 23.1 Noise diagram.

For voice-band transmission, the bandwidth (or frequency response) of a telecommunication channel is defined as the limiting frequencies where loop loss is down by 10 dB from its 1000 Hz value [IEEE, 1984]. The field measurement of bandwidth usually does not proceed with measurement of the 10-dB points. Instead, simple slope tests are made; these provide an indirect but reliable indicator of the transmission channel bandwidth. If the slope, as defined subsequently, is within predetermined limits, the bandwidth of the channel can be assumed to be acceptable.

Slope tests are loss measurements at 404, 1004, and 2804 Hz (this is also called the *three-tone slope*). The loss at the reference frequency of 1004 Hz is subtracted from the loss at 404 Hz to give the low-frequency slope, and from 2804 Hz to give the high-frequency slope.

23.2 Noise

Noise is any interfering signal on the telecommunication channel. There must be a noise source, a coupling mechanism, and a receptor. The relationship among the three is illustrated in Fig. 23.1. Noise sources are either manmade or natural. Practically any piece of electrical or electronic equipment can be a manmade noise source, and power lines are perhaps the most pervasive of all of these. Natural noise comes from lightning and other atmospherics and random thermal motion of electrons and galactic sources, as well as electrostatic discharges.

Noise is coupled by radiation, induction, and conduction. The predominant coupling mode in analog voiceband subscriber loops is by induction from nearby power lines. The other coupling modes exist to some extent, too, depending on the situation. For example, noise can be conducted into the loop through insulation faults or poor or faulty grounding methods.

23.3 Crosstalk

Crosstalk falls in two categories:

1. unintelligible crosstalk (babble)
2. intelligible crosstalk

The latter is most disturbing because it removes any impression of privacy. It can be caused by a single disturbing channel with enough coupling to spill into adjacent circuits. Unintelligible crosstalk is usually

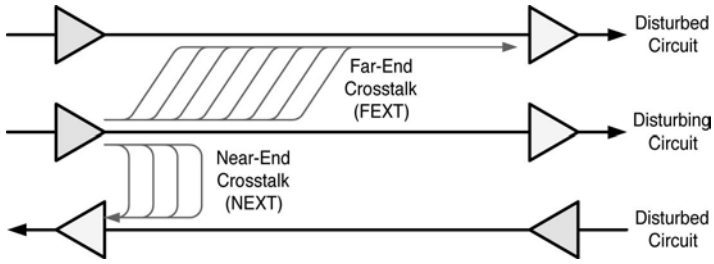


FIGURE 23.2 Crosstalk.

caused by a large number of disturbing channels, none of which is of sufficient magnitude to be understood, or extraneous modulation products in carrier transmission systems.

Crosstalk can be further categorized as near-end and far-end. As these names imply, near-end crosstalk is caused by crosstalk interference at the near end of a circuit with respect to the listener; far-end crosstalk is crosstalk interference at the far end, as shown in Fig. 23.2.

Crosstalk of any kind is caused by insufficient shielding, excessively large disparity between signal levels in adjacent circuits, unbalanced lines, or overloaded analog carrier transmission systems or interfaces. Crosstalk is a statistical quantity because the number of sources and coupling paths is usually too large to quantify.

23.4 Circuit Noise

The noise that appears across the two conductors (tip and ring) of a loop, heard by the subscriber, is called **circuit noise** (also called *message circuit noise, noise metallic*, or *differential noise*). The noise can be due to random thermal motion of electrons (known as *white noise* or *Gaussian noise*) or static from lightning storms, but on subscriber loops its most likely source is interference from power line induction. For the purposes of this discussion, then, circuit noise and interference are assumed to be the same.

The total noise power on a loop is related to the noise bandwidth. Since particular frequencies differently affect the various services (for example, voice, data, and radio studio material), filter frequency response curves have been designed to restrict the frequency response of the noise measuring sets with which objective tests are made. This frequency response restriction is called *weighting*.

Noise, in voice applications, is described in terms of decibels above a noise reference when measured with a noise meter containing a special weighting filter. There are four common filters that are used to provide the necessary weighting for analog loop measurements:

- C-message
- 3-kHz flat (3.4 kHz flat and D-filter in newer test equipment)
- 15-kHz flat

The most common filter is called a C-message filter and measurements are based on decibels with respect to reference noise, C-message weighted (dBrnC). The noise reference is 1 pW (-90 dBm); noise with reference power of 1 pW will read 0 dBrnC on a properly calibrated meter.

C-message weighting is primarily used to measure noise that affects voice transmission when common telephone instruments are used, but it also is used to evaluate the effects of noise on analog voiceband data circuits. It weights the various frequencies according to their perceived annoyance such that frequencies below 600 or 700 Hz and above 3000 Hz have less importance (that is, they are attenuated and do not affect the measurement).

The 3-kHz-flat weighting curve is used on voice circuits, too, but all frequencies within the 3000-Hz bandwidth carry equal importance. This filter rolls off above 3000 Hz and approximates the response of common voiceband modems. It generally is used to investigate problems caused by power induction at

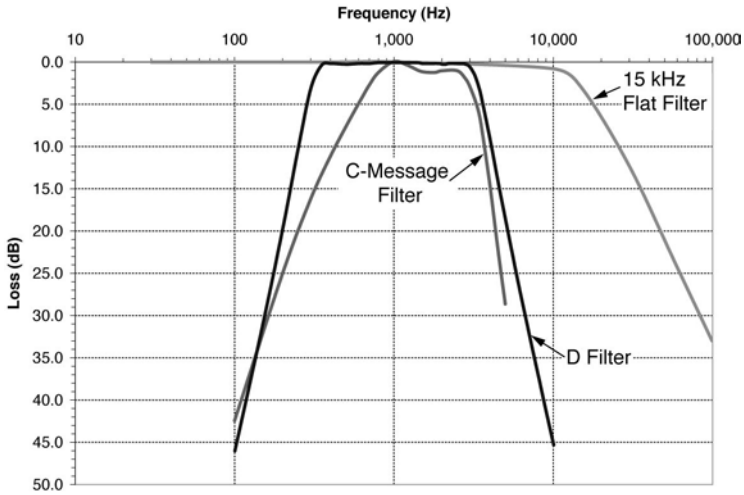


FIGURE 23.3 Weighting curve comparison.

the lower power harmonic frequencies or by higher interfering frequencies. Frequencies in these ranges can affect data transmission as well as voice frequency signaling equipment. The 3-kHz-flat filter has been replaced in modern test equipment with the 3.4-kHz-flat filter. The 3.4-kHz-flat filter better approximates the frequency response of modern high-speed analog voiceband modem loop interface circuits.

The 15-kHz-flat weighting curve is also used to test circuits between radio and television studios and remote transmitter sites. It has a flat response from 20 Hz to 15 kHz.

A comparison of the various weighting curves is shown in Fig. 23.3.

23.5 Impulse Noise

Data circuits are particularly sensitive to **impulse noise**. Impulse noise, heard as clicks, is usually defined as a voltage increase of 12 dB or more above the background [root mean square (rms)] noise lasting 10 ms or less. Its main source is from telephone set rotary dials, switching transients in electromechanical switching systems, maintenance activity on adjacent circuits, or electrical system switching transients. It is less of a problem with modern telecommunication systems because there are fewer rotary dials and very few electromechanical switching systems left in the PSTN.

Impulse noise objectives vary with the type and makeup of the circuit. Usually, a threshold is established and counts are made of any impulse noise that exceeds that threshold in a given time period. When impulse noise tests are made on a single circuit, the usual specification is 15 counts in 15 min. When a group of circuits is being tested, shorter time intervals are used.

On **subscriber loops**, the background noise threshold historically has been 59 dBrnC when measured at the central office [Bellcore, 1986]. However, this threshold is obsolete and no industry standards exist that define current requirements. High-speed analog voiceband modems, such as the 56 kb/s V.90 modems, are very susceptible to impulse noise and use forward error correction and other means to minimize errors.

23.6 Attenuation Distortion

Attenuation distortion is the change in circuit loss with frequency. It is also known as *frequency response*. Ideally, attenuation should be constant throughout the frequency band of interest. Unfortunately, this is not usually the case. Unless it is excessive, however, attenuation distortion is not noticeable to the human ear. Attenuation distortion manifests itself on voice calls by changing the sound of the talker's voice as

it is heard by the listener. The change may be dramatic enough to render the voice unrecognizable. On an analog voiceband data circuit, attenuation distortion can manifest itself in the form of errors through loss of signal energy at critical frequencies.

The inherent attenuation distortion in subscriber loops used in regular switched service is not objectionable except in extreme cases. On voice transmissions, excessive low-frequency slope degrades voice quality, whereas excessive high-frequency slope degrades intelligibility. On data transmissions using *phase-shift keying* (PSK) methods and its derivatives, such as *quadrature amplitude modulation* (QAM), both low- and high-frequency slope affects performance, whereas only high-frequency slope affects low-speed modems that use *frequency-shift keying* (FSK) [CCITT, 1989]. Some slope is considered necessary for stability at band edges.

Attenuation distortion is frequently specified in terms of the three tone slope as discussed previously. The slope objectives for loops used in special switched services [for example, private branch exchange (PBX) trunks and foreign exchange lines] are similar to regular switched loop objectives. It is not an extremely critical parameter on regular switched subscriber loops.

23.7 Envelope Delay Distortion

Envelope delay distortion (EDD), also called group delay distortion, is distortion in the rate of change of phase shift with frequency of a signal. Ideally, the rate of change should be constant with frequency, and it is approximately so in the voice band with nonloaded cables. With loaded cables this is not the case, especially near the cutoff frequency.

Envelope delay distortion is defined as the difference, in time units such as microseconds, between the maximum and minimum envelope delay within the frequency band of interest. If the difference is zero then, by this definition, there is no EDD, but this is hardly ever the case in practical systems.

Voice-band signals (not just voice signals but all signals in the voice band, including analog voiceband data signals) are made up of many frequencies. Each particular frequency propagates at a different velocity (called *phase velocity*) due to the facility's inherent transmission characteristics. This causes phase delay. If the relationship between the resulting phase shift and frequency is nonlinear, the facility will cause delay distortion. EDD at the upper edge of the voice band can cause *near singing* (the "rain-barrel effect") in the conversation. Excessive EDD at the lower edge can cause *speech blurring* [CCITT, 1989].

EDD results from any part of the circuit where a nonlinear relationship exists between phase shift and frequency. Very little EDD results from the loop itself. Usually this occurs at the terminal and intermediate equipment where filters in analog-to-digital and digital-to-analog converters are located. Some data transmission modulation techniques are more susceptible to intersymbol interference caused by EDD than others, which can explain why some modems of a given speed give better performance than others.

23.8 Line Conditioning

When the network consisted primarily of analog transmission facilities, attenuation and envelope delay distortion were major impairments. These were controlled through the use of line conditioning devices such as amplifiers and equalizers, which were installed at the ends of the transmission facilities. When required, equalization was almost always applied to the receiver side of a four-wire circuit. Sometimes, however, predistortion equalization was provided on the transmit side. The loop had little impact on performance unless it was very long.

In modern telecommunication networks, the transmission facilities are based on digital technologies such as optical fibers and digital microwave radios that do not introduce significant distortion. Any analog signal to be transported in today's network requires analog-to-digital (A/D) and digital-to-analog

TABLE 23.1 Analog Voiceband Data Transmission Objectives

Transmission Parameter	Group 1	Group 2		Group 3
		Tier 1	Tier 2	
Loss Deviation (dB)	±1	±1	±1	±1
Attenuation Distortion (dB)				
404–2804 Hz	-2.0/+5.0	-0.8/+1.5	-0.8/+3.5	-1.5/+9.5
304–3004 Hz	-2.0/+6.0	-0.8/+1.5	-0.8/+4.5	-2.5/+11.5
304–3204 Hz	-2.0/+8.0	Not specified	Not specified	Not specified
Envelope Delay Distortion (μs)				
804–2604 Hz	£550	£150	£650	£1700
604–2804 Hz	£800	£200	Not specified	Not specified
504–2804 Hz	Not specified	£550	Not specified	Not specified
504–3004 Hz	£1500	£2950	Not specified	Not specified
Intermodulation Distortion (dB)				
R2	≥49	≥32	≥32	≥26
R3	≥50	≥42	≥42	≥34
Signal-to-C-Notched Noise Ratio (dB)	≥33	≥32	≥32	≥26
Notched Impulse Noise Threshold (dBrnC)	63	65	65	69
Phase Jitter (∞Peak-Peak)				
20–300 Hz	Not applicable	Not applicable	Not applicable	£8
4–300 Hz	Not applicable	Not applicable	Not applicable	£13
Frequency Offset (Hz)	Not applicable	Not applicable	Not applicable	±1

(D/A) conversion at the ends. The analog signals usually are carried to the A/D conversion points by twisted pair loops. If the loops are long, the analog signals may first need to be equalized to offset a loop's high-frequency rolloff characteristic, which affects both attenuation and envelope delay distortion to some extent. Line conditioning is not used on switched voice frequency loops.

The inherent loss of a loop used in most special access (private line) applications can be reduced by adding gain with amplifiers. This has to be done carefully to prevent the loop from singing (oscillating) if two-wire conversion is used at any point in the circuit.

With the divestiture of AT&T in 1984, the specification and provisioning of end-to-end transmission facilities became much more difficult because the public network moved from being mostly one-dimensional to having multi-network operator and multi-dimensional characteristics. Any circuit that crossed exchange boundaries was a concatenation of elements provided by different network operators. A given network operator provided lines with predetermined characteristics according to the facilities owned and operated by them and them only. As a result, Bellcore (now Telcordia) developed standardized *voice grade* (VG) circuit types and transmission performance requirements that applied within the Bell Operating Company's service areas. These requirements recognized the widespread use at that time of analog transmission facilities throughout the Bell System and, thus, were very detailed as to analog transmission parameters. About 10 years later, the T1 Committee of the Alliance for Telecommunications Industry Solutions (ATIS), through the American National Standards Institute (ANSI), published a more general set of specifications that recognized the more widespread use of digital transmission facilities throughout the public network. Table 23.1 shows the characteristics specified by the ANSI standard [ANSI, 1994]. In this table, the attenuation distortion is shown with respect to the attenuation at 1004 Hz. The envelope delay distortion is given in terms of the difference between the maximum and minimum envelope delay (in microseconds) within the frequency band shown. All values are for circuits from network operator demarcation point (network interface) to network operator demarcation point. The ANSI requirements also specify limiting values for other transmission impairments, some of which do not apply to facilities composed entirely of digital transmission systems (Group 1 and 2 in the table).

23.9 Other Impairments to Analog Transmission

Intermodulation distortion (IMD), also known as nonlinear distortion, is an impairment that arises due to the nonlinear characteristics of loop interfaces. IMD is the power generated at extraneous frequencies (intermodulation products) when a multi-tone signal is applied to a circuit. Loop facilities composed entirely of twisted pair cables do not contribute to IMD but the cables always are connected to interfaces. In particular, nonlinearities in the interface electronics lead to generation of undesirable sum and difference signals related to the original desired signals. As with other transmission impairments, IMD manifests itself as analog voiceband modem errors and reduced throughput or connection speed, but IMD does not significantly impair voice quality. Where there is only a single A/D – D/A conversion process in an analog circuit, the ratio of desired signal to undesired signal due to IMD is around 40 to 50 dB.

Signal-to-Noise Ratio (SNR) is also important to modem transmission. In this case, the signal must be high enough above the noise level so the modem receiver can detect and recover the signal with low probability of error. SNR can be specified for both background noise and impulse noise. Typical values for SNR are greater than 30 dB. SNR is not as important on voice circuits as the absolute noise level (high noise during speech is not as disturbing as the same noise level during speech pauses).

Jitter and **Transients** can wreak havoc on high-speed voiceband modem signals and thus must be controlled. They have little effect on voice communications. Jitter is a small but significant periodic movement of the signal in either amplitude (amplitude jitter) or phase (phase jitter) from its desired value. Jitter in analog telephone channels is caused by analog interfaces (A/D and D/A converters) and transmission systems such as analog wireless transmission systems. Digital transmission systems do not introduce jitter on analog channels. Transients are departures of the signal amplitude or phase that exceed a threshold and are followed by quiet intervals. Transients can be caused by fading in wireless transmission systems (both digital and analog) and protection line switching in optical fiber and twisted pair transmission systems.

Defining Terms

Analog telephone channel: A telecommunications channel suitable for transmission of a band-limited, time continuous signal.

Voiceband modem: MOdulator-DEModulator, a device used for transmitting digital data over an analog voiceband channel.

Circuit noise: The noise heard by the subscriber or detected by an analog voiceband modem that appears across the pair of conductors in the loop.

Crosstalk: Interference coupled into the present telephone connection from adjacent channels or connections.

Envelope delay distortion (EDD): Distortion in the rate of change of phase shift as a function of frequency.

Impulse noise: Short voltage increases of 12 dB or more usually caused by electromechanical switching transients, maintenance activity, or electrical system switching transients.

Intermodulation distortion (IMD): The power generated at extraneous frequencies (intermodulation products) when a multi-tone signal is applied to a channel.

Intersymbol interference (ISI): Interference arising in modem signal detection from the edges of neighboring pulses.

Jitter: Relatively small but significant periodic movements of a signal's amplitude or phase from its desired value.

kb/s: Kilobits (1000 bits) per second.

Line conditioning: The use of amplifiers and equalizers or other devices to control the transmission characteristics of metallic twisted cable pairs.

Public switched telephone network (PSTN): Traditionally, the network developed to provide dial-up voice communications but now meant to include all telecommunication facilities available to the public.

Signal-to-noise ratio (SNR): The ratio of signal power to noise power in a channel. Analog voiceband modems require adequate SNR to achieve high throughput and low error rate.

Subscriber loop: The subscriber loop is the transmission and signaling channel between a telephone subscriber's terminal equipment and the network.

Transient: Sudden change in the amplitude or phase of a received signal that lasts at least 4 ms.

Voice band: The usable bandwidth of a telephone voice channel, often taken to be 300–3400 Hz.

References

AT&T. 1975. *Telecommunications Transmission Engineering,* Vols. 1–3, Western Electric Company, Inc. Winston-Salem, NC.

TELCORDIA. 2000. *Telcordia Notes on the Networks*. Telcordia. Morristown, NJ (www.telcordia.com).

CCITT. 1989. *Telephone Transmission Quality, Series P Recommendations,* CCITT Blue Book, Vol. V, Consultative Committee on International Telephony and Telegraphy. Geneva, Switzerland (www.itu.int).

IEEE. 1984. *IEEE Standard Telephone Loop Performance Characteristics.* IEEE Standard 820-1984 (R1999), Institute of Electrical and Electronics Engineers, New York (www.ieee.org).

ANSI. 1994. *Network Performance—Point-to-Point Voice-Grade Special Access Network Voiceband Data Transmission Objectives,* ANSI T1.512-1994, American National Standards Institute, New York (www.ansi.org).

Further Information

See *Subscriber Loop Signaling and Transmission Handbook: Analog* by Whitham D. Reeve, IEEE Press, 1992, *Subscriber Loop Signaling and Transmission Handbook: Digital* by Whitham D. Reeve, IEEE Press, 1995, and the *Telcordia Notes on the Networks*, 2000.

24

Baseband Signalling and Pulse Shaping

Michael L. Honig
Northwestern University

Melbourne Barton
Telcordia Technologies

Many physical communications channels, such as radio channels, accept a continuous-time waveform as input. Consequently, a sequence of source bits, representing data or a digitized analog signal, must be converted to a continuous-time waveform at the transmitter. In general, each successive group of bits taken from this sequence is mapped to a particular continuous-time pulse. In this chapter we discuss the basic principles involved in selecting such a pulse for channels that can be characterized as linear and time invariant with finite bandwidth.

24.1 Communications System Model

Figure 24.1a shows a simple block diagram of a communications system. The sequence of source bits $\{b_i\}$ are grouped into sequential blocks (vectors) of m bits $\{\boldsymbol{b}_i\}$, and each binary vector $\boldsymbol{b}_i$ is mapped to one of 2^m pulses, $p(\boldsymbol{b}_i; t)$, which is transmitted over the channel. The transmitted signal as a function of time can be written as

$$s(t) = \sum_i p(\boldsymbol{b}_i; t - iT) \tag{24.1}$$

where $1/T$ is the rate at which each group of m bits, or pulses, is introduced to the channel. The information (bit) rate is therefore m/T.

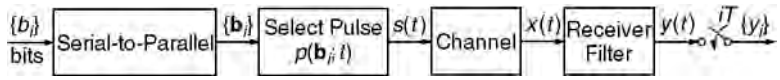


FIGURE 24.1a Communication system model. The source bits are grouped into binary vectors, which are mapped to a sequence of pulse shapes.

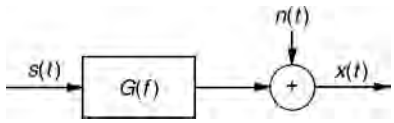


FIGURE 24.1b Channel model consisting of a linear, time-invariant system (transfer function) followed by additive noise.

The channel in Fig. 24.1a can be a radio link, which may distort the input signal $s(t)$ in a variety of ways. For example, it may introduce pulse dispersion (due to finite bandwidth) and multipath, as well as additive background noise. The output of the channel is denoted as $x(t)$, which is processed by the receiver to determine estimates of the source bits. The receiver can be quite complicated; however, for the purpose of this discussion, it is sufficient to assume only that it contains a front-end filter and a sampler, as shown in Fig. 24.1a. This assumption is valid for a wide variety of detection strategies. The purpose of the receiver filter is to remove noise outside of the transmitted frequency band and to compensate for the channel frequency response.

A commonly used channel model is shown in Fig. 24.1b and consists of a linear, time-invariant filter, denoted as $G(f)$, followed by additive noise $n(t)$. The channel output is, therefore,

$$x(t) = [g(t) * s(t)] + n(t) \tag{24.2}$$

where $g(t)$ is the channel impulse response associated with $G(f)$, and the asterisk denotes convolution,

$$g(t) * s(t) = \int_{-\infty}^{\infty} g(t-\tau)s(\tau)\,\mathrm{d}\tau$$

This channel model accounts for all linear, time-invariant channel impairments, such as finite bandwidth and time-invariant multipath. It does not account for time-varying impairments, such as rapid fading due to time-varying multipath. Nevertheless, this model can be considered valid over short time periods during which the multipath parameters remain constant.

In Figs. 24.1a, and 24.1b, it is assumed that all signals are **baseband signals**, which means that the frequency content is centered around $f = 0$ (DC). The channel passband, therefore, partially coincides with the transmitted spectrum. In general, this condition requires that the transmitted signal be modulated by an appropriate carrier frequency and demodulated at the receiver. In that case, the model in Figs. 24.1a, and 24.1b still applies; however, *baseband-equivalent* signals must be derived from their modulated (passband) counterparts. *Baseband signalling* and *pulse shaping* refers to the way in which a group of source bits is mapped to a baseband transmitted pulse.

As a simple example of baseband signalling, we can take $m = 1$ (map each source bit to a pulse), assign a 0 bit to a pulse $p(t)$, and a 1 bit to the pulse $-p(t)$. Perhaps the simplest example of a baseband pulse is the *rectangular* pulse given by $p(t) = 1$, $0 < t \le T$, and $p(t) = 0$ elsewhere. In this case, we can write the transmitted signal as

$$s(t) = \sum_i A_i p(t - iT) \tag{24.3}$$

where each symbol A_i takes on a value of +1 or -1, depending on the value of the ith bit, and $1/T$ is the *symbol rate*, namely, the rate at which the symbols A_i are introduced to the channel.

The preceding example is called *binary* **pulse amplitude modulation (PAM)**, since the data symbols A_i are binary valued, and they amplitude modulate the transmitted pulse $p(t)$. The information rate (bits per second) in this case is the same as the symbol rate $1/T$. As a simple extension of this signalling technique, we can increase m and choose A_i from one of $M = 2^m$ values to transmit at bit rate m/T. This is known as M-ary PAM. For example, letting $m = 2$, each pair of bits can be mapped to a pulse in the set $\{p(t), -p(t), 3p(t), -3p(t)\}$.

In general, the transmitted symbols $\{A_i\}$, the baseband pulse $p(t)$, and channel impulse response $g(t)$ can be *complex valued*. For example, each successive pair of bits might select a symbol from the set $\{1, -1, j, -j\}$, where $j = \sqrt{-1}$. This is a consequence of considering the baseband equivalent of passband modulation (that is, generating a transmitted spectrum which is centered around a carrier frequency f_c). Here we are not concerned with the relation between the passband and baseband equivalent models and simply point out that the discussion and results in this chapter apply to complex-valued symbols and pulse shapes.

As an example of a signalling technique which is not PAM, let $m = 1$ and

$$\begin{aligned} p(0;t) &= \begin{cases} \sqrt{2}\sin(2\pi f_1 t) & 0 < t < T \\ 0 & \text{elsewhere} \end{cases} \\ p(1;t) &= \begin{cases} \sqrt{2}\sin(2\pi f_2 t) & 0 < t < T \\ 0 & \text{elsewhere} \end{cases} \end{aligned} \tag{24.4}$$

where f_1 and $f_2 \neq f_1$ are fixed frequencies selected so that f_1T and f_2T (number of cycles for each bit) are multiples of 1/2. These pulses are *orthogonal*, namely,

$$\int_0^T p(1;t)p(0;t)\,\mathrm{d}t = 0$$

This choice of pulse shapes is called binary **frequency-shift keying (FSK)**.

Another example of a set of orthogonal pulse shapes for $m = 2$ bits/T is shown in Fig. 24.2. Because these pulses may have as many as three transitions within a symbol period, the transmitted spectrum occupies roughly four times the transmitted spectrum of binary PAM with a rectangular pulse shape. The spectrum is, therefore, spread across a much larger band than the smallest required for reliable transmission, assuming a data rate of $2/T$. This type of signalling is referred to as **spread-spectrum**.

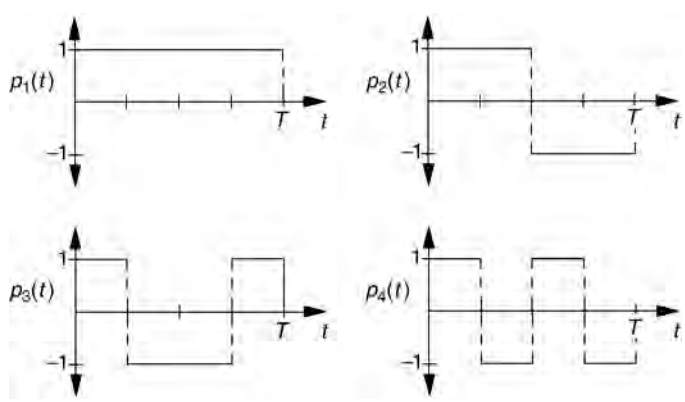


FIGURE 24.2 Four orthogonal spread-spectrum pulse shapes.

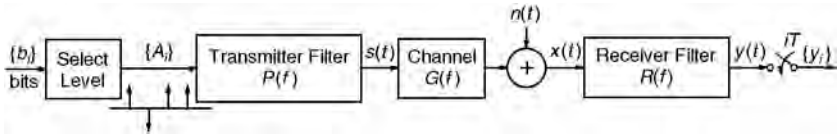


FIGURE 24.3 Baseband model of a pulse amplitude modulation system.

Spread-spectrum signals are more robust with respect to interference from other transmitted signals than are narrowband signals.[1]

24.2 Intersymbol Interference and the Nyquist Criterion

Consider the transmission of a PAM signal illustrated in Fig. 24.3. The source bits $\{b_i\}$ are mapped to a sequence of levels $\{A_i\}$, which modulate the transmitter pulse $p(t)$. The channel input is, therefore, given by Eq. (24.3) where $p(t)$ is the impulse response of the transmitter *pulse-shaping filter* $P(f)$ shown in Fig. 24.3. The input to the transmitter filter $P(f)$ is the modulated sequence of delta functions $\sum_i A_i\delta(t - iT)$. The channel is represented by the transfer function $G(f)$ (plus noise), which has impulse response $g(t)$, and the receiver filter has transfer function $R(f)$ with associated impulse response $r(t)$.

Let $h(t)$ be the overall impulse response of the combined transmitter, channel, and receiver, which has transfer function $H(f) = P(f)G(f)R(f)$. We can write $h(t) = p(t) * g(t) * r(t)$. The output of the receiver filter is then

$$y(t) = \sum_i A_i h(t - iT) + \tilde{n}(t) \tag{24.5}$$

where $\tilde{n}(t) = r(t) * n(t)$ is the output of the filter $R(f)$ with input $n(t)$. Assuming that samples are collected at the output of the filter $R(f)$ at the symbol rate $1/T$, we can write the kth sample of $y(t)$ as

$$\begin{aligned} y(kT) &= \sum_i A_i h(kT - iT) + \tilde{n}(kT) \\ &= A_k h(0) + \sum_{i \neq k} A_i h(kT - iT) + \tilde{n}(kT) \end{aligned} \tag{24.6}$$

The first term on the right-hand side of Eq. (24.6) is the kth transmitted symbol scaled by the system impulse response at $t = 0$. If this were the only term on the right side of Eq. (24.6), we could obtain the source bits without error by scaling the received samples by $1/h(0)$. The second term on the right-hand side of Eq. (24.6) is called **intersymbol interference**, which reflects the view that neighboring symbols interfere with the detection of each desired symbol.

One possible criterion for choosing the transmitter and receiver filters is to minimize intersymbol interference. Specifically, if we choose $p(t)$ and $r(t)$ so that

$$h(kT) = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases} \tag{24.7}$$

[1]This example can also be viewed as coded binary PAM. Namely, each pair of two source bits are mapped to 4 coded bits, which are transmitted via binary PAM with a rectangular pulse. The current IS-95 air interface uses an extension of this signalling method in which groups of 6 b are mapped to 64 orthogonal pulse shapes with as many as 63 transitions during a symbol.

then the kth received sample is

$$y(kT) = A_k + \tilde{n}(kT) \tag{24.8}$$

In this case, the intersymbol interference has been eliminated. This choice of $p(t)$ and $r(t)$ is called a **zero-forcing** solution, since it forces the intersymbol interference to zero. Depending on the type of detection scheme used, a zero-forcing solution may not be desirable. This is because the probability of error also depends on the noise intensity, which generally increases when intersymbol interference is suppressed. It is instructive, however, to examine the properties of the zero-forcing solution.

We now view Eq. (24.7) in the frequency domain. Since $h(t)$ has Fourier transform

$$H(f) = P(f)G(f)R(f) \tag{24.9}$$

where $P(f)$ is the Fourier transform of $p(t)$, the bandwidth of $H(f)$ is limited by the bandwidth of the channel $G(f)$. We will assume that $G(f) = 0, |f| > W$. The sampled impulse response $h(kT)$ can, therefore, be written as the inverse Fourier transform

$$h(kT) = \int_{-W}^{W} H(f)e^{j2\pi fkT}\mathrm{d}f$$

Through a series of manipulations, this integral can be rewritten as an inverse discrete Fourier transform,

$$h(kT) = T\int_{-1/(2T)}^{1/(2T)} H_{\mathrm{eq}}(e^{j2\pi fT})e^{j2\pi fkT}\mathrm{d}f \tag{24.10a}$$

where

$$\begin{aligned} H_{\mathrm{eq}}(e^{j2\pi fT}) &= \frac{1}{T}\sum_k H\left(f+\frac{k}{T}\right) \\ &= \frac{1}{T}\sum_k P\left(f+\frac{k}{T}\right)G\left(f+\frac{k}{T}\right)R\left(f+\frac{k}{T}\right) \end{aligned} \tag{24.10b}$$

This relation states that $H_{\mathrm{eq}}(z)$, $z = e^{j2\pi fT}$, is the discrete Fourier transform of the sequence $\{h_k\}$, where $h_k = h(kT)$. Sampling the impulse response $h(t)$ therefore changes the transfer function $H(f)$ to the *aliased* frequency response $H_{\mathrm{eq}}(e^{j2\pi fT})$. From Eqs. (24.10), and (24.6) we conclude that $H_{\mathrm{eq}}(z)$ is the transfer function that relates the sequence of input data symbols $\{A_i\}$ to the sequence of received samples $\{y_i\}$, where $y_i = y(iT)$, in the absence of noise. This is illustrated in Fig. 24.4. For this reason, $H_{\mathrm{eq}}(z)$ is called the **equivalent discrete-time transfer function** for the overall system transfer function $H(f)$.

Since $H_{\mathrm{eq}}(e^{j2\pi fT})$ is the discrete Fourier transform of the sequence $\{h_k\}$, the time-domain, or sequence condition (24.7) is equivalent to the frequency-domain condition

$$H_{\mathrm{eq}}(e^{j2\pi fT}) = 1 \tag{24.11}$$

This relation is called the **Nyquist criterion**. From Eqs. (24.10b) and (24.11) we make the following observations.

1. To satisfy the Nyquist criterion, the channel bandwidth W must be at least $1/(2T)$. Otherwise, $G(f + n/T) = 0$ for f in some interval of positive length for all n, which implies that $H_{\mathrm{eq}}(e^{j2\pi fT}) = 0$ for f in the same interval.

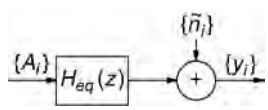


FIGURE 24.4 Equivalent discrete-time channel for the PAM system shown in Fig. 24.3 [$y_i = y(iT)$, $\tilde{n}_i = \tilde{n}(iT)$].

2. For the minimum bandwidth $W = 1/(2T)$, Eqs. (24.10b) and (24.11) imply that $H(f) = T$ for $|f| < 1/(2T)$ and $H(f) = 0$ elsewhere. This implies that the system impulse response is given by

$$h(t) = \frac{\sin(\pi t/T)}{\pi t/T} \tag{24.12}$$

(Since $\int_{-\infty}^{\infty} h^2(t)\,dt = T$, the transmitted signal $s(t) = \sum_i A_i\, h(t - iT)$ has power equal to the symbol variance $E[|A_i|^2]$.) The impulse response in Eq. (24.12) is called a *minimum bandwidth* or Nyquist pulse. The frequency band $[-1/(2T), 1/(2T)]$ [i.e., the passband of $H(f)$] is called the **Nyquist band**.

3. Suppose that the channel is bandlimited to twice the Nyquist bandwidth. That is, $G(f) = 0$ for $|f| > 1/T$. The condition (24.11) then becomes

$$H(f) + H\left(f - \frac{1}{T}\right) + H\left(f + \frac{1}{T}\right) = T \tag{24.13}$$

Assume for the moment that $H(f)$ and $h(t)$ are both real valued, so that $H(f)$ is an even function of $f[H(f) = H(-f)]$. This is the case when the receiver filter is the matched filter (see Section 24.3). We can then rewrite Eq. (24.13) as

$$H(f) + H\left(\frac{1}{T} - f\right) = T, \quad 0 < f < \frac{1}{2T} \tag{24.14}$$

which states that $H(f)$ must have odd symmetry about $f = 1/(2T)$. This is illustrated in Fig. 24.5, which shows two different functions $H(f)$ that satisfy the Nyquist criterion.

4. The pulse shape $p(t)$ enters into Eq. (24.11) only through the product $P(f)R(f)$. Consequently, either $P(f)$ or $R(f)$ can be fixed, and the other filter can be adjusted or adapted to the particular channel. Typically, the pulse shape $p(t)$ is fixed, and the receiver filter is adapted to the (possibly time-varying) channel.

Raised Cosine Pulse

Suppose that the channel is ideal with transfer function

$$G(f) = \begin{cases} 1, & |f| < W \\ 0, & |f| > W \end{cases} \tag{24.15}$$

To maximize bandwidth efficiency, Nyquist pulses given by Eq. (24.12) should be used where $W = 1/(2T)$. This type of signalling, however, has two major drawbacks. First, Nyquist pulses are noncausal and of infinite duration. They can be approximated in practice by introducing an appropriate delay and truncating the pulse. The pulse, however, decays very slowly, namely, as $1/t$, so that the truncation window must be wide.

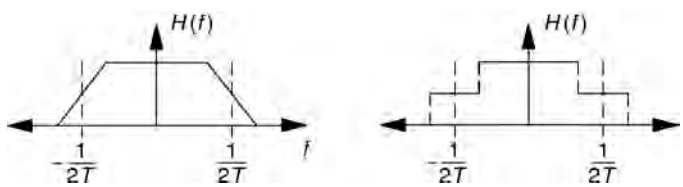


FIGURE 24.5 Two examples of frequency responses that satisfy the Nyquist criterion.

This is equivalent to observing that the ideal bandlimited frequency response given by Eq. (24.15) is difficult to approximate closely. The second drawback, which is more important, is the fact that this type of signalling is not robust with respect to sampling jitter. Namely, a small sampling offset ε produces the output sample

$$y(kT+\varepsilon) = \sum_i A_i \frac{\sin[\pi(k-i+\varepsilon/T)]}{\pi(k-i+\varepsilon/T)} \tag{24.16}$$

Since the Nyquist pulse decays as $1/t$, this sum is not guaranteed to converge. A particular choice of symbols $\{A_i\}$ can, therefore, lead to very large intersymbol interference, no matter how small the offset. Minimum bandwidth signalling is therefore impractical.

The preceding problem is generally solved in one of two ways in practice:

1. The pulse bandwidth is increased to provide a faster pulse decay than $1/t$.
2. A *controlled* amount of intersymbol interference is introduced at the transmitter, which can be subtracted out at the receiver.

The former approach sacrifices bandwidth efficiency, whereas the latter approach sacrifices power efficiency. We will examine the latter approach in Section 24.5. The most common example of a pulse, which illustrates the first technique, is the **raised cosine pulse**, given by

$$h(t) = \left[\frac{\sin(\pi t/T)}{\pi t/T}\right]\left[\frac{\cos(\alpha\pi t/T)}{1-(2\alpha t/T)^2}\right] \tag{24.17}$$

which has Fourier transform

$$H(f) = \begin{cases} T & 0 \le |f| \le \frac{1-\alpha}{2T} \\ \frac{T}{2}\left\{1+\cos\left[\frac{\pi T}{\alpha}\left(|f|-\frac{1-\alpha}{2T}\right)\right]\right\} & \frac{1-\alpha}{2T} \le |f| \le \frac{1+\alpha}{2T} \\ 0 & |f| > \frac{1+\alpha}{2T} \end{cases} \tag{24.18}$$

where $0 \le \alpha \le 1$.

Plots of $p(t)$ and $P(f)$ are shown in Figs. 24.6a, and 24.6b for different values of α. It is easily verified that $h(t)$ satisfies the Nyquist criterion (24.7) and, consequently, $H(f)$ satisfies Eq. (24.11). When $\alpha = 0$, $H(f)$ is the Nyquist pulse with minimum bandwidth $1/(2T)$, and when $\alpha > 0$, $H(f)$ has bandwidth $(1+\alpha)/(2T)$ with a raised cosine rolloff. The parameter α, therefore, represents the additional or **excess bandwidth** as a fraction of the minimum bandwidth $1/(2T)$. For example, when $\alpha = 1$, we say that the pulse is a raised cosine pulse with 100% excess bandwidth. This is because the pulse bandwidth $1/T$ is twice the minimum bandwidth. Because the raised cosine pulse decays as $1/t^3$, performance is robust with respect to sampling offsets.

The raised cosine frequency response (24.18) applies to the combination of transmitter, channel, and receiver. If the transmitted pulse shape $p(t)$ is a raised cosine pulse, then $h(t)$ is a raised cosine pulse only if the combined receiver and channel frequency response is constant. Even with an ideal (transparent) channel, however, the optimum (matched) receiver filter response is generally not constant in the presence of additive Gaussian noise. An alternative is to transmit the *square-root raised cosine* pulse shape, which has frequency response $P(f)$ given by the square-root of the raised cosine frequency response in Eq. (24.18). Assuming an ideal channel, setting the receiver frequency response $R(f) = P(f)$ then results in an overall raised cosine system response $H(f)$.

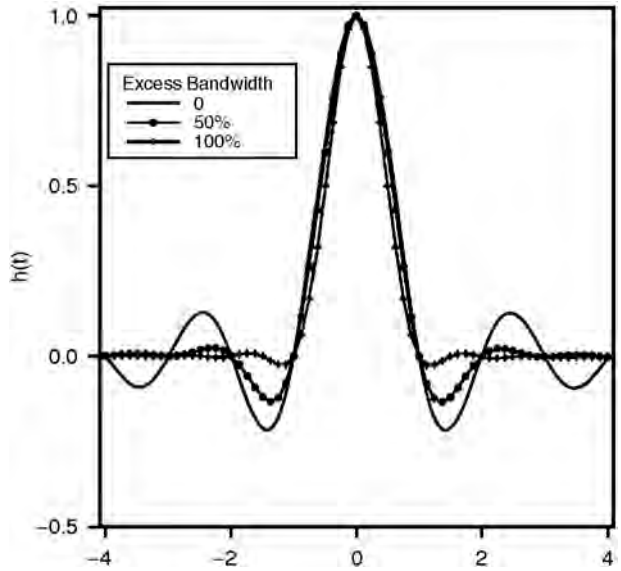


FIGURE 24.6a Raised cosine pulse.

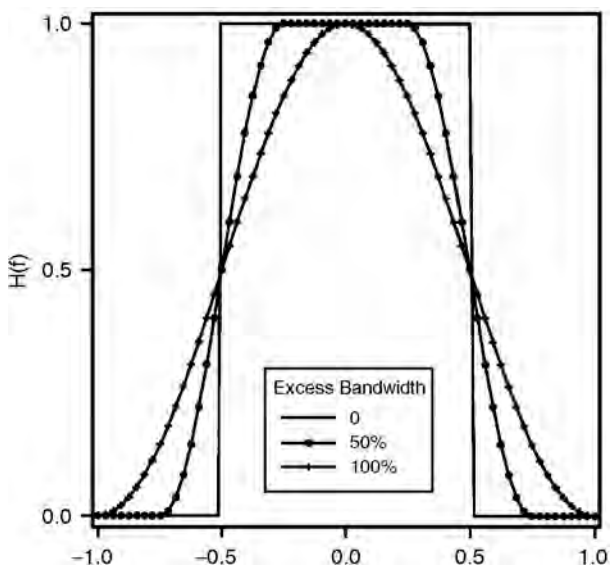


FIGURE 24.6b Raised cosine spectrum.

24.3 Nyquist Criterion with Matched Filtering

Consider the transmission of an isolated pulse $A_0\delta(t)$. In this case the input to the receiver in Fig. 24.3 is

$$x(t) = A_0\,\tilde{g}(t) + n(t) \tag{24.19}$$

where $\tilde{g}(t)$ is the inverse Fourier transform of the combined transmitter-channel transfer function $\tilde{G}(f) = P(f)G(f)$. We will assume that the noise $n(t)$ is white with spectrum $N_0/2$. The output of the receiver

filter is then

$$y(t) = r(t) * x(t) = A_0[r(t) * \tilde{g}(t)] + [r(t) * n(t)] \tag{24.20}$$

The first term on the right-hand side is the desired signal, and the second term is noise. Assuming that $y(t)$ is sampled at $t = 0$, the ratio of signal energy to noise energy, or signal-to-noise ratio (SNR) at the sampling instant, is

$$SNR = \frac{E[|A_0|^2]\left|\int_{-\infty}^{\infty} r(-t)\tilde{g}(t)\,\mathrm{d}t\right|^2}{\frac{N_0}{2}\int_{-\infty}^{\infty} |r(t)|^2 \mathrm{d}t} \tag{24.21}$$

The receiver impulse response that maximizes this expression is $r(t) = \tilde{g}^*(-t)$ [complex conjugate of $\tilde{g}(-t)$], which is known as the **matched filter** impulse response. The associated transfer function is $R(f) = \tilde{G}^*(-f)$.

Choosing the receiver filter to be the matched filter is optimal in more general situations, such as when detecting a sequence of channel symbols with intersymbol interference (assuming the additive noise is Gaussian). We, therefore, reconsider the Nyquist criterion when the receiver filter is the matched filter. In this case, the baseband model is shown in Fig. 24.7, and the output of the receiver filter is given by

$$y(t) = \sum_i A_i h(t - iT) + \tilde{n}(t) \tag{24.22}$$

where the baseband pulse $h(t)$ is now the impulse response of the filter with transfer function $|\tilde{G}(f)|^2 = |P(f)G(f)|^2$. This impulse response is the *autocorrelation* of the impulse response of the combined transmitter-channel filter $\tilde{G}(f)$,

$$h(t) = \int_{-\infty}^{\infty} \tilde{g}^*(s)\tilde{g}(s+t)\,\mathrm{d}s \tag{24.23}$$

With a matched filter at the receiver, the equivalent discrete-time transfer function is

$$\begin{aligned} H_{\mathrm{eq}}(e^{j2\pi fT}) &= \frac{1}{T}\sum_k \left|\tilde{G}\left(f - \frac{k}{T}\right)\right|^2 \\ &= \frac{1}{T}\sum_k \left|P\left(f - \frac{k}{T}\right)G\left(f - \frac{k}{T}\right)\right|^2 \end{aligned} \tag{24.24}$$

which relates the sequence of transmitted symbols $\{A_k\}$ to the sequence of received samples $\{y_k\}$ in the absence of noise. Note that $H_{\mathrm{eq}}(e^{j2\pi fT})$ is positive, real valued, and an even function of f. If the channel is bandlimited to twice the Nyquist bandwidth, then $H(f) = 0$ for $|f| > 1/T$, and the Nyquist condition is given by Eq. (24.14) where $H(f) = |G(f)P(f)|^2$. The aliasing sum in Eq. (24.10b) can therefore be described as a folding operation in which the channel response $|H(f)|^2$ is folded around the Nyquist frequency $1/(2T)$. For this reason, $H_{\mathrm{eq}}(e^{j2\pi fT})$ with a matched receiver filter is often referred to as the folded channel spectrum.

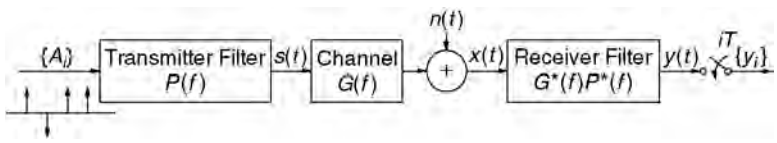


FIGURE 24.7 Baseband PAM model with a matched filter at the receiver.

24.4 Eye Diagrams

One way to assess the severity of distortion due to intersymbol interference in a digital communications system is to examine the **eye diagram**. The eye diagram is illustrated in Figs. 24.8a and 24.8b for a raised cosine pulse shape with 25% excess bandwidth and an ideal bandlimited channel. Figure 24.8a shows the data signal at the receiver,

$$y(t) = \sum_i A_i h(t - iT) + \tilde{n}(t) \tag{24.25}$$

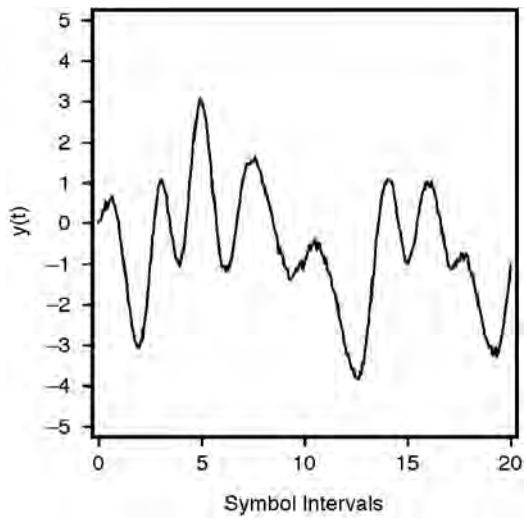


FIGURE 24.8a Received signal $y(t)$.

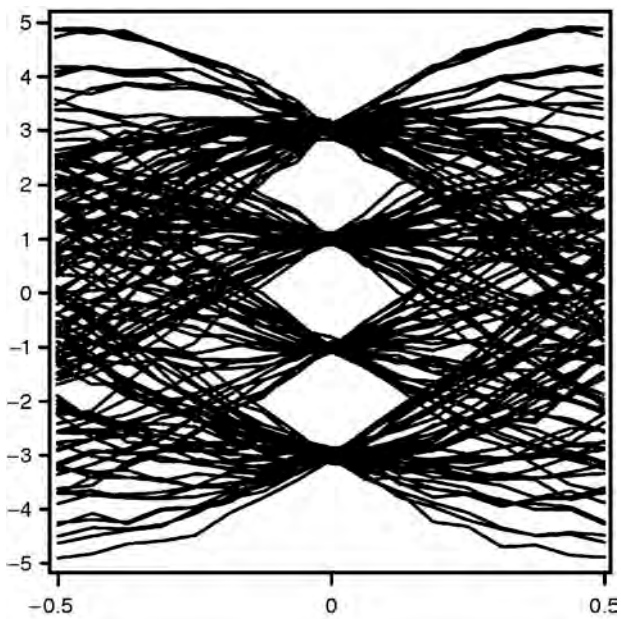


FIGURE 24.8b Eye diagram for received signal shown in Fig. 24.8a.

where $h(t)$ is given by Eq. (24.17), $\alpha = 1/4$, each symbol A_i is independently chosen from the set $\{\pm 1, \pm 3\}$, where each symbol is equally likely, and $\tilde{n}(t)$ is bandlimited white Gaussian noise. (The received *SNR* is 30 dB.) The eye diagram is constructed from the time-domain data signal $y(t)$ as follows (assuming nominal sampling times at kT, $k = 0, 1, 2, \ldots$):

1. Partition the waveform $y(t)$ into successive segments of length T starting from $t = T/2$.
2. Translate each of these waveform segments $[y(t), (k + 1/2)T \le t \le (k + 3/2)T, k = 0, 1, 2, \ldots]$ to the interval $[-T/2, T/2]$, and superimpose.

The resulting picture is shown in Fig. 24.8(b) for the $y(t)$ shown in Fig. 24.8(a). (Partitioning $y(t)$ into successive segments of length iT, $i > 1$, is also possible. This would result in i successive eye diagrams.) The number of eye openings is one less than the number of transmitted signal levels. In practice, the eye diagram is easily viewed on an oscilloscope by applying the received waveform $y(t)$ to the vertical deflection plates of the oscilloscope and applying a sawtooth waveform at the symbol rate $1/T$ to the horizontal deflection plates. This causes successive symbol intervals to be translated into one interval on the oscilloscope display.

Each waveform segment $y(t)$, $(k + 1/2)T \le t \le (k + 3/2)T$, depends on the particular sequence of channel symbols surrounding A_k. The number of channel symbols that affects a particular waveform segment depends on the extent of the intersymbol interference, shown in Eq. (24.6). This, in turn, depends on the duration of the impulse response $h(t)$. For example, if $h(t)$ has most of its energy in the interval $0 < t < mT$, then each waveform segment depends on approximately m symbols. Assuming binary transmission, this implies that there are a total of 2^m waveform segments that can be superimposed in the eye diagram. (It is possible that only one sequence of channel symbols causes significant intersymbol interference, and this sequence occurs with very low probability.) In current digital wireless applications the impulse response typically spans only a few symbols.

The eye diagram has the following important features which measure the performance of a digital communications system.

Vertical Eye Opening

The vertical openings at any time t_0, $-T/2 \le t_0 \le T/2$, represent the separation between signal levels with worst-case intersymbol interference, assuming that $y(t)$ is sampled at times $t = kT + t_0$, $k = 0, 1, 2, \ldots$. It is possible for the intersymbol interference to be large enough so that this vertical opening between some, or all, signal levels disappears altogether. In that case, the eye is said to be closed. Otherwise, the eye is said to be open. A closed eye implies that if the estimated bits are obtained by thresholding the samples $y(kT)$, then the decisions will depend primarily on the intersymbol interference rather than on the desired symbol. The probability of error will, therefore, be close to 1/2. Conversely, wide vertical spacings between signal levels imply a large degree of immunity to additive noise. In general, $y(t)$ should be sampled at the times $kT + t_0$, $k = 0, 1, 2, \ldots$, where t_0 is chosen to maximize the vertical eye opening.

Horizontal Eye Opening

The width of each opening indicates the sensitivity to timing offset. Specifically, a very narrow eye opening indicates that a small timing offset will result in sampling where the eye is closed. Conversely, a wide horizontal opening indicates that a large timing offset can be tolerated, although the error probability will depend on the vertical opening.

Slope of the Inner Eye

The slope of the inner eye indicates sensitivity to timing jitter or variance in the timing offset. Specifically, a very steep slope means that the eye closes rapidly as the timing offset increases. In this case, a significant amount of jitter in the sampling times significantly increases the probability of error.

The shape of the eye diagram is determined by the pulse shape. In general, the faster the baseband pulse decays, the wider the eye opening. For example, a rectangular pulse produces a box-shaped eye diagram (assuming binary signalling). The minimum bandwidth pulse shape Eq. (24.12) produces an eye diagram which is closed for all t except for $t = 0$. This is because, as shown earlier, an arbitrarily small timing offset can lead to an intersymbol interference term that is arbitrarily large, depending on the data sequence.

24.5 Partial-Response Signalling

To avoid the problems associated with Nyquist signalling over an ideal bandlimited channel, bandwidth and/or power efficiency must be compromised. Raised cosine pulses compromise bandwidth efficiency to gain robustness with respect to timing errors. Another possibility is to introduce a controlled amount of intersymbol interference at the transmitter, which can be removed at the receiver. This approach is called **partial-response (PR) signalling**. The terminology reflects the fact that the sampled system impulse response does not have the full response given by the Nyquist condition Eq. (24.7).

To illustrate PR signalling, suppose that the Nyquist condition Eq. (24.7) is replaced by the condition

$$h_k = \begin{cases} 1 & k = 0, 1 \\ 0 & \text{all other } k \end{cases} \tag{24.26}$$

The kth received sample is then

$$y_k = A_k + A_{k-1} + \tilde{n}_k \tag{24.27}$$

so that there is intersymbol interference from one neighboring transmitted symbol. For now we focus on the spectral characteristics of PR signalling and defer discussion of how to detect the transmitted sequence $\{A_k\}$ in the presence of intersymbol interference. The equivalent discrete-time transfer function in this case is the discrete Fourier transform of the sequence in Eq. (24.26),

$$\begin{aligned} H_{\text{eq}}(e^{j2\pi fT}) &= \frac{1}{T}\sum_k H\left(f + \frac{k}{T}\right) \\ &= 1 + e^{-j2\pi fT} = 2e^{-j\pi fT}\cos(\pi fT) \end{aligned} \tag{24.28}$$

As in the full-response case, for Eq. (24.28) to be satisfied, the *minimum* bandwidth of the channel $G(f)$ and transmitter filter $P(f)$ is $W = 1/(2T)$. Assuming $P(f)$ has this minimum bandwidth implies

$$H(f) = \begin{cases} 2Te^{-j\pi fT}\cos(\pi fT) & |f| < 1/(2T) \\ 0 & |f| > 1/(2T) \end{cases} \tag{24.29a}$$

and

$$h(t) = T\{\text{sinc}\,(t/T) + \text{sinc}\,[(t-T)/T]\} \tag{24.29b}$$

where sinc $x = (\sin \pi x)/(\pi x)$. This pulse is called a *duobinary* pulse and is shown along with the associated $H(f)$ in Fig. 24.9. [Notice that $h(t)$ satisfies Eq. (24.26).] Unlike the ideal bandlimited frequency response, the transfer function $H(f)$ in Eq. (24.29a) is continuous and is, therefore, easily approximated by a physically realizable filter. Duobinary PR was first proposed by Lender [7] and later generalized by Kretzmer [6].

The main advantage of the duobinary pulse Eq. (24.29b), relative to the minimum bandwidth pulse Eq. (24.12), is that signalling at the Nyquist symbol rate is feasible with zero excess bandwidth. Because the pulse decays much more rapidly than a Nyquist pulse, it is robust with respect to timing errors.

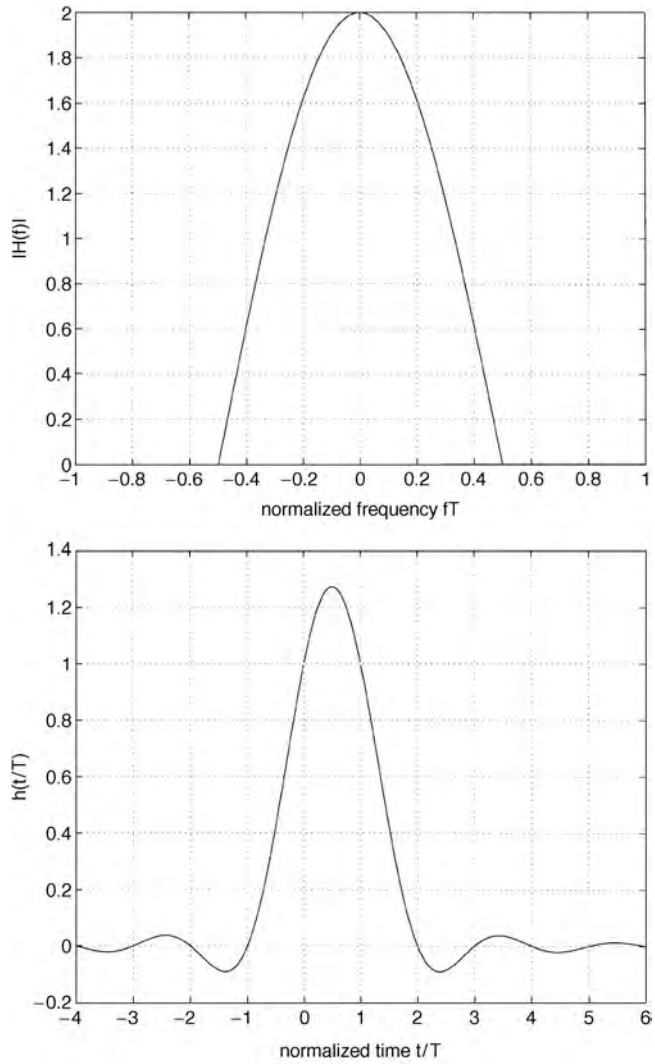


FIGURE 24.9 Duobinary frequency response and minimum bandwidth pulse.

Selecting the transmitter and receiver filters so that the overall system response is duobinary is appropriate in situations where the channel frequency response $G(f)$ is near zero or has a rapid rolloff at the Nyquist band edge $f = 1/(2T)$.

As another example of PR signalling, consider the *modified* duobinary partial response

$$h_k = \begin{cases} 1 & k = -1 \\ -1 & k = 1 \\ 0 & \text{all other } k \end{cases} \tag{24.30}$$

which has equivalent discrete-time transfer function

$$\begin{aligned} H_{\text{eq}}(e^{j2\pi fT}) &= e^{j2\pi fT} - e^{-j2\pi fT} \\ &= j2\sin(2\pi fT) \end{aligned} \tag{24.31}$$

With zero excess bandwidth, the overall system response is

$$H(f) = \begin{cases} j2T\sin(2\pi fT) & |f| < 1/(2T) \\ 0 & |f| > 1/(2T) \end{cases} \tag{24.32a}$$

and

$$h(t) = T\{\text{sinc}\,[(t+T)/T] - \text{sinc}\,[(t+T)/T]\} \tag{24.32b}$$

These functions are plotted in Fig. 24.10. This pulse shape is appropriate when the channel response $G(f)$ is near zero at both DC ($f = 0$) and at the Nyquist band edge. This is often the case for wire (twisted-pair) channels where the transmitted signal is coupled to the channel through a transformer. Like duobinary PR, modified duobinary allows minimum bandwidth signalling at the Nyquist rate.

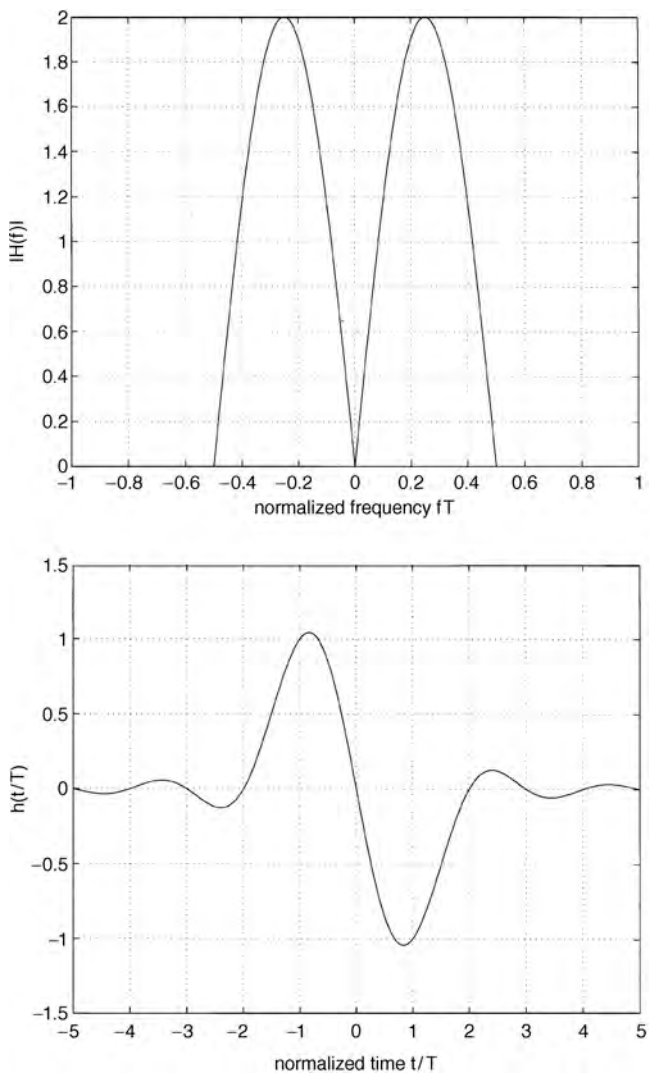


FIGURE 24.10 Modified duobinary frequency response and minimum bandwidth pulse.

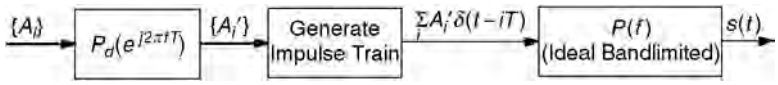


FIGURE 24.11 Generation of PR signal.

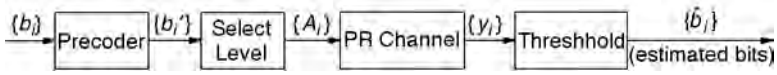


FIGURE 24.12 Precoding for a PR channel.

A particular partial response is often identified by the polynomial

$$\sum_{k=0}^{K} h_k D^k$$

where D (for delay) takes the place of the usual z^{-1} in the z transform of the sequence $\{h_k\}$. For example, duobinary is also referred to as $1 + D$ partial response.

In general, more complicated system responses than those shown in Figs. 24.9 and 24.10 can be generated by choosing more nonzero coefficients in the sequence $\{h_k\}$. This complicates detection, however, because of the additional intersymbol interference that is generated.

Rather than modulating a PR pulse $h(t)$, a PR signal can also be generated by filtering the sequence of transmitted levels $\{A_i\}$. This is shown in Fig. 24.11. Namely, the transmitted levels are first passed through a discrete-time (digital) filter with transfer function $P_d(e^{j2\pi fT})$ (where the subscript d indicates discrete). [Note that $P_d(e^{j2\pi fT})$ can be selected to be $H_{eq}(e^{j2\pi fT})$.] The outputs of this filter form the PAM signal, where the pulse shaping filter $P(f) = 1$, $|f| < 1/(2T)$ and is zero elsewhere. If the transmitted levels $\{A_k\}$ are selected independently and are identically distributed, then the transmitted spectrum is $\sigma_A^2 |P_d(e^{j2\pi fT})|^2$ for $|f| < 1/(2T)$ and is zero for $|f| > 1/(2T)$, where $\sigma_A^2 = E[|A_k|^2]$.

Shaping the transmitted spectrum to have nulls coincident with nulls in the channel response potentially offers significant performance advantages. By introducing intersymbol interference, however, PR signalling increases the number of received signal levels, which increases the complexity of the detector and may reduce immunity to noise. For example, the set of received signal levels for duobinary signalling is $\{0, \pm 2\}$ from which the transmitted levels $\{\pm 1\}$ must be estimated. The performance of a particular PR scheme depends on the channel characteristics as well as the type of detector used at the receiver. We now describe a simple suboptimal detection strategy.

Precoding

Consider the received signal sample Eq. (24.27) with duobinary signalling. If the receiver has correctly de-coded the symbol A_{k-1}, then in the absence of noise A_k can be decoded by subtracting A_{k-1} from the received sample y_k. If an error occurs, however, then subtracting the preceding symbol estimate from the received sample will cause the error to propagate to successive detected symbols. To avoid this problem, the transmitted levels can be **precoded** in such a way as to compensate for the intersymbol interference introduced by the overall partial response.

We first illustrate precoding for duobinary PR. The sequence of operations is illustrated in Fig. 24.12. Let $\{b_k\}$ denote the sequence of source bits where $b_k \in \{0, 1\}$. This sequence is transformed to the sequence $\{b'_k\}$ by the operation

$$b'_k = b_k \oplus b'_{k-1} \tag{24.33}$$

TABLE 24.1 Example of Precoding for Duobinary PR

$\{b_i\}$:		1	0	0	1	1	1	0	0	1	0
$\{b'_i\}$:	0	1	1	1	0	1	0	0	0	1	1
$\{A_i\}$:	−1	1	1	1	−1	1	−1	−1	−1	1	1
$\{y_i\}$:		0	2	2	0	0	0	−2	−2	0	2

where $\oplus$ denotes modulo 2 addition (exclusive OR). The sequence $\{b'_k\}$ is mapped to the sequence of binary transmitted signal levels $\{A_k\}$ according to

$$A_k = 2b'_k - 1 \tag{24.34}$$

That is, $b'_k = 0$ $(b'_k = 1)$ is mapped to the transmitted level $A_k = -1$ $(A_k = 1)$. In the absence of noise, the received symbol is then

$$y_k = A_k + A_{k-1} = 2(b'_k + b'_{k-1} - 1) \tag{24.35}$$

and combining Eqs. (24.33) and (24.35) gives

$$b_k = \left(\frac{1}{2}y_k + 1\right) \bmod 2 \tag{24.36}$$

That is, if $y_k = \pm 2$, then $b_k = 0$, and if $y_k = 0$, then $b_k = 1$. Precoding, therefore, enables the detector to make *symbol-by-symbol* decisions that do not depend on previous decisions. Table 24.1 shows a sequence of transmitted bits $\{b_i\}$, precoded bits $\{b'_i\}$, transmitted signal levels $\{A_i\}$, and received samples $\{y_i\}$.

The preceding precoding technique can be extended to multilevel PAM and to other PR channels. Suppose that the PR is specified by

$$H_{\text{eq}}(D) = \sum_{k=0}^{K} h_k D^k$$

where the coefficients are integers and that the source symbols $\{b_k\}$ are selected from the set $\{0, 1,\ldots, M-1\}$. These symbols are transformed to the sequence $\{b'_k\}$ via the precoding operation

$$b'_k = \left(b_k - \sum_{i=1}^{K} h_i b'_{k-i}\right) \bmod M \tag{24.37}$$

Because of the modulo operation, each symbol b'_k is also in the set $\{0, 1,\ldots, M-1\}$. The kth transmitted signal level is given by

$$A_k = 2b'_k - (M-1) \tag{24.38}$$

so that the set of transmitted levels is $\{-(M-1),\ldots,(M-1)\}$ (i.e., a shifted version of the set of values assumed by b_k). In the absence of noise, the received sample is

$$y_k = \sum_{i=0}^{K} h_i A_{k-i} \tag{24.39}$$

and it can be shown that the kth source symbol is given by

$$b_k = \frac{1}{2}(y_k + (M-1) \cdot H_{\text{eq}}(1)) \bmod M \tag{24.40}$$

Precoding the symbols $\{b_k\}$ in this manner, therefore, enables symbol-by-symbol decisions at the receiver. In the presence of noise, more sophisticated detection schemes (e.g., maximum likelihood) can be used with PR signalling to obtain improvements in performance.

24.6 Additional Considerations

In many applications, bandwidth and intersymbol interference are not the only important considerations for selecting baseband pulses. Here we give a brief discussion of additional practical constraints that may influence this selection.

Average Transmitted Power and Spectral Constraints

The constraint on average transmitted power varies according to the application. For example, low-average power is highly desirable for mobile wireless applications that use battery-powered transmitters. In many applications (e.g., digital subscriber loops, as well as digital radio), constraints are imposed to limit the amount of interference, or crosstalk, radiated into neighboring receivers and communications systems. Because this type of interference is frequency dependent, the constraint may take the form of a spectral mask that specifies the maximum allowable transmitted power as a function of frequency. For example, crosstalk in wireline channels is generally caused by capacitive coupling and increases as a function of frequency. Consequently, to reduce the amount of crosstalk generated at a particular transmitter, the pulse shaping filter generally attenuates high frequencies more than low frequencies.

In radio applications where signals are assigned different frequency bands, constraints on the transmitted spectrum are imposed to limit *adjacent-channel interference*. This interference is generated by transmitters assigned to adjacent frequency bands. Therefore, a constraint is needed to limit the amount of *out-of-band power* generated by each transmitter, in addition to an overall average power constraint. To meet this constraint, the transmitter filter in Fig. 24.3 must have a sufficiently steep rolloff at the edges of the assigned frequency band. (Conversely, if the transmitted signals are time multiplexed, then the duration of the system impulse response must be contained within the assigned time slot.)

Peak-to-Average Power

In addition to a constraint on average transmitted power, a *peak-power* constraint is often imposed as well. This constraint is important in practice for the following reasons:

1. The dynamic range of the transmitter is limited. In particular, saturation of the output amplifier will "clip" the transmitted waveform.
2. Rapid fades can severely distort signals with high peak-to-average power.
3. The transmitted signal may be subjected to nonlinearities. Saturation of the output amplifier is one example. Another example that pertains to wireline applications is the companding process in the voice telephone network [5]. Namely, the compander used to reduce quantization noise for pulse-code modulated voice signals introduces amplitude-dependent distortion in data signals.

The preceding impairments or constraints indicate that the transmitted waveform should have a low **peak-to-average power ratio (PAR)**. For a transmitted waveform $x(t)$, the PAR is defined as

$$\text{PAR} = \frac{\max |x(t)|^2}{E\{|x(t)|^2\}}$$

where E(·) denotes expectation. Using binary signalling with rectangular pulse shapes minimizes the PAR. However, this compromises bandwidth efficiency. In applications where PAR should be low, binary signalling with rounded pulses are often used. Operating RF power amplifiers with power back-off can also reduce PAR, but leads to inefficient amplification.

For an **orthogonal frequency division multiplexing (OFDM)** system, it is well known that the transmitted signal can exhibit a very high PAR compared to an equivalent single-carrier system. Hence more sophisticated approaches to PAR reduction are required for OFDM. Some proposed approaches are described in [8] and references therein. These include altering the set of transmitted symbols and setting aside certain OFDM tones specifically to minimize PAR.

Channel and Receiver Characteristics

The type of channel impairments encountered and the type of detection scheme used at the receiver can also influence the choice of a transmitted pulse shape. For example, a constant amplitude pulse is appropriate for a fast fading environment with noncoherent detection. The ability to track channel characteristics, such as phase, may allow more bandwidth efficient pulse shapes in addition to multilevel signalling.

High-speed data communications over time-varying channels requires that the transmitter and/or receiver adapt to the changing channel characteristics. Adapting the transmitter to compensate for a time-varying channel requires a feedback channel through which the receiver can notify the transmitter of changes in channel characteristics. Because of this extra complication, adapting the receiver is often preferred to adapting the transmitter pulse shape. However, the following examples are notable exceptions.

1. The current IS-95 air interface for direct-sequence code-division multiple access adapts the transmitter power to control the amount of interference generated and to compensate for channel fades. This can be viewed as a simple form of adaptive transmitter pulse shaping in which a single parameter associated with the pulse shape is varied.
2. Multitone modulation divides the channel bandwidth into small subbands, and the transmitted power and source bits are distributed among these subbands to maximize the information rate. The received signal-to-noise ratio for each subband must be transmitted back to the transmitter to guide the allocation of transmitted bits and power [1].

In addition to multitone modulation, *adaptive precoding* (also known as Tomlinson–Harashima precoding [4,11]) is another way in which the transmitter can adapt to the channel frequency response. Adaptive precoding is an extension of the technique described earlier for partial-response channels. Namely, the equivalent discrete-time channel impulse response is measured at the receiver and sent back to the transmitter, where it is used in a precoder. The precoder compensates for the intersymbol interference introduced by the channel, allowing the receiver to detect the data by a simple threshhold operation. Both multitone modulation and precoding have been used with wireline channels (voiceband modems and digital subscriber loops).

Complexity

Generation of a bandwidth-efficient signal requires a filter with a sharp cutoff. In addition, bandwidth-efficient pulse shapes can complicate other system functions, such as timing and carrier recovery. If sufficient bandwidth is available, the cost can be reduced by using a rectangular pulse shape with a simple detection strategy (low-pass filter and threshold).

Tolerance to Interference

Interference is one of the primary channel impairments associated with digital radio. In addition to adjacent-channel interference described earlier, *cochannel interference* may be generated by other transmitters assigned to the same frequency band as the desired signal. Cochannel interference can be controlled through frequency (and perhaps time slot) assignments and by pulse shaping. For example,

assuming fixed average power, increasing the bandwidth occupied by the signal lowers the power spectral density and decreases the amount of interference into a narrowband system that occupies part of the available bandwidth. Sufficient bandwidth spreading, therefore, enables wideband signals to be overlaid on top of narrowband signals without disrupting either service.

Probability of Intercept and Detection

The broadcast nature of wireless channels generally makes eavesdropping easier than for wired channels. A requirement for most commercial as well as military applications is to guarantee the privacy of user conversations (low probability of intercept). An additional requirement, in some applications, is that determining whether or not communications is taking place must be difficult (low probability of detection). Spread-spectrum waveforms are attractive in these applications since spreading the pulse energy over a wide frequency band decreases the power spectral density and, hence, makes the signal less visible. Power-efficient modulation combined with coding enables a further reduction in transmitted power for a target error rate.

24.7 Examples

We conclude this chapter with a brief description of baseband pulse shapes used in existing and emerging standards for digital mobile cellular and Personal Communications Services (PCS).

Global System for Mobile Communications (GSM)

The European GSM standard for digital mobile cellular communications operates in the 900-MHz frequency band and is based on time-division multiple access (TDMA) [9]. The U.S. version operates at 1900 MHz, and is called PCS-1900. A special variant of binary FSK is used called *Gaussian minimum-shift keying (GMSK)*. The GMSK modulator is illustrated in Fig. 24.13. The input to the modulator is a binary PAM signal $s(t)$, given by Eq. (24.3), where the pulse $p(t)$ is a Gaussian function and $|s(t)| < 1$. This waveform frequency modulates the carrier f_c, so that the (passband) transmitted signal is

$$w(t) = K\cos\left[2\pi f_c t + 2\pi f_d \int_{-\infty}^{t} s(\tau)\,\mathrm{d}\tau\right]$$

The maximum frequency deviation from the carrier is $f_d = 1/(2T)$, which characterizes minimum-shift keying. This technique can be used with a noncoherent receiver that is easy to implement. Because the transmitted signal has a constant envelope, the data can be reliably detected in the presence of rapid fades that are characteristic of mobile radio channels.

U.S. Digital Cellular (IS-136)

The IS-136 air interface (formerly IS-54) operates in the 800 MHz band and is based on TDMA [3]. There is also a 1900 MHz version of IS-136. The baseband signal is given by Eq. (24.3) where the symbols are complex-valued, corresponding to quadrature phase modulation. The pulse has a square-root raised cosine spectrum with 35% excess bandwidth.

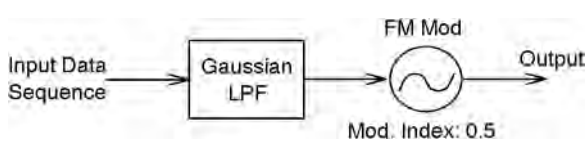


FIGURE 24.13 Generation of GMSK signal; LPF is low-pass filter.

Interim Standard-95

The IS-95 air interface for digital mobile cellular uses spread-spectrum signalling (CDMA) in the 800-MHz band [10]. There is also a 1900 MHz version of IS-95. The baseband transmitted pulse shapes are analogous to those shown in Fig. 24.2, where the number of square pulses (chips) per bit is 128. To improve spectral efficiency the (wideband) transmitted signal is filtered by an approximation to an ideal low-pass response with a small amount of excess bandwidth. This shapes the chips so that they resemble minimum bandwidth pulses.

Personal Access Communications System (PACS)

Both PACS and the Japanese personal handy phone (PHP) system are TDMA systems which have been proposed for personal communications systems (PCS), and operate near 2 GHz [2]. The baseband signal is given by Eq. (24.3) with four complex symbols representing four-phase quadrature modulation. The baseband pulse has a square-root raised cosine spectrum with 50% excess bandwidth.

Defining Terms

Baseband signal: A signal with frequency content centered around DC.

Equivalent discrete-time transfer function: A discrete-time transfer function (z transform) that relates the transmitted amplitudes to received samples in the absence of noise.

Excess bandwidth: That part of the baseband transmitted spectrum which is not contained within the Nyquist band.

Eye diagram: Superposition of segments of a received PAM signal that indicates the amount of intersymbol interference present.

Frequency-shift keying: A digital modulation technique in which the transmitted pulse is sinusoidal, where the frequency is determined by the source bits.

Intersymbol interference: The additive contribution (interference) to a received sample from transmitted symbols other than the symbol to be detected.

Matched filter: The receiver filter with impulse response equal to the time-reversed, complex conjugate impulse response of the combined transmitter filter-channel impulse response.

Nyquist band: The narrowest frequency band that can support a PAM signal without intersymbol interference (the interval $[-1/(2T), 1/(2T)]$ where $1/T$ is the symbol rate).

Nyquist criterion: A condition on the overall frequency response of a PAM system that ensures the absence of intersymbol interference.

Orthogonal frequency division multiplexing (OFDM): Modulation technique in which the transmitted signal is the sum of low-bit-rate narrowband digital signals modulated on orthogonal carriers.

Partial-response signalling: A signalling technique in which a controlled amount of intersymbol interference is introduced at the transmitter in order to shape the transmitted spectrum.

Precoding: A transformation of source symbols at the transmitter that compensates for intersymbol interference introduced by the channel.

Pulse amplitude modulation (PAM): A digital modulation technique in which the source bits are mapped to a sequence of amplitudes that modulate a transmitted pulse.

Raised cosine pulse: A pulse shape with Fourier transform that decays to zero according to a raised cosine; see Eq. (24.18). The amount of excess bandwidth is conveniently determined by a single parameter (α).

Spread spectrum: A signalling technique in which the pulse bandwidth is many times wider than the Nyquist bandwidth.

Zero-forcing criterion: A design constraint which specifies that intersymbol interference be eliminated.

References

1. Bingham, J.A.C., Multicarrier modulation for data transmission: an idea whose time has come. *IEEE Commun. Mag.,* 28(May), 5–14, 1990.
2. Cox, D.C., Wireless personal communications: what is it? *IEEE Personal Comm.,* 2(2), 20–35, 1995.
3. Electronic Industries Association/Telecommunications Industry Association. Recommended minimum performance standards for 800 MHz dual-mode mobile stations. Incorp. EIA/TIA 19B, EIA/TIA Project No. 2216, Mar., 1991.
4. Harashima, H. and Miyakawa, H., Matched-transmission technique for channels with intersymbol interference. *IEEE Trans. on Commun.,* COM-20(Aug.), 774–780, 1972.
5. Kalet, I. and Saltzberg, B.R., QAM transmission through a companding channel—signal constellations and detection. *IEEE Trans. on Comm.,* 42(2–4), 417–429, 1994.
6. Kretzmer, E.R., Generalization of a technique for binary data communication. *IEEE Trans. Comm. Tech.,* COM-14 (Feb.), 67, 68, 1966.
7. Lender, A., The duobinary technique for high-speed data transmission. *AIEE Trans. on Comm. Electronics,* 82 (March), 214–218, 1963.
8. Muller, S.H. and Huber, J.B., A comparison of peak power reduction schemes for OFDM. *Proc. GLOBECOM '97,* (Mon.), 1–5, 1997.
9. Rahnema, M., Overview of the GSM system and protocol architecture. *IEEE Commun. Mag.,* (April), 92–100, 1993.
10. Telecommunication Industry Association. Mobile station-base station compatibility standard for dual-mode wideband spread spectrum cellular system. TIA/EIA/IS-95-A. May, 1995.
11. Tomlinson, M., New automatic equalizer employing modulo arithmetic. *Electron. Lett.,* 7 (March), 138, 139, 1971.

Further Information

Baseband signalling and pulse shaping is fundamental to the design of any digital communications system and is, therefore, covered in numerous texts on digital communications. For more advanced treatments see E.A. Lee and D.G. Messerschmitt, *Digital Communication*, Kluwer 1994, and J.G. Proakis, *Digital Communications*, McGraw-Hill 1995.

25

Channel Equalization

John G. Proakis
Northeastern University

25.1 Characterization of Channel Distortion

Many communication channels, including telephone channels and some radio channels, may be generally characterized as band-limited linear filters. Consequently, such channels are described by their frequency response $C(f)$, which may be expressed as

$$C(f) = A(f)e^{j\theta(f)} \tag{25.1}$$

where $A(f)$ is called the *amplitude response* and $\theta(f)$ is called the *phase response*. Another characteristic that is sometimes used in place of the phase response is the *envelope delay* or *group delay*, which is defined as

$$\tau(f) = -\frac{1}{2\pi}\frac{d\theta(f)}{df} \tag{25.2}$$

A channel is said to be nondistorting or ideal if, within the bandwidth W occupied by the transmitted signal, $A(f) = \text{const}$ and $\theta(f)$ is a linear function of frequency [or the envelope delay $\tau(f) = \text{const}$]. On the other hand, if $A(f)$ and $\tau(f)$ are not constant within the bandwidth occupied by the transmitted signal, the channel distorts the signal. If $A(f)$ is not constant, the distortion is called *amplitude distortion* and if $\tau(f)$ is not constant, the distortion on the transmitted signal is called *delay distortion*.

As a result of the amplitude and delay distortion caused by the nonideal channel frequency response characteristic $C(f)$, a succession of pulses transmitted through the channel at rates comparable to the bandwidth W are smeared to the point that they are no longer distinguishable as well-defined pulses at the receiving terminal. Instead, they overlap and, thus, we have **intersymbol interference (ISI)**. As an example of the effect of delay distortion on a transmitted pulse, Fig. 25.1(a) illustrates a band-limited pulse having zeros periodically spaced in time at points labeled $\pm T$, $\pm 2T$, etc. If information is conveyed by the pulse amplitude, as in pulse amplitude modulation (PAM), for example, then one can transmit a sequence of pulses, each of which has a peak at the periodic zeros of the other pulses. Transmission of the pulse through a channel modeled as having a linear envelope delay characteristic $\tau(f)$ [quadratic

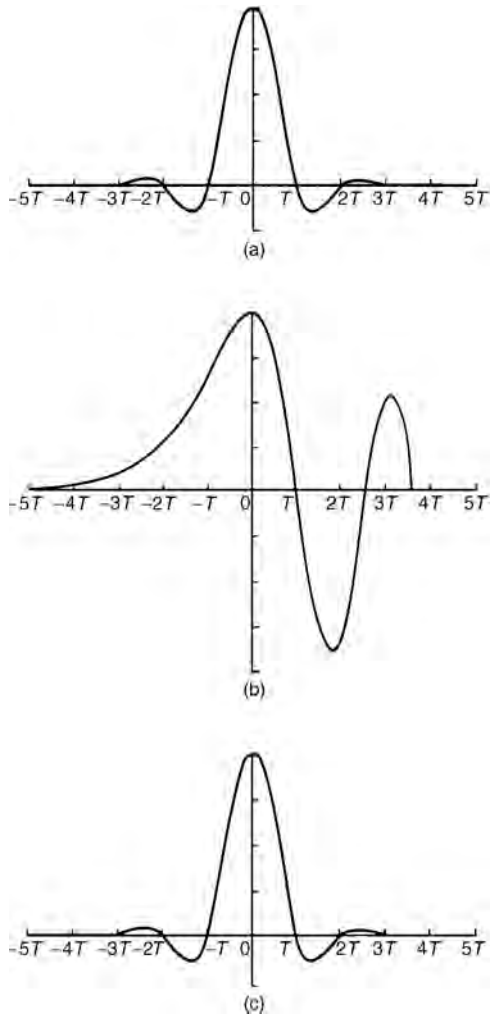


FIGURE 25.1 Effect of channel distortion: (a) channel input, (b) channel output, (c) equalizer output.

phase $\theta(f)$], however, results in the received pulse shown in Fig. 25.1(b) having zero crossings that are no longer periodically spaced. Consequently, a sequence of successive pulses would be smeared into one another, and the peaks of the pulses would no longer be distinguishable. Thus, the channel delay distortion results in intersymbol interference. As will be discussed in this chapter, it is possible to compensate for the nonideal frequency response characteristic of the channel by use of a filter or equalizer at the demodulator. Figure 25.1(c) illustrates the output of a linear equalizer that compensates for the linear distortion in the channel.

The extent of the intersymbol interference on a telephone channel can be appreciated by observing a frequency response characteristic of the channel. Figure 25.2 illustrates the measured average amplitude and delay as a function of frequency for a medium-range (180–725 mile) telephone channel of the switched telecommunications network as given by Duffy and Tratcher [17]. We observe that the usable band of the channel extends from about 300 Hz to about 3000 Hz. The corresponding impulse response of the average channel is shown in Fig. 25.3. Its duration is about 10 ms. In comparison, the transmitted symbol rates on such a channel may be of the order of 2500 pulses or symbols per second. Hence, intersymbol interference might extend over 20–30 symbols.

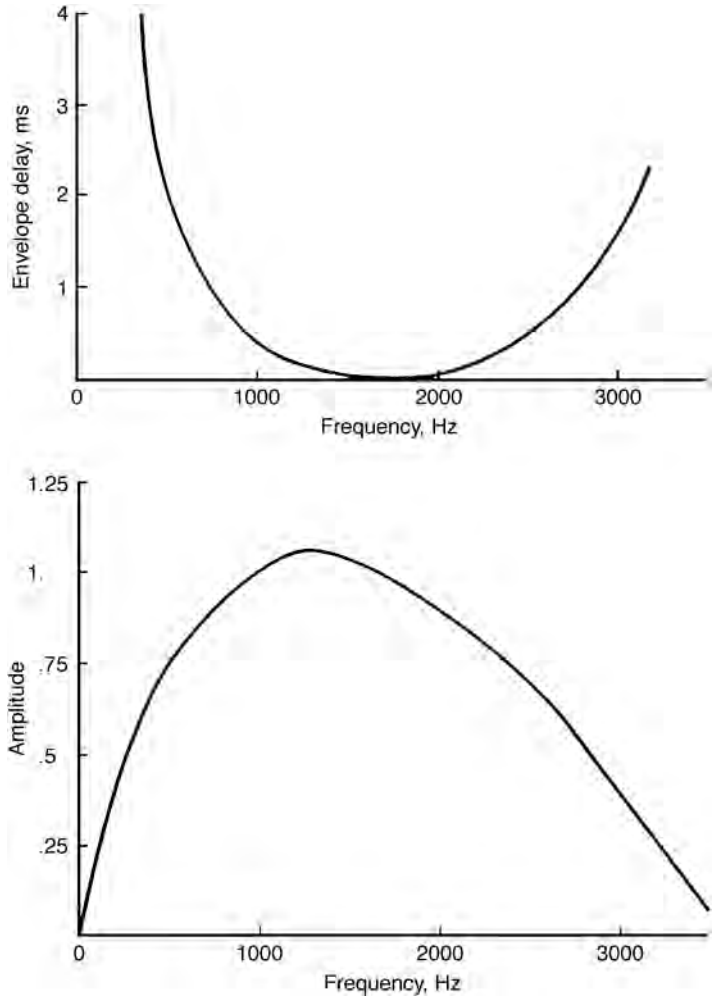


FIGURE 25.2 Average amplitude and delay characteristics of medium-range telephone channel.

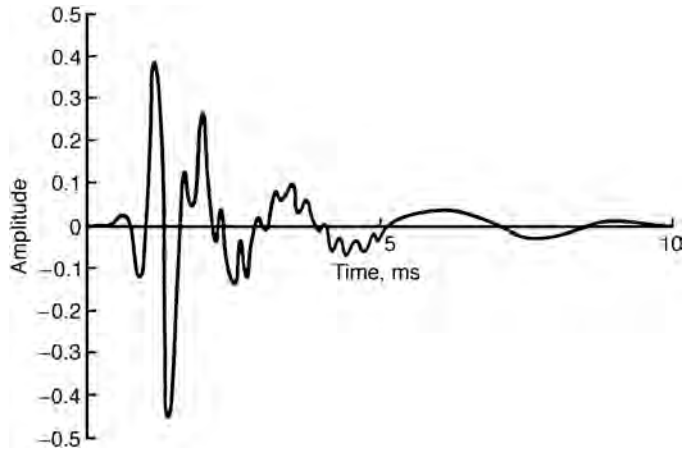


FIGURE 25.3 Impulse response of average channel with amplitude and delay shown in Fig. 25.2.

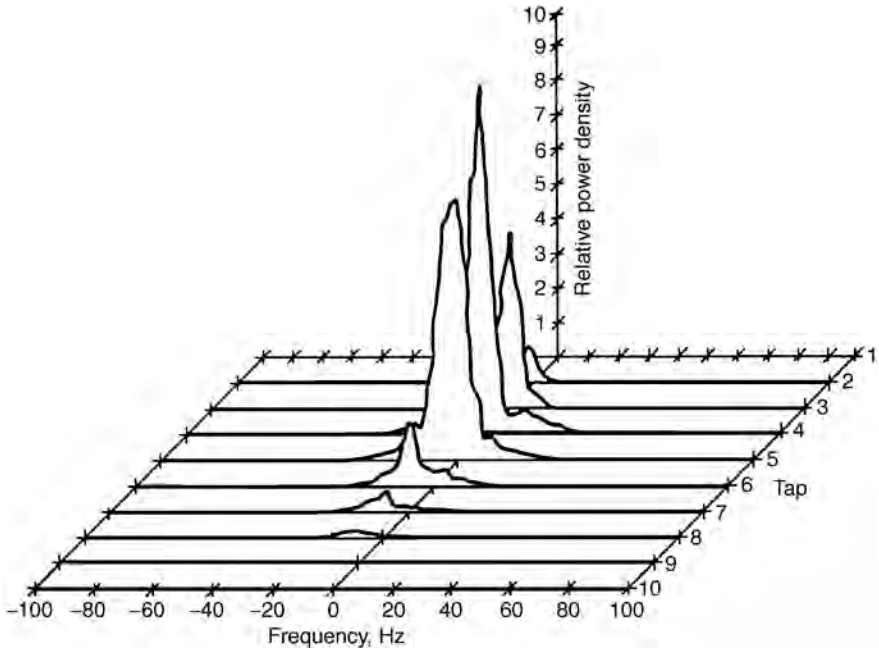


FIGURE 25.4 Scattering function of a medium-range tropospheric scatter channel.

Besides telephone channels, there are other physical channels that exhibit some form of time dispersion and, thus, introduce intersymbol interference. Radio channels, such as short-wave ionospheric propagation (HF), tropospheric scatter, and mobile cellular radio are three examples of time-dispersive wireless channels. In these channels, time dispersion and, hence, intersymbol interference is the result of multiple propagation paths with different path delays. The number of paths and the relative time delays among the paths vary with time and, for this reason, these radio channels are usually called time-variant multipath channels. The time-variant multipath conditions give rise to a wide variety of frequency response characteristics. Consequently, the frequency response characterization that is used for telephone channels is inappropriate for time-variant multipath channels. Instead, these radio channels are characterized statistically in terms of the scattering function, which, in brief, is a two-dimensional representation of the average received signal power as a function of relative time delay and Doppler frequency (see Proakis, 2001).

For illustrative purposes, a scattering function measured on a medium-range (150 mile) tropospheric scatter channel is shown in Fig. 25.4. The total time duration (multipath spread) of the channel response is approximately 0.7 μs on the average, and the spread between half-power points in Doppler frequency is a little less than 1 Hz on the strongest path and somewhat larger on the other paths. Typically, if one is transmitting at a rate of 10^7 symbols/s over such a channel, the multipath spread of 0.7 μs will result in intersymbol interference that spans about seven symbols.

25.2 Characterization of Intersymbol Interference

In a digital communication system, channel distortion causes intersymbol interference, as illustrated in the preceding section. In this section, we shall present a model that characterizes the ISI. The digital modulation methods to which this treatment applies are PAM, phase-shift keying (PSK), and quadrature amplitude modulation (QAM). The transmitted signal for these three types of modulation may be expressed as

$$\begin{aligned} s(t) &= u_c(t)\cos 2\pi f_c t - u_s(t)\sin 2\pi f_c t \\ &= \mathrm{Re}[u(t)e^{j2\pi f_c t}] \end{aligned} \tag{25.3}$$

where $\upsilon(t) = \upsilon_c(t) + j\upsilon_x(t)$ is called the *equivalent low-pass signal*, f_c is the carrier frequency, and Re[] denotes the real part of the quantity in brackets.

In general, the equivalent low-pass signal is expressed as

$$u(t) = \sum_{n=0} I_n g_T(t - nT) \tag{25.4}$$

where $g_T(t)$ is the basic pulse shape that is selected to control the spectral characteristics of the transmitted signal, $\{I_n\}$ is the sequence of transmitted information symbols selected from a signal constellation consisting of M points, and T is the signal interval ($1/T$ is the symbol rate). For PAM, PSK, and QAM, the values of I_n are points from M-ary signal constellations. Figure 25.5 illustrates the signal constellations for the case of $M = 8$ signal points. Note that for PAM, the signal constellation is one-dimensional. Hence, the equivalent low-pass signal $\upsilon(t)$ is real valued, i.e., $\upsilon_s(t) = 0$ and $\upsilon_c(t) = \upsilon(t)$. For M-ary ($M > 2$) PSK and QAM, the signal constellations are two-dimensional and, hence, $\upsilon(t)$ is complex valued.

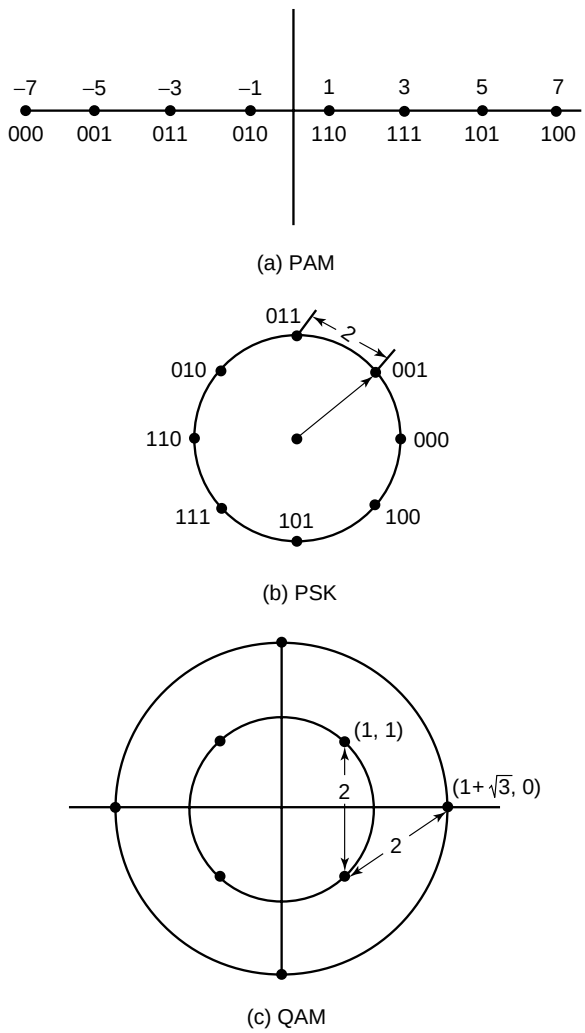


FIGURE 25.5 $M = 8$ signal constellations for PAM, PSK, and QAM.

The signal $s(t)$ is transmitted over a bandpass channel that may be characterized by an equivalent low-pass frequency response $C(f)$. Consequently, the equivalent low-pass received signal can be represented as

$$r(t) = \sum_{n=0}^{\infty} I_n h(t - nT) + w(t) \tag{25.5}$$

where $h(t) = g_T(t) * c(t)$, $c(t)$ is the impulse response of the equivalent low-pass channel, the asterisk denotes convolution, and $w(t)$ represents the additive noise in the channel.

To characterize the ISI, suppose that the received signal is passed through a receiving filter and then sampled at the rate $1/T$ samples/s. In general, the optimum filter at the receiver is matched to the received signal pulse $h(t)$. Hence, the frequency response of this filter is $H^*(f)$. We denote its output as

$$y(t) = \sum_{n=0}^{\infty} I_n x(t - nT) + v(t) \tag{25.6}$$

where $x(t)$ is the signal pulse response of the receiving filter, i.e., $X(f) = H(f)H^*(f) = |H(f)|^2$, and $v(t)$ is the response of the receiving filter to the noise $w(t)$. Now, if $y(t)$ is sampled at times $t = kT$, $k = 0, 1, 2, \ldots$, we have

$$\begin{aligned} y(kT) \equiv y_k &= \sum_{n=0}^{\infty} I_n x(kT - nT) + v(kT) \\ &= \sum_{n=0}^{\infty} I_n x_{k-n} + v_k, \quad k = 0, 1, \ldots \end{aligned} \tag{25.7}$$

The sample values $\{y_k\}$ can be expressed as

$$y_k = x_0 \left(I_k + \frac{1}{x_0} \sum_{\substack{n=0 \\ n \neq k}}^{\infty} I_n x_{k-n} \right) + v_k, \quad k = 0, 1, \ldots \tag{25.8}$$

The term x_0 is an arbitrary scale factor, which we arbitrarily set equal to unity for convenience. Then

$$y_k = I_k + \sum_{\substack{n=0 \\ n \neq k}}^{\infty} I_n x_{k-n} + v_k \tag{25.9}$$

The term I_k represents the desired information symbol at the kth sampling instant, the term

$$\sum_{\substack{n=0 \\ n \neq k}}^{\infty} I_n x_{k-n} \tag{25.10}$$

represents the ISI, and v_k is the additive noise variable at the kth sampling instant.

The amount of ISI and noise in a digital communications system can be viewed on an oscilloscope. For PAM signals, we can display the received signal $y(t)$ on the vertical input with the horizontal sweep

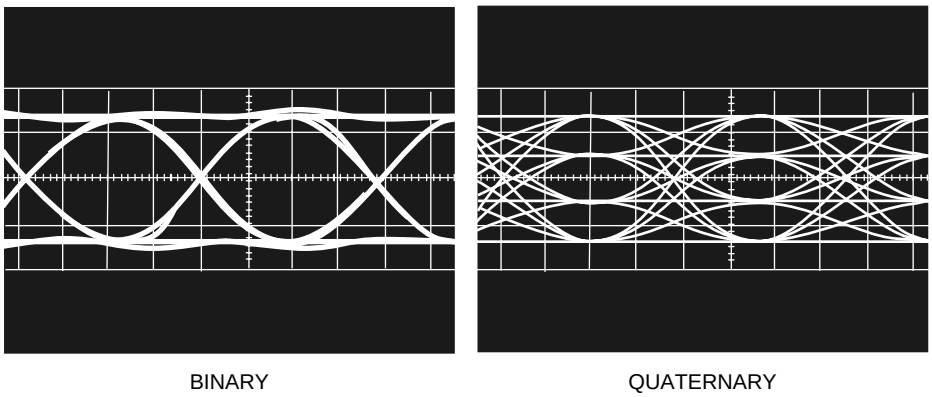


FIGURE 25.6 Examples of eye patterns for binary and quaternary amplitude shift keying (or PAM).

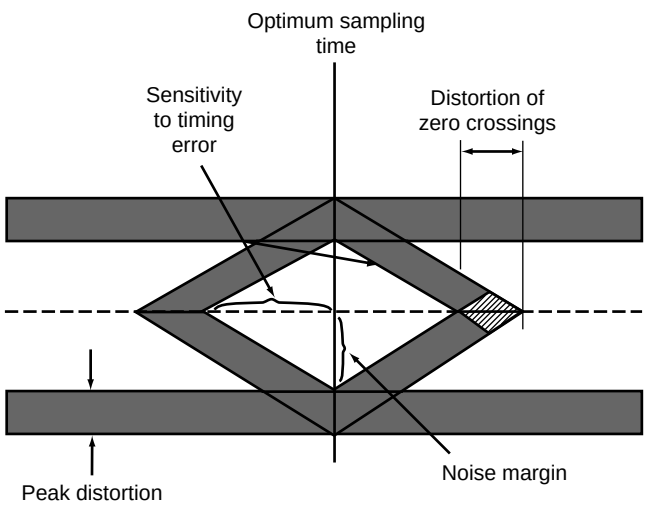


FIGURE 25.7 Effect of intersymbol interference on eye opening.

rate set at $1/T$. The resulting oscilloscope display is called an *eye pattern* because of its resemblance to the human eye. For example, Fig. 25.6 illustrates the eye patterns for binary and four-level PAM modulation. The effect of ISI is to cause the eye to close, thereby reducing the margin for additive noise to cause errors. Figure 25.7 graphically illustrates the effect of ISI in reducing the opening of a binary eye. Note that intersymbol interference distorts the position of the *zero* crossings and causes a reduction in the eye opening. Thus, it causes the system to be more sensitive to a synchronization error.

For PSK and QAM it is customary to display the eye pattern as a two-dimensional scatter diagram illustrating the sampled values $\{y_k\}$ that represent the decision variables at the sampling instants. Figure 25.8 illustrates such an eye pattern for an 8-PSK signal. In the absence of intersymbol interference and noise, the superimposed signals at the sampling instants would result in eight distinct points corresponding to the eight transmitted signal phases. Intersymbol interference and noise result in a deviation of the received samples $\{y_k\}$ from the desired 8-PSK signal. The larger the intersymbol interference and noise, the larger the scattering of the received signal samples relative to the transmitted signal points.

In practice, the transmitter and receiver filters are designed for zero ISI at the desired sampling times $t = kT$. Thus, if $G_T(f)$ is the frequency response of the transmitter filter and $G_R(f)$ is the frequency

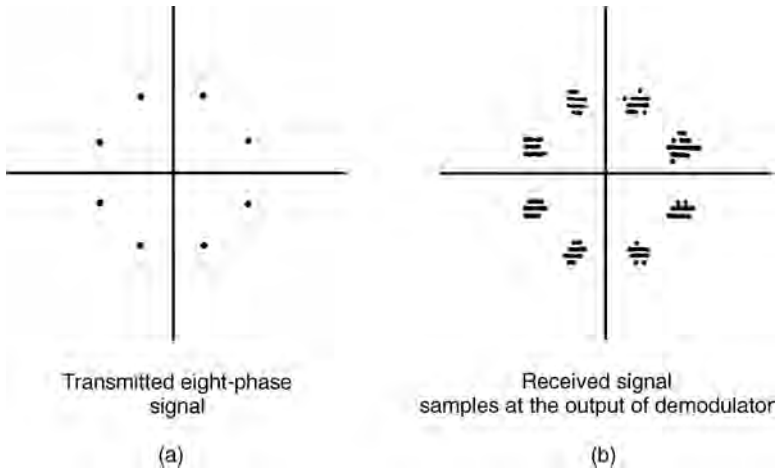


FIGURE 25.8 Two-dimensional digital eye patterns.

response of the receiver filter, then the product $G_T(f)G_R(f)$ is designed to yield zero ISI. For example, the product $G_T(f)G_R(f)$ may be selected as

$$G_T(f)G_R(f) = X_{rc}(f) \tag{25.11}$$

where $X_{rc}(f)$ is the raised-cosine frequency response characteristic, defined as

$$X_{rc}(f) = \begin{cases} T, & 0 \le |f| \le (1-\alpha)/2T \\ \frac{T}{2}\left[1+\cos\frac{\pi T}{\alpha}\left(|f|-\frac{1-\alpha}{2T}\right)\right], & \frac{1-\alpha}{2T} \le |f| \le \frac{1+\alpha}{2T} \\ 0, & |f| > \frac{1+\alpha}{2T} \end{cases} \tag{25.12}$$

where α is called the *rolloff* factor, which takes values in the range $0 \le \alpha \le 1$, and $1/T$ is the symbol rate. The frequency response $X_{rc}(f)$ is illustrated in Fig. 25.9(a) for $\alpha = 0$, 1/2, and 1. Note that when $\alpha = 0$, $X_{rc}(f)$ reduces to an ideal brick wall physically nonrealizable frequency response with bandwidth occupancy $1/2T$. The frequency $1/2T$ is called the *Nyquist frequency*. For $\alpha > 0$, the bandwidth occupied by the desired signal $X_{rc}(f)$ beyond the Nyquist frequency $1/2T$ is called the *excess bandwidth* and is usually expressed as a percentage of the Nyquist frequency. For example, when $\alpha = 1/2$, the excess bandwidth is 50%, and when $\alpha = 1$, the excess bandwidth is 100%. The signal pulse $x_{rc}(t)$ having the raised-cosine spectrum is

$$x_{rc}(t) = \frac{\sin \pi t/T}{\pi t/T} \quad \frac{\cos(\pi\alpha t/T)}{1-4\alpha^2 t^2/T^2} \tag{25.13}$$

Figure 25.9(b) illustrates $x_{rc}(t)$ for $\alpha = 0$, 1/2, and 1. Note that $x_{rc}(t) = 1$ at $t = 0$ and $x_{rc}(t) = 0$ at $t = kT$, $k = \pm 1, \pm 2, \ldots$. Consequently, at the sampling instants $t = kT$, $k \ne 0$, there is no ISI from adjacent symbols when there is no channel distortion. In the presence of channel distortion, however, the ISI given by Eq. (25.10) is no longer zero, and a channel equalizer is needed to minimize its effect on system performance.

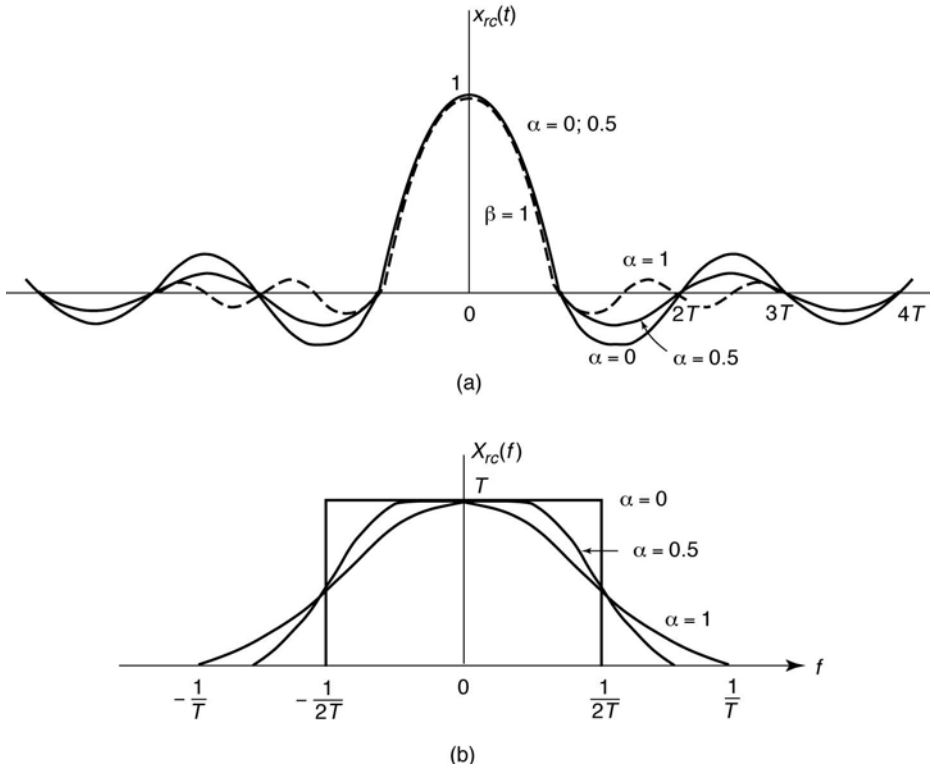


FIGURE 25.9 Pulses having a raised-cosine spectrum.

25.3 Linear Equalizers

The most common type of channel equalizer used in practice to reduce ISI is a linear transversal filter with adjustable coefficients $\{c_i\}$, as shown in Fig. 25.10.

On channels whose frequency response characteristics are unknown but are time invariant, we may measure the channel characteristics and adjust the parameters of the equalizer; once adjusted, the parameters remain fixed during the transmission of data. Such equalizers are called **preset equalizers**. On the other hand, **adaptive equalizers** update their parameters on a periodic basis during the transmission of data and, thus, they are capable of tracking a slowly time-varying channel response.

First, let us consider the design characteristics for a linear equalizer from a frequency domain viewpoint. Figure 25.11 shows a block diagram of a system that employs a linear filter as a **channel equalizer**.

The demodulator consists of a receiver filter with frequency response $G_R(f)$ in cascade with a channel equalizing filter that has a frequency response $G_E(f)$. As indicated in the preceding section, the receiver filter response $G_R(f)$ is matched to the transmitter response, i.e., $G_R(f) = G_T^*(f)$, and the product $G_R(f)G_T(f)$ is usually designed so that there is zero ISI at the sampling instants as, for example, when $G_R(f)G_T(f) = X_{rc}(f)$.

For the system shown in Fig. 25.11, in which the channel frequency response is not ideal, the desired condition for zero ISI is

$$G_T(f)C(f)G_R(f)G_E(f) = X_{rc}(f) \tag{25.14}$$

where $X_{rc}(f)$ is the desired raised-cosine spectral characteristic. Since $G_T(f)G_R(f) = X_{rc}(f)$ by design, the frequency response of the equalizer that compensates for the channel distortion is

$$G_E(f) = \frac{1}{C(f)} = \frac{1}{|C(f)|}e^{-j\theta_c(f)} \tag{25.15}$$

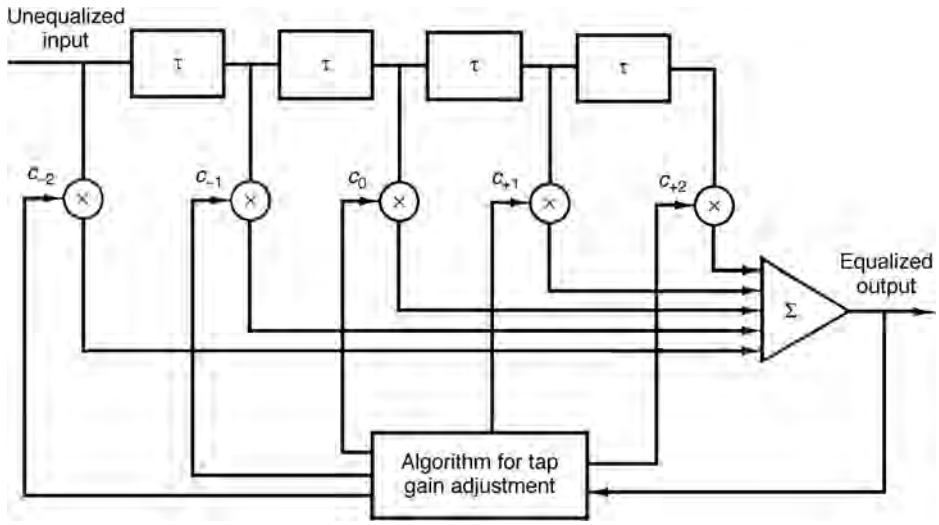


FIGURE 25.10 Linear transversal filter.

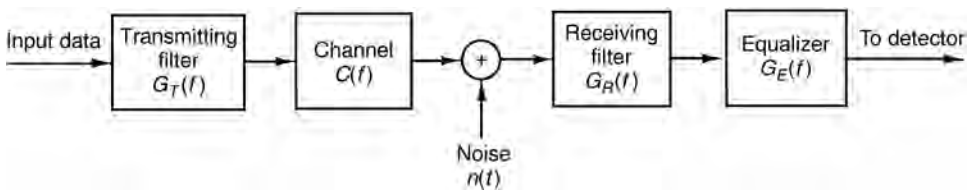


FIGURE 25.11 Block diagram of a system with an equalizer.

Thus, the amplitude response of the equalizer is $|G_E(f)| = 1/|C(f)|$ and its phase response is $\theta_E(f) = -\theta_c(f)$. In this case, the equalizer is said to be the *inverse channel filter* to the channel response.

We note that the inverse channel filter completely eliminates ISI caused by the channel. Since it forces the ISI to be zero at the sampling instants $t = kT$, $k = 0, 1, \ldots$, the equalizer is called a **zero-forcing equalizer**. Hence, the input to the detector is simply

$$z_k = I_k + \eta_k, \quad k = 0, 1, \ldots \tag{25.16}$$

where η_k represents the additive noise and I_k is the desired symbol.

In practice, the ISI caused by channel distortion is usually limited to a finite number of symbols on either side of the desired symbol. Hence, the number of terms that constitute the ISI in the summation given by Eq. (25.10) is finite. As a consequence, in practice, the channel equalizer is implemented as a finite duration impulse response (FIR) filter, or transversal filter, with adjustable tap coefficients $\{c_n\}$, as illustrated in Fig. 25.10. The time delay τ between adjacent taps may be selected as large as T, the symbol interval, in which case the FIR equalizer is called a **symbol-spaced equalizer**. In this case, the input to the equalizer is the sampled sequence given by Eq. (25.7). We note that when the symbol rate $1/T < 2W$, however, frequencies in the received signal above the folding frequency $1/T$ are aliased into frequencies below $1/T$. In this case, the equalizer compensates for the aliased channel-distorted signal.

On the other hand, when the time delay τ between adjacent taps is selected such that $1/\tau \geq 2W > 1/T$, no aliasing occurs, and, hence, the inverse channel equalizer compensates for the true channel distortion. Since $\tau < T$, the channel equalizer is said to have *fractionally spaced taps* and it is called a **fractionally spaced equalizer**. In practice, τ is often selected as $\tau = T/2$. Notice that, in this case, the sampling rate at the output of the filter $G_R(f)$ is $2/T$.

The impulse response of the FIR equalizer is

$$g_E(t) = \sum_{n=-N}^{N} c_n \delta(t - n\tau) \tag{25.17}$$

and the corresponding frequency response is

$$G_E(f) = \sum_{n=-N}^{N} c_n e^{-j2\pi f n \tau} \tag{25.18}$$

where $\{c_n\}$ are the $(2N+1)$ equalizer coefficients and N is chosen sufficiently large so that the equalizer spans the length of the ISI, i.e., $2N+1 \geq L$, where L is the number of signal samples spanned by the ISI. Since $X(f) = G_T(f)C(f)G_R(f)$ and $x(t)$ is the signal pulse corresponding to $X(f)$, then the equalized output signal pulse is

$$q(t) = \sum_{n=-N} c_n x(t - n\tau) \tag{25.19}$$

The zero-forcing condition can now be applied to the samples of $q(t)$ taken at times $t = mT$. These samples are

$$q(mT) = \sum_{n=-N}^{N} c_n x(mT - n\tau), \quad m = 0, \pm 1, \ldots, \pm N \tag{25.20}$$

Since there are $2N+1$ equalizer coefficients, we can control only $2N+1$ sampled values of $q(t)$. Specifically, we may force the conditions

$$q(mT) = \sum_{n=-N}^{N} c_n x(mT - n\tau) = \begin{cases} 1, & m = 0 \\ 0, & m = \pm 1, \pm 2, \ldots, \pm N \end{cases} \tag{25.21}$$

which may be expressed in matrix form as $\boldsymbol{Xc} = \boldsymbol{q}$, where $\boldsymbol{X}$ is a $(2N+1) \times (2N+1)$ matrix with elements $\{x(mT - n\tau)\}$, $\boldsymbol{c}$ is the $(2N+1)$ coefficient vector, and $\boldsymbol{q}$ is the $(2N+1)$ column vector with one nonzero element. Thus, we obtain a set of $2N+1$ linear equations for the coefficients of the zero-forcing equalizer.

We should emphasize that the FIR zero-forcing equalizer does not completely eliminate ISI because it has a finite length. As N is increased, however, the residual ISI can be reduced, and in the limit as $N \to \infty$, the ISI is completely eliminated.

Example 25.1

Consider a channel distorted pulse $x(t)$, at the input to the equalizer, given by the expression

$$x(t) = \frac{1}{1 + \left(\frac{2t}{T}\right)^2}$$

where $1/T$ is the symbol rate. The pulse is sampled at the rate $2/T$ and equalized by a zero-forcing equalizer. Determine the coefficients of a five-tap zero-forcing equalizer.

Solution 25.1

According to Eq. (25.21), the zero-forcing equalizer must satisfy the equations

$$q(mT) = \sum_{n=-2}^{2} c_n x(mT - nT/2) = \begin{cases} 1, & m = 0 \\ 0, & m = \pm 1, \pm 2 \end{cases}$$

The matrix $\boldsymbol{X}$ with elements $x(mT - nT/2)$ is given as

$$\boldsymbol{X} = \begin{bmatrix} \frac{1}{5} & \frac{1}{10} & \frac{1}{17} & \frac{1}{26} & \frac{1}{37} \\ 1 & \frac{1}{2} & \frac{1}{5} & \frac{1}{10} & \frac{1}{17} \\ \frac{1}{5} & \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{5} \\ \frac{1}{17} & \frac{1}{10} & \frac{1}{5} & \frac{1}{2} & 1 \\ \frac{1}{37} & \frac{1}{26} & \frac{1}{17} & \frac{1}{10} & \frac{1}{5} \end{bmatrix} \tag{25.22}$$

The coefficient vector $\boldsymbol{c}$ and the vector $\boldsymbol{q}$ are given as

$$\boldsymbol{c} = \begin{bmatrix} c_{-2} \\ c_{-1} \\ c_0 \\ c_1 \\ c_2 \end{bmatrix} \quad \boldsymbol{q} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \tag{25.23}$$

Then, the linear equations $\boldsymbol{Xc} = \boldsymbol{q}$ can be solved by inverting the matrix $\boldsymbol{X}$. Thus, we obtain

$$\boldsymbol{c}_{\text{opt}} = \boldsymbol{X}^{-1}\boldsymbol{q} = \begin{bmatrix} -2.2 \\ 4.9 \\ -3 \\ 4.9 \\ -2.2 \end{bmatrix} \tag{25.24}$$

One drawback to the zero-forcing equalizer is that it ignores the presence of additive noise. As a consequence, its use may result in significant noise enhancement. This is easily seen by noting that in a frequency range where $C(f)$ is small, the channel equalizer $G_E(f) = 1/C(f)$ compensates by placing a large gain in that frequency range. Consequently, the noise in that frequency range is greatly enhanced. An alternative is to relax the zero ISI condition and select the channel equalizer characteristic such that the combined power in the residual ISI and the additive noise at the output of the equalizer is minimized. A channel equalizer that is optimized based on the minimum mean square error (MMSE) criterion accomplishes the desired goal.

To elaborate, let us consider the noise corrupted output of the FIR equalizer, which is

$$z(t) = \sum_{n=-N}^{N} c_n y(t - n\tau) \tag{25.25}$$

where $y(t)$ is the input to the equalizer, given by Eq. (25.6). The equalizer output is sampled at times $t = mT$. Thus, we obtain

$$z(mT) = \sum_{n=-1}^{N} c_n y(mT - n\tau) \tag{25.26}$$

The desired response at the output of the equalizer at $t = mT$ is the transmitted symbol I_m. The error is defined as the difference between I_m and $z(mT)$. Then, the mean square error (MSE) between the actual output sample $z(mT)$ and the desired values I_m is

$$\begin{aligned} MSE &= E|z(mT) - I_m|^2 \\ &= E\left[\left|\sum_{N=-N}^{N} c_n y(mT - n\tau) - I_m\right|^2\right] \\ &= \sum_{n=-N}^{N}\sum_{k=-N}^{N} c_n c_k R_y(n-k) - 2\sum_{k=-N}^{N} c_k R_{IY}(k) + E(|I_m|^2) \end{aligned} \tag{25.27}$$

where the correlations are defined as

$$\begin{aligned} R_Y(n-k) &= E[y^*(mT - n\tau)y(mT - k\tau)] \\ R_{IY}(k) &= E[y(mT - k\tau)I_m^*] \end{aligned} \tag{25.28}$$

and the expectation is taken with respect to the random information sequence $\{I_m\}$ and the additive noise.

The minimum MSE solution is obtained by differentiating Eq. (25.27) with respect to the equalizer coefficients $\{c_n\}$. Thus, we obtain the necessary conditions for the minimum MSE as

$$\sum_{n=-N}^{N} c_n R_Y(n-k) = R_{IY}(k), \quad k = 0, \pm 1, 2, \ldots, \pm N \tag{25.29}$$

These are the $(2N + 1)$ linear equations for the equalizer coefficients. In contrast to the zero-forcing solution already described, these equations depend on the stastical properties (the autocorrelation) of the noise as well as the ISI through the autocorrelation $R_Y(n)$.

In practice, the autocorrelation matrix $R_Y(n)$ and the crosscorrelation vector $R_{IY}(n)$ are unknown *a priori*. These correlation sequences can be estimated, however, by transmitting a test signal over the channel and using the time-average estimates

$$\begin{aligned} \hat{R}_Y(n) &= \frac{1}{K}\sum_{k=1}^{K} y^*(kT - n\tau)y(kT) \\ \hat{R}_{IY}(n) &= \frac{1}{K}\sum_{k=1}^{K} y(kT - n\tau)I_k^* \end{aligned} \tag{25.30}$$

in place of the ensemble averages to solve for the equalizer coefficients given by Eq. (25.29).

Adaptive Linear Equalizers

We have shown that the tap coefficients of a linear equalizer can be determined by solving a set of linear equations. In the zero-forcing optimization criterion, the linear equations are given by Eq. (25.21). On the other hand, if the optimization criterion is based on minimizing the MSE, the optimum equalizer coefficients are determined by solving the set of linear equations given by Eq. (25.29).

In both cases, we may express the set of linear equations in the general matrix form

$$\boldsymbol{B}\boldsymbol{c} = \boldsymbol{d} \tag{25.31}$$

where $\boldsymbol{B}$ is a $(2N+1)\times(2N+1)$ matrix, $\boldsymbol{c}$ is a column vector representing the $2N+1$ equalizer coefficients, and $\boldsymbol{d}$ is a $(2N+1)$-dimensional column vector. The solution of Eq. (25.31) yields

$$\boldsymbol{c}_{\text{opt}} = \boldsymbol{B}^{-1}\boldsymbol{d} \tag{25.32}$$

In practical implementations of equalizers, the solution of Eq. (25.31) for the optimum coefficient vector is usually obtained by an iterative procedure that avoids the explicit computation of the inverse of the matrix $\boldsymbol{B}$. The simplest iterative procedure is the method of steepest descent, in which one begins by choosing arbitrarily the coefficient vector $\boldsymbol{c}$, say $\boldsymbol{c}_0$. This initial choice of coefficients corresponds to a point on the criterion function that is being optimized. For example, in the case of the *MSE* criterion, the initial guess $\boldsymbol{c}_0$ corresponds to a point on the quadratic *MSE* surface in the $(2N+1)$-dimensional space of coefficients. The gradient vector, defined as $\boldsymbol{g}_0$, which is the derivative of the MSE with respect to the $2N+1$ filter coefficients, is then computed at this point on the criterion surface, and each tap coefficient is changed in the direction opposite to its corresponding gradient component. The change in the jth tap coefficient is proportional to the size of the jth gradient component.

For example, the gradient vector denoted as $\boldsymbol{g}_k$, for the MSE criterion, found by taking the derivatives of the MSE with respect to each of the $2N+1$ coefficients, is

$$\boldsymbol{g}_k = \boldsymbol{B}\boldsymbol{c}_k - \boldsymbol{d}, \quad k = 0, 1, 2, \ldots \tag{25.33}$$

Then the coefficient vector $\boldsymbol{c}_k$ is updated according to the relation

$$\boldsymbol{c}_{k+1} = \boldsymbol{c}_k - \Delta\boldsymbol{g}_k \tag{25.34}$$

where Δ is the *step-size parameter* for the iterative procedure. To ensure convergence of the iterative procedure, Δ is chosen to be a small positive number. In such a case, the gradient vector $\boldsymbol{g}_k$ converges toward zero, i.e., $\boldsymbol{g}_k \to \boldsymbol{0}$ as $k \to \infty$, and the coefficient vector $\boldsymbol{c}_k \to \boldsymbol{c}_{\text{opt}}$ as illustrated in Fig. 25.12 based on two-dimensional optimization. In general, convergence of the equalizer tap coefficients to $\boldsymbol{c}_{\text{opt}}$ cannot

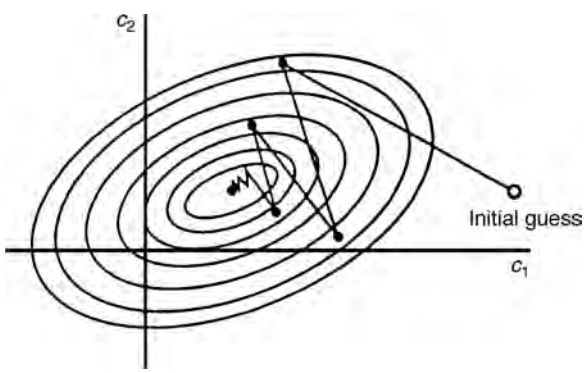


FIGURE 25.12 Example of convergence characteristics of a gradient algorithm.

be attained in a finite number of iterations with the steepest-descent method. The optimum solution $\boldsymbol{c}_{\text{opt}}$, however, can be approached as closely as desired in a few hundred iterations. In digital communication systems that employ channel equalizers, each iteration corresponds to a time interval for sending one symbol and, hence, a few hundred iterations to achieve convergence to $\boldsymbol{c}_{\text{opt}}$ corresponds to a fraction of a second.

Adaptive channel equalization is required for channels whose characteristics change with time. In such a case, the ISI varies with time. The channel equalizer must track such time variations in the channel response and adapt its coefficients to reduce the ISI. In the context of the preceding discussion, the optimum coefficient vector $\boldsymbol{c}_{\text{opt}}$ varies with time due to time variations in the matrix $\boldsymbol{B}$ and, for the case of the *MSE* criterion, time variations in the vector $\boldsymbol{d}$. Under these conditions, the iterative method described can be modified to use estimates of the gradient components. Thus, the algorithm for adjusting the equalizer tap coefficients may be expressed as

$$\hat{\boldsymbol{c}}_{k+1} = \hat{\boldsymbol{c}}_k - \Delta\hat{\boldsymbol{g}}_k \tag{25.35}$$

where $\hat{\boldsymbol{g}}_k$ denotes an estimate of the gradient vector $\boldsymbol{g}_k$ and $\hat{\boldsymbol{c}}_k$ denotes the estimate of the tap coefficient vector.

In the case of the MSE criterion, the gradient vector $\boldsymbol{g}_k$ given by Eq. (25.33) may also be expressed as

$$\boldsymbol{g}_k = -E(e_k \boldsymbol{y}_k^*)$$

An estimate $\hat{\boldsymbol{g}}_k$ of the gradient vector at the kth iteration is computed as

$$\hat{\boldsymbol{g}}_k = -e_k \boldsymbol{y}_k^* \tag{25.36}$$

where e_k denotes the difference between the desired output from the equalizer at the kth time instant and the actual output $z(kT)$, and $\boldsymbol{y}_k$ denotes the column vector of $2N+1$ received signal values contained in the equalizer at time instant k. The *error signal* e_k is expressed as

$$e_k = I_k - z_k \tag{25.37}$$

where $z_k = z(kT)$ is the equalizer output given by Eq. (25.26) and I_k is the desired symbol. Hence, by substituting Eq. (25.36) into Eq. (25.35), we obtain the adaptive algorithm for optimizing the tap coefficients (based on the *MSE* criterion) as

$$\hat{\boldsymbol{c}}_{k+1} = \hat{\boldsymbol{c}}_k + \Delta e_k \boldsymbol{y}_k^* \tag{25.38}$$

Since an estimate of the gradient vector is used in Eq. (25.38), the algorithm is called a stochastic **gradient algorithm**; it is also known as the **LMS algorithm**.

A block diagram of an adaptive equalizer that adapts its tap coefficients according to Eq. (25.38) is illustrated in Fig. 25.13. Note that the difference between the desired output I_k and the actual output z_k from the equalizer is used to form the error signal e_k. This error is scaled by the step-size parameter Δ, and the scaled error signal Δe_k multiplies the received signal values $\{y(kT - n\tau)\}$ at the $2N+1$ taps. The products $\Delta e_k y^*(kT - n\tau)$ at the $(2N+1)$ taps are then added to the previous values of the tap coefficients to obtain the updated tap coefficients, according to Eq. (25.38). This computation is repeated as each new symbol is received. Thus, the equalizer coefficients are updated at the symbol rate.

Initially, the adaptive equalizer is trained by the transmission of a known pseudorandom sequence $\{I_m\}$ over the channel. At the demodulator, the equalizer employs the known sequence to adjust its coefficients. Upon initial adjustment, the adaptive equalizer switches from a **training mode** to a **decision-directed mode**,

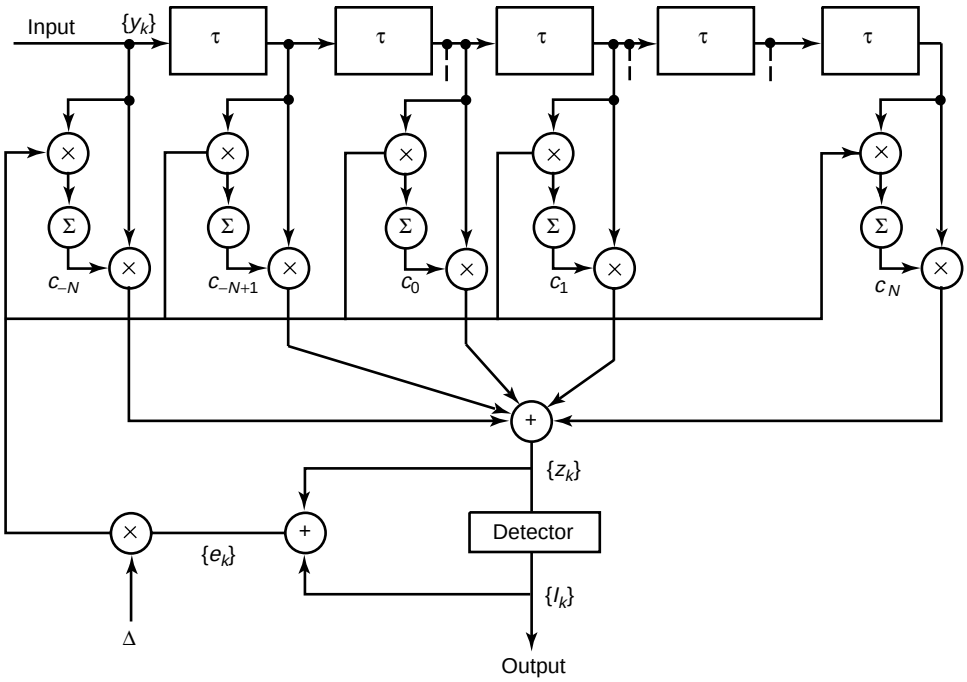


FIGURE 25.13 Linear adaptive equalizer based on the MSE criterion.

in which case the decisions at the output of the detector are sufficiently reliable so that the error signal is formed by computing the difference between the detector output and the equalizer output, i.e.,

$$e_k = \tilde{I}_k - z_k \tag{25.39}$$

where $\tilde{I}_k$ is the output of the detector. In general, decision errors at the output of the detector occur infrequently and, consequently, such errors have little effect on the performance of the tracking algorithm given by Eq. (25.38).

A rule of thumb for selecting the step-size parameter so as to ensure convergence and good tracking capabilities in slowly varying channels is

$$\Delta = \frac{1}{5(2N+1)P_R} \tag{25.40}$$

where P_R denotes the received signal-plus-noise power, which can be estimated from the received signal (see Proakis, 2001).

The convergence characteristic of the stochastic gradient algorithm in Eq. (25.38) is illustrated in Fig. 25.14. These graphs were obtained from a computer simulation of an 11-tap adaptive equalizer operating on a channel with a rather modest amount of ISI. The input signal-plus-noise power P_R was normalized to unity. The rule of thumb given in Eq. (25.40) for selecting the step size gives $\Delta = 0.018$. The effect of making Δ too large is illustrated by the large jumps in MSE as shown for $\Delta = 0.115$. As Δ is decreased, the convergence is slowed somewhat, but a lower MSE is achieved, indicating that the estimated coefficients are closer to $\boldsymbol{c}_{\text{opt}}$.

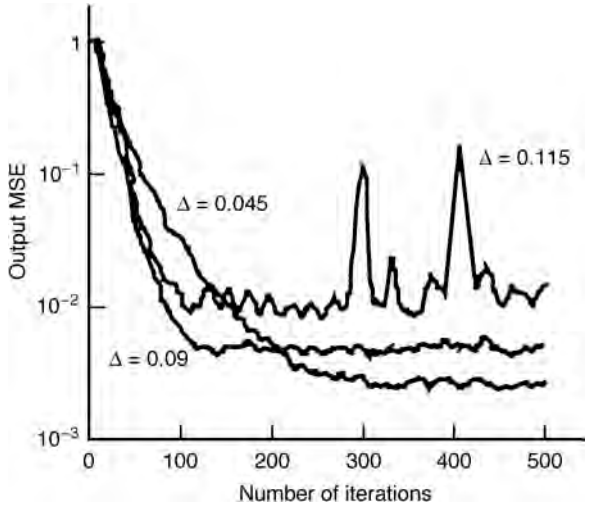


FIGURE 25.14 Initial convergence characteristics of the LMS algorithm with different step sizes.

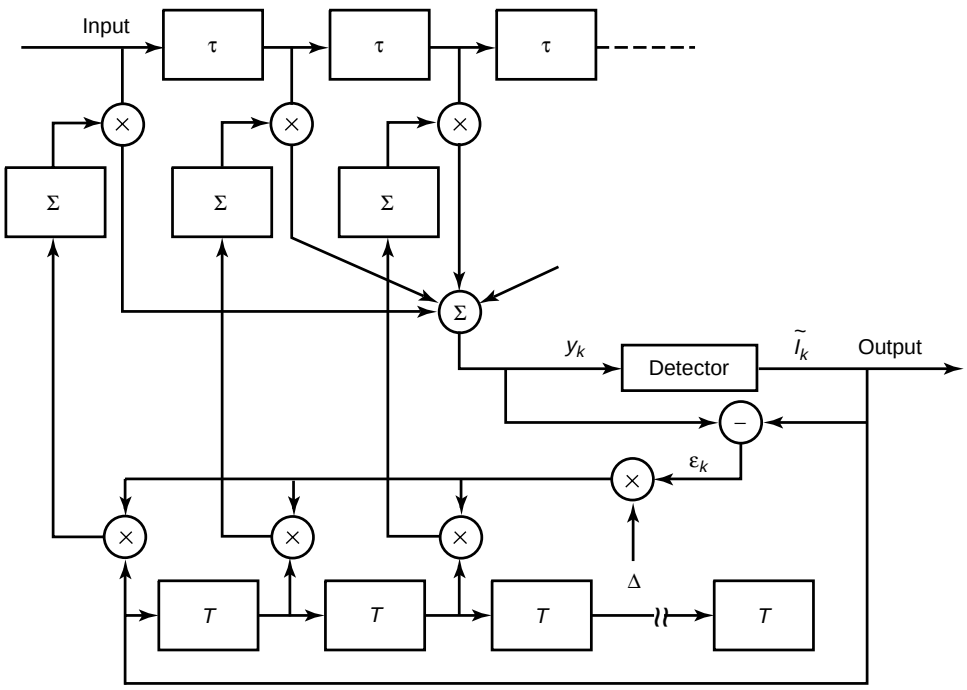


FIGURE 25.15 An adaptive zero-forcing equalizer.

Although we have described in some detail the operation of an adaptive equalizer that is optimized on the basis of the MSE criterion, the operation of an adaptive equalizer based on the zero-forcing method is very similar. The major difference lies in the method for estimating the gradient vectors g_k at each iteration. A block diagram of an adaptive zero-forcing equalizer is shown in Fig. 25.15. For more details on the tap coefficient update method for a zero-forcing equalizer, the reader is referred to the papers by Lucky [2,3] and the text by Proakis [2001].

25.4 Decision-Feedback Equalizer

The linear filter equalizers described in the preceding section are very effective on channels, such as wire line telephone channels, where the ISI is not severe. The severity of the ISI is directly related to the spectral characteristics and not necessarily to the time span of the ISI. For example, consider the ISI resulting from the two channels that are illustrated in Fig. 25.16. The time span for the ISI in channel A is 5 symbol intervals on each side of the desired signal component, which has a value of 0.72. On the other hand, the time span for the ISI in channel B is one symbol interval on each side of the desired signal component, which has a value of 0.815. The energy of the total response is normalized to unity for both channels.

In spite of the shorter ISI span, channel B results in more severe ISI. This is evidenced in the frequency response characteristics of these channels, which are shown in Fig. 25.17. We observe that channel B has a spectral null (the frequency response $C(f) = 0$ for some frequencies in the band $|f| \leq W$) at $f = 1/2T$, whereas this does not occur in the case of channel A. Consequently, a linear equalizer will introduce a large gain in its frequency response to compensate for the channel null. Thus, the noise in channel B will be enhanced much more than in channel A. This implies that the performance of the linear equalizer for channel B will be significantly poorer than that for channel A. This fact is borne out by the computer simulation results for the performance of the two linear equalizers shown in Fig. 25.18. Hence, the basic limitation of a linear equalizer is that it performs poorly on channels having spectral nulls. Such channels are often encountered in radio communications, such as ionospheric transmission at frequencies below 30 MHz, and mobile radio channels, such as those used for cellular radio communications.

A **decision-feedback equalizer (DFE)** is a nonlinear equalizer that employs previous decisions to eliminate the ISI caused by previously detected symbols on the current symbol to be detected. A simple block diagram for a DFE is shown in Fig. 25.19. The DFE consists of two filters. The first filter is called a *feedforward filter* and it is generally a fractionally spaced FIR filter with adjustable tap coefficients. This filter is identical in form to the linear equalizer already described. Its input is the received filtered signal

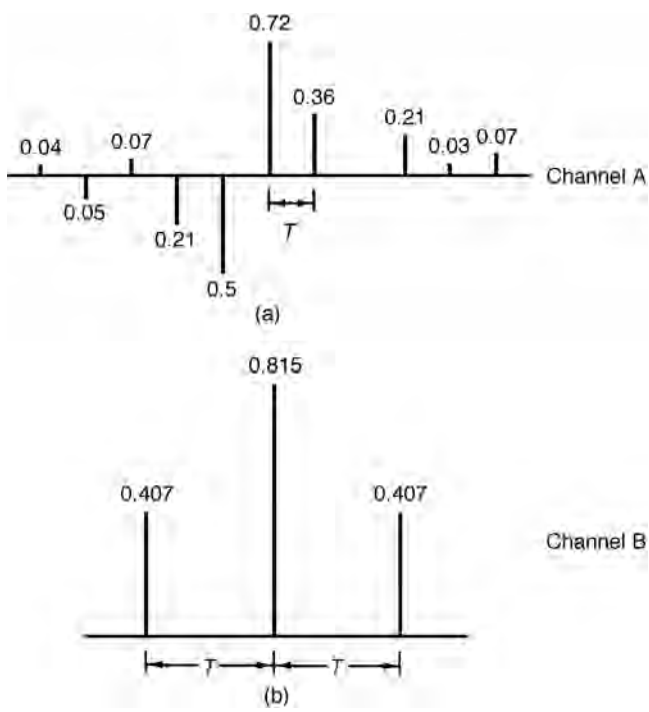


FIGURE 25.16 Two channels with ISI.

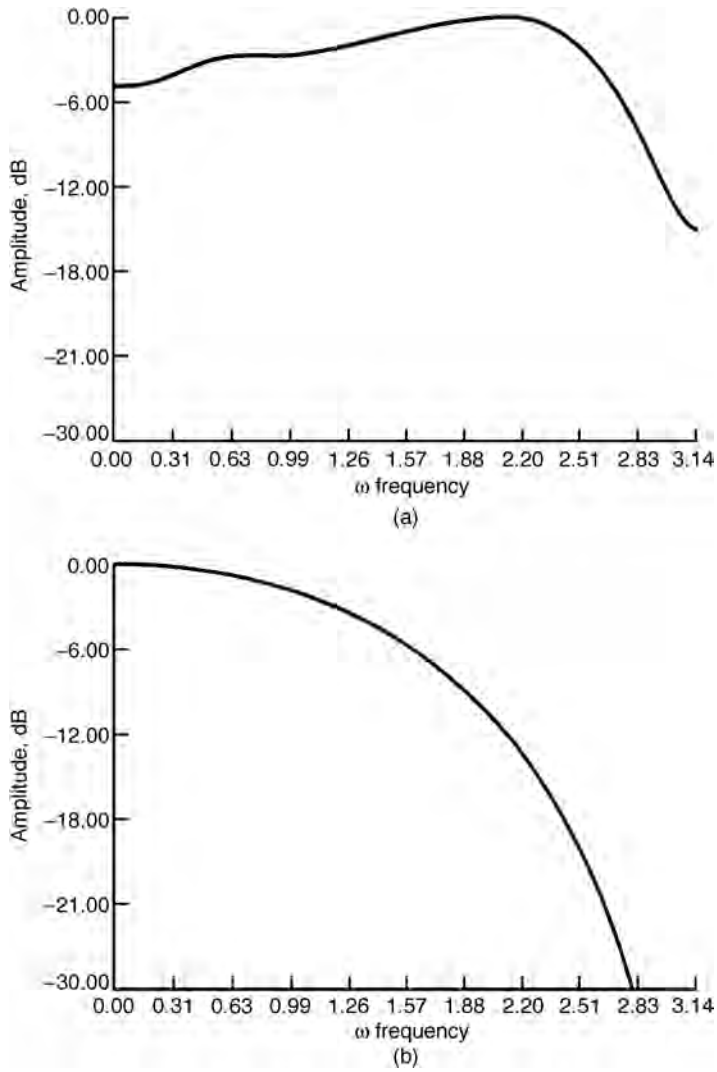


FIGURE 25.17 Amplitude spectra for (a) channel A shown in Fig. 25.16(a) and (b) channel B shown in Fig. 25.16(b).

$y(t)$ sampled at some rate that is a multiple of the symbol rate, e.g., at rate $2/T$. The second filter is a *feedback filter*. It is implemented as an FIR filter with symbol-spaced taps having adjustable coefficients. Its input is the set of previously detected symbols. The output of the feedback filter is subtracted from the output of the feedforward filter to form the input to the detector. Thus, we have

$$z_m = \sum_{n=-N_1}^{0} c_n y(mT - n\tau) - \sum_{n=1}^{N_2} b_n \bar{I}_{m-n} \tag{25.41}$$

where $\{c_n\}$ and $\{b_n\}$ are the adjustable coefficients of the feedforward and feedback filters, respectively, $\bar{I}_{m-n}$, $n = 1, 2, \ldots, N_2$ are the previously detected symbols, $N_1 + 1$ is the length of the feedforward filter, and N_2 is the length of the feedback filter. Based on the input z_m, the detector determines which of the possible transmitted symbols is closest in distance to the input signal I_m. Thus, it makes its decision and outputs $\tilde{I}_m$. What makes the **DFE** nonlinear is the nonlinear characteristic of the detector that provides the input to the feedback filter.

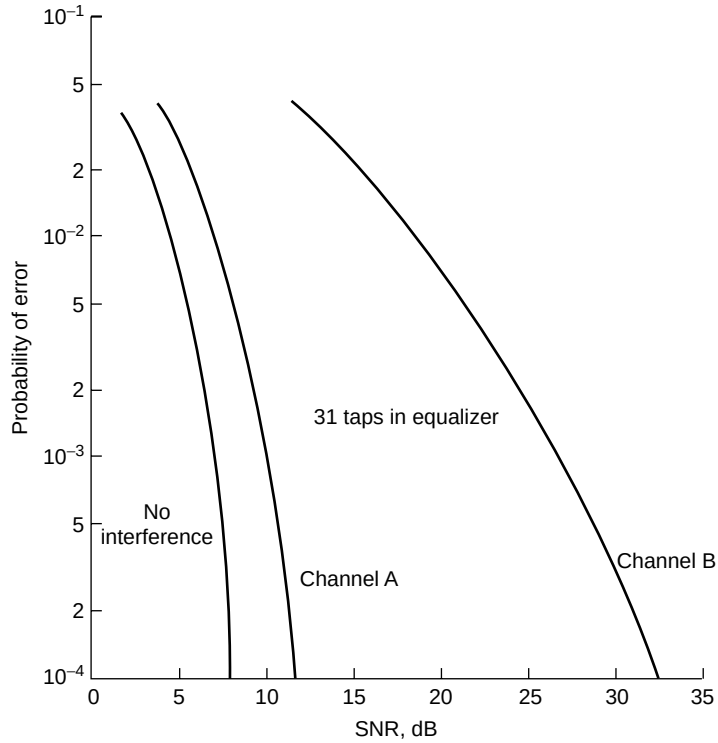


FIGURE 25.18 Error-rate performance of linear MSE equalizer.

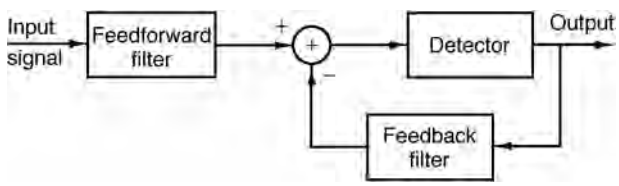


FIGURE 25.19 Block diagram of DFE.

The tap coefficients of the feedforward and feedback filters are selected to optimize some desired performance measure. For mathematical simplicity, the MSE criterion is usually applied, and a stochastic gradient algorithm is commonly used to implement an adaptive DFE. Figure 25.20 illustrates the block diagram of an adaptive DFE whose tap coefficients are adjusted by means of the LMS stochastic gradient algorithm. Figure 25.21 illustrates the probability of error performance of the DFE, obtained by computer simulation, for binary PAM transmission over channel B. The gain in performance relative to that of a linear equalizer is clearly evident.

We note that the decision errors from the detector that are fed back to the feedback filter result in a loss in the performance of the DFE, as illustrated in Fig. 25.21. This loss can be avoided by placing the feedback filter of the DFE at the transmitter and the feedforward filter at the receiver. Thus, the problem of error propagation due to incorrect decisions in the feedback filter is completely eliminated.

This approach is especially suitable for wireline channels, where the channel characteristics do not vary significantly with time. The linear fractionally-spaced feedforward filter of the DFE, placed at the receiver, compensates for the ISI that results from any small time variations in the channel response. The synthesis

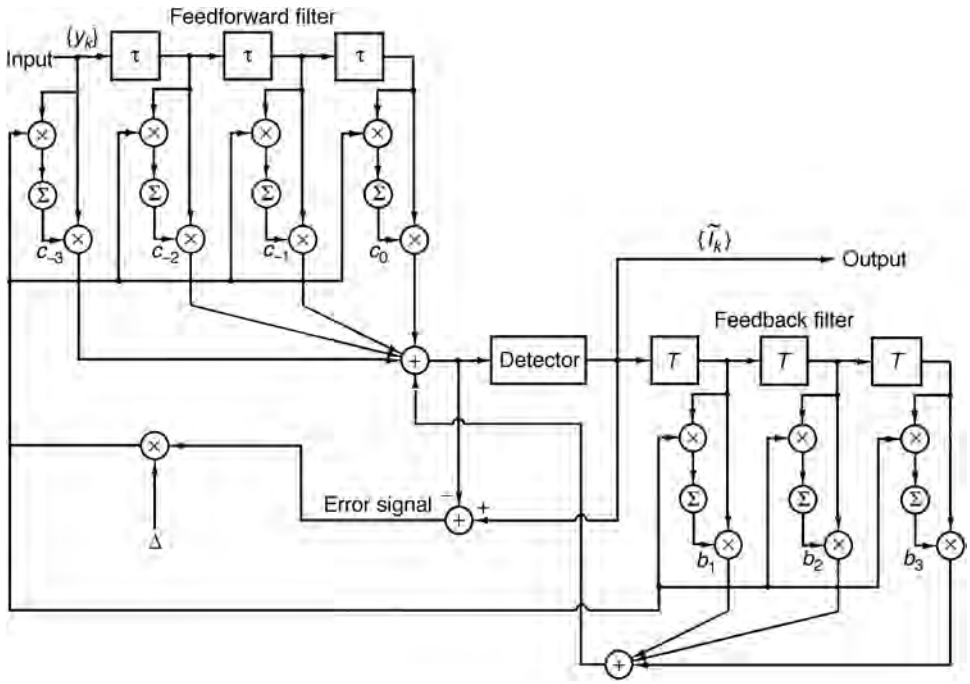


FIGURE 25.20 Adaptive DFE.

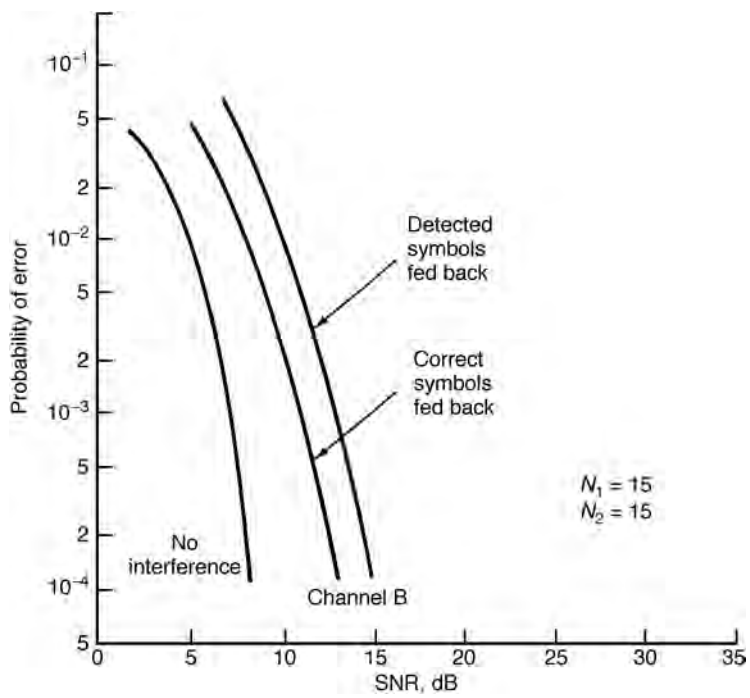


FIGURE 25.21 Performance of DFE with and without error propagation.

of the feedback filter of the DFE at the transmitter side is usually performed after the response of the channel is measured at the receiver by the transmission of a channel probe signal and the receiver sends to the transmitter the coefficients of the feedback filter.

One problem with this approach to implementing the DFE is that the signal points at the transmitter, after subtracting the tail (postcursors) of the ISI, generally have a larger dynamic range than the original signal constellation and, consequently, require a larger transmitter power. This problem is solved by precoding the information symbols prior to transmission as described by Tomlinson [5] and Harashima and Miyakawa [6].

25.5 Maximum-Likelihood Sequence Detection

Although the DFE outperforms a linear equalizer, it is not the optimum equalizer from the viewpoint of minimizing the probability of error in the detection of the information sequence $\{I_k\}$ from the received signal samples $\{y_k\}$ given in Eq. (25.5). In a digital communication system that transmits information over a channel that causes ISI, the optimum detector is a maximum-likelihood symbol sequence detector which produces at its output the most probable symbol sequence $\{\tilde{I}_k\}$ for the given received sampled sequence $\{y_k\}$. That is, the detector finds the sequence $\{\tilde{I}_k\}$ that maximizes the *likelihood function*.

$$\Lambda(\{I_k\}) = \ln p(\{y_k\}|\{I_k\}) \tag{25.42}$$

where $p(\{y_k\}|\{I_k\})$ is the joint probability of the received sequence $\{y_k\}$ conditioned on $\{I_k\}$. The sequence of symbols $\{\tilde{I}_k\}$ that maximizes this joint conditional probability is called the **maximum-likelihood sequence detector**.

An algorithm that implements maximum-likelihood sequence detection (MLSD) is the Viterbi algorithm, which was originally devised for decoding convolutional codes. For a description of this algorithm in the context of sequence detection in the presence of ISI, the reader is referred to the paper by Forney [1] and the text by Proakis, 2001.

The major drawback of MLSD for channels with ISI is the exponential behavior in computational complexity as a function of the span of the ISI. Consequently, MLSD is practical only for channels where the ISI spans only a few symbols and the ISI is severe in the sense that it causes a severe degradation in the performance of a linear equalizer or a decision-feedback equalizer. For example, Fig. 25.22 illustrates the error probability performance of the Viterbi algorithm for a binary PAM signal transmitted through channel B (see Fig. 25.16). For purposes of comparison, we also illustrate the probability of error for a DFE. Both results were obtained by computer simulation. We observe that the performance of the maximum likelihood sequence detector is about 4.5 dB better than that of the DFE at an error probability of 10^{-4}. Hence, this is one example where the ML sequence detector provides a significant performance gain on a channel with a relatively short ISI span.

The MLSD implemented efficiently by use of the Viterbi algorithm is widely used in mobile cellular communications systems, such as the GSM system, where the span of the ISI is limited to five or six symbols.

The performance advantage of the MLSD, compared with the linear equalizer and the DFE, in channels with severe ISI, has motivated a significant amount of research on methods for reducing the complexity of MLSD while retaining its superior performance characteristics. One approach to the design of reduced complexity MLSD is focused on methods that reduce the length of the ISI span by pre-processing the received signal prior to the maximum-likelihood detector. Falconer and Magee [7] and Beare [8] used a linear equalizer (LE) to reduce the span of the ISI to some small specified length prior to the Viterbi detector. Lee and Hill [9] employed a DFE in place of the LE. Thus, the large ISI span in the channel is reduced to a sufficiently small length, called the *desired impulse response*, so that the complexity of the Viterbi detector following the LE or DFE is manageable. The choice of the desired impulse response is tailored to the ISI characteristics of the channel. This approach to reducing the complexity of the Viterbi detector has proved very effective in high-density magnetic recording systems, as illustrated in the papers by Siegel and Wolf [10], Tyner and Proakis [11], Moon and Carley [12] and Proakis [13].

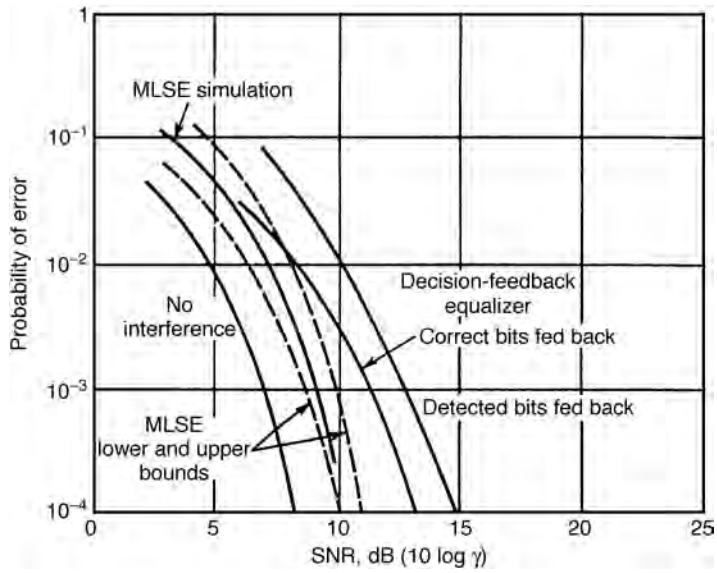


FIGURE 25.22 Comparison of performance between MLSE and decision-feedback equalization for channel B of Fig. 25.16.

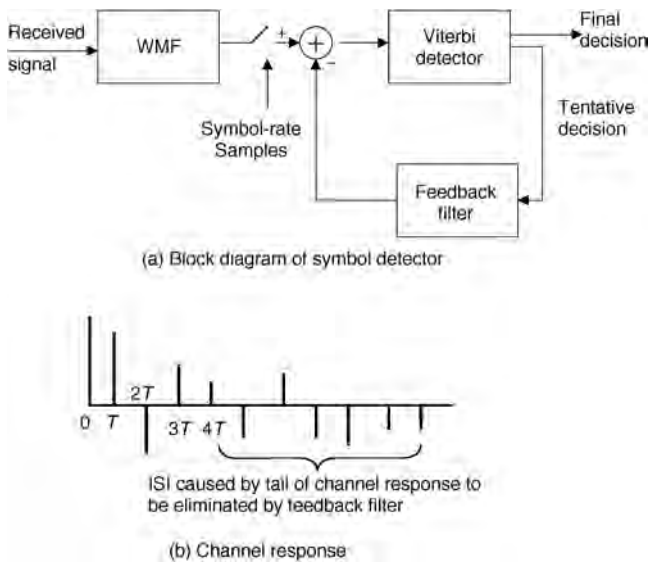


FIGURE 25.23 Reduced complexity MLSD using feedback from the Viterbi detector.

From a performance viewpoint, a more effective approach for reducing the computational complexity of the MLSD is to employ decision feedback within the Viterbi detector to reduce the effective length of the ISI. A prefilter, called a *whitened matched filter* (*WMF*) precedes the Viterbi detector and reduces the channel to one that has a minimum phase characteristic, as shown in Fig. 25.23. Preliminary decisions from the Viterbi detector are fed back through a filter, as shown in Fig. 25.23, to synthesize the tail in the ISI caused by the channel response and, thus, to cancel the ISI at the input to the Viterbi detector. The preliminary decisions from the Viterbi detector can be obtained by using the most probable surviving sequence in the Viterbi detector. This approach to tail cancellation had been called *global feedback* by

Bergmans et al. [14]. Alternatively, Bergmans et al. [14] propose that one can use preliminary decisions corresponding to each surviving sequence to cancel the ISI in the tail of the corresponding surviving sequence. Thus, the ISI can be perfectly cancelled when the correct sequence is among the surviving sequences in the Viterbi detector, even if it is not the most probable sequence at any instant in time in the detection process. Bergmans et al. [14] named this approach to tail cancellations as using *local feedback*. Simulation results given in the paper by Bergmans et al. [14] indicate that local feedback gives superior performance compared with global feedback.

25.6 Maximum *A Posteriori* Probability Detector and Turbo Equalization

For a channel with ISI, the MLSD minimizes the probability of error in the detection of a sequence of transmitted information symbols. Instead of using the MLSD to recover the information symbols, we may use a detector that makes optimum decisions on a symbol-by-symbol basis based on the computation of the maximum *a posteriori* probability (MAP) for each symbol. Consequently, the MAP detector is optimum in the sense that it minimizes the probability of a symbol error. Basically, the MAP criterion for the information symbol I_k at the kth time interval involves the computation of the *a posteriori* probabilities $P(I_k = S_i/y_{k+D}, y_{k+D-1}, —, y_1)$ where $\{y_k\}$ is the observed received sequence, D is a delay parameter that is chosen to be equal to or exceed the span of the ISI, and $\{S_i, 1 \le i \le M\}$ is the set of M possible points in the signal constellation. A decision is made of the symbol corresponding to the largest *a posteriori* probability.

A computationally efficient iterative detection algorithm based on the MAP criterion is described in a paper by Bahl et al. [15] and is usually called the BCJR algorithm. In general, the computational complexity of the BCJR algorithm is greater than that of the Viterbi algorithm, so the latter is preferable in implementing an optimum detector for a channel with ISI.

An equalizer based on the MAP criterion is particularly suitable in a receiver that couples the equalizer to a decoder that performs iterative decoding. To elaborate, suppose that the transmitter of a digital communication system employs a binary systematic convolutional encoder followed by a block interleaver and modulator. The channel is a linear time-dispersive channel that introduces ISI having a finite span. In such a case, we view the channel with ISI as an inner encoder in a serially (cascade) concatenated code, as shown in Fig. 25.24. By treating the system as a serially concatenated coded system, we can apply recently developed iterative decoding techniques which provide a significant improvement in performance, compared to a system that performs the equalization and decoding operations separately.

The basic configuration of the iterative equalizer and decoder is shown in Fig. 25.25. The input to the MAP equalizer is the sequence of received signal samples $\{y_k\}$ from the receiver filter, which may be

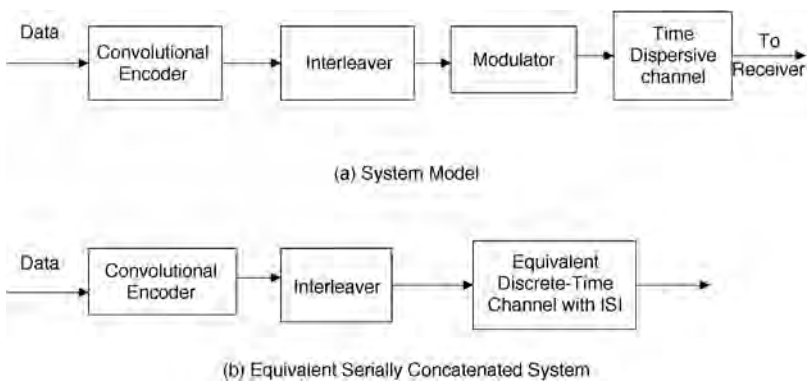


FIGURE 25.24 Channel with ISI viewed as a serially concatenated coded system.

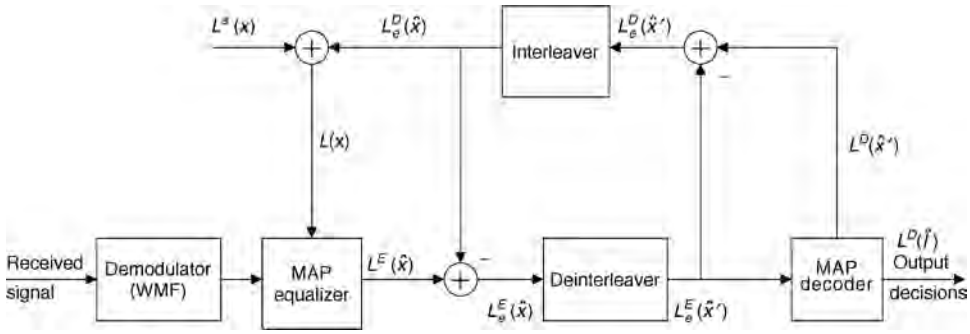


FIGURE 25.25 Iterative MAP equalization and decoding.

implemented as a WMF. The MAP equalizer employs the BCJR algorithm to compute the *a posteriori* probabilities for each of the possible M signal points and feeds to the outer decoder, after deinterleaving, the so-called *extrinsic information*, denoted by $L_e^E(\hat{x}')$. The outer decoder uses this information in computing the corresponding probabilities for each of the coded bits based on the MAP criterion. Thus, the outer decoder for the convolutional code also employs the BCJR algorithm to compute the *a posteriori* probabilities for the information bits. Then, the MAP decoder feeds back to the MAP equalizer, after interleaving, the so-called extrinsic information, denoted as $L_e^D(\hat{x})$. This information is used by the MAP equalizer in making another pass through the data and generally achieves better estimates of the coded bits. This iterative process generally continues for several iterations (usually four to eight) until little additional improvement is achieved with additional iterations.

By using a MAP equalizer in conjunction with an iterative MAP decoder, it is possible to completely eliminate the performance loss due to the ISI. The paper by Bauch et al. [16] provides a description of the computations performed by the MAP equalizer and MAP decoder and illustrates the performance gain that is achieved by joint MAP equalization and iterative decoding. The implementation of a MAP equalizer used in conjunction with an iterative MAP decoder has been called *turbo equalization*. While it yields superior performance compared to the other equalization techniques described above, turbo equalization has the disadvantage of a significantly higher computational complexity. However, we envision that this equalization technique will eventually be implemented in future communication systems that transmit information through time dispersive channels.

25.7 Conclusions

Channel equalizers are widely used in digital communication systems to mitigate the effects of ISI caused by channel distortion. Linear equalizers are widely used for high-speed modems that transmit data over telephone channels. For wireless (radio) transmission, such as in mobile cellular communications and interoffice communications, the multipath propagation of the transmitted signal results in severe ISI. Such channels require more powerful equalizers to combat the severe ISI. The decision-feedback equalizer and the MLSD are two nonlinear channel equalizers that are suitable for radio channels with severe ISI.

Turbo equalization is a newly developed method that couples a MAP equalizer to a MAP decoder to achieve superior performance compared with the conventional equalization methods such as the linear equalizer, the DFE, or MLSD, which operate independently of the decoder.

Defining Terms

Adaptive equalizer: A channel equalizer whose parameters are updated automatically and adaptively during transmission of data.

Channel equalizer: A device that is used to reduce the effects of channel distortion in a received signal.

Decision-directed mode: Mode for adjustment of the equalizer coefficient adaptively based on the use of the detected symbols at the output of the detector.

Decision-feedback equalizer (DFE): An adaptive equalizer that consists of a feedforward filter and a feedback filter, where the latter is fed with previously detected symbols that are used to eliminate the intersymbol interference due to the tail in the channel impulse response.

Fractionally spaced equalizer: A tapped-delay line channel equalizer in which the delay between adjacent taps is less than the duration of a transmitted symbol.

Intersymbol interference: Interference in a received symbol from adjacent (nearby) transmitted symbols caused by channel distortion in data transmission.

LMS algorithm: See stochastic gradient algorithm.

Maximum-likelihood sequence detector: A detector for estimating the most probable sequence of data symbols by maximizing the likelihood function of the received signal.

Preset equalizer: A channel equalizer whose parameters are fixed (time-invariant) during transmission of data.

Stochastic gradient algorithm: An algorithm for adaptively adjusting the coefficients of an equalizer based on the use of (noise-corrupted) estimates of the gradients.

Symbol-spaced equalizer: A tapped-delay line channel equalizer in which the delay between adjacent taps is equal to the duration of a transmitted symbol.

Training mode: Mode for adjustment of the equalizer coefficients based on the transmission of a known sequence of transmitted symbols.

Zero-forcing equalizer: A channel equalizer whose parameters are adjusted to completely eliminate intersymbol interference in a sequence of transmitted data symbols.

References

1. Forney, G.D., Jr. 1972. Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference. *IEEE Trans. Inform. Theory,* IT 18(May):363–378.
2. Lucky, R.W. 1965. Automatic equalization for digital communications. *Bell Syst. Tech. J.,* 44(April): 547–588.
3. Lucky, R.W. 1966. Techniques for adaptive equalization of digital communication. *Bell Syst. Tech. J.,* 45(Feb.):255–286.
4. Proakis, J.G. *Digital Communications,* 4th ed. McGraw-Hill, New York, 2001.
5. Tomlinson, M. 1971. A new automatic equalizer employing modulo arithmetic. *Electr. Lett.,* 7: 138–139.
6. Harashima, H. and Miyakawa, H. 1972. Matched-transmission technique for channels with intersymbol interference. *IEEE Trans. Commun.,* COM-20:774–780.
7. Falconer, D.D. and Magee, F.R. 1973. Adaptive channel memory truncation for maximum likelihood sequence estimation. *Bell Syst. Tech. J.,* 52:1541–1562.
8. Beare, C.T. 1978. The choice of the desired impulse response in combined linear-Viterbi algorithm equalizers. *IEEE Trans. Commun.,* 26:1301–1307.
9. Lee, W.U. and Hill, F.S. 1977. A maximum-likelihood sequence estimator with decision-feedback equalizer. *IEEE Trans. Commun.,* 25:971–979.
10. Siegel, P.H. and Wolf, J.K. 1991. Modulation and coding for information storage. *IEEE Commun. Mag.,* 30:68–86.
11. Tyner, D.J. and Proakis, J.G. 1993. Partial response equalizer performance in digital magnetic recording channels. *IEEE Trans. Magnetics,* 29:4194–4208.
12. Moon, J. and Carley, L.R. 1988. Partial response signaling in a magnetic recording channel. *IEEE Trans. Magnetics,* 24:2973–2975.
13. Proakis, J.G. 1998. Equalization techniques for high-density magnetic recording. *IEEE Signal Processing Mag.,* 15:73–82.

14. Bergmans, J.W.M., Rajput, S.A., and Van DeLaar, F.A.M. 1987. On the use of decision feedback for simplifying the Viterbi detector. *Philips J. Research,* 42(4):399–428.
15. Bahl, L.R., Cocke, J., Jelinek, F., and Raviv, J. 1974. Optimal decoding of linear codes for minimizing symbol error rate. *IEEE Trans. Inform. Theory,* IT-20:284–287.
16. Bauch, G., Khorram, H., and Hagenauer, J. 1997. Iterative equalization and decoding in mobile communications systems. *Proc. European Personal Mobile Commun. Conf. (EPMCC'97),* 307–312.

Further Information

For a comprehensive treatment of adaptive equalization techniques and their performance characteristics, the reader may refer to the book by Proakis, 2001. The two papers by Lucky [2,3] provide a treatment on linear equalizers based on the zero-forcing criterion. Additional information on decision-feedback equalizers may be found in the journal papers "An Adaptive Decision-Feedback Equalizer" by D.A. George, R.R. Bowen, and J.R. Storey, *IEEE Transactions on Communications Technology,* Vol. COM-19, pp. 281–293, June 1971, and "Feedback Equalization for Fading Dispersive Channels" by P. Monsen, *IEEE Transactions on Information Theory,* Vol. IT-17, pp. 56–64, January 1971. A thorough treatment of channel equalization based on maximum-likelihood sequence detection is given in the paper by Forney [1].

26

Pulse-Code Modulation Codec-Filters

Michael D. Floyd
Motorola Semiconductor Products

Garth D. Hillman
Motorola Semiconductor Products

26.1 Introduction and General Description of a Pulse-Code Modulation (PCM) Codec-Filter

This chapter introduces the reader to the pulse-code modulation (PCM) codec-filter function and where it is used in the telephone network. A PCM codec-filter was originally used in the switching offices and network for digitizing and reconstructing the human voice. Over the last five years, linear PCM codec-filters have been used in high-speed modems at the subscriber's premise.

The name codec is an acronym from coder for the analog-to-digital converter (ADC) used to digitize voice, and decoder for the digital-to-analog converter (DAC) used for reconstructing voice. A codec is a single device that does both the ADC and DAC conversions. A PCM codec-filter includes the bandlimiting filter for the ADC and the reconstruction smoothing filter for the output of the DAC, in addition to the ADC and DAC functions. PCM codec-filter is often referred to as a PCM codec, Combo™ (from National Semiconductor), Monocircuit, or Cofidec.

PCM codec-filters were developed to transmit telephone conversations over long distances with improved performance and at lower cost. Digitizing the voice channel is relatively economical compared to expensive low-noise analog transmission equipment.

Multiple digitized voice channels can be multiplexed into one higher data rate digital channel without concern for interference or crosstalk of the analog voice information. Digitized voice data can also be received and retransmitted without attenuation and noise degradation. This digitized data could be transmitted via T1, microwave, satellite, fiber optics, RF carrier, etc. without loss to the digitized voice channel.

26.2 Where PCM Codec-Filters are Used in the Telephone Network

With the advancements in microeletronics, the PCM codec-filter function has evolved from the early 1960s' technology of passive resistor-capacitor-inductor filters with discrete analog-to-digital converter and discrete digital-to-analog converter implementations to the fully integrated devices that were introduced in the late 1970s and early 1980s. As monolithic devices, the PCM codec-filter has experienced performance improvements and cost reductions by taking advantage of the improvements in integrated circuit technology developed within the semiconductor industry. Through this evolutionary progression, the cost reductions of the PCM codec-filter function has increased their applicability to include switching systems for telephone central offices (CO), private branch exchanges (PBX), and key systems. The transmission applications have increased also to include digital loop carriers, pair gain multiplexers, telephone loop extenders, integrated services digital network (ISDN) terminals, digital cellular telephones, and digital cordless telephones. New applications have developed, including voice recognition equipment, voice storage, voice mail, and digital tapeless answering machines.

Figure 26.1 is a simplified diagram showing some of the services that the local telephone service provider has to offer. The telephone network must be able to operate in extreme weather conditions and

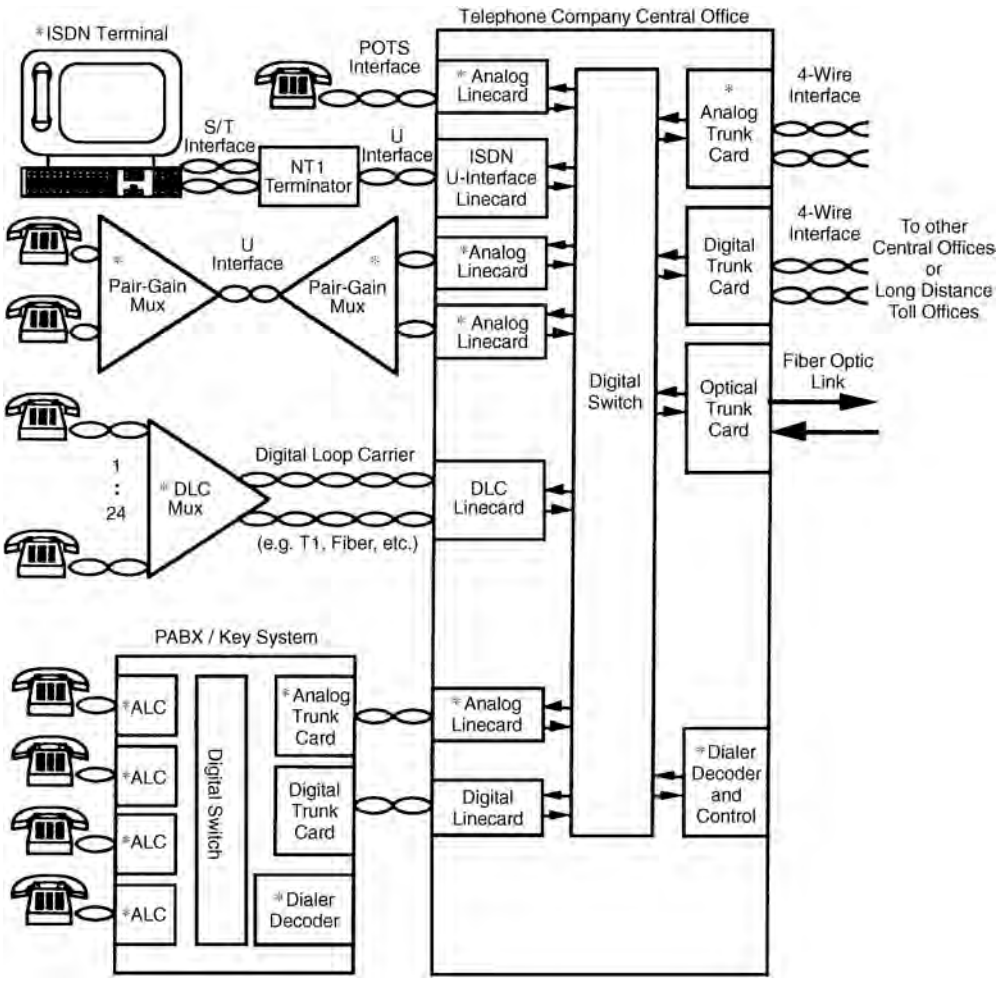


FIGURE 26.1 Public switching telephone network (PSTN).

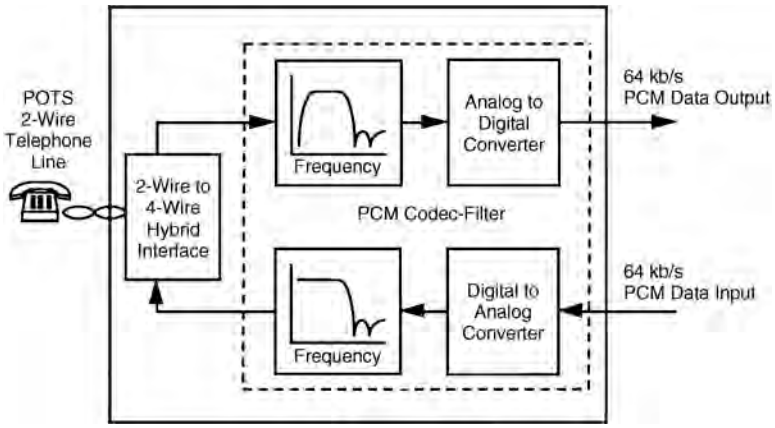


FIGURE 26.2 Analog telephone linecard.

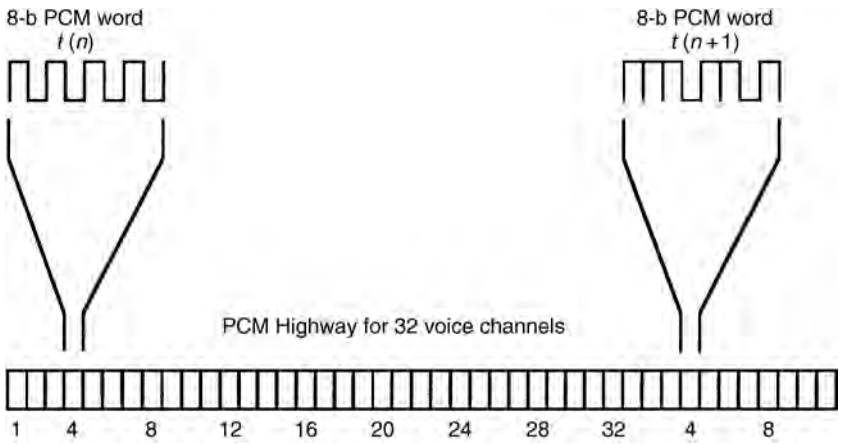


FIGURE 26.3 Multiplexed PCM highway.

during the loss of AC power by emergency battery backup. This requires low-power dissipation of the PCM codec-filter. Complementary metal-oxide-semiconductor (CMOS) integrated circuit (IC) fabrication technology has proven to be a reliable low-power semiconductor solution for this type of complex analog and digital very large-scale integration (VLSI) circuit. A PCM codec-filter is used in each location with an asterisk (*).

The analog linecard is the interface between the switch and the telephone wires that reach from the central office or service provider to the subscriber's telephone. The analog linecard also provides the digital interface for the switching or transmission equipment. The accepted name for analog telephone service is plain old telephone service (POTS). Figure 26.2 is a simplified block diagram for an analog linecard showing the functions of the PCM codec-filter.

The PCM codec-filter converts voice into 8-b PCM words at a conversion rate of 8 kilosamples per second. This PCM data is serially shifted out of the PCM codec-filter in the form of a serial 8-b word. This results in a 64-kb/s digital bit stream. This 64 kb/s PCM data from an analog linecard is typically multiplexed onto a PCM data bus or highway for transfer to a digital switch, where it will be routed to the appropriate output PCM highway, which routes the data to another linecard, completing the signal path for the conversation (refer to Fig. 26.1).

Transmission equipment will have multiple PCM codec-filters shifting their PCM words onto a single conductor. This is referred to as a serial bus or PCM highway. Figure 26.3 shows the 8-b PCM words

from a PCM codec-filter for one voice channel being multiplexed into a time slot on a PCM highway that can accommodate 32 voice channels. For more information on PCM switching, see the chapters in this handbook on switching systems and common channel signalling. For more information on hybrid interfaces, see the chapters in this handbook on POTS, analog hierarchy, and analog telephone channels and conditioning.

26.3 Design of Voice PCM Codec-Filters: Analog Transmission Performance and Voice Quality for Intelligibility

Filtering

The pass-band frequency response of the voice channel is roughly 300–3000 Hz. The 300–3000 Hz spectrum is where most of the energy in the human voice is located. Voice contains spectral energy below 300 Hz and above 3 kHz, but its absence is not detrimental to intelligibility. The frequency response within the passband of 300 Hz–3 kHz must be tightly controlled, because the public switching telephone network (PSTN) is allowed to have as many as seven ADC/DAC conversions end-to-end. The cumulative effects of the network plus private equipment such as PBX or key systems could audibly distort the frequency response of the voice channel with less stringent requirements. Figure 26.4 shows the typical half-channel pass-band frequency response requirement of +/-0.25 dB.

The 3-kHz bandwidth for voice signals determines the sample rate for the ADC and DAC. In a sampling environment, Nyquist theory states that to properly sample a continuous signal, it must be sampled at a frequency higher than twice the signal's highest frequency component. Minimizing the sample rate for digitizing voice is a priority since it determines the system data rates which are directly proportional to cost. The voice bandwidth of 3 kHz plus the filter transition band sampled at 8 kHz represents the best compromise that meets the Nyquist criterion.

The amount of attenuation required for frequencies of 4 kHz and higher is dictated by the required signal-to-distortion ratio of about 30 dB for acceptable voice communication as determined by the telephone industry. Frequencies in the filter transition band of 3.4–4.6 kHz must be attenuated enough such that their reconstructed alias frequencies will have a combined attenuation of about 30 dB. This permits typical filter transition band attenuation of 15 dB at 4 kHz for both the input filter for the ADC and the output reconstruction filter for the DAC. All frequencies above 4.6 kHz must be satisfactorily attenuated before the ADC conversion to prevent aliasing in-band. The requirement for out-of-band attenuation is specified by the country of interest, but it typically ranges from 25 to 32 dB for frequencies of 4600 Hz and higher.

The requirement for limiting frequencies below 300 Hz is also regionally dependent and application dependent, ranging from flat response to 20 dB or more attenuation at 50 and 60 Hz. The telephone line is susceptible to 50/60-Hz power line coupling, which must be attenuated from the signal by a high-pass filter before the ADC. Attenuation of power line frequencies is desirable to prevent 60 Hz noise during voice applications. Figure 26.4 shows a typical frequency response requirement for North American telephone equipment.

The digital-to-analog conversion process reconstructs a pulse-amplitude modulated (PAM) staircase version of the desired in-band signal, which has spectral images of the in-band signal modulated about the sample frequency and its harmonics. These high-frequency spectral images are called aliasing components which need to be attenuated to meet performance specifications. The low-pass filter used to attenuate these aliasing components is typically called a reconstruction or smoothing filter. The low-pass filter characteristics of the reconstruction filter are similar to the low-pass filter characteristics of the input antialiasing filter for the ADC.

The accuracy of these filters requires op amps with both high gain and low noise on the same monolithic substrate with precision matched capacitors which are used in the switched-capacitor filter structures. For more information on filter requirements and sampling theory, refer to the chapters in this handbook on analog modulation, sampling, and pulse-code modulation.

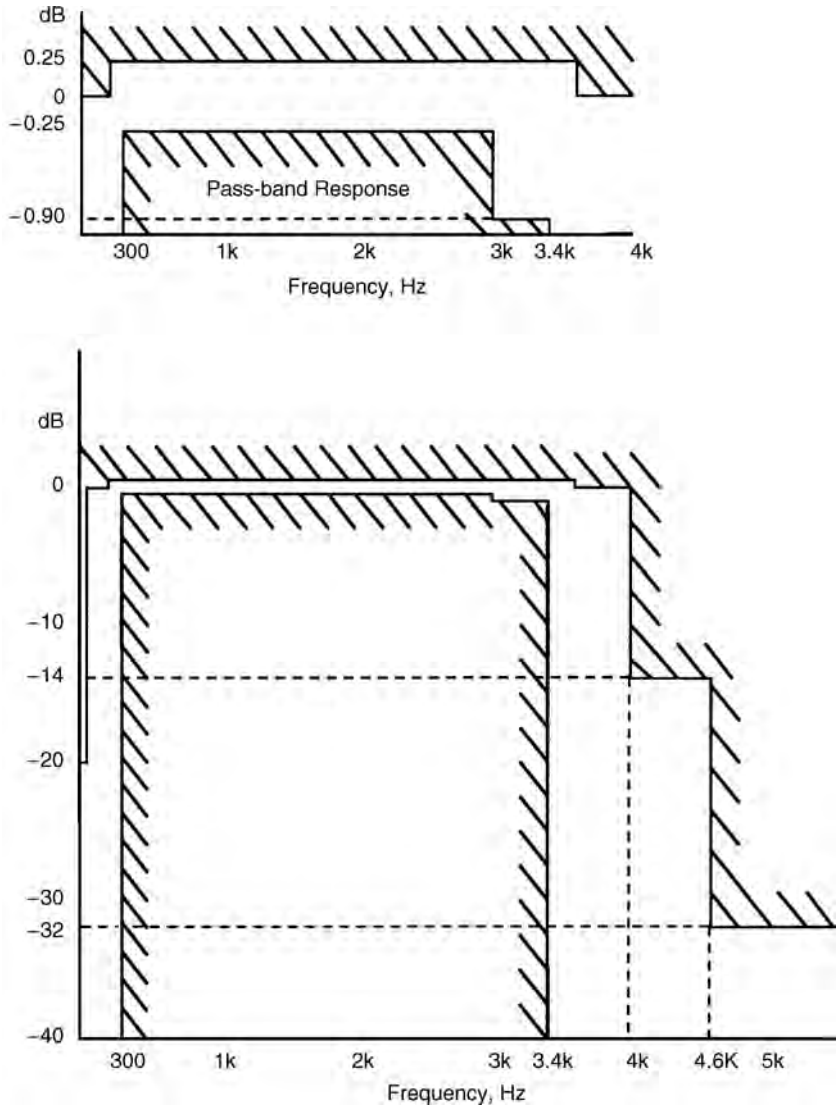


FIGURE 26.4 Typical frequency response requirement for the input filter: pass-band and stop band.

Quantizing Distortion

To digitize voice intelligibly requires a signal-to-distortion ratio of about 30 dB over a dynamic range of about 40 dB. This may be accomplished with a linear 13-b ADC and DAC, but will far exceed the required signal-to-distortion ratio for voice at amplitudes greater than 40 dB below the peak overload amplitude.

This excess performance comes at the expense of increased bits per sample which directly translates into system cost, as stated earlier. Figure 26.5 shows the signal-to-quantization distortion ratio performance of a 13-b linear DAC compared to a companded mu-law DAC. A high-speed modem application is discussed later in this chapter, which requires linear 13 b or better performance.

Two methods of data reduction are implemented, both of which use compressing the 13-b linear codes during the analog-to-digital conversion process and expanding the codes back during the digital-to-analog conversion process. These compression and expansion schemes are referred to as companding. The two companding schemes are: mu-255 law, primarily used in North America and Japan, and A-law, primarily used in Europe. Companding effectively trades the signal to distortion performance at larger

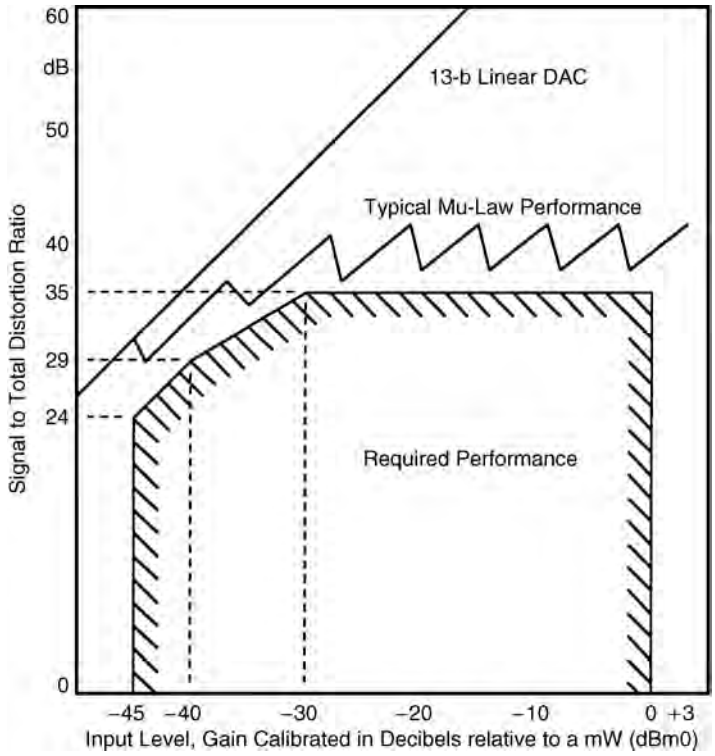


FIGURE 26.5 Signal-to-quantization distortion performance for mu-law and 13-b linear DACs compared to the telephone network requirements.

amplitudes for fewer bits for each sample. These companding schemes map the 13-b linear codes into pseudologarithmic 8-b codes. These companding schemes follow a segmented or piecewise-linear input–output transfer curve formatted as sign bit, three chord bits, and four step bits. For a given chord, all 16 of the steps have the same voltage weighting. As the voltage of the analog input increases, the four step bits increment and carry over to the three chord bits, which increment. When the chord bits increment, the step bits double their voltage weighting. This results in an effective resolution of six bits (1 sign bit + 1 effective chord bit + 4 step bits) with a sinusoidal stimulus across a 42-dB dynamic range (seven chords above zero, by 6 dB per chord, i.e., effectively 1 b). This satisfies the original requirement of 30-dB signal-to-distortion over a 40-dB dynamic range. Figure 26.6 is a graphical explanation of mu-law companding.

Note that to minimize die size area, which further cost reduces the PCM codec-filter, compressing ADC designs and expanding DAC designs are used instead of linear ADC/DAC designs with read-only memory (ROM) conversion lookup tables.

A-law companding is very similar to mu-law with three differences. The first difference is that A-law has equal voltage weighting per step for the two positive or negative chords near zero volts. In mu-law, the step weighting voltage for chord zero is one-half the step weighting voltage for A-law chord zero. This reduces the A-law resolution for small signals near the voltage origin. The second difference is that the smallest voltage of the DAC curve does not include zero volts and instead produces a positive half-step for positive zero and a negative half-step for negative zero. This is in contrast to mu-law companding, which produces a zero voltage output for both positive and negative zero PCM codes. Both mu-law and A-law have symmetric transfer curves about the zero volts origin, but mu-law has redundancy at zero volts and, therefore, effectively an unused code. The third difference between the two companding schemes deals with the data bit format. A-law inverts the least significant bit (LSB) and every other bit, leaving the sign bit unchanged relative to a conventional binary code. For example, applying the A-law data format formula to binary zero, s0000000, results in s1010101, the A-law code for +/- zero. Mu-law

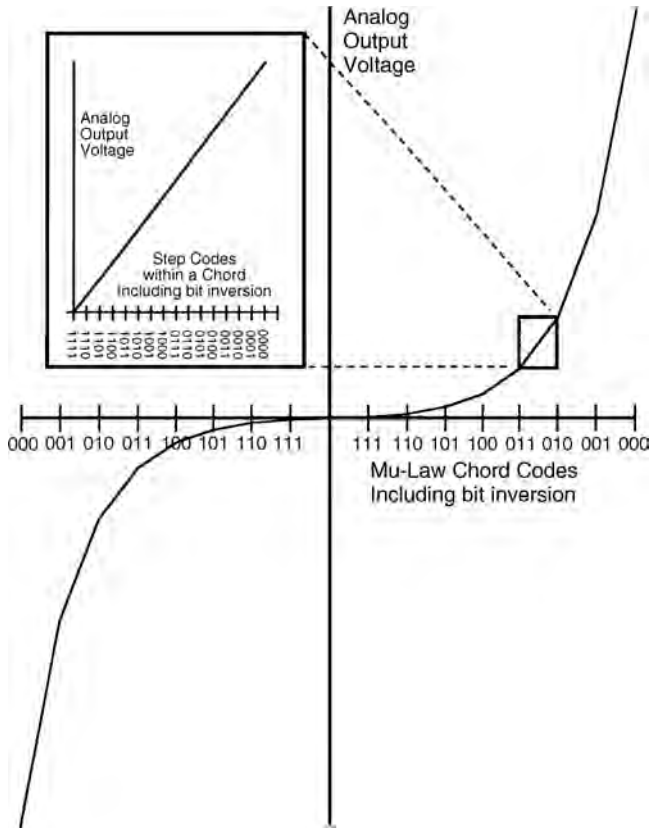


FIGURE 26.6 Mu-law companding showing piecewise-linear approximation of a logarithmic curve.

	Mu-Law			A-Law		
Level	Sign Bit	Chord Bits	Step Bits	Sign Bit	Chord Bits	Step Bits
+Full scale	1	000	0000	1	010	1010
+Zero	1	111	1111	1	101	0101
-Zero	0	111	1111	0	101	0101
-Full scale	0	000	0000	0	010	1010

FIGURE 26.7 Full scale and zero codes for mu-law and A-law.

maintains a positive polarity for the sign bit with all of the magnitude bits inverted. Figure 26.7 shows the zero codes and full-scale codes for both mu-law and A-law companding schemes.

Mu-law and A-law companding schemes are recognized around the world. Many equipment manufacturers build products for distribution worldwide. This distribution versatility is facilitated by the PCM codec-filter being programmable for either mu-law or A-law. With the deployment of ISDN, the PCM codec-filter is located at the subscriber's location. The need to place international telephone calls dictates the need for a PCM codec-filter that is both mu-law and A-law compatible.

Gain Calibration: Transmission Level Point

The test tone amplitude for analog channels is 3 dB down from the sinusoidal clip level. This test level for a POTS line is generally 1 mW, which is 0 dBm. The telephone line historically was referred to as having a 0 dB transmission level point (TLP). (With the improvements in transmission quality and fewer

	Mu-Law			A-Law		
Phase	Sign Bit	Chord Bits	Step Bits	Sign Bit	Chord Bits	Step Bits
$\pi/8$	0	001	1110	0	011	0100
$3\pi/8$	0	000	1011	0	010	0001
$5\pi/8$	0	000	1011	0	010	0001
$7\pi/8$	0	001	1110	0	011	0100
$9\pi/8$	1	001	1110	1	011	0100
$11\pi/8$	1	000	1011	1	010	0001
$13\pi/8$	1	000	1011	1	010	0001
$15\pi/8$	1	001	1110	1	011	0100

FIGURE 26.8 PCM codes for digital milliwatt.

losses, telephone lines often have attenuation pads to maintain talker/listener appropriate levels.) Amplitudes that have been TLP calibrated are given unit abbreviations that have a zero added to them, for example, dBm0. The test signal may experience gain or attenuation as it is transmitted through the network. As an example, the absolute magnitude of one electrical node for the voice channel could result in a measured value of +6 dBm when a 0 dBm0 test signal is applied. This node would be gain calibrated as having a +6 dB TLP. This would result in absolute measurements for noise having this same +6 dB of gain. To properly reference the noise, all measurements must be gain calibrated against the 0 dBm0 level, which, in this example, would result in 6 dB being subtracted from all measurements. All signals in this channel, including distortion and noise, would be attenuated by 6 dB by the time they get to the subscriber's telephone.

For PCM channels, the 0 dBm0 calibration level is nominally 3 dB down from maximum sinusoidal clip level; specifically, it is 3.17 dB down for mu-law and 3.14 dB down for A-law. These minor deviations from an ideal 3 dB result from using a simple code series, called the digital milliwatt, for generating the digital calibration signal. The digital milliwatt is a series of eight PCM codes that are repeated to generate the 1 kHz calibration tone. Figure 26.8 shows the eight PCM codes which are repeated to generate a digital milliwatt for both mu-law and A-law in the DAC.

The digital milliwatt is used to calibrate the gains of telephone networks using voice grade PCM codec-filters.

Idle Channel Noise

Idle channel noise is the noise when there is no signal applied to the voice channel, which is different from the quantization distortion caused by the ADC/DAC process. This noise can be from any circuit element in the channel, in addition to the PCM codec-filter. Noise may be coupled in from other circuits including, but not limited to, hybrid transformer interfaces, power supply induced noise, digital circuitry, radio frequency radiation, and resistive noise sources.

The PCM codec-filter itself has many opportunities to contribute to the idle channel noise level. The input high-pass filter often is third order, whereas both the input and the output low-pass filters are typically fifth-order elliptic designs. The potential for noise generation is proportional to the order of the filter. The ADC and DAC arrays have many components that are controlled by sequencing digital circuits. The power supply rejection capability of PCM codec-filters once dominated the noise level of the channel. The power supply rejection ratio of these devices has been improved to a level where the device is virtually immune to power supply noise within the allowable power supply voltage range. This performance was attained by differential analog circuit designs in combination with tightly controlled matching between on-chip components.

Noise measurements require a different decibel unit as they usually involve some bandwidth or filter conditioning. One such unit commonly used (especially in North America) is decibels above reference noise (dBrn). The reference noise level is defined as 1 pW or -90 dBm. Telephone measurements typically refer to dBrnC, which is the noise level measured through a C-message weighting filter (a filter that

simulates the frequency response of the human ear's noise sensitivity). European systems use a related term called dBmp, which is the dBm level noise measured through a psophometric filter. Noise measurements made by either dBrnC or dBmp filter weightings have units to show that the noise has been gain referenced to 0 dBm0 by adding a zero, hence dBrnC0 and dBm0p. Two examples are shown to illustrate the use of these units:

1. Mu-law: If 0 dB TLP = +6 dB, then a noise measurement of 20 dBrnC equals 14 dBrnC0.
2. A-law: If 0 dB TLP = +4LPdB, then a noise measurement of -70 dBmp equals -74 dBm0p.

The examples are representative of typical idle channel noise measurements for a full-duplex digital channel. Idle channel noise measuring 14 dBrnc0 at a 0 dB TLP output is about 123-μV root mean square (rms). This low-noise level is very suceptible to degradation by the noise sources mentioned earlier.

The idle channel noise of the voice channel directly impacts the dynamic range, which is the ratio of maximum power the channel can handle to this noise level. Typical dynamic range for a mu-law PCM codec-filter is about 78 dB. The term dynamic range is similar to signal-to-noise ratio for linear circuits. The signal-to-noise plus distortion ratio for a companded channel is generally limited by the nonlinear companding except at very low-signal amplitudes.

Gain Tracking: Gain Variations as a Function of Level

The quantization curves, as discussed earlier, are not linear and may cause additional nonlinear distortions. The concern is that the gain through the channel may be different if the amplitude of the signal is changed. Gain variations as a function of level or how well the gain tracks with level is also called gain tracking. This is a type of distortion, but could easily be missed by a tone stimulus signal-to-distortion test alone. Gain tracking errors can cause ringing in the 2-wire–4-wire hybrid circuits if the gain at any level gets too large. Gain tracking performance is dominated by the IC fabrication technology, more specifically, the matching consistency of capacitors and resistors on a monolithic substrate. Figure 26.9 shows the half-channel gain tracking performance recommended for a PCM digital interface.

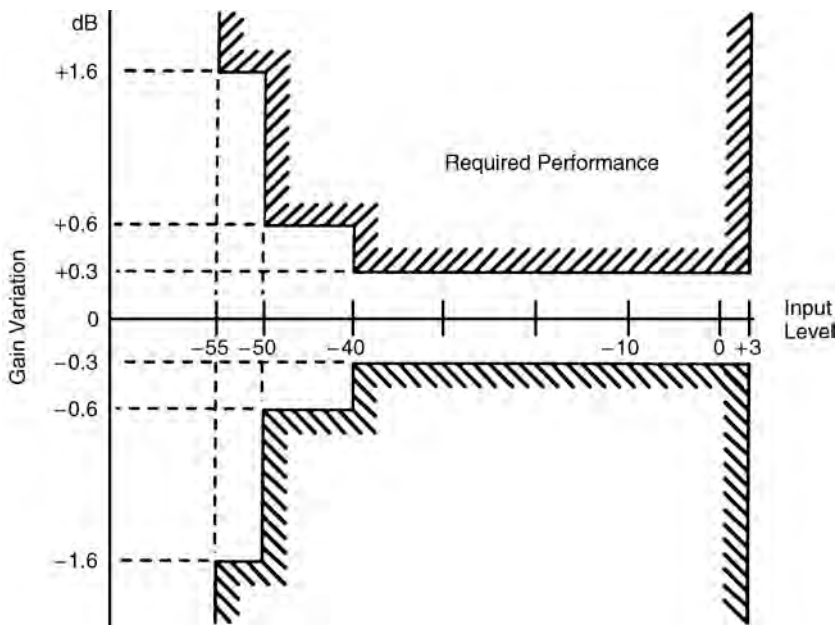


FIGURE 26.9 Variation of gain with input level using a tone stimulus.

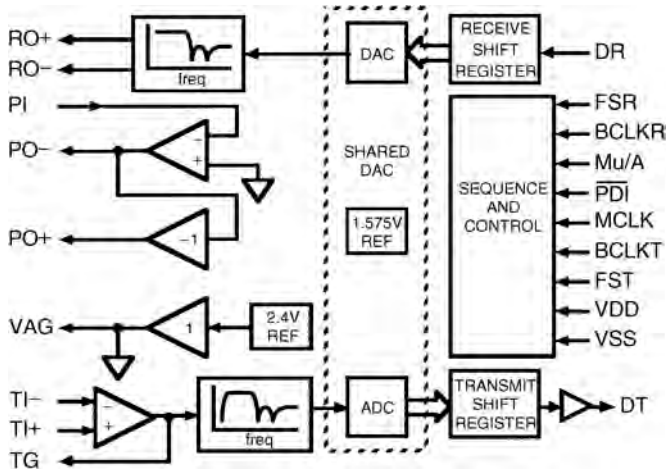


FIGURE 26.10 MC145480 5 V PCM codec-filter.

Figure 26.10 is a block diagram for a monolithic MC145480 5 V PCM codec-filter manufactured by Motorola Inc. The MC145480 uses CMOS for reliable low-power operation. The operating temperature range is from -40° to +85° C, with a typical power dissipation of 23 mW. All analog signal processing circuitry is differential utilizing Motorola's six-sigma quality IC fabrication process. This device includes both antialiasing and reconstruction switched capacitor filters, compressing ADC, expanding DAC, a precision voltage reference, shift registers, and additional features to facilitate interfacing to both the analog and digital circuitry on a linecard.

26.4 Linear PCM Codec-Filter for High-Speed Modem Applications

As discussed in previous sections of this chapter, the original PCM codec-filters were developed to facilitate the digital switching of voice signals in the central office of the PSTN. The signal processing in the digital domain consisted of simply moving digitized samples of voice from one time slot to another to accomplish the switching function. As such, the A/D and D/A conversion processes could be nonlinear, and companding was employed.

With the advent of the personal computer, the facsimile machine, and most recently the Internet, the need to transmit data originating at the subscriber over the PSTN at the highest possible rates has grown to the point where the theoretical limits as defined by Shannon's law [Shannon, 1948] are being approached. For data transmission, the signal processing is, in a sense, the inverse of that which spawned the original PCM codecs in that the initial signal is data that must be modulated to look like a voice signal in order to be transmitted over the PSTN. At data rates above 2400 b/s, the modulation is done using sophisticated digital signal processing techniques, which have resulted in the development of a whole new class of linear PCM codec-filters.

Codec-filters for V.32 (9.6 kb/s), V.32bis (14.4 kb/s), and V.34 (28.8 kb/s) modems must be highly linear with a signal to distortion much greater than 30 dB. This linearity and signal to distortion is required to implement the echo cancellation function with sufficient precision to resolve the voltages corresponding to the data points in the dense constellations of high-speed data modems. Please refer to the chapters in this handbook on signal space and echo cancellation for more information. The critical signal-to-noise plus distortion plots for two popular commercially available codec-filters used in V.32bis (T7525) and V.34bis (STLC 7545) modems are shown in Fig. 26.11. Both of these codec-filters are implemented using linear high ratio oversampling sigma–delta (SD) A/D and D/A conversion technology and digital

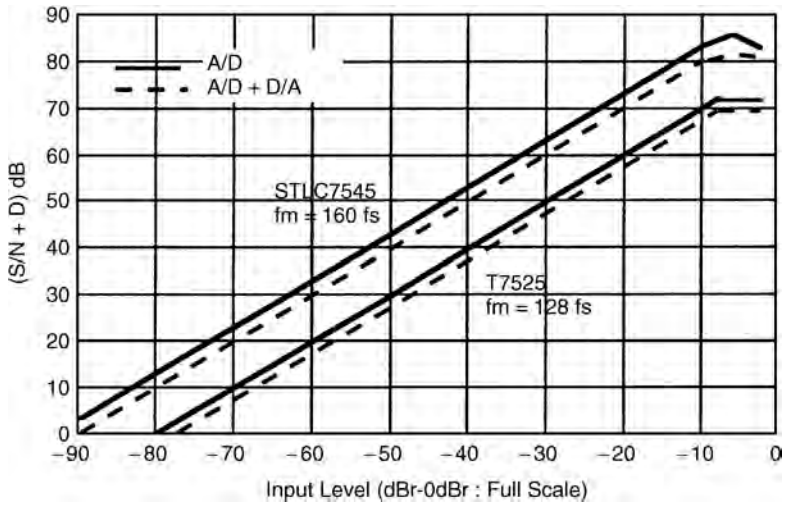


FIGURE 26.11 Linear codec dynamic range, where fm is the modulator frequency and fs is the sample frequency.

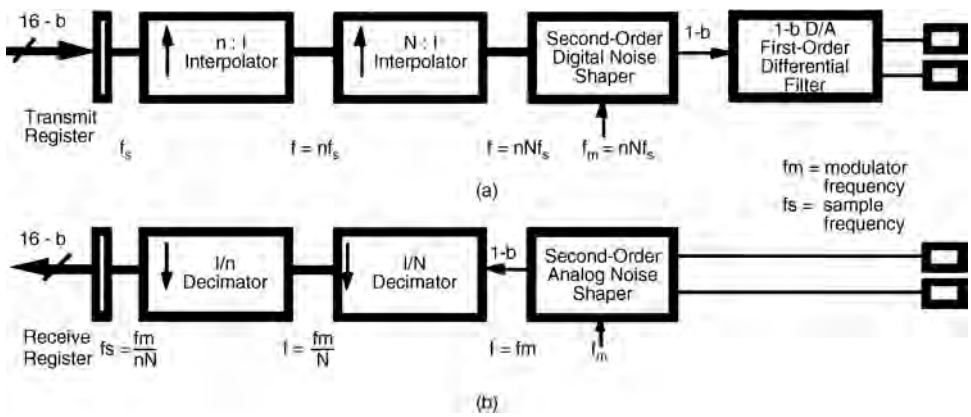


FIGURE 26.12 Generalized SD PCM modem codec-filter: (a) transmit channel and (b) receive channel.

filters. This can be contrasted with the nonlinear companding converters and switched capacitor filters of voice grade PCM codec-filters.

Sigma–delta conversion technology is based on coarsely sampling a signal at a high rate and filtering the resulting noise [Candy and Temes, 1992; Park, 1990]. A generalized block diagram of sigma–delta PCM codec-filters for modem applications is shown in Fig. 26.12. The transmit and receive signal paths are the inverse of each other. The second-order modulators basically consist of two integrators and a comparator with two feedback loops. The action of the modulators is to low-pass filter the signal and high-pass filter the noise so that greater than 12 b of resolution (70 dB) can be achieved in-band. If it is assumed that the input signal is sufficiently active to make the quantization noise random, then it can be shown that [Candy and Oconnell, 1981], for a simple first-order modulator,

$$No^2 = (1/fo)^3$$

where:

No = net rms noise in band

fo = oversampling frequency

That is, the in-band noise power decreases in proportion to the cube of the oversampling frequency resulting in improved in-band signal-to-noise performance. The high-frequency quantization noise in the receive channel is removed by the digital decimation filters, whereas the high-frequency alias noise due to the interpolation process in the transmit channel is removed by interpolation filters. The transmit and receive channels are synchronous.

The decimation and interpolation operations are done in two stages: a low-ratio (n) stage and a high-ratio (N) stage. The high-ratio stage is implemented with cascaded comb filters, which are simply moving average filters having the following frequency response:

$$H(z) = 1/N((1 - z^{-N})/(1 - z^{-1}))^{i}$$

where i is the number of cascaded filters. For modem applications, the order i of the cascade is generally three. Comb filters are essentially linear phase finite impulse response (FIR) filters with unity coefficients and, therefore, they do not require a multiplier to implement, which makes them silicon area efficient. The low-ratio stage is usually implemented with cascaded second-order infinite impulse response (IIR) filters. These filters do require a multiplier to implement. The frequency response of the IIR filters compensate for the $\sin x/x$ droop of the comb filters and shape the pass-band to meet telephone specifications. The total oversampling product, $nN = fo$, is typically greater than or equal to 128 with $N = 32$ or 64. For V.32bis modems, fs is fixed at 9.6 kHz, whereas for V.34bis modems, fs varies from 9.6 to ~14 kHz.

26.5 Concluding Remarks

The PCM voice codec-filter has continuously evolved since the late 1950s. With each improvement in performance and each reduction in cost, the number of applications has increased. The analog IC technology incorporated in today's PCM voice codec-filters represents one of the true bargains in the semiconductor industry.

Sigma–delta-based PCM codec-filters have become the standard for implementing high-speed modems because they can offer sufficient resolution at the lowest cost. Sigma–delta codecs are cost/performance effective because they are 90% digital and digital circuitry scales with integrated circuit (IC) process density improvements and associated speed enhancements. It is interesting to note that although the analog portion of the sigma–delta codec-filter is small and does not require precision or matched components like other converter technologies, it (the D/A output differential filter) limits the ultimate performance of the codec-filter.

References

Bellamy, J., *Digital Telephony,* John Wiley and Sons, New York, 1991.

Bellcore, Functional criteria for digital loop carrier systems, Tech. Ref. Doc., TR-TSY-000057, Issue 1, April, 1987.

Candy, J.C. and Oconnell, J.B., The structure of quantization noise from sigma–delta modulation, *IEEE Trans. Comm.,* COM-29 (Sept.), 1316, 1981.

Candy, J.C. and Temes, G.C., *Oversampling Delta–Sigma Data Converters Theory, Design and Simulation,* IEEE Press, New York, 1992.

ITU, ITU-T digital networks, transmission systems and multiplexing equipment, CCITT Red Book, G.711, G.712, G.713, G.714, International Telecommunications Union—Telecommunications, 1984.

Park, S., Principles of sigma-delta modulation for analog-to-digital converters, Motorola Digital Signal Processing Division Application Note, Austin, TX, 1990.

Shannon, C.E., A mathematical theory of communication, *Bell Syst. Tech. J.,* 27 (July), 379, 1948.

27

Digital Hierarchy

B. P. Lathi
California State University

Maynard A. Wright
Acterna

27.1 Introduction

Following the introduction of digital encoding and transmission of voice signals in the early 1960s, multiplexing schemes for increasing the number of channels that could be transported were developed by the Consultative Committee on International Telephony and Telegraphy (CCITT) in Europe and by the Bell System in North America. The initial development of such systems was aimed at coaxial cable and millimeter waveguide transmission systems. The advent of low-loss optical fiber and efficient modulation schemes for digital radio greatly increased the need to transport digital channels at rates higher than the primary multiplexing rate.

Separate multiplexing schemes, called digital hierarchies, were developed by CCITT and by the Bell System. During the period in which the two hierarchies were conceived, there was no reliable way to distribute highly accurate clock signals to each central office that might serve as a multiplexing point. Local clock signals were, therefore, provided by relatively inaccurate crystal oscillators, which were the only economical signal sources available. Each hierarchy was, therefore, designed with the expectation that the clocks controlling the various stages of the multiplexing and demultiplexing processes would not be accurate and that means would be required for compensating for the asynchronicities among the various signals.

The mechanism chosen to accomplish synchronization of the multiplexing process was positive bit stuffing, as described elsewhere in this handbook. The CCITT development was termed the plesiochronous digital hierarchy (PDH) and, in North America, the scheme was designated the asynchronous digital hierarchy. The two differ significantly. The North American digital hierarchy will be described in the sections that follow.

The advent of accurate and stable clock sources has made possible the deployment of multiplexing schemes that anticipate that most of the signals involved will be synchronous. The North American synchronous multiplexing scheme is the Synchronous Optical Network (SONET) and is described elsewhere in this handbook. The International Telecommunications Union—Telecommunications Standardization Sector (ITU-T) version is termed the synchronous digital hierarchy (SDH). In spite of the development of these synchronous multiplexing hierarchies, the asynchronous hierarchies remain important because there are substantial numbers of asynchronous legacy systems and routes that must be maintained and that must support further growth.

27.2 North American Asynchronous Digital Hierarchy

Following dissolution of the Bell System in 1984, Committee T1—Telecommunications was formed to take responsibility for maintenance and development of North American telecommunications standards. The Alliance for Telecommunications Industry Solutions (ATIS) serves as the secretariat for Committee T1. Committee T1 maintains standard T1.107 [T1.107, 1995] that describes the North American asynchronous digital hierarchy.

The capacities of the bit streams of the various multiplexing levels in the North American digital hierarchy are divided into overhead and payload. Overhead functions include framing, error checking, and maintenance functions, which will be described separately for each level. Payload capacity is used to carry signals delivered from the next lower level or delivered directly to the multiplexer by a customer.

The various rates in the digital hierarchy are termed digital signal n (DSn), where n is the specified level in the hierarchy. The various rates are summarized in Table 27.1.

In addition, T1.107 provides standards for multiplexing various sub-DS0 digital data signals into a 64 kb/s DS0.

The relationships among the various signals are illustrated in Fig. 27.1. Note that DS1C is a dead-end rate. It is useful for channeling paired cable plant but it cannot be multiplexed into higher level signals. DM in Fig. 27.1 represents digital multiplexer.

Digital Signal Level 0 (DS0)

The DS0 signal is a 64-kb/s signal that is usually organized in 8-b bytes and which may be developed by encoding analog signals as described elsewhere in this handbook. Alternatively, the DS0 may be built from digital signals that are synchronous with the DS1 clock. There is no provision for handling asynchronous signals at the DS0 rate, although lower rate asynchronous signals are sometimes oversampled at 64 kb/s to accommodate them to the rate of the DS0 channel rate or to the rate of a lower rate synchronous channel derived from the DS0.

TABLE 27.1 Rates in the North American Digital Hierarchy

DS0	64-kb/s signal that may contain digital data or $\mu = 255$ encoded analog signals
DS1	1.544 Mb/s signal that byte-interleaves 24 DS0s
DS1C	3.152 Mb/s signal that bit-interleaves 2 DS1s
DS2	6.312 Mb/s signal that bit-interleaves 4 DS1s
DS3	44.736 Mb/s signal that bit-interleaves 7 DS2s

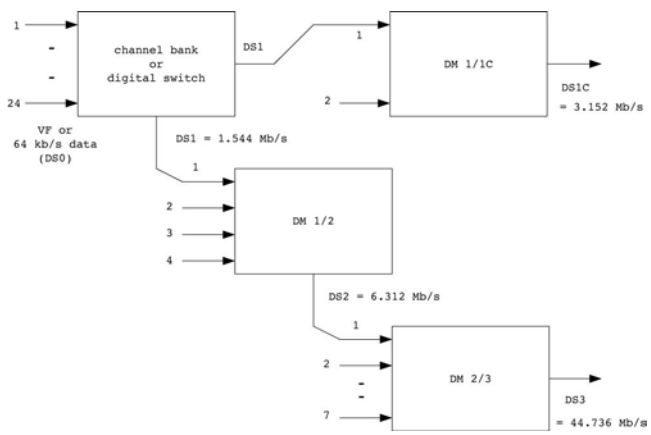


FIGURE 27.1 Multiplexing in the North America digital hierarchy.

TABLE 27.2 DS1 SF Overhead Bit Assignments

Frame Number:	1	2	3	4	5	6	7	8	9	10	11	12
Ft bits	1	–	0	–	1	–	0	–	1	–	0	–
Fs bits	–	0	–	0	–	1	–	1	–	1	–	0
Composite pattern	1	0	0	0	1	1	0	1	1	1	0	0

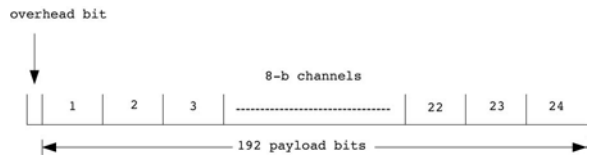


FIGURE 27.2 DS1 frame structure.

Digital Signal Level 1 (DS1)

The DS1 signal is built by interleaving 1.536 Mb/s of payload capacity with 8 kb/s of overhead. Following each overhead bit, 8-b bytes from each of the 24 DS0 channels to be multiplexed are byte interleaved. The resulting structure appears in Fig. 27.2. The DS0 bytes (channels) are numbered sequentially from 1 to 24 in Fig. 27.2, and that scheme is the standard method for interleaving DS0 channels. Other schemes, which employ a nonsequential ordering of DS0 channels, have been used in the past.

This 193-b pattern is called a frame and is repeated 8000 times per second to coincide with the sampling rate for encoding analog signals. Each 8-b encoded sample of a particular channel may be transmitted in the appropriate channel position of a single frame.

Digital Signal Level 1 Overhead: Superframe Format

In the earliest 1.544 Mb/s systems, the framing pattern was simply used to identify the beginning of each frame so that the channels could be properly demultiplexed by the receiver. With the advent of robbed-bit signaling, it became necessary to identify the particular frames in which robbed-bit signaling occurs. A superframe (SF) of twelve frames was developed in which robbed-bit signaling occurs in the 6th and 12th frames. The overhead capacity in the superframe format is organized into two streams. The pattern carried by the overhead bits (Ft bits—terminal framing) in odd-numbered frames identifies the beginning of the frame structure and serves to locate the overhead stream. The pattern in the even-numbered frames (Fs bits—signaling framing) locates the beginning of each superframe so that the robbed-bit signaling can be properly demultiplexed. The SF framing pattern is shown in Table 27.2.

Note that each frame in which the logical state of the Fs bit changes is a signaling frame that may carry robbed-bit signaling.

There are two bits per superframe devoted to signaling. They are denoted A and B and may be used to carry four-state signaling, although two-state signaling (on- and off-hook indications) is more common.

Digital Signal Level 1: Extended Superframe Format

Improvements in electronics and framing algorithms have made it possible for receivers to frame rapidly and efficiently on patterns with less information than is contained in the 8-kb/s superframe overhead stream. The extended superframe format (ESF) uses a 2-kb/s framing pattern to locate a 24-frame extended superframe. The frame structure remains as for the SF format, and the ESF format therefore provides 6 kb/s of overhead capacity which is not devoted to framing.

A communications channel [data link or (DL)] consumes 4 kb/s of overhead. This channel is used to convey performance monitoring information from one DS1 terminal location to the other (scheduled messages) and to send alarm messages and loopback requests (unscheduled messages). The use of this channel is detailed in T1.403 [ANSI, 1999]. When no message is being sent in the DL, it is usually filled with unconcatenated high-level data link control (HDLC) flags (01111110) although some older

TABLE 27.3 DS1 Extended Superframe Overhead Bit Assignments

Frame Number:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
FAS				0				0				1				0				1				1
DL	D		D		D		D		D		D		D		D		D		D		D		D	
CRC		C				C				C				C				C				C		

equipment may fill with an all ones pattern. HDLC frames for carrying performance information from DS1 path terminations, as well as from intermediate points, are defined in T1.403 as the performance report message (PRM), the network performance report message (NPRM), and the supplementary performance report message (SPRM).

The remaining 2 kb/s of overhead in the ESF format carry a cyclic redundancy code with a 6-b remainder (CRC-6) channel to provide error checking. The divisor polynomial is $X^6 + X + 1$. Division is carried out over all 4632 b of an extended superframe with all the overhead bits set to logical ones. The 6-b remainder resulting from the division is written into the CRC-6 bits of the following extended superframe.

Table 27.3 contains a summary of the overhead bit assignments in the ESF format where FAS represents the frame alignment signal. The bits forming the 4 kb/s DL are represented by D and the six individual bit positions for the CRC-6 remainder from the previous extended superframe are shown as C.

Digital Signal Level 1C Format

Although, as mentioned previously, DS1C cannot be multiplexed to higher levels in the hierarchy, it has proved useful for channelizing interoffice cable pairs with more channels than can be carried by a comparable DS1 system. A transmission facility for DS1C, the T1C line, was introduced and deployed during the 1970s. Although lightwave interoffice transmission has significantly diminished the importance of DS1C, a discussion of that rate provides a good starting point for discussing the higher levels in the hierarchy.

The DS1C format multiplexes two 1.544 Mb/s DS1 signals into a single 3.152 Mb/s stream. Unlike the other levels in the digital hierarchy, DS1C signals cannot be multiplexed to higher levels. Two DS1Cs together exceed the capacity of a single DS2, making such multiplexing impossible and making DS1C an orphan rate.

As is the case for all the hierarchical rates above DS1, the DS1C format is organized into M-frames and M-subframes. The length of a single M-frame is 1272 b. Overhead bits occur every 53 b, with 52 payload bits interspersed between adjacent overhead bits. Each M-frame is divided into 4 M-subframes of 318 b each. Overhead assignments are shown in Table 27.4.

Digital Signal Level 1C Frame Alignment

The information in Table 27.4 is reorganized into columns by M-subframe in Table 27.5.

Note that certain of the overhead bits, including the F1 and F2 bits, recur in the same position in every M-subframe. Typical framing practice, therefore, is to frame on the F bits to locate the boundaries of the M-subframes and to then frame on the first bits of the M-subframes to locate the boundaries of the M-frame structure.

Digital Signal Level 1C X Bit

The X bit provides a communications channel between DS1C terminals that runs at just under 2500 b/s. The usual use of the X bit channel is to provide a remote alarm indication (RAI) to the distant terminal. When no alarm condition exists, the X-bit is set to logical one.

TABLE 27.4 DS1C Overhead Bit Assignments

M-frame Bit Number	M-subframe Number	Overhead Bit Assignment	Logical Value of Bit
0	1	M1	0
53	1	C1	first DS1 stuff control
106	1	F1	0
159	1	C2	first DS1 stuff control
212	1	C3	first DS1 stuff control
265	1	F2	1
318	2	M2	1
371	2	C1	second DS1 stuff control
424	2	F1	0
477	2	C2	second DS1 stuff control
530	2	C3	second DS1 stuff control
583	2	F2	1
636	3	M3	1
689	3	C1	first DS1 stuff control
742	3	F1	0
795	3	C2	first DS1 stuff control
848	3	C3	first DS1 stuff control
901	3	F2	1
954	4	X	X
1007	4	C1	second DS1 stuff control
1060	4	F1	0
1113	4	C2	second DS1 stuff control
1166	4	C3	second DS1 stuff control
1219	4	F2	1

TABLE 27.5 DS1C M-subframe Structure

M-subframe 1	M-subframe 2	M-subframe 3	M-subframe 4	Overhead Bit Assignment
0	318	636	954	M1/M2/M3/X
53	371	689	1007	C1
106	424	742	1060	F1
159	477	795	1113	C2
212	530	848	1166	C3
265	583	901	1219	F2

Digital Signal Level 1C Bit Stuffing and Stuffing Control Bits

Of the 1272 b in a DS1C M-frame, 24 are overhead bits. The bandwidth available for transporting payload is therefore:

$$\begin{aligned} PCmax &= (Mfb\ ts - OHb\ ts)/(Mfb\ ts * rate) \\ &= (1272 - 24)/(1272 * 3.152\ \text{Mb/s}) \\ &= 3.092\ \text{Mb/s} \end{aligned}$$

where:

PCmax = maximum payload capacity of the DS1C signal
Mfbits = number of bits per M-frame, 1272
OHbits = number of overhead bits per M-frame, 24
rate = DS1C bit rate, 3.152 Mb/s

Note that this is more than the payload required to transport two DS1s running at their nominal rates, which is:

$$\begin{aligned} PCreq &= 2 * 1.544 \text{ Mb/s} \\ &= 3.088 \text{ Mb/s} \end{aligned}$$

Each M-subframe, however, contains a payload bit that is designated as a stuff bit. It can be used to carry a payload bit or it may be passed over and left unused. If it is unused in every M-subframe, the bandwidth made available for the payload will be:

$$\begin{aligned} PCm\ n &= (Mfb\ ts - OHb\ ts - Sb\ ts)/(Mfb\ ts * rate) \\ &= (1272 - 24 - 4)/(1272 * 3.152 \text{ Mb/s}) \\ &= 3.083 \text{ Mb/s} \end{aligned}$$

where:

PCmin = minimum payload capacity of the DS1C signal
Sbits = number of stuff bits (opportunities) per M-frame

If all the stuff bits are skipped, the capacity of the DS1C channel is less than the amount required by two DS1s.

Note that, by either using or skipping stuff bits, the actual payload capacity of the DS1C signal may be varied between the extremes represented by *PCmin* and *PCmax* to handle the rate of the DS1 signal to be transported. The range available between *PCmin* and *PCmax* exceeds the range of rates allowed for DS1 signals by T1.102 [ANSI, 1999].

Stuff bits for the first DS1 occur in the first and third M-subframes and for the second DS1 in the second and fourth M-subframes. The stuff bit for a particular M-subframe is always the third time slot allocated to the DS1 involved following overhead bit C3. For DS1 number 1, this is the fifth bit after C3, and for DS1 number 2, it is the sixth bit after C3. For a particular M-subframe, stuffing will be performed if the C bits (C1, C2, and C3) for that M-subframe are all set to logical ones. If the C bits are set to logical zeroes, no stuffing will occur in that M-subframe. The use of three C bits allows for majority voting by the receiver where one of the C bits may have been corrupted by a line error. This makes the process much more robust to such errors.

Digital Level 1C Payload

The two DS1s which are to be multiplexed are bit interleaved together to form the DS1C payload. Prior to bit interleaving, DS1 number 2 is logically inverted. Prior to inserting the interleaved payload into the DS1C overhead structure, the payload is scrambled in a single-stage scrambler. The output of the scrambler is the modulo-2 sum of the current input bit and the previous output bit.

Higher Rate Formats

As does DS1C, the DS2 and DS3 formats use positive bit stuffing to reconcile the rates of the signals being multiplexed. Both use the same M-frame and M-subframe structure with overhead bits assigned to the same tasks as for DS1C. Each of the rates uses bit interleaving to insert subsidiary bit streams into their payload bits. A synopsis of the characteristics of the various rates appears in Table 27.6. Each rate will be discussed in the sections that follow.

The Digital Signal Level 2 Rate

The DS2 rate is summarized in Table 27.6. It operates in a manner very similar to DS1C except that the rate is higher and four DS1s may be carried by a single DS2. A transmission system for carrying DS2 signals over paired copper cable called T2 was once available (1970s) but was never widely deployed.

TABLE 27.6 Characteristics of Levels in the Digital Hierarchy

	DS1C	DS2	DS3
Transmission rate, Mb/s	3.152	6.312	44.736
Subsidiary format	DS1	DS1	DS2
Number of subsidiaries multiplexed	2	4	7
Ratio of payload bits to overhead bits	52:1	48:1	84:1
M-frame length, b	1272	1176	4760
Number of M-subframes per M-Frame	4	4	7
M-subframe length, b	318	294	680
Number of X bits per M-frame	1	1	2
Number of C bits per M-frame	12	12	21
Number of M bits per M-frame	3	3	3
Number of F bits per M-frame	2	2	4
Number of P bits per M-frame	0	0	2
Number of stuff opportunities per tributary channel per M-frame	2	1	1

DS2 serves today primarily as a bridge between DS1 and DS3 and is only rarely found outside the confines of a single unit of equipment.

The Digital Signal Level 3 Rate

DS3 is heavily used as an interface to lightwave and digital radio systems. DS3 operates in much the same way as DS1C and DS2. An additional feature of DS3 is the pair of parity bits carried by each M-frame. They are used to transmit a parity error indication for the preceding M-frame. If the modulo-2 sum of all the information bits in the preceding M-frame is one, then both P bits are set to one. If the modulo-2 sum of the information bits is zero, then the P bits are set to zero. The two P bits of an M-frame are always set to the same value.

The same is true of the two X-bits in an M-frame. They are used as an alarm channel as for DS1C but are always set to the same value.

C-Bit Parity Digital Signal Level 3

A DS3 that operates using positive bit stuffing to multiplex its constituent DS2s is known as an M23 application. Another DS3 application, C-bit parity, is also defined in T1.107 [ANSI, 1995a].

Since the DS2 rate is almost never used except internal to a multiplexer as a stepping stone from DS1 to DS3, it is possible to lock its rate to a particular value which is slaved to the DS3 signal generator in the multiplexer. If the rate chosen for the DS2s provides for either no bit stuffing or for stuffing at every opportunity, the receiver can be made to know that and will be able to demultiplex the DS2s without reading the stuffing information that is normally carried by the C-bits.

The C-bit parity format operates the DS2 with stuffing at every opportunity and so frees up the control bits for other uses. There are 21 C bits per M-frame, which provides a channel running at approximately 197 kb/s. The 21 C bits per M-frame, are assigned as shown in Table 27.7.

The C-bit parity identifier is always set to logical one and is used as a tag to identify the DS3 signal as C-bit parity formatted. Note that the C-bit parity identifier is necessary but not sufficient for this purpose because it may be counterfeited by a DS2 in DS3 timeslot number one running at minimum rate, which, therefore, requires stuffing at every opportunity.

The Far End Alarm and Control Channel (FEAC) carries alarm and status information from one DS3 terminal to another and may be used as a channel to initiate DS1 and DS3 maintenance loopbacks at the distant DS3 terminal.

The path parity (CP) bits are set to the same value as are the P-bits at the terminal that generates them. The CP-bits are not to be changed by intermediate network elements and, therefore, provide a more reliable end-to-end parity indication than do the DS3 P-bits.

TABLE 27.7 Assignment of C Bits in C-bit Parity Application

C bit	Application	C bit	Application
1	C-bit parity identifier	12	FEBE
2	N = 1 (future network use)	13	DL (data link)
3	FEAC	14	DL (data link)
4	application specific	15	DL (data link)
5	application specific	16	application specific
6	application specific	17	application specific
7	CP (path parity)	18	application specific
8	CP (path parity)	19	application specific
9	CP (path parity)	20	application specific
10	FEBE	21	application specific
11	FEBE		

The Far-End Block Error (FEBE) bits are set to all ones (111) only when no framing bit or CP-bit error has been detected in the incoming signal by the terminal generating the outgoing FEBE. When errors are detected, the FEBE-bits are set to any combination of 1s and 0s except 111.

The data link (DL) bits are used as a 28.2 kb/s data channel between the two DS3 terminals. Messages carried by this channel use the LAPD format. When no messages are carried, LAPD idle code (flags), which consists of repetitions of 01111110, is sent. The DL is used to carry messages identifying the DS3 path, the source of a DS3 idle signal, or the source of a DS3 test signal.

A network performance report message (NPRM) for DS3 similar to that defined for DS1 in T1.403 [ANSI, 1999] is under development by Working Group T1E1.2 of Committee T1 at the time of this writing. The NPRM will be transported by the C-bit parity data link.

Unchannelized Digital Signal Level 3

T1.107 also provides for the use of the DS3 payload for direct transport of data at the payload rate of 44.210 Mb/s.

Defining Terms

Alarm Indication Signal (AIS): A signal that is transmitted in the direction of a failure to indicate that a network element has detected the failure. AIS provides a "keep alive" signal to equipment downstream from the failure and prevents multiple network elements from issuing redundant and confusing alarms about the same failure.

Alliance for Telecommunications Industry Solutions (ATIS): The body that serves as the secretariat for Committee T1—Telecommunications.

Committee T1—Telecommunications: An accredited standards development body that develops and maintains standards for telecommunications in North America. Note that the "T1" in the committee name above has no connection with the T1 line that operates at the DS1 rate.

Cyclic redundancy code (CRC) with an n-bit remainder (CRC-n): CRC codes provide highly reliable error checking of blocks of transmitted information.

DL: Data link for transporting messages across a digital path using certain of the overhead bits as a data channel. The DL is sometimes called the FDL, for facility data link.

DSn: DSn stands for digital signal level n and refers to one of the levels (rates) in the North American digital hierarchy that are discussed in this chapter.

Extended superframe (ESF): A DS1 superframe that is 24 frames in length and that makes more efficient use of the overhead bits than the older SF format.

Frame alignment signal (FAS): Serves to allow a receiver to locate significant repetitive points within the received bit stream so that the information may be extracted.

FEAC: Far end alarm and control channel used in C-bit parity DS3. The FEAC uses repeated 16-bit code words to send status messages alarm signals and requests for loopback.

High level data link control (HDLC): A format used to implement layer 2 (the data link layer) of the ISO seven-layer model.

Link access procedure—D channel (LAPD): A subset of HDLC used for transporting messages over the data links (DL) in the North American digital hierarchy.

Network performance report message (NPRM): An HDLC frame that carries DS1 performance information along the extended superframe data link. The NPRM is intended to allow transmission of performance information from points intermediate to the DS1 path terminations in an ESF DL that is already carrying performance information from the path terminations in PRMs.

Overhead: Bits in a digital signal that do not carry the information signals the signal is intended to transport but that perform housekeeping functions such as framing, error detection, and the transport of maintenance data from one digital terminal to the other.

Payload: The aggregate of the information bits the digital signal is intended to transport.

Performance report message (PRM): An HDLC frame that is intended to carry DS1 performance information along the extended superframe data link.

Remote alarm indication (RAI): A Signal that indicates to the terminal at one end of a digital path that the terminal at the other end has detected a failure in the incoming signal.

SF—DS1 superframe format: A format in which the superframe is twelve frames in length and in which all the overhead bits are used for framing.

SPRM: A PRM that has been modified to use spare bits in the HDLC information field to carry performance information from a point intermediate to the DS1 path terminations.

Superframe: The 12-frame superframe format for DS1 known as SF or an aggregation of frames that provides a longer repetitive structure than a frame in any format. The latter definition of the term is more often called a multiframe at rates other than DS1.

References

ANSI T1.107-1995, American National Standard for Telecommunications—Digital Hierarchy—Formats Specifications.

ANSI T1.403-1999, American National Standard for Telecommunications—Network-to-Customer Installation—DS1 Metallic Interface.

ANSI T1.102-1993, American National Standard for Telecommunications—Digital Hierarchy—Electrical Interfaces.

Further Information

For further information on the North American digital hierarchy, see B. P. Lathi, *Modern Digital and Analog Communication Systems*, 3rd ed., Oxford University Press, 1998. Information supplementing the standards may be found in various technical requirements of Telcordia Technologies, Inc. including GR-499 and TR-TSY-000009.

28

Line Coding

Joseph L. LoCicero
Illinois Institute of Technology

Bhasker P. Patel
Illinois Institute of Technology

28.1 Introduction

The terminology **line coding** originated in telephony with the need to transmit digital information across a copper telephone *line*; more specifically, binary data over a digital repeatered line. The concept of line coding, however, readily applies to any transmission line or channel. In a digital communication system, there exists a known set of symbols to be transmitted. These can be designated as $\{m_i\}$, $i = 1, 2, ., N$, with a probability of occurrence $\{p_i\}$, $i = 1, 2, ., N$, where the sequentially transmitted symbols are generally assumed to be statistically independent. The conversion or *coding* of these abstract symbols into real, temporal waveforms to be transmitted in baseband is the process of line coding. Since the most common type of line coding is for binary data, such a waveform can be succinctly termed a direct format for serial bits. The concentration in this section will be line coding for binary data.

Different channel characteristics, as well as different applications and performance requirements, have provided the impetus for the development and study of various types of line coding [1,2]. For example, the channel might be AC coupled and, thus, could not support a line code with a DC component or large DC content. Synchronization or timing recovery requirements might necessitate a discrete component at the data rate. The channel bandwidth and **crosstalk** limitations might dictate the type of line coding employed. Even such factors as the complexity of the encoder and the economy of the decoder could determine the line code chosen. Each line code has its own distinct properties. Depending on the application, one property may be more important than the other. In what follows, we describe, in general, the most desirable features that are considered when choosing a line code.

It is commonly accepted [1,2,5,8] that the dominant considerations effecting the choice of a line code are: (1) timing, (2) DC content, (3) power spectrum, (4) performance monitoring, (5) probability of error, and (6) transparency. Each of these are detailed in the following paragraphs.

(1) *Timing:* The waveform produced by a line code should contain enough timing information such that the receiver can synchronize with the transmitter and decode the received signal properly. The timing content should be relatively independent of source statistics, i.e., a long string of **1**s or **0**s should not result in loss of timing or jitter at the receiver.

(2) *DC content:* Since the repeaters used in telephony are AC coupled, it is desirable to have zero DC in the waveform produced by a given line code. If a signal with significant DC content is used in AC coupled lines, it will cause **DC wander** in the received waveform. That is, the received signal baseline will vary with time. Telephone lines do not pass DC due to AC coupling with transformers and capacitors to eliminate DC ground loops. Because of this, the telephone channel causes a droop in constant signals. This causes DC wander. It can be eliminated by DC restoration circuits, feedback systems, or with specially designed line codes.

(3) *Power spectrum:* The power spectrum and bandwidth of the transmitted signal should be matched to the frequency response of the channel to avoid significant distortion. Also, the power spectrum should be such that most of the energy is contained in as small bandwidth as possible. The smaller the bandwidth, the higher the transmission efficiency.

(4) *Performance monitoring:* It is very desirable to detect errors caused by a noisy transmission channel. The error detection capability in turn allows performance monitoring while the channel is in use (i.e., without elaborate testing procedures that require suspending use of the channel).

(5) *Probability of error:* The average error probability should be as small as possible for a given transmitter power. This reflects the reliability of the line code.

(6) *Transparency:* A line code should allow all the possible patterns of **1**s and **0**s. If a certain pattern is undesirable due to other considerations, it should be mapped to a unique alternative pattern.

28.2 Common Line Coding Formats

A line coding format consists of a formal definition of the line code that specifies how a string of binary digits are converted to a line code waveform. There are two major classes of binary line codes: **level codes** and **transition codes**. Level codes carry information in their voltage level, which may be high or low for a full bit period or part of the bit period. Level codes are usually instantaneous since they typically encode a binary digit into a distinct waveform, independent of any past binary data. However, some level codes do exhibit memory. Transition codes carry information in the change in level appearing in the line code waveform. Transition codes may be instantaneous, but they generally have memory, using past binary data to dictate the present waveform. There are two common forms of level line codes: one is called **return to zero (RZ)** and the other is called **nonreturn to zero (NRZ)**. In RZ coding, the level of the pulse returns to zero for a portion of the bit interval. In NRZ coding, the level of the pulse is maintained during the entire bit interval.

Line coding formats are further classified according to the polarity of the voltage levels used to represent the data. If only one polarity of voltage level is used, i.e., positive or negative (in addition to the zero level), then it is called **unipolar** signalling. If both positive and negative voltage levels are being used, with or without a zero voltage level, then it is called **polar** signalling. The term **bipolar** signalling is used by some authors to designate a specific line coding scheme with positive, negative, and zero voltage levels. This will be described in detail later in this section. The formal definition of five common line codes is given in the following along with a representative waveform, the *power spectral density* (PSD), the probability of error, and a discussion of advantages and disadvantages. In some cases specific applications are noted.

Unipolar NRZ (Binary On-Off Keying)

In this line code, a binary **1** is represented by a non-zero voltage level and a binary **0** is represented by a zero voltage level as shown in Fig. 28.1(a). This is an instantaneous level code. The PSD of this code with equally likely **1**s and **0**s is given by [5,8]

$$S_1(f) = \frac{V^2 T}{4}\left(\frac{\sin \pi f T}{\pi f T}\right)^2 + \frac{V^2}{4}\delta(f) \tag{28.1}$$

where V is the binary **1** voltage level, $T = 1/R$ is the bit duration, and R is the bit rate in bits per second. The spectrum of unipolar NRZ is plotted in Fig. 28.2a. This PSD is a two-sided even spectrum, although only half of the plot is shown for efficiency of presentation. If the probability of a binary **1** is p, and the probability of a binary **0** is $(1 - p)$, then the PSD of this code, in the most general case, is $4p(1 - p)S_1(f)$. Considering the frequency of the first spectral null as the bandwidth of the waveform, the bandwidth of unipolar NRZ is R in hertz. The error rate performance of this code, for equally likely data, with additive white Gaussian noise (AWGN) and optimum, i.e., matched filter, detection is given by [1,5]

$$P_e = \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{E_b}{2N_0}}\right) \tag{28.2}$$

where E_b/N_0 is a measure of the signal-to-noise ratio (SNR) of the received signal. In general, E_b is the energy per bit and $N_0/2$ is the two-sided PSD of the AWGN. More specifically, for unipolar NRZ, E_b is the energy in a binary **1**, which is V^2T. The performance of the unipolar NRZ code is plotted in Fig. 28.3.

The principle advantages of unipolar NRZ are ease of generation, since it requires only a single power supply, and a relatively low bandwidth of R Hz. There are quite a few disadvantages of this line code. A loss of synchronization and timing jitter can result with a long sequence of **1**s or **0**s because no pulse transition is present. The code has no error detection capability and, hence, performance cannot be monitored. There is a significant DC component as well as a DC content. The error rate performance is not as good as that of polar line codes.

Unipolar RZ

In this line code, a binary **1** is represented by a nonzero voltage level during a portion of the bit duration, usually for half of the bit period, and a zero voltage level for rest of the bit duration. A binary **0** is represented by a zero voltage level during the entire bit duration. Thus, this is an instantaneous level code. Figure 28.1(b) illustrates a unipolar RZ waveform in which the **1** is represented by a nonzero voltage level for half the bit period. The PSD of this line code, with equally likely binary digits, is given by [5,6,8]

$$S_2(f) = \frac{V^2 T}{16}\left(\frac{\sin \pi f T/2}{\pi f T/2}\right)^2 + \frac{V^2}{4\pi^2}\left[\frac{\pi^2}{4}\delta(f) + \sum_{n=-\infty}^{\infty}\frac{1}{(2n+1)^2}\delta(f-(2n+1)R)\right] \tag{28.3}$$

where again V is the binary **1** voltage level, and $T = 1/R$ is the bit period. The spectrum of this code is drawn in Fig. 28.2a. In the most general case, when the probability of a **1** is p, the continuous portion of the PSD in Eq. (28.3) is scaled by the factor $4p(1 - p)$ and the discrete portion is scaled by the factor $4p^2$. The first null bandwidth of unipolar RZ is $2R$ Hz. The error rate performance of this line code is the same as that of the unipolar NRZ, provided we increase the voltage level of this code such that the energy in binary **1**, E_b, is the same for both codes. The probability of error is given by Eq. (28.2) and identified in Fig. 28.3. If the voltage level and bit period are the same for unipolar NRZ and unipolar RZ, then the energy in a binary **1** for unipolar RZ will be $V^2T/2$ and the probability of error is worse by 3 dB.

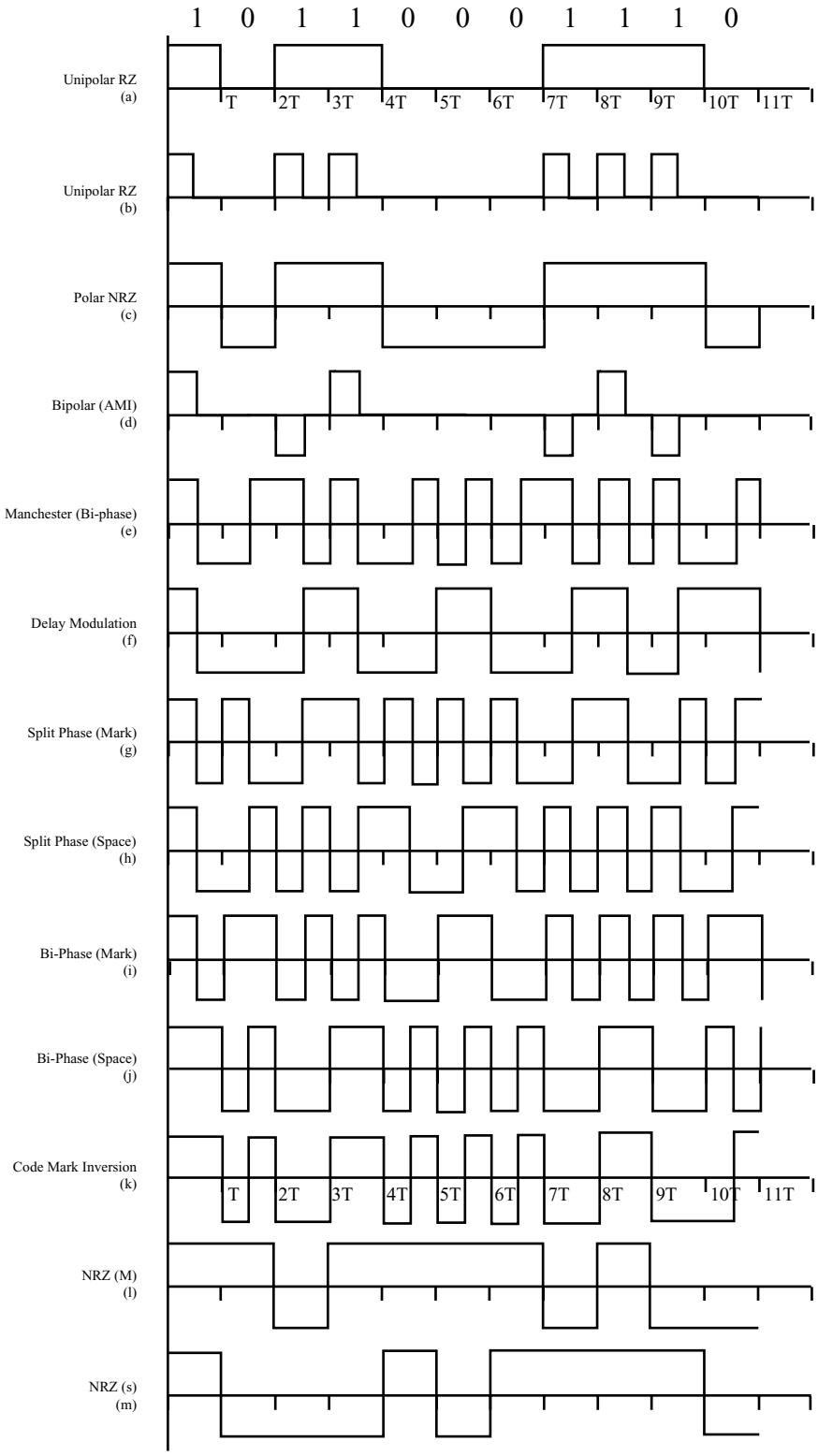


FIGURE 28.1 Waveforms for different line codes.

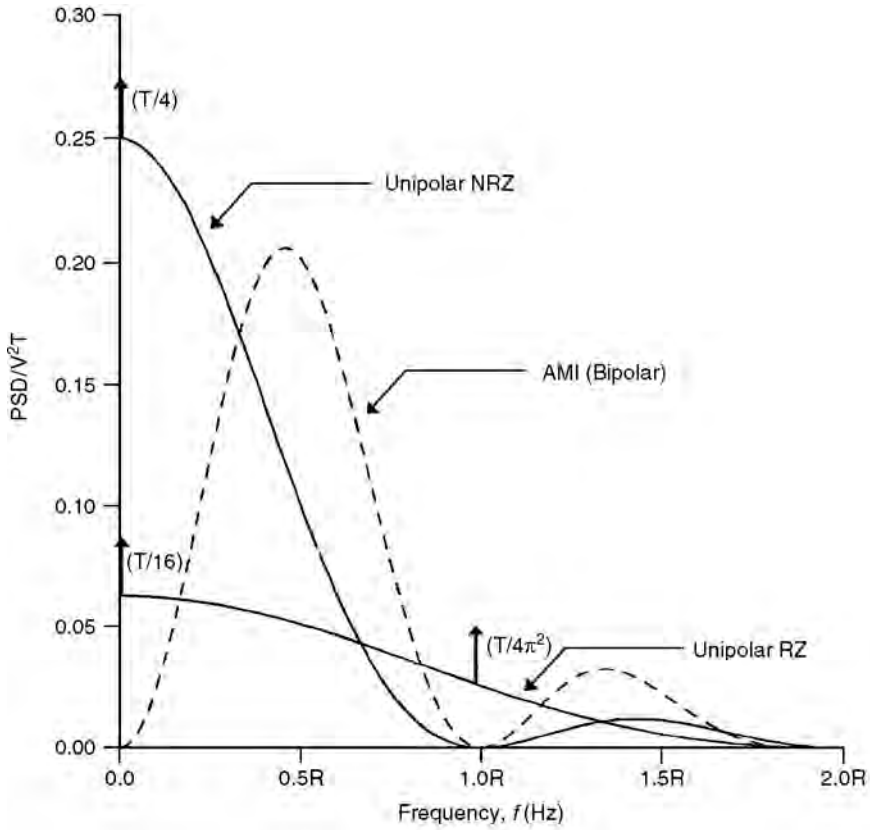


FIGURE 28.2a Power spectral density of different line codes, where $R = 1/T$ is the bit rate.

The main advantages of unipolar RZ are, again, ease of generation since it requires a single power supply and the presence of a discrete spectral component at the symbol rate, which allows simple timing recovery. A number of disadvantages exist for this line code. It has a nonzero DC component and nonzero DC content, which can lead to DC wander. A long string of **0**s will lack pulse transitions and could lead to loss of synchronization. There is no error detection capability and, hence, performance monitoring is not possible. The bandwidth requirement ($2R$ Hz) is higher than that of NRZ signals. The error rate performance is worse than that of polar line codes.

Unipolar NRZ as well as unipolar RZ are examples of pulse/no-pulse type of signalling. In this type of signalling, the pulse for a binary **0**, $g_2(t)$, is zero and the pulse for a binary **1** is specified generically as $g_1(t) = g(t)$. Using $G(f)$ as the Fourier transform of $g(t)$, the PSD of pulse/no-pulse signalling is given as [6,7,10]

$$S_{\mathrm{PNP}}(f) = p(1-p)R|G(f)|^2 + p^2R^2\sum_{n=-\infty}^{\infty}|G(nR)|^2\delta(f-nR) \tag{28.4}$$

where p is the probability of a binary **1**, and R is the bit rate.

Polar NRZ

In this line code, a binary **1** is represented by a positive voltage $+V$ and a binary **0** is represented by a negative voltage $-V$ over the full bit period. This code is also referred to as NRZ (L), since a bit is represented by maintaining a level (L) during its entire period. A polar NRZ waveform is shown in Fig. 28.1(c). This is again an instantaneous level code. Alternatively, a **1** may be represented by a $-V$

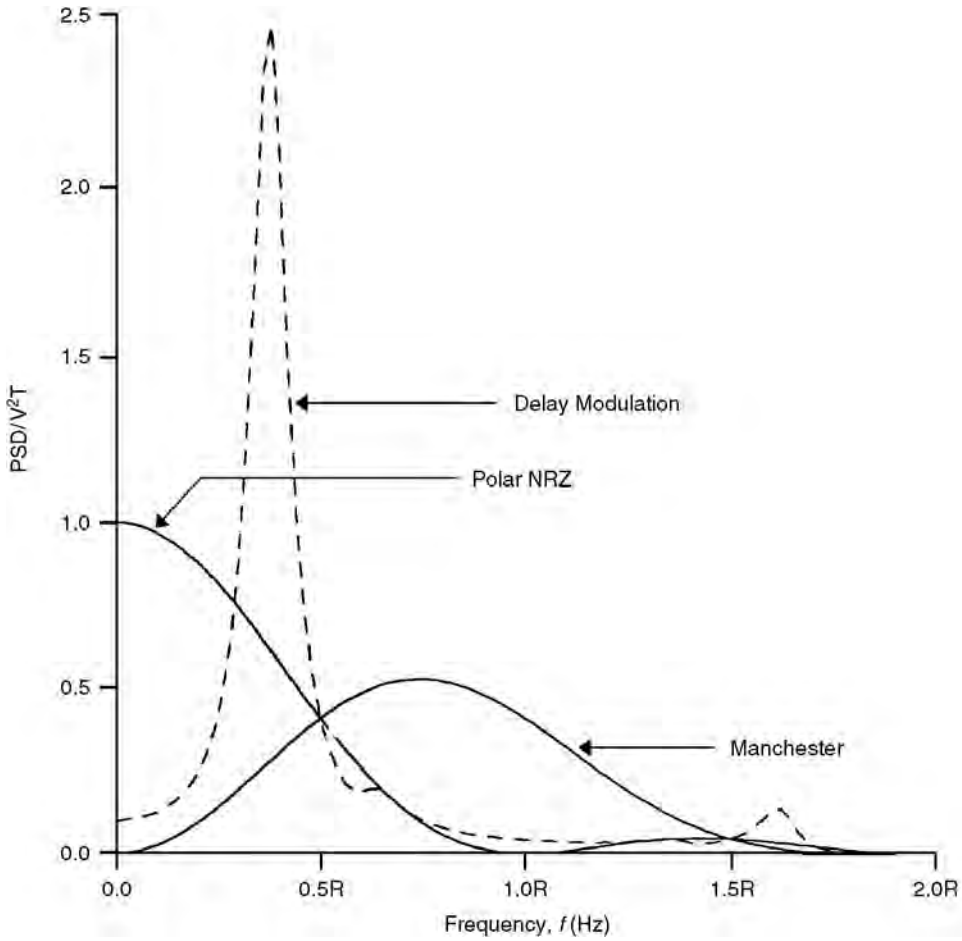


FIGURE 28.2b Power spectral density of different line codes, where $R = 1/T$ is the bit rate.

voltage level and a **0** by a $+V$ voltage level, without changing the spectral characteristics and performance of the line code. The PSD of this line code with equally likely bits is given by [5,8]

$$S_3(f) = V^2 T\left(\frac{\sin \pi f T}{\pi f T}\right)^2 \tag{28.5}$$

This is plotted in Fig. 28.2b. When the probability of a **1** is p, and p is not 0.5, a DC component exists, and the PSD becomes [10]

$$S_{3p}(f) = 4V^2 Tp(1-p)\left(\frac{\sin \pi f T}{\pi f T}\right)^2 + V^2(1-2p)^2\,\delta(f) \tag{28.6}$$

The first null bandwidth for this line code is again R Hz, independent of p. The probability of error of this line code when $p = 0.5$ is given by [1,5]

$$P_e = \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \tag{28.7}$$

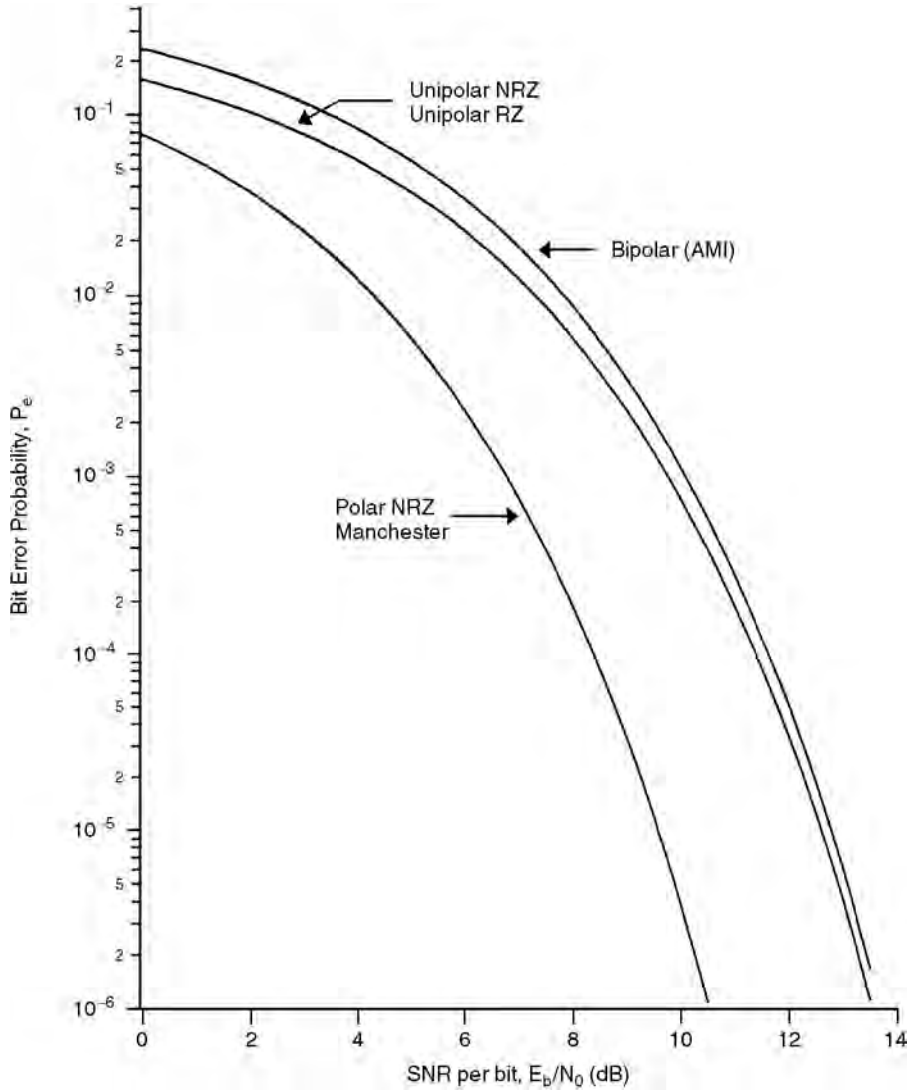


FIGURE 28.3 Bit error probability for different line codes.

The performance of polar NRZ is plotted in Fig. 28.3. This is better than the error performance of the unipolar codes by 3 dB.

The advantages of polar NRZ include a low-bandwidth requirement, R Hz, comparable to unipolar NRZ, very good error probability, and greatly reduced DC because the waveform has a zero DC component when $p = 0.5$ even though the DC content is never zero. A few notable disadvantages are that there is no error detection capability, and that a long string of **1**s or **0**s could result in loss of synchronization, since there are no transitions during the string duration. Two power supplies are required to generate this code.

Polar RZ [Bipolar, Alternate Mark Inversion (AMI), or Pseudoternary]

In this scheme, a binary **1** is represented by alternating the positive and negative voltage levels, which return to zero for a portion of the bit duration, generally half the bit period. A binary **0** is represented by a zero voltage level during the entire bit duration. This line coding scheme is often called **alternate mark inversion** (**AMI**) since **1**s (marks) are represented by alternating positive and negative pulses. It is

also called *pseudoternary* since three different voltage levels are used to represent binary data. Some authors designate this line code as bipolar RZ (BRZ). An AMI waveform is shown in Fig. 28.1(d). Note that this is a level code with memory. The AMI code is well known for its use in telephony. The PSD of this line code with memory is given by [1,2,7]

$$S_{4p}(f) = 2p(1-p)R|G(f)|^2\left(\frac{1-\cos 2\pi f T}{1+(2p-1)^2+2(2p-1)\cos 2\pi f T}\right) \tag{28.8}$$

where $G(f)$ is the Fourier transform of the pulse used to represent a binary **1**, and p is the probability of a binary **1**. When $p = 0.5$ and square pulses with amplitude $\pm V$ and duration $T/2$ are used to represent binary **1**s, the PSD becomes

$$S_4(f) = \frac{V^2 T}{4}\left(\frac{\sin \pi f T/2}{\pi f T/2}\right)^2 \sin^2(\pi f T) \tag{28.9}$$

This PSD is plotted in Fig. 28.2a. The first null bandwidth of this waveform is R Hz. This is true for RZ rectangular pulses, independent of the value of p in Eq. (28.8). The error rate performance of this line code for equally likely binary data is given by [5]

$$P_e \approx \frac{3}{4}\,\mathrm{erfc}\left(\sqrt{\frac{E_b}{2N_0}}\right), \quad E_b/N_0 > 2 \tag{28.10}$$

This curve is plotted in Fig. 28.3 and is seen to be no more than 0.5 dB worse than the unipolar codes.

The advantages of polar RZ (or AMI, as it is most commonly called) outweigh the disadvantages. This code has no DC component and zero DC content, completely avoiding the DC wander problem. Timing recovery is rather easy since squaring, or full-wave rectifying, this type of signal yields a unipolar RZ waveform with a discrete component at the bit rate, R Hz. Because of the alternating polarity pulses for binary **1**s, this code has error detection and, hence, performance monitoring capability. It has a low-bandwidth requirement, R Hz, comparable to unipolar NRZ. The obvious disadvantage is that the error rate performance is worse than that of the unipolar and polar waveforms. A long string of **0**s could result in loss of synchronization, and two power supplies are required for this code.

Manchester Coding (Split Phase or Digital Biphase)

In this coding, a binary **1** is represented by a pulse that has positive voltage during the first-half of the bit duration and negative voltage during second-half of the bit duration. A binary **0** is represented by a pulse that is negative during the first-half of the bit duration and positive during the second-half of the bit duration. The negative or positive midbit transition indicates a binary **1** or binary **0**, respectively. Thus, a Manchester code is classified as an instantaneous transition code; it has no memory. The code is also called diphase because a square wave with a 0∞ phase is used to represent a binary **1** and a square wave with a phase of 180∞ used to represent a binary **0**; or vice versa. This line code is used in Ethernet local area networks (LANs). The waveform for Manchester coding is shown in Fig. 28.1(e). The PSD of a Manchester waveform with equally likely bits is given by [5,8]

$$S_5(f) = V^2 T\left(\frac{\sin \pi f T/2}{\pi f T/2}\right)^2 \sin^2(\pi f T/2) \tag{28.11}$$

where $\pm V$ are used as the positive/negative voltage levels for this code. Its spectrum is plotted in Fig. 28.2b. When the probability p of a binary **1** is not equal to one-half, the continuous portion of the PSD is reduced in amplitude and discrete components appear at integer multiples of the bit rate, $R = 1/T$. The resulting

PSD is [6,10]

$$S_{5p}(f) = V^2 T 4p(1-p)\left(\frac{\sin \pi f T/2}{\pi f T/2}\right)^2 \sin^2 \frac{\pi f T}{2} + V^2(1-2p)^2 \sum_{n=-\infty, n\pi 0}^{\infty} \left(\frac{2}{n\pi}\right)^2 \delta(f - nR \qquad (28.12)$$

The first null bandwidth of the waveform generated by a Manchester code is $2R$ Hz. The error rate performance of this waveform when $p = 0.5$ is the same as that of polar NRZ, given by Eq. (28.9), and plotted in Fig. 28.3.

The advantages of this code include a zero DC content on an individual pulse basis, so no pattern of bits can cause DC buildup; midbit transitions are always present making it is easy to extract timing information; and it has good error rate performance, identical to polar NRZ. The main disadvantage of this code is a larger bandwidth than any of the other common codes. Also, it has no error detection capability and, hence, performance monitoring is not possible.

Polar NRZ and Manchester coding are examples of the use of pure polar signalling where the pulse for a binary **0**, $g_2(t)$, is the negative of the pulse for a binary **1**, i.e., $g_2(t) = -g_1(t)$. This is also referred to as an antipodal signal set. For this broad type of polar binary line code, the PSD is given by [10]

$$S_{BP}(f) = 4p(1-p)R|G(f)|^2 + (2p-1)^2 R^2 \sum_{n=-\infty}^{\infty} |G(nR)|^2 \delta(f - nR) \qquad (28.13)$$

where $|G(f)|$ is the magnitude of the Fourier transform of either $g_1(t)$ or $g_2(t)$.

A further generalization of the PSD of binary line codes can be given, wherein a continuous spectrum and a discrete spectrum is evident. Let a binary **1**, with probability p, be represented by $g_1(t)$ over the $T = 1/R$ second bit interval; and let a binary **0**, with probability $1 - p$, be represented by $g_2(t)$ over the same T second bit interval. The two-sided PSD for this general binary line code is [10]

$$S_{GB}(f) = p(1-p)R|G_1(f) - G_2(f)|^2 + R^2 \sum_{n=-\infty}^{\infty} |pG_1(nR) + (1-p)G_2(nR)|^2 \delta(f - nR) \qquad (28.14)$$

where the Fourier transform of $g_1(t)$ and $g_2(t)$ are given by $G_1(f)$ and $G_2(f)$, respectively.

28.3 Alternate Line Codes

Most of the line codes discussed thus far were instantaneous level codes. Only AMI had memory, and Manchester was an instantaneous transition code. The alternate line codes presented in this section all have memory. The first four are transition codes, where binary data is represented as the presence or absence of a transition, or by the direction of transition, i.e., positive to negative or vice versa. The last four codes described in this section are level line codes with memory.

Delay Modulation (Miller Code)

In this line code, a binary **1** is represented by a transition at the midbit position, and a binary **0** is represented by no transition at the midbit position. If a **0** is followed by another **0**, however, the signal transition also occurs at the end of the bit interval, that is, between the two **0**s. An example of delay modulation is shown in Fig. 28.1(f). It is clear that delay modulation is a transition code with memory.

This code achieves the goal of providing good timing content without sacrificing bandwidth. The PSD of the Miller code for equally likely data is given by [10]

$$S_6(f) = \frac{V^2 T}{2(\pi f T)^2(17 + 8\cos 2\pi f T)}(23 - 2\cos \pi f T - 22\cos 2\pi f T - 12\cos 3\pi f T + 5\cos 4\pi f T + 12\cos 5\pi f T + 2\cos 6\pi f T - 8\cos 7\pi f T + 2\cos 8\pi f T) \tag{28.15}$$

This spectrum is plotted in Fig. 28.2b. The advantages of this code are that it requires relatively low bandwidth and most of the energy is contained in less than 0.5*R*. However, there is no distinct spectral null within the 2*R*-Hz band. It has low DC content and no DC component. It has very good timing content, and carrier tracking is easier than Manchester coding. Error rate performance is comparable to that of the common line codes. One important disadvantage is that it has no error detection capability and, hence, performance cannot be monitored.

Split Phase (Mark)

This code is similar to Manchester in the sense that there are always midbit transitions. Hence, this code is relatively easy to synchronize and has no DC. Unlike Manchester, however, split phase (mark) encodes a binary digit into a midbit transition dependent on the midbit transition in the previous bit period [12]. Specifically, a binary **1** produces a reversal of midbit transition relative to the previous midbit transition. A binary **0** produces no reversal of the midbit transition. Certainly this is a transition code with memory. An example of a split phase (mark) coded waveform is shown in Fig. 28.1(g), where the waveform in the first bit period is chosen arbitrarily. Since this method encodes bits differentially, there is no 180º -phase ambiguity associated with some line codes. This phase ambiguity may not be an issue in most baseband links but is important if the line code is modulated. Split phase (space) is very similar to split phase (mark), where the role of the binary **1** and binary **0** are interchanged. An example of a split phase (space) coded waveform is given in Fig. 28.1(h); again, the first bit waveform is arbitrary.

Biphase (Mark)

This code, designated as Bi ϕ-M, is similar to a Miller code in that a binary **1** is represented by a midbit transition, and a binary **0** has no midbit transition. However, this code always has a transition at the beginning of a bit period [10]. Thus, the code is easy to synchronize and has no DC. An example of Bi ϕ-M is given in Fig. 28.1 (i), where the direction of the transition at $t = 0$ is arbitrarily chosen. Biphase (space) or Bi ϕ-S is similar to Bi ϕ-M, except the role of the binary data is reversed. Here a binary **0** (space) produces a midbit transition, and a binary **1** does not have a midbit transition. A waveform example of Bi ϕ-S is shown in Fig. 28.1(j). Both Bi ϕ-S and Bi ϕ-M are transition codes with memory.

Code Mark Inversion (CMI)

This line code is used as the interface to a Consultative Committee on International Telegraphy and Telephony (CCITT) multiplexer and is very similar to Bi ϕ-S. A binary **1** is encoded as an NRZ pulse with alternate polarity, +*V* or -*V*. A binary **0** is encoded with a definitive midbit transition (or square wave phase) [1]. An example of this waveform is shown in Fig. 28.1(k) where a negative to positive transition (or 180∞ phase) is used for a binary **0**. The voltage level of the first binary **1** in this example is chosen arbitrarily. This example waveform is identical to Bi ϕ-S shown in Fig. 28.1(j), except for the last bit. CMI has good synchronization properties and has no DC.

NRZ (I)

This type of line code uses an inversion (I) to designate binary digits, specifically, a change in level or no change in level. There are two variants of this code, NRZ mark (M) and NRZ space (S) [5,12]. In NRZ (M), a change of level is used to indicate a binary **1**, and no change of level is used to indicate a binary **0**. In NRZ (S) a change of level is used to indicate a binary **0**, and no change of level is used to indicate a binary **1**. Waveforms for NRZ (M) and NRZ (S) are depicted in Fig. 28.1(l) and Fig. 28.1(m), respectively, where the voltage level of the first binary **1** in the example is chosen arbitrarily. These codes are level codes with memory. In general, line codes that use differential encoding, like NRZ (I), are insensitive to 180∞ phase ambiguity. Clock recovery with NRZ (I) is not particularly good, and dc wander is a problem as well. Its bandwidth is comparable to polar NRZ.

Binary *N* Zero Substitution (BNZS)

The common bipolar code AMI has many desirable properties of a line code. Its major limitation, however, is that a long string of zeros can lead to loss of synchronization and timing jitter because there are no pulses in the waveform for relatively long periods of time. **Binary *N* zero substitution (BNZS)** attempts to improve AMI by substituting a special code of length *N* for all strings of *N* zeros. This special code contains pulses that look like binary **1**s but purposely produce violations of the AMI pulse convention. Two consecutive pulses of the same polarity violate the AMI pulse convention, independent of the number of zeros between the two consecutive pulses. These violations can be detected at the receiver and the special code replaced by *N* zeros. The special code contains pulses facilitating synchronization even when the original data has a long string of zeros. The special code is chosen such that the desirable properties of AMI coding are retained despite the AMI pulse convention violations, i.e., DC balance and error detection capability. The only disadvantage of BNZS compared to AMI is a slight increase in crosstalk due to the increased number of pulses and, hence, an increase in the average energy in the code.

Choosing different values of *N* yields different BNZS codes. The value of *N* is chosen to meet the timing requirements of the application. In telephony, there are three commonly used BNZS codes: B6ZS, B3ZS, and B8ZS. All BNZS codes are level codes with memory.

In a B6ZS code, a string of six consecutive zeros is replaced by one of two the special codes according to the rule:

If the last pulse was positive (+), the special code is: 0 + - 0 - +.
If the last pulse was negative (-), the special code is: 0 - + 0 + -.

Here a zero indicates a zero voltage level for the bit period; a plus designates a positive pulse, and a minus indicates a negative pulse.

This special code causes two AMI pulse violations: in its second bit position and in its fifth bit position. These violations are easily detected at the receiver and zeros resubstituted. If the number of consecutive zeros is 12, 18, 24,…, the substitution is repeated 2, 3, 4,… times. Since the number of violations is even, the B6ZS waveform is the same as the AMI waveform outside the special code, i.e., between special code sequences.

There are four pulses introduced by the special code that facilitates timing recovery. Also, note that the special code is DC balanced. An example of the B6ZS code is given as follows, where the special code is indicated by the bold characters.

Original data:	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	1	1
B6ZS format:	0	+	**0**	+	-	**0**	-	+	-	+	0	-	**0**	-	+	**0**	+	-	+	-

The computation of the PSD of a B6ZS code is tedious. Its shape is given in Fig. 28.4, for comparison purposes with AMI, for the case of equally likely data.

TABLE 28.1 B3ZS Substitution Rules

Number of *B* Pulses Since Last Violation	Polarity of Last *B* Pulse	Substitution Code	Substitution Code Form
Odd	Negative (-)	**00−**	**00***V*
Odd	Positive (+)	**00+**	**00***V*
Even	Negative (-)	**+0+**	*B***0***V*
Even	Positive (+)	**−0−**	*B***0***V*

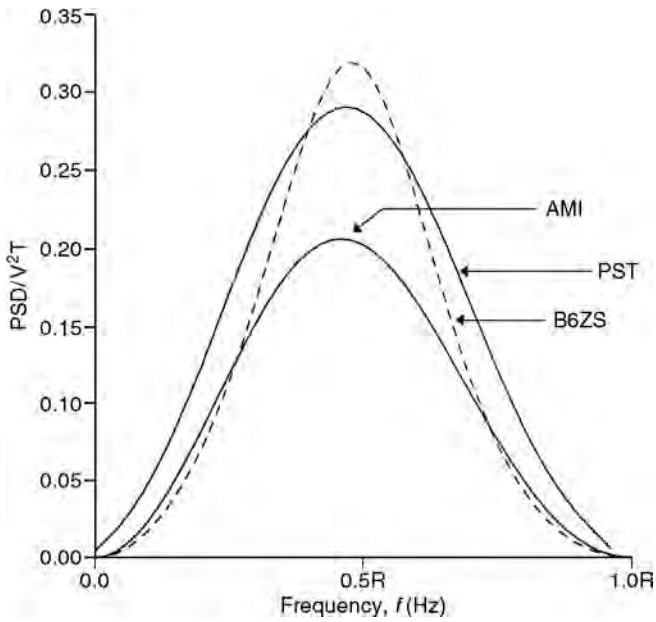


FIGURE 28.4 Power spectral density of different line codes, where $R = 1/T$ is the bit rate.

In a B3ZS code, a string of three consecutive zeros is replaced by either *B***0***V* or **00V**, where *B* denotes a pulse obeying the AMI (bipolar) convention and *V* denotes a pulse violating the AMI convention. *B***0***V* or **00***V* is chosen such that the number of bipolar (*B*) pulses between the violations is odd. The B3ZS rules are summarized in Table 28.1.

Observe that the violation always occurs in the third bit position of the substitution code, and so it can be easily detected and zero replacement made at the receiver. Also, the substitution code selection maintains DC balance. There is either one or two pulses in the substitution code, facilitating synchronization. The error detection capability of AMI is retained in B3ZS because a single channel error would make the number of bipolar pulses between violations even instead of being odd. Unlike B6ZS, the B3ZS waveform between violations may not be the same as the AMI waveform. B3ZS is used in the digital signal-3 (DS-3) signal interface in North America and also in the long distance-4 (LD-4) coaxial transmission system in Canada. Next is an example of a B3ZS code, using the same symbol meaning as in the B6ZS code.

Original data:	1	0	0	1	0	0	0	1	1	0	0	0	0	1	0	0	0	1
B3ZS format:																		
Even No. of *B* pulses:	+	0	0	-	+	**0**	+	-	+	-	**0**	-	0	+	**0**	**0**	+	-
Odd No. of *B* pulses:	+	0	0	-	**0**	**0**	-	+	-	+	**0**	+	0	-	**0**	**0**	-	+

The last BNZS code considered here uses $N = 8$. A B8ZS code is used to provide transparent channels for the Integrated Services Digital Network (ISDN) on T1 lines and is similar to the B6ZS code. Here a

TABLE 28.2 HDB3 Substitution Rules

Number of *B* Pulses Since Last Violation	Polarity of Last *B* Pulse	Substitution Code	Substitution Code Form
Odd	Negative (-)	**000−**	**000***V*
Odd	Positive (+)	**000+**	**000***V*
Even	Negative (-)	**+00+**	*B***00***V*
Even	Positive (+)	**−00−**	*B***00***V*

string of eight consecutive zeros is replaced by one of two special codes according to the following rule:

If the last pulse was positive (+), the special code is: 0 0 0 + - 0 - +.
If the last pulse was negative (-), the special code is: 0 0 0 - + 0 + -.

There are two bipolar violations in the special codes, at the fourth and seventh bit positions. The code is DC balanced, and the error detection capability of AMI is retained. The waveform between substitutions is the same as that of AMI. If the number of consecutive zeros is 16, 24,..., then the substitution is repeated 2, 3,..., times.

High-Density Bipolar *N* (HDBN)

This coding algorithm is a CCITT standard recommended by the Conference of European Posts and Telecommunications Administrations (CEPT), a European standards body. It is quite similar to BNZS coding. It is thus a level code with memory. Whenever there is a string of $N + 1$ consecutive zeros, they are replaced by a special code of length $N + 1$ containing AMI violations. Specific codes can be constructed for different values of N. A specific **high-density bipolar *N* (HDBN)** code, HDB3, is implemented as a CEPT primary digital signal. It is very similar to the B3ZS code. In this code, a string of four consecutive zeros is replaced by either *B***00***V*. *B***00***V* or **000***V*. *B***00***V* or **000*V*** is chosen such that the number of bipolar (*B*) pulses between violations is odd. The HDB3 rules are summarized in Table 28.2.

Here the violation always occurs in the fourth bit position of the substitution code, so that it can be easily detected and zero replacement made at the receiver. Also, the substitution code selection maintains DC balance. There are either one or two pulses in the substitution code facilitating synchronization. The error detection capability of AMI is retained in HDB3 because a single channel error would make the number of bipolar pulses between violations even instead of being odd.

Ternary Coding

Many line coding schemes employ three symbols or levels to represent only one bit of information, like AMI. Theoretically, it should be possible to transmit information more efficiently with three symbols, specifically the maximum efficiency is $\log_2 3 = 1.58$ bits per symbol. Alternatively, the redundancy in the code signal space can be used to provide better error control. Two examples of ternary coding are described next [1,2]: **pair selected ternary (PST)** and **4 binary 3 ternary (4B3T)**. The PST code has many of the desirable properties of line codes, but its transmission efficiency is still 1 bit per symbol. The 4B3T code also has many of the desirable properties of line codes, and it has increased transmission efficiency.

In the PST code, two consecutive bits, termed a binary pair, are grouped together to form a word. These binary pairs are assigned codewords consisting of two ternary symbols, where each ternary symbol can be +, -, or 0, just as in AMI. There are nine possible ternary codewords. Ternary codewords with identical elements, however, are avoided, i.e., ++, --, and 00. The remaining six codewords are transmitted using two modes called + mode and - mode. The modes are switched whenever a codeword with a single pulse is transmitted. The PST code and mode switching rules are summarized in Table 28.3.

PST is designed to maintain DC balance and include a strong timing component. One drawback of this code is that the bits must be framed into pairs. At the receiver, an *out-of-frame* condition is signalled

TABLE 28.3 PST Codeword Assignment and Mode Switching Rules

Binary Pair	Ternary Codewords +Mode	Ternary Codewords -Mode	Mode Switching
11	+ -	+ -	No
10	+ 0	- 0	Yes
01	0 +	0 -	Yes
00	- +	- +	No

TABLE 28.4 Modified PST Codeword Assignment and Mode Switching Rules

Binary Pair	Ternary Codewords + Mode	Ternary Codewords - Mode	Mode Switching
11	+ 0	0 -	Yes
10	0 -	0 -	No
01	- +	- +	No
00	0 +	- 0	Yes

TABLE 28.5 4B3T Codeword Assignment

Binary Words	Ternary Codewords Column 1	Ternary Codewords Column 2	Ternary Codewords Column 3
0000	- - -		+ + +
0001	- - 0		+ + 0
0010	- 0 -		+ 0 +
0011	0 - -		0 + +
0100	- - +		+ + -
0101	- + -		+ - +
0110	+ - -		- + +
0111	- 0 0		+ 0 0
1000	0 - 0		0 + 0
1001	0 0 -		0 0 +
1010		0 + -	
1011		0 - +	
1100		+ 0 -	
1101		- 0 +	
1110		+ - 0	
1111		- + 0	

when unused ternary codewords (++, --, and 00) are detected. The mode switching property of PST provides error detection capability. PST can be classified as a level code with memory.

If the original data for PST coding contains only **1**s or **0**s, an alternating sequence of +- +- … is transmitted. As a result, an out-of-frame condition cannot be detected. This problem can be minimized by using the modified PST code as shown in Table 28.4.

It is tedious to derive the PSD of a PST coded waveform. Again, Fig. 28.4 shows the PSD of the PST code along with the PSD of AMI and B6ZS for comparison purposes, all for equally likely binary data. Observe that PST has more power than AMI and, thus, a larger amount of energy per bit, which translates into slightly increased crosstalk.

In 4B3T coding, words consisting of four binary digits are mapped into three ternary symbols. Four bits imply $2^4 = 16$ possible binary words, whereas three ternary symbols allow $3^3 = 27$ possible ternary codewords. The binary-to-ternary conversion in 4B3T insures DC balance and a strong timing component. The specific codeword assignment is as shown in Table 28.5.

There are three types of codewords in Table 28.5, organized into three columns. The codewords in the first column have negative DC, codewords in the second column have zero DC, and those in the third column have positive DC. The encoder monitors the integer variable

$$I = N_p - N_n \tag{28.16}$$

where N_p is the number of positive pulses transmitted and N_n are the number of negative pulses transmitted. Codewords are chosen according to following rule:

If $I < 0$, choose the ternary codeword from columns 1 and 2.
If $I > 0$, choose the ternary codeword from columns 2 and 3.
If $I = 0$, choose the ternary word from column 2, and from column 1 if the previous $I > 0$ or from column 3 if the previous $I < 0$.

Note that the ternary codeword 000 is not used, but the remaining 26 codewords are used in a complementary manner. For example, the column 1 codeword for 0001 is −−0, whereas the column 3 codeword is ++0. The maximum transmission efficiency for the 4B3T code is 1.33 bits per symbol compared to 1 bit per symbol for the other line codes. The disadvantages of 4B3T are that framing is required and that performance monitoring is complicated. The 4B3T code is used in the T148 span line developed by ITT Telecommunications. This code allows transmission of 48 channels using only 50% more bandwidth than required by T1 lines, instead of 100% more bandwidth.

28.4 Multilevel Signalling, Partial Response Signalling, and Duobinary Coding

Ternary coding, such as 4B3T, is an example of the use of more than two levels to improve the transmission efficiency. To increase the transmission efficiency further, more levels and/or more signal processing is needed. Multilevel signalling allows an improvement in the transmission efficiency at the expense of an increase in the error rate, i.e., more transmitter power will be required to maintain a given probability of error. In partial response signalling, intersymbol interference is deliberately introduced by using pulses that are wider and, hence, require less bandwidth. The controlled amount of interference from each pulse can be removed at the receiver. This improves the transmission efficiency, at the expense of increased complexity. **Duobinary coding**, a special case of partial response signalling, requires only the minimum theoretical bandwidth of $0.5R$ Hz. In what follows, these techniques are discussed in slightly more detail.

Multilevel Signalling

The number of levels that can be used for a line code is not restricted to two or three. Since more levels or symbols allow higher transmission efficiency, multilevel signalling can be considered in bandwidth-limited applications. Specifically, if the signalling rate or baud rate is R_s and the number of levels used is L, the equivalent transmission bit rate R_b is given by

$$R_b = R_s \log_2 [L] \tag{28.17}$$

Alternatively, multilevel signalling can be used to reduce the baud rate, which in turn can reduce crosstalk for the same equivalent bit rate. The penalty, however, is that the SNR must increase to achieve the same error rate. The T1G carrier system of AT&T uses multilevel signalling with $L = 4$ and a baud rate of 3.152 mega-symbols/s to double the capacity of the T1C system from 48 channels to 96 channels. Also, a four

level signalling scheme at 80-kB is used to achieve 160 kb/s as a basic rate in a digital subscriber loop (DSL) for ISDN.

Partial Response Signalling and Duobinary Coding

This class of signalling is also called *correlative* coding because it purposely introduces a controlled or correlated amount of intersymbol interference in each symbol. At the receiver, the known amount of interference is effectively removed from each symbol. The advantage of this signalling is that wider pulses can be used requiring less bandwidth, but the SNR must be increased to realize a given error rate. Also, errors can propagate unless *precoding* is used.

There are many commonly used partial response signalling schemes, often described in terms of the delay operator D, which represents one signalling interval delay. For example, in (1 + D) signalling the current pulse and the previous pulse are added. The T1D system of AT&T uses (1 + D) signalling with precoding, referred to as duobinary signalling, to convert binary (two level) data into ternary (three level) data at the same rate. This requires the minimum theoretical channel bandwidth without the deleterious effects of intersymbol interference and avoids error propagation. Complete details regarding duobinary coding are found in Lender [9] and Schwartz [11]. Some partial response signalling schemes, such as $(1 - D)$, are used to shape the bandwidth rather than control it. Another interesting example of duobinary coding is a $(1 - D^2)$, which can be analyzed as the product $(1 - D)(1 + D)$. It is used by GTE in its modified T carrier system. AT&T also uses $(1 - D^2)$ with four input levels to achieve an equivalent data rate of 1.544 Mb/s in only a 0.5-MHz bandwidth.

28.5 Bandwidth Comparison

We have provided the PSD expressions for most of the commonly used line codes. The actual bandwidth requirement, however, depends on the pulse shape used and the definition of bandwidth itself. There are many ways to define bandwidth, for example, as a percentage of the total power or the sidelobe suppression relative to the main lobe. Using the first null of the PSD of the code as the definition of bandwidth, Table 28.6 provides a useful bandwidth comparison.

The notable omission in Table 28.6 is delay modulation (Miller code). It does not have a first null in the $2R$-Hz band, but most of its power is contained in less than $0.5R$ Hz.

28.6 Concluding Remarks

An in-depth presentation of line coding, particularly applicable to telephony, has been included in this chapter. The most desirable characteristics of line codes were discussed. We introduced five common line codes and eight alternate line codes. Each line code was illustrated by an example waveform. In most cases expressions for the PSD and the probability of error were given and plotted. Advantages and disadvantages of all codes were included in the discussion, and some specific applications were noted. Line codes for optical fiber channels and networks built around them, such as fiber distributed data interface (FDDI), were not included in this section. A discussion of line codes for optical fiber channels and other new developments in this topic area can be found in [1,3,4].

TABLE 28.6 First Null Bandwidth Comparison

Bandwidth	Codes	
	Unipolar NRZ	BNZS
R	Polar NRZ	HDBN
	Polar RZ (AMI)	PST
$2R$	Unipolar RZ	Split Phase
	Manchester	CMI

Defining Terms

Alternate mark inversion (AMI): A popular name for bipolar line coding using three levels: zero, positive, and negative.

Binary *N* zero substitution (BNZS): A class of coding schemes that attempts to improve AMI line coding.

Bipolar: A particular line coding scheme using three levels: zero, positive, and negative.

Crosstalk: An unwanted signal from an adjacent channel.

DC wander: The DC level variation in the received signal due to a channel that cannot support DC.

Duobinary coding: A coding scheme with binary input and ternary output requiring the minimum theoretical channel bandwidth.

4 Binary 3 Ternary (4B3T): A line coding scheme that maps four binary digits into three ternary symbols.

High-density bipolar *N* (HDBN): A class of coding schemes that attempts to improve AMI.

Level codes: Line codes carrying information in their voltage levels.

Line coding: The process of converting abstract symbols into real, temporal waveforms to be transmitted through a baseband channel.

Nonreturn to zero (NRZ): A signal that stays at a nonzero level for the entire bit duration.

Pair selected ternary (PST): A coding scheme based on selecting a pair of three level symbols.

Polar: A line coding scheme using both polarity of voltages, with or without a zero level.

Return to zero (RZ): A signal that returns to zero for a portion of the bit duration.

Transition codes: Line codes carrying information in voltage level transitions.

Unipolar: A line coding scheme using only one polarity of voltage, in addition to a zero level.

References

1. Bellamy, J., *Digital Telephony,* John Wiley & Sons, New York, NY, 1991.
2. Bell Telephone Laboratories Technical Staff Members. *Transmission Systems for Communications,* 4th ed., Western Electric Company, Technical Publications, Winston-Salem, NC, 1970.
3. Bic, J.C., Duponteil, D., and Imbeaux, J.C., *Elements of Digital Communication,* John Wiley & Sons, New York, NY, 1991.
4. Bylanski, P., *Digital Transmission Systems,* Peter Peregrinus, Herts, England, 1976.
5. Couch, L.W., *Modern Communication Systems: Principles and Applications,* Prentice-Hall, Englewood Cliffs, NJ, 1994.
6. Feher, K., *Digital Modulation Techniques in an Interference Environment,* EMC Encyclopedia Series, Vol. IX. Don White Consultants, Germantown, MD, 1977.
7. Gibson, J.D., *Principles of Analog and Digital Communications,* MacMillan Publishing, New York, NY, 1993.
8. Lathi, B.P., *Modern Digital and Analog Communication Systems,* Holt, Rinehart and Winston, Philadelphia, PA, 1989.
9. Lender, A., Duobinary Techniques for High Speed Data Transmission, *IEEE Trans. Commun. Electron.*, CE-82, 214–218, May 1963.
10. Lindsey, W.C. and Simon, M.K., *Telecommunication Systems Engineering,* Prentice-Hall, Englewood Cliffs, NJ, 1973.
11. Schwartz, M., *Information Transmission, Modulation, and Noise,* McGraw-Hill, New York, NY, 1980.
12. Stremler, F.G., *Introduction to Communication Systems,* Addison-Wesley Publishing, Reading, MA, 1990.

29

Telecommunications Network Synchronization

Madihally J. Narasimha
Stanford University

29.1 Introduction

Today's telecommunications network comprises customer premises equipment and telephone central offices interconnected by suitable transmission facilities. Although analog technology still exists in the customer loop, digital time-division multiplex (TDM) technology is more prevalent in the central office switching and transmission systems.

A digital switch located within the interoffice network terminates TDM signals originating from many other offices. It performs a combination of time-slot interchange and space switching to accomplish the interconnect function amongst the individual channels of the multiplexed signals. In order to accomplish this function without impairments, the average rates of all of the TDM signals terminating on the switch have to be synchronized to within some achievable bound. Furthermore, the reference clock of the switch itself must also be synchronized to the common rate of the incoming signals. These synchronization requirements are also applicable to a digital cross-connect system since it realizes the channel interconnection function in a similar manner.

Synchronous multiplexers, such as synchronous optical network (SONET) and synchronous digital hierarchy (SDH) terminals, used in fiber optic systems, generate the high-speed output signal by interleaving the time slots of the lower speed input signals. Again, to accomplish this function properly, the rates of the incoming lines and that of the multiplexer clock have to be synchronized. Primary rate multiplexers (known as channel banks) also employ synchronous time interleaving of the 64-kb/s, tributary signals, often generated by other network elements (digital switches and signalling transfer points, for example) within the same office. For unimpaired information transfer in this case, the network elements that terminate these 64-kb/s signals have to be synchronized in both frequency and phase (bit and byte synchronization!).

Network synchronization is the background technology that enables the operating clocks of the various network elements throughout the network to be synchronized. Robust and accurate synchronization networks are essential to the proper functioning of these elements and the reliable transfer of information between them. The growth of data services and the deployment of SONET and SDH transmission equipment has further emphasized the need for such networks.

29.2 Synchronization Distribution Networks

The goal of the synchronization distribution network is to provide reference clocks, traceable to a highly accurate clock called a primary reference source (PRS), to all of the network elements. Since the transport of timing signals over long distances incurs many impairments, the *interoffice* distribution of synchronization references is generally more difficult compared to the task of distributing these reference clocks to the network elements within an office (*intraoffice* distribution).

One method of achieving synchronization in the interoffice network is to designate a single PRS as the master clock for the entire network and transport this clock to every office using a **master–slave** hierarchical discipline. This method is impractical for large networks because of noise accumulation through many levels of cascaded slave clocks, delay variations caused by the rerouting of clock distribution paths under failure conditions, and geopolitical problems. At the other extreme is the **plesiochronous** mode of operation where each office has its own PRS, and no interoffice clock distribution is necessary. This strategy is expensive to implement now, although it can be envisioned in the future because of the projected availability of affordable PRSs. Therefore, a combination of the two techniques is typically employed in practice. As shown in Fig. 29.1, the network is divided into small synchronization regions, each of which has a duplicated set of PRSs as the master clock. Within each region, the timing distribution is accomplished by following a master-slave hierarchy of **stratum clocks**. A stratum 1 clock, normally implemented with cesium beam technology, is required to have a long term accuracy of better than 1×10^{-11}, completely autonomous of other references. The PRS is also required to have the same long-term accuracy as a stratum 1 clock. However, it can be realized either as an autonomous cesium clock or as a nonautonomous clock disciplined by precision global positioning system (GPS) or LORAN-C radio-navigational signals. Stratum 2, 3, and 4 clocks have progressively lower accuracy and performance requirements. Synchronization references are passed from higher performance master clocks to equivalent or lower performance slave clocks. Path redundancy is achieved by providing primary and secondary reference sources at the slave clocks.

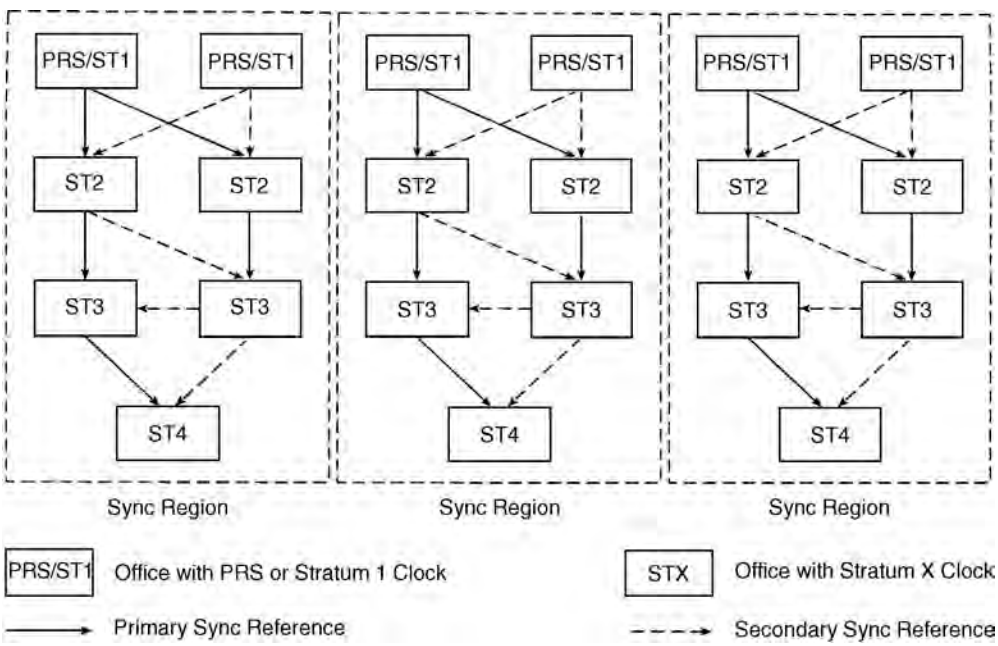


FIGURE 29.1 Network synchronization distribution plan.

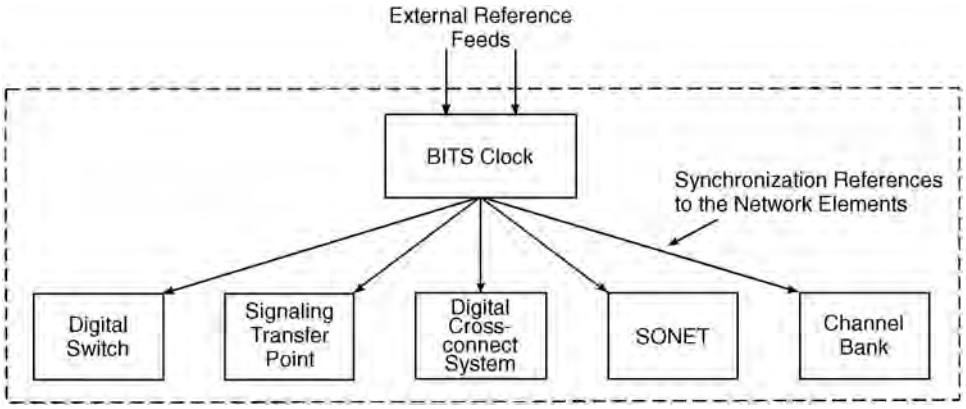


FIGURE 29.2 Intraoffice synchronization distribution plan.

Whereas the interoffice plan for distributing synchronization follows the regionalized master–slave hierarchical strategy previously described, the distribution of reference clocks within an office is realized with a discipline known as the building integrated timing supply (BITS) plan. As shown in Fig. 29.2, this plan calls for the deployment of a properly stratified slave clock called the BITS in each office, which distributes reference timing to all of the network elements. It is the only clock in the office that is directly synchronized to an external reference traceable to a PRS. For robustness, it accepts two external reference feeds and contains duplicated stratum oscillators. If both the external reference feeds are disrupted, the BITS enters the holdover mode of operation where the output reference timing signals are generated using the data acquired during the normal (synchronized) mode.

The master–slave synchronization distribution scheme functions well within a (small) region if the underlying network topology is of the star or tree type. However, self-healing ring topologies are popular with fiber optic networks using the SONET and SDH technology, especially in the loop environment. Feeding multiple synchronization references to slave clocks, or directly to the network elements at the nodes where the BITS plan is not implemented, in such networks can lead to **timing loops** under failure conditions [Bellcore, 1992]. Timing loops are undesirable since they lead to isolation of the clocks within the loop from the PRS and can also cause frequency instabilities due to reference feedback. Providing only a single-synchronization reference source at the nodes is one method of avoiding the possibility of timing loops [Bellcore, 1992]. However, this compromises the path redundancy provided by dual feeds. An alternate procedure is the use of synchronization messages embedded in the overhead channels of the reference signals. Here the nodes indicate the synchronization status (e.g., reference traceable to stratum 1, traceability unknown, etc.) of the transmitted signals to neighboring nodes. The slave clocks (or the network elements) at the other nodes can then avoid the inadvertent creation of timing loops by deciding intelligently, based on the embedded messages, whether or not to switch to the secondary synchronization reference source upon failure of the primary.

29.3 Effect of Synchronization Impairments

The severity of impact on network traffic due to disruptions in the synchronization distribution system or accumulation of phase noise depends on many factors. These include buffering techniques employed to manage the incoming phase variations at the network elements, the type of information transported through the traffic network, and the architecture of the synchronization distribution system itself.

Figure 29.3 shows the situation where the reference clocks of the two interconnected digital switches (or two digital cross-connect systems) are operating at different frequencies due to disruptions in the synchronization distribution network, for example. The receiver at switch 2 has a 125-μs slip buffer to

TABLE 29.1 Impact of Slips on Network Service

Service	Observed Effect
PCM voice	No noticeable degradation for moderate slip rates
Voice-band data	Serious degradation with some modems needing up to 6 s to recover from a slip
Group III fax	Each slip can wipe out 0.08 in. of vertical space
Encrypted voice	Serious degradation requiring retransmission of the encryption key
Compressed video	Serious degradation resulting in freeze frames or missing lines
Digital data (e.g., SS7)	Loss of messages resulting in requests for retransmission and degraded throughput

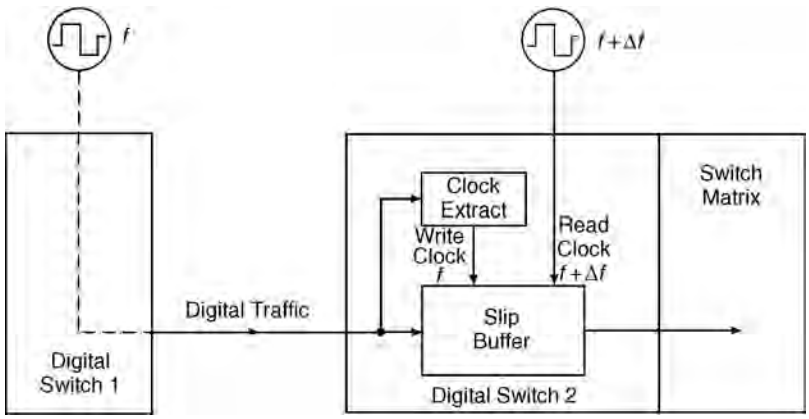


FIGURE 29.3 Slips in the digital network.

absorb the phase variations of the incoming line with respect to its clock. Since the rates of the write and read clocks at the slip buffer are different, however, an eventual overflow or underflow of the buffer is inevitable. This results in a frame of data either being repeated or deleted at the output of the buffer. Such an event is called a controlled slip. The duration between slips is given by $(125/\varepsilon)$ μs, where ε is the fractional frequency deviation $|\Delta f|/f$. As an example, the duration between slips for plesiochronous operation ($\varepsilon = \pm 10^{-11}$) is about 72 days. The impact of slips on various network services is delineated in Table 29.1.

Slips occur infrequently when there are no synchronization disruptions since the 125-μs slip buffers can readily absorb the phase noise caused by the existing transmission systems. However, excessive phase noise either in the incoming lines or in the reference clock at a synchronous multiplexer can cause impairments to the transported data. To understand this, consider the operation of a SONET multiplexer, shown in Fig. 29.4, that combines three 51.84-Mb/s synchronous transport system level 1 (STS-1) signals to form a 155.52-Mb/s STS-3 signal. Data is written into the small-capacity (8-16 bytes) receive buffer in the pointer processor at the rate of the incoming line. If the buffer is near its normal half-full capacity, it is read out at the rate of the local reference clock. However, if the buffer shows signs of filling up or emptying (determined by the upper and lower threshold settings), its read clock is modified by either augmenting or deleting eight pulses from the reference clock. This event, known as a *pointer adjustment*, results in a phase jump of eight unit intervals for the data read out of the buffer and multiplexed onto the STS-3 signal. Occasional pointer adjustments occur even in a synchronized network because of short-term phase variations of the incoming signals, and they cause no harm since data is not lost. However, existing asynchronous equipment may not function properly when receiving

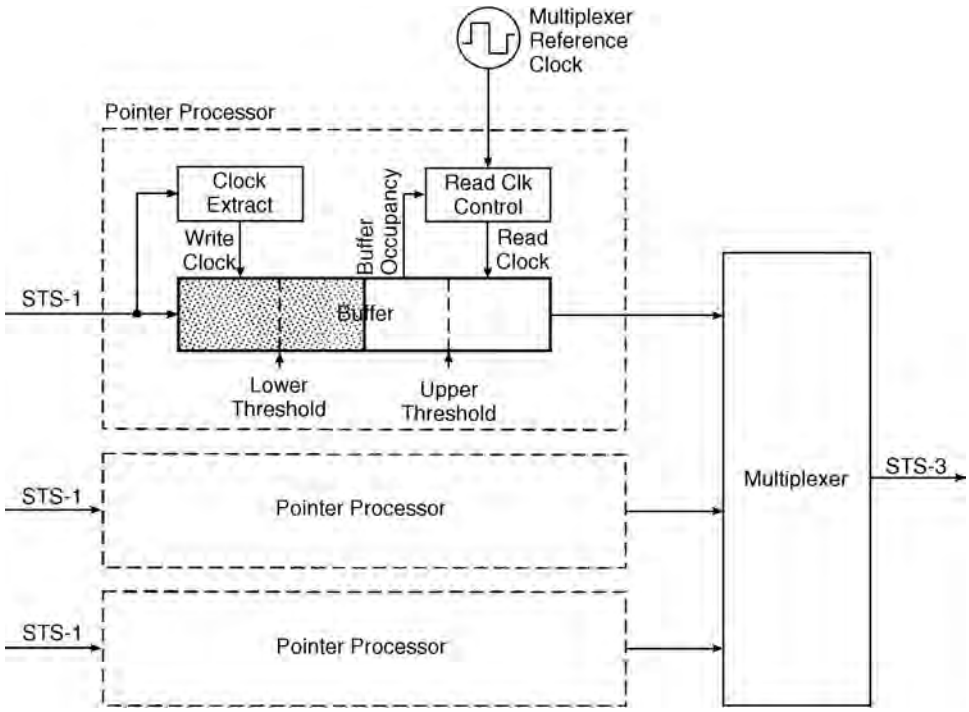


FIGURE 29.4 SONET multiplexer.

severely jittered signals, caused by frequent pointer adjustments, at gateway points between the synchronous and the asynchronous networks. Furthermore, the rate of pointer adjustments is not allowed to exceed certain limits in SONET and SDH networks. Severe phase noise may cause these limits to be exceeded resulting in a buffer overflow or underflow situation, which leads to loss or repetition of data as in the case of a slip.

29.4 Characterization of Synchronization Impairments

The synchronization performance of a telecommunications signal is typically determined from measured values of the deviations in time between when the signal transitions actually occur and when they are ideally expected. The master clock that provides the ideal time positions should be significantly more accurate than the signal under test. The measurement yields the raw phase error signal, also known as the *time delay* or simply *phase*, in units of time as a function of elapsed time. Postprocessing of this raw data is, however, necessary to extract useful parameters that help define the underlying synchronization impairment model.

The current synchronization impairment model used for the telecommunications signal is based mainly on the source of the impairments. It refers to the higher frequency components of the phase (error) oscillations as **jitter** and to the lower frequency components as **wander**, with 10 Hz being the demarcation frequency. Jitter is produced mainly by regenerative repeaters and asynchronous multiplexers, and normal levels of it can readily be buffered and filtered out. However, excessive jitter is a potential source of bit errors in the digital network. Wander, on the other hand, has many causes. These include temperature cycling effects in cables, waiting time effects in asynchronous multiplexers, and the effects of frequency and phase quantization in slave clocks which employ narrowband filters in their servocontrol loops. Since it contains very low-frequency components, wander cannot be completely filtered out and,

hence, is passed on. Excessive wander adversely affects the slip performance of the network and can also lead to an unacceptable rate of pointer adjustments in SONET and SDH multiplexers. Furthermore, wander on an input reference signal compromises the holdover performance of a slave clock during reference interruptions.

Although jitter and wander are adequate descriptors of synchronization impairments in the existing digital network, a more comprehensive model is necessary to characterize the phase variations caused by the newly introduced SONET and SDH equipment and to appropriately specify the limits on timing noise at network interfaces. Recently, there has been renewed interest in adapting the traditional noise model used with precision oscillators and clocks in time and frequency metrology to the telecommunications applications. This model permits the characterization of the complex timing impairments with a handful of parameters classified into systematic components (phase offset, frequency offset, and frequency drift) and stochastic power-law noise components (white PM, flicker PM, white FM, flicker FM, and random walk FM).

Maximum time interval error (MTIE) and time variance (TVAR) are two parameters [ANSI, 1994] for specifying the synchronization performance of telecommunications signals that are defined to capture the essential features of the traditional model. MTIE is effective in characterizing peak-to-peak phase movements, which are primarily due to systematic components, whereas TVAR is useful in characterizing the stochastic power-law noise components of the phase noise.

The algorithm for computing the MTIE from N time delay (phase error) samples x_i, $i = 1, 2, \ldots, N$, measured at time intervals τ_0, is illustrated in Fig. 29.5. For a given observation interval S spanning n samples, the peak-to-peak phase excursion values are noted for all possible positions of the observation window that can be accommodated by the collected data. The maximum of all such values yields MTIE(S). This computation can be expressed as

$$\mathrm{MTIE}(S) = \max_{j=1,\ldots,N-n+1} x_{ppj}(S)$$

where

$$x_{ppj}(S) = \max x_i - \min x_i, \quad i = j, \ldots, j+n-1$$

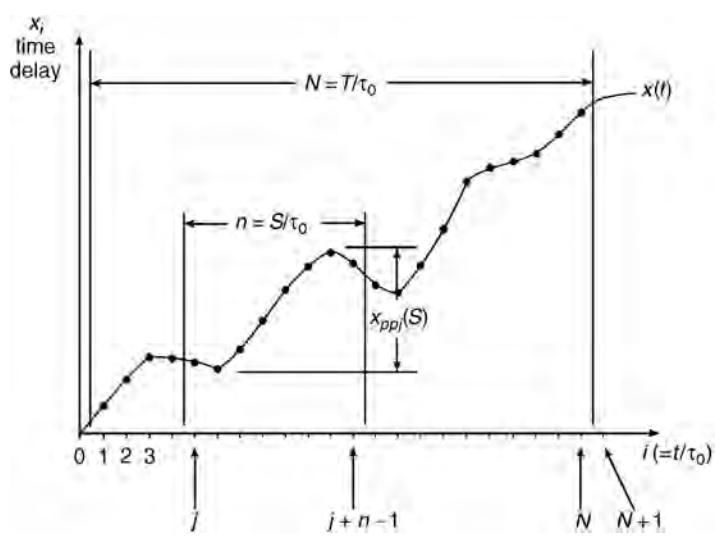


FIGURE 29.5 Computation of the MTIE from measured time delay samples.

The calculation of TVAR for the same set of time delay samples is expressed by the equation:

$$\mathrm{TVAR}(\tau) = \frac{1}{6n^2(N-3n+1)} \sum_{j=1}^{N-3n+1} \left[\sum_{j=0}^{n-1} (x_{j+2n+k} - 2x_{j+n+k} + x_{j+k}) \right]^2$$

where the independent variable $\tau = n\tau_0$ and spanning n samples is known as the *integration time*. This calculation may be viewed as a spectral estimation process with a nonuniform bandpass filter bank. The magnitude response of the bandpass filter corresponding to the integration time τ at frequency f is given by

$$\left| \sqrt{\frac{8}{3}} \frac{\sin^3(\pi\tau f)}{n \sin(\pi\tau_0 f)} \right|$$

A standard variance estimate of the filtered samples then yields TVAR(τ). The square root of TVAR is denoted by time deviation (TDEV).

The graphs of MTIE plotted as a function of the observation interval S and TDEV plotted as a function of the integration time τ are primarily used to characterize the synchronization performance of a telecommunications signal.

29.5 Synchronization Standards

Standards for synchronization in telecommunications systems are set by the American National Standards Institute (ANSI) in the U.S. and by the Consultative Committee on International Telephony and Telegraphy (CCITT), now known as the International Telecommunications Union (ITU), for international applications. In addition, Bellcore issues technical references and technical advisories that specify the requirements from the viewpoint of the regional Bell operating companies (RBOCs). These documents are listed in the References. Some highlights of these standards are reviewed here.

The various standards deal with two categories of specifications: the characteristics of the clocks used in the synchronization distribution network, and the synchronization performance of the reference signals at network interfaces.

The **free-run accuracy**, the **holdover stability**, and the **pull-in range** requirements of the stratum clocks employed in the synchronized network are delineated in Table 29.2. (The ITU uses a different terminology for the stratum levels. Refer to Annex D of the revised ANSI standard T1.101 for the differences.) In addition to the pull-in range requirements, the slave clocks should be able to tolerate certain amounts of jitter, wander, and **phase transients** at their reference inputs. Moreover, there are other performance constraints, besides accuracy and stability, on the output signal. These include wander generation, wander transfer, jitter generation, and phase transients. These are detailed in the Bellcore [1993b] Technical Reference TR-NWT-00 1244.

As an illustration, Fig. 29.6 shows the MTIE specification mask for a PRS. Also shown are the measured performance curves of typical PRS clocks based on GPS and LORAN-C receiver technology.

TABLE 29.2 Performance Requirements of Stratum Clocks

Stratum	Free-Run Accuracy	Holdover Stability	Pull-In Range
1	$\pm 1 \times 10^{-11}$	—	—
2	$\pm 1.6 \times 10^{-8}$	$\pm 1 \times 10^{-10}$/day	$\pm 1.6 \times 10^{-8}$
3E	$\pm 4.6 \times 10^{-6}$	$\pm 1 \times 10^{-8}$/day	$\pm 4.6 \times 10^{-6}$
3	$\pm 4.6 \times 10^{-6}$	$\pm 3.7 \times 10^{-7}$/day	$\pm 4.6 \times 10^{-6}$
4/4E	$\pm 32 \times 10^{-6}$	—	$\pm 32 \times 10^{-6}$

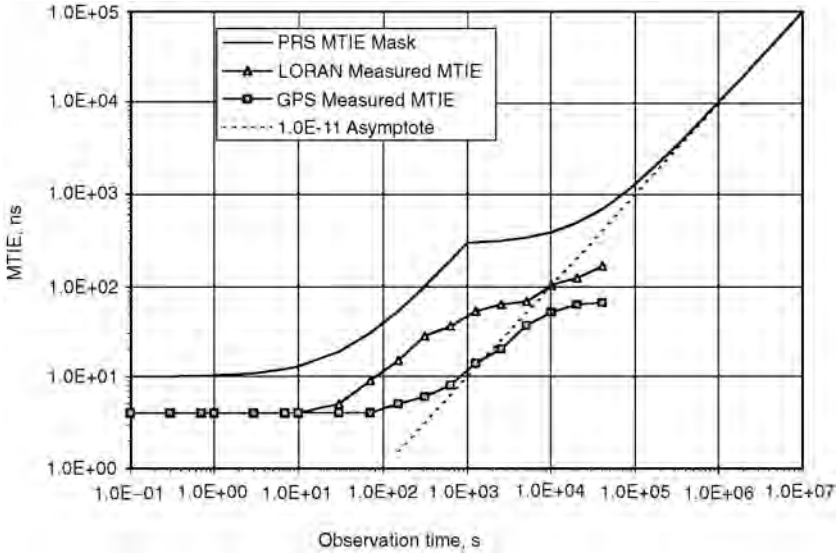


FIGURE 29.6 MTIE specification mask and measured performance of GPS and LORAN PRSs.

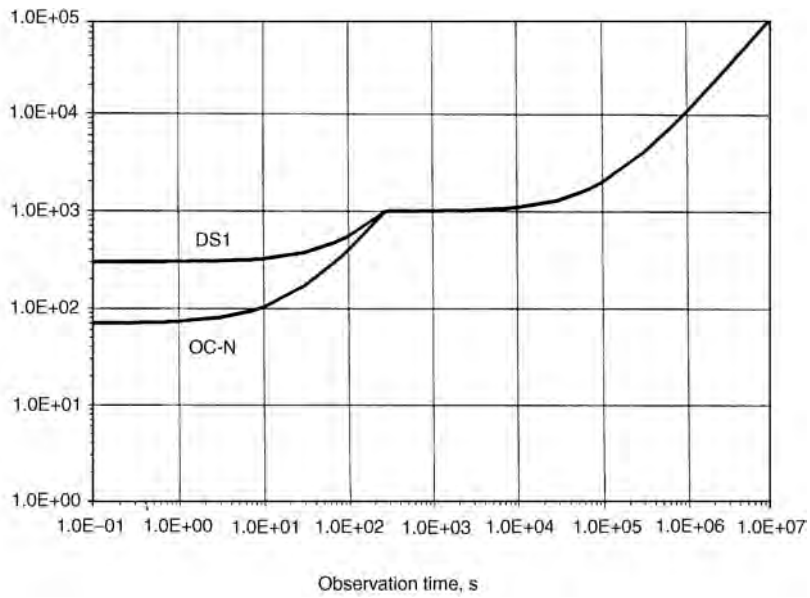


FIGURE 29.7 MTIE specification for synchronization reference signals at network interfaces.

The performance requirements of synchronization references at network interfaces depend on whether optical rate signals [i.e., optical carrier level N (OC-N)] or primary rate electrical signals [e.g., digital signal level 1 (DS1)] are being considered. The specifications for optical interfaces are tighter because SONET and SDH pointer adjustments are sensitive to short-term phase noise. The MTIE and TDEV specification masks for DS1 and OC-N signals are shown in Figs. 29.7 and 29.8.

These reference signals also have to satisfy certain constraints on the magnitude of phase transients. The revised ANSI Standard T1.101-1987 [ANSI, 1994] specifies these details.

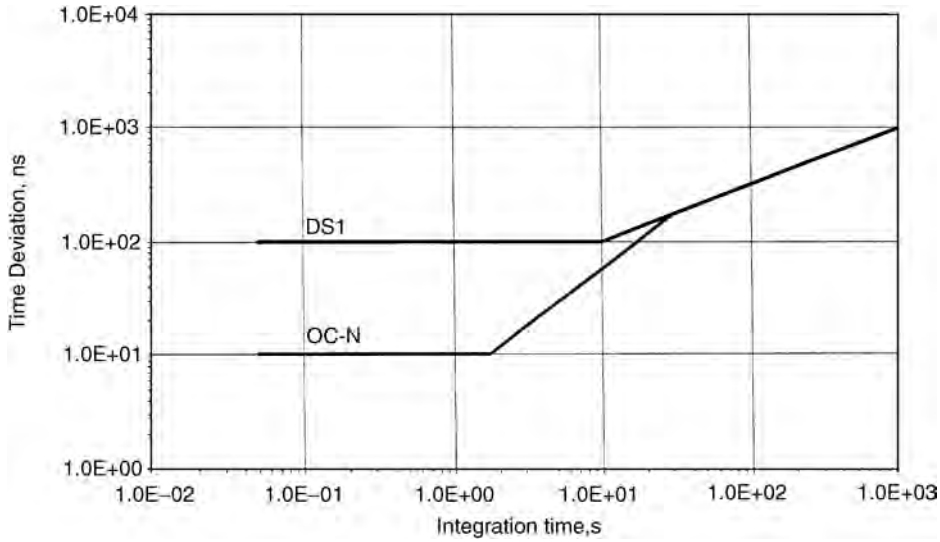


FIGURE 29.8 TDEV specification for synchronization reference signals at network interfaces.

29.6 Summary and Conclusions

The deployment of digital switching exchanges and cross-connect equipment in telecommunications systems created the necessity for robust synchronization distribution networks. These networks were originally designed to guarantee satisfactory slip performance for an end-to-end connection. However, recently introduced high-speed synchronous multiplexing and transmission systems based on SONET and SDH technology have emphasized the need for enhancing their accuracy and reliability. A disciplined approach to the design of such networks is delineated in many of the synchronization standards issued so far.

Our current understanding of the sources of synchronization impairments in telecommunications networks has not yet reached a mature point. The adaption of traditional clock noise models, such as those used in time and frequency metrology, to describe these impairments is a step in the right direction. However, much work is necessary to first evaluate reliably the parameters of this model from the phase error measurements, and then to identify the actual sources of impairments from them. As newer transport technologies such as asynchronous transfer mode (ATM) are introduced, the distribution of accurate synchronization references will be more complicated, and the impact of synchronization disruptions on network services will be harder to predict. These and other issues are currently being investigated by the various standards organizations.

Defining Terms

Free-run accuracy: The maximum long-term (20 years) deviation limit of a clock from the nominal frequency with no external frequency reference.

Holdover stability: The maximum rate of change of the clock frequency with respect to time upon loss of all input frequency references.

Jitter: The short-term variations of the significant instants (e.g., zero level crossings) of a digital signal from their ideal positions in time, where short term implies phase variations of frequency greater than or equal to 10 Hz.

Master–slave hierarchy: The hierarchical method where synchronization references are distributed from offices with higher performance master stratum clocks to offices with the same or lower performance slave stratum clocks.

Phase transient: Perturbations in phase of limited duration (typically several time constants of the slave clock which produces it) seen at synchronization interfaces.

Plesiochronous: Two signals are plesiochronous if their corresponding significant instants (e.g., zero level crossings) occur at nominally the same rate, any variation in rate being constrained within specified limits.

Pull-in range: Measure of the maximum reference frequency deviation from the nominal rate that can be overcome by a slave clock to pull itself into synchronization.

Stratum clocks: A classification of clocks in the synchronization network based on performance. Stratum 1 is the highest and stratum 4 is the lowest level of performance.

Timing loop: The situation where a slaved clock receives input reference timing from itself via a chain of other slaved clocks.

Wander: The long-term variations of the significant instants (e.g., zero level crossings) of a digital signal from their ideal positions in time, where long term implies phase variations of low frequency (less than 10 Hz).

References

ANSI. 1987. Synchronization interface standards for digital networks. ANSI Standard T1.101-1987, American National Standards Institute.

ANSI. 1994. Revision of ANSI Standard T1. 101 - 1987. ANSI T1. 101 - 1994, American National Standards Institute.

Bellcore. 1992. SONET synchronization planning guidelines. Special Rep. SR-NWT-002224, Issue 1, Feb.

Bellcore. 1993a. Digital network synchronization plan. Tech. Advisory TA-NWT-000436, Issue 2, June.

Bellcore. 1993b. Clocks for the synchronized network: Common generic criteria. Tech. Ref. TR-NWT-001244, Issue 1, June.

CCITT. 1988a. Timing requirements at the outputs of primary reference clocks suitable for plesiochronous operation of international digital links. Recommendation G.811, Consultative Committee on International Telephony and Telegraphy, Blue Book, Melbourne, Nov.

CCITT. 1988b. Timing requirements at the outputs of slave clocks suitable for plesiochronous operation of international digital links. Recommendation G.812, Consultative Committee on International Telephony and Telegraphy, Blue Book, Melbourne, Nov.

CCITT. 1992. Timing characteristics of slave clocks suitable for operation of SDH equipment. Draft Recommendation G.81s, Consultative Committee on International Telephony and Telegraphy, Geneva, June.

Zampetti, G. 1992. Synopsis of timing measurement techniques used in telecommunications, in *Proceedings of the 24th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting,* McLean, VA, 313–324.

Further Information

The ANSI, Bellcore, and ITU (CCITT) documents listed under the references are the most suitable sources for the synchronization standards. Information regarding the revisions to these standards and newer standards being contemplated is available from the recent contributions to the T1X1.3 working group of the ANSI and to the SG XIII/WP6 of the ITU. Zampetti's paper [1992] provides a synopsis of the performance parameters MTIE and TVAR for characterizing synchronization impairments in the network and also proposes an additional parameter to gain further insights.

30

Echo Cancellation

Giovanni Cherubini
IBM Research, Zurich Research Laboratory

30.1 Introduction

Full-duplex data transmission over a single twisted-pair cable permits the simultaneous flow of information in two directions when the same frequency band is used. Examples of applications of this technique are found in digital communications systems that operate over the telephone network. In a digital subscriber loop, at each end of the full-duplex link, a circuit known as a hybrid separates the two directions of transmission. To avoid signal reflections at the near- and far-end hybrid, a precise knowledge of the line impedance would be required. Since the line impedance depends on line parameters that, in general, are not exactly known, an attenuated and distorted replica of the transmit signal leaks to the receiver input as an echo signal. Data-driven adaptive echo cancellation mitigates the effects of impedance mismatch.

A similar problem is caused by crosstalk in transmission systems over voice-grade unshielded twisted-pair cables for local-area network applications, where multipair cables are used to physically separate the two directions of transmission. Crosstalk is a statistical phenomenon due to randomly varying differential capacitive and inductive coupling between adjacent two-wire transmission lines. At the rates of several megabits per second that are usually considered for local-area network applications, near-end crosstalk (NEXT) represents the dominant disturbance; hence, adaptive NEXT cancellation must be performed to ensure reliable communications.

In voiceband data modems, the model for the echo channel is considerably different from the echo model adopted in baseband transmission. The transmitted signal is a passband signal obtained by quadrature amplitude modulation (QAM), and the far-end echo may exhibit significant carrier-phase jitter and carrier-frequency shift, which are caused by signal processing at intermediate points in the telephone network. Therefore, a digital adaptive echo canceller for voiceband modems needs to embody algorithms that account for the presence of such additional impairments.

In this chapter, we describe the echo channel models and adaptive echo canceller structures that are obtained for various digital communications systems, which are classified according to the employed modulation techniques. We also address the tradeoffs between complexity, speed of adaptation, and accuracy of cancellation in adaptive echo cancellers.

30.2 Echo Cancellation for Pulse-Amplitude Modulation (PAM) Systems

The model of a full-duplex baseband data transmission system employing pulse-amplitude modulation (PAM) and adaptive echo cancellation is shown in Fig. 30.1. To describe system operations, we consider one end of the full-duplex link. The configuration of an echo canceller for a PAM transmission system is shown in Fig. 30.2. The transmitted data consist of a sequence $\{a_n\}$ of independent and identically distributed (i.i.d.) real-valued symbols from the M-ary alphabet $\mathcal{A} = \{\pm 1, \pm 3, \ldots, \pm(M-1)\}$. The sequence $\{a_n\}$ is converted into an analog signal by a digital-to-analog (D/A) converter. The conversion to a staircase signal by a zero-order hold D/A converter is described by the frequency response $H_{\mathrm{D/A}}(f) = T\sin(\pi f T)/(\pi f T)$, where T is the modulation interval. The D/A converter output is filtered by the analog transmit filter and is input to the channel through the hybrid.

The signal $x(t)$ at the output of the low-pass analog receive filter has three components, namely, the signal from the far-end transmitter $r(t)$, the echo $u(t)$, and additive Gaussian noise $w(t)$. The signal $x(t)$ is given by

$$\begin{aligned} x(t) &= r(t) + u(t) + w(t) \\ &= \sum_{n=-\infty}^{\infty} a_n^R h(t - nT) + \sum_{n=-\infty}^{\infty} a_n h_E(t - nT) + w(t) \end{aligned} \tag{30.1}$$

where $\{a_n^R\}$ is the sequence of symbols from the remote transmitter, and $h(t)$ and $h_E(t) = \{h_{\mathrm{D/A}} \otimes g_E\}(t)$ are the impulse responses of the overall channel and the echo channel, respectively. In the expression of $h_E(t)$, the function $h_{\mathrm{D/A}}(t)$ is the inverse Fourier transform of $H_{\mathrm{D/A}}(f)$, and the operator $\otimes$ denotes convolution. The signal obtained after echo cancellation is processed by a detector that outputs the sequence of estimated symbols $\{\hat{a}_n^R\}$. In the case of full-duplex PAM data transmission over multi-pair cables for

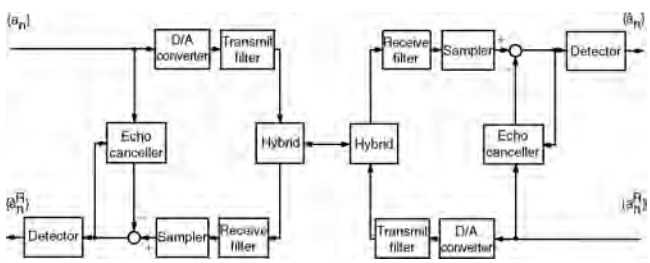


FIGURE 30.1 Model of a full-duplex PAM transmission system.

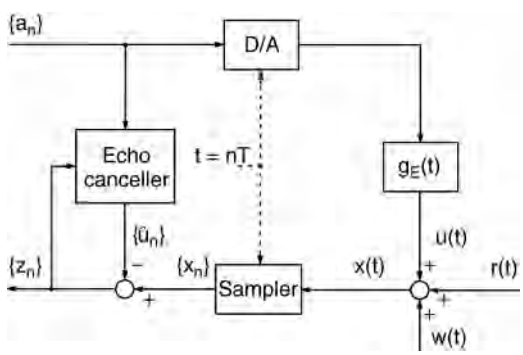


FIGURE 30.2 Configuration of an echo canceller for a PAM transmission system.

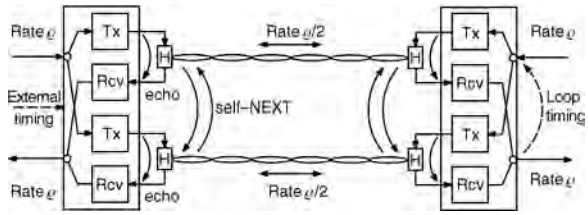


FIGURE 30.3 Model of a dual-duplex transmission system.

local-area network applications, where NEXT represents the main disturbance, the configuration of a digital NEXT canceller is also obtained as shown in Fig. 30.2, with the echo channel replaced by the crosstalk channel. For these applications, however, instead of *mono-duplex* transmission, where one pair is used to transmit only in one direction and the other pair to transmit only in the reverse direction, *dual-duplex* transmission may be adopted. Bi-directional transmission at rate ϱ over two pairs is then accomplished by full-duplex transmission of data streams at $\varrho/2$ over each of the two pairs. The lower modulation rate and/or spectral efficiency required per pair for achieving an aggregate rate equal to ϱ represents an advantage of dual-duplex over mono-duplex transmission. Dual-duplex transmission requires two transmitters and two receiver at each end of a link, as well as separation of the simultaneously transmitted and received signals on each pair, as illustrated in Fig. 30.3. In dual-duplex transceivers, it is therefore necessary to suppress echoes returning from the hybrids and impedance discontinuities in the cable, as well as self NEXT, by adaptive digital echo and NEXT cancellation [3]. Although a dual-duplex scheme might appear to require higher implementation complexity than a mono-duplex scheme, it turns out that the two schemes are equivalent in terms of the number of multiply-and-add operations per second that are needed to perform the various filtering operations.

One of the transceivers in a full-duplex link will usually employ an externally provided reference clock for its transmit and receive operations. The other transceiver will extract timing from the receive signal and use this timing for its transmitter operations. This is known as *loop timing*, also illustrated in Fig. 30.3. If signals were transmitted in opposite directions with independent clocks, signals received from the remote transmitter would generally shift in phase relative to the also received echo signals. To cope with this effect, some form of interpolation would be required that can significantly increase the transceiver complexity [2].

In general, we consider baseband signalling techniques such that the signal at the output of the overall channel has non-negligible excess bandwidth, i.e., non-negligible spectral components at frequencies larger than half of the modulation rate, $|f| \geq 1/2T$. Therefore, to avoid aliasing, the signal $x(t)$ is sampled at twice the modulation rate or at a higher sampling rate. Assuming a sampling rate equal to m/T, $m > 1$, the ith sample during the nth modulation interval is given by

$$\begin{aligned} x\left[(nm+i)\frac{T}{m}\right] &= x_{nm+i} = r_{nm+i} + u_{nm+i} + w_{nm+i}, \quad i = 0,\ldots,m-1 \\ &= \sum_{k=-\infty}^{\infty} h_{km+i}\, a_{n-k}^{R} + \sum_{k=-\infty}^{\infty} h_{E,km+i}\, a_{n-k} + w_{nm+i} \end{aligned} \tag{30.2}$$

where $\{h_{nm+i}, i = 0,\ldots,m-1\}$ and $\{h_{E,nm+i}, i = 0,\ldots,m-1\}$ are the discrete-time impulse responses of the overall channel and the echo channel, respectively, and $\{w_{nm+i}, i = 0,\ldots,m-1\}$ is a sequence of Gaussian noise samples with zero mean and variance σ_w^2. Equation (30.2) suggests that the sequence of samples $\{x_{nm+i}, i = 0,\ldots,m-1\}$ be regarded as a set of m interleaved sequences, each with a sampling rate equal to the modulation rate. Similarly, the sequence of echo samples $\{u_{nm+i}, i = 0,\ldots,m-1\}$ can be regarded as a set of m interleaved sequences that are output by m independent echo channels with discrete-time impulse responses $\{h_{E,nm+i}\}$, $i = 0,\ldots,m-1$, and an identical sequence $\{a_n\}$ of input symbols [8]. Hence, echo cancellation can be performed by m interleaved echo cancellers, as shown in Fig. 30.4. Since the

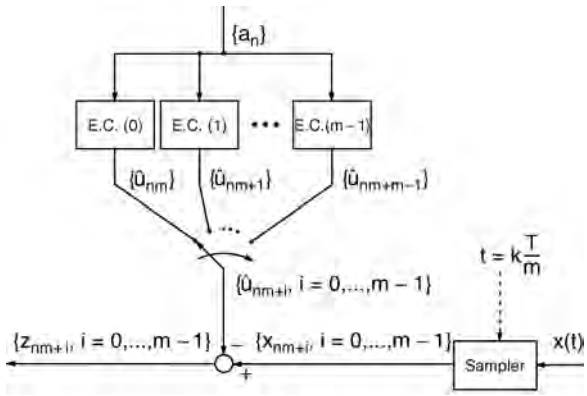


FIGURE 30.4 A set of m interleaved echo cancellers.

performance of each canceller is independent of the other $m - 1$ units, in the remaining part of this section we will consider the operations of a single echo canceller.

The echo canceller generates an estimate $\hat{u}_n$ of the echo signal. If we consider a transversal filter realization, $\hat{u}_n$ is obtained as the inner product of the vector of filter coefficients at time $t = nT$, $\boldsymbol{c}_n = (c_{n,0}, \ldots, c_{n,N-1})'$ and the vector of signals stored in the echo canceller delay line at the same instant, $\boldsymbol{a}_n = (a_n, \ldots, a_{n-N+1})'$, expressed by

$$\hat{u}_n = \boldsymbol{c}_n' \boldsymbol{a}_n = \sum_{k=0}^{N-1} c_{n,k}\, a_{n-k} \tag{30.3}$$

where $\boldsymbol{c}_n'$ denotes the transpose of the vector $\boldsymbol{c}_n$. The estimate of the echo is subtracted from the received signal. The result is defined as the cancellation error signal

$$z_n = x_n - \hat{u}_n = x_n - \boldsymbol{c}_n' \boldsymbol{a}_n \tag{30.4}$$

The echo attenuation that must be provided by the echo canceller to achieve proper system operation depends on the application. For example, for the Integrated Services Digital Network (ISDN) U-Interface transceiver, the echo attenuation must be larger than 55 dB [11]. It is then required that the echo signals outside of the time span of the echo canceller delay line be negligible, i.e., $h_{E,n} \approx 0$ for $n < 0$ and $n > N - 1$. As a measure of system performance, we consider the mean square error ε_n^2 at the output of the echo canceller at time $t = nT$, defined by

$$\varepsilon_n^2 = E\{z_n^2\} \tag{30.5}$$

where $\{z_n\}$ is the error sequence and $E\{\cdot\}$ denotes the expectation operator. For a particular coefficient vector $\boldsymbol{c}_n$, substitution of Eq. (30.4) into Eq. (30.5) yields

$$\varepsilon_n^2 = E\{x_n^2\} - 2\boldsymbol{c}_n' \boldsymbol{q} + \boldsymbol{c}_n' \boldsymbol{R} \boldsymbol{c}_n \tag{30.6}$$

where $\boldsymbol{q} = E\{x_n \boldsymbol{a}_n\}$ and $\boldsymbol{R} = E\{\boldsymbol{a}_n \boldsymbol{a}_n'\}$. With the assumption of i.i.d. transmitted symbols, the correlation matrix $\boldsymbol{R}$ is diagonal. The elements on the diagonal are equal to the variance of the transmitted symbols, $\sigma_a^2 = (M^2 - 1)/3$. The minimum mean square error is given by

$$\varepsilon_{\min}^2 = E\{x_n^2\} - \boldsymbol{c}_{\text{opt}}' \boldsymbol{R} \boldsymbol{c}_{\text{opt}} \tag{30.7}$$

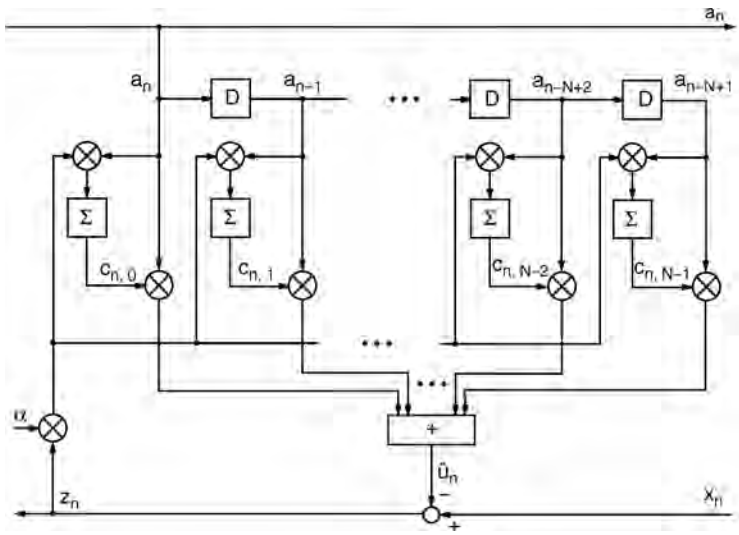


FIGURE 30.5 Block diagram of an adaptive transversal filter echo canceller.

where the optimum coefficient vector is $\boldsymbol{c}_{\text{opt}} = \boldsymbol{R}^{-1}\boldsymbol{q}$. We note that proper system operation is achieved only if the transmitted symbols are uncorrelated with the symbols from the remote transmitter. If this condition is satisfied, the optimum filter coefficients are given by the values of the discrete-time echo channel impulse response, i.e., $\boldsymbol{c}_{\text{opt},k} = h_{E,k}$, $k = 0,\ldots,N-1$.

By the decision-directed stochastic gradient algorithm, also known as the least mean square (LMS) algorithm, the coefficients of the echo canceller converge in the mean to $\boldsymbol{c}_{\text{opt}}$. The LMS algorithm for an N-tap adaptive linear transversal filter is formulated as follows:

$$\boldsymbol{c}_{n+1} = \boldsymbol{c}_n - \frac{1}{2}\alpha\nabla_c\{z_n^2\} = \boldsymbol{c}_n + \alpha z_n \boldsymbol{a}_n \tag{30.8}$$

where α is the adaptation gain and

$$\nabla_c\{z_n^2\} = \left(\frac{\partial z_n^2}{\partial c_{n,0}},\ldots,\frac{\partial z_n^2}{\partial c_{n,N-1}}\right)' = -2z_n\boldsymbol{a}_n$$

is the gradient of the squared error with respect to the vector of coefficients. The block diagram of an adaptive transversal filter echo canceller is shown in Fig. 30.5.

If we define the vector $\boldsymbol{p}_n = \boldsymbol{c}_{\text{opt}} - \boldsymbol{c}_n$, the mean square error can be expressed as

$$\varepsilon_n^2 = \varepsilon_{\min}^2 + \boldsymbol{p}_n'\boldsymbol{R}\boldsymbol{p}_n \tag{30.9}$$

where the term $\boldsymbol{p}_n'\boldsymbol{R}\boldsymbol{p}_n$ represents an excess mean square distortion due to the misadjustment of the filter settings. The analysis of the convergence behavior of the excess mean square distortion was first proposed for adaptive equalizerss [14] and later extended to adaptive echo cancellers [10]. Under the assumption that the vectors $\boldsymbol{p}_n$ and $\boldsymbol{a}_n$ are statistically independent, the dynamics of the mean square error are given by

$$E\{\varepsilon_n^2\} = \varepsilon_0^2[1-\alpha\sigma_a^2(2-\alpha N\sigma_a^2)]^n + \frac{2\varepsilon_{\min}^2}{2-\alpha N\sigma_a^2} \tag{30.10}$$

where ε_0^2 is determined by the initial conditions. The mean square error converges to a finite steady-state value ε_∞^2 if the stability condition $0 < \alpha < 2/(N\sigma_a^2)$ is satisfied. The optimum adaptation gain that yields fastest convergence at the beginning of the adaptation process is $\alpha_{\text{opt}} = 1/(N\sigma_a^2)$. The corresponding time constant and asymptotic mean square error are $\tau_{\text{opt}} = N$ and $\varepsilon_\infty^2 = 2\varepsilon_{\min}^2$, respectively.

We note that a fixed adaptation gain equal to α_{opt} could not be adopted in practice, since after echo cancellation the signal from the remote transmitter would be embedded in a residual echo having approximately the same power. If the time constant of the convergence mode is not a critical system parameter, an adaptation gain smaller than α_{opt} will be adopted to achieve an asymptotic mean square error close to $\varepsilon_{\min}^2$. On the other hand, if fast convergence is required, a variable gain will be chosen.

Several techniques have been proposed to increase the speed of convergence of the LMS algorithm. In particular, for echo cancellation in data transmission, the speed of adaptation is reduced by the presence of the signal from the remote transmitter in the cancellation error. To mitigate this problem, the data signal can be adaptively removed from the cancellation error by a decision-directed algorithm [6].

Modified versions of the LMS algorithm have been also proposed to reduce system complexity. For example, the sign algorithm suggests that only the sign of the error signal be used to compute an approximation of the stochastic gradient [5]. An alternative means to reduce the implementation complexity of an adaptive echo canceller consists of the choice of a filter structure with a lower computational complexity than the transversal filter.

At high data rates, very large-scale integration (VLSI) technology is needed for the implementation of transceivers for full-duplex data transmission. High-speed echo cancellers and near-end crosstalk cancellers that do not require multiplications represent an attractive solution because of their low complexity. As an example of an architecture suitable for VLSI implementation, we consider echo cancellation by a distributed-arithmetic filter, where multiplications are replaced by table lookup and shift-and-add operations [13]. By segmenting the echo canceller into filter sections of shorter lengths, various tradeoffs concerning the number of operations per modulation interval and the number of memory locations needed to store the lookup tables are possible. Adaptivity is achieved by updating the values stored in the lookup tables by the LMS algorithm.

To describe the principles of operations of a istributed-arithmetic echo canceller, we assume that the number of elements in the alphabet of input symbols is a power of two, $M = 2^W$. Therefore, each symbol is represented by the vector $(a_n^{(0)}, \ldots, a_n^{(W-1)})$, where $a_n^{(i)}$, $i = 0, \ldots, W-1$, are independent binary random variables, i.e.,

$$a_n = \sum_{w=0}^{W-1} (2a_n^{(w)} - 1)2^w = \sum_{w=0}^{W-1} b_n^{(w)} 2^w \tag{30.11}$$

where $b_n^{(w)} = (2a_n^{(w)} - 1) \in \{-1, +1\}$. By substituting Eq. (30.11) into Eq. (30.1) and segmenting the delay line of the echo canceller into L sections with $K = N/L$ delay elements each, we obtain

$$\hat{u}_n = \sum_{\ell=0}^{L-1} \sum_{w=0}^{W-1} 2^w \left[\sum_{k=0}^{K-1} b_{n-\ell K-k}^{(w)} c_{n,\ell K+k} \right] \tag{30.12}$$

Equation (30.12) suggests that the filter output can be computed using a set of $L2^K$ values that are stored in L tables with 2^K memory locations each. The binary vectors $\boldsymbol{a}_{n,\ell}^{(w)} = (a_{n-(\ell+1)K+1}^{(w)}, \ldots, a_{n-\ell K}^{(w)})$, $w = 0, \ldots, W-1$, $\ell = 0, \ldots, L-1$, determine the addresses of the memory locations where the values that are needed to compute the filter output are stored. The filter output is obtained by WL table lookup and shift-and-add operations.

We observe that $\boldsymbol{a}_{n,\ell}^{(w)}$ and its binary complement $\bar{\boldsymbol{a}}_{n,\ell}^{(w)}$ select two values that differ only in their sign. This symmetry is exploited to halve the number of values to be stored. To determine the output of a

distributed-arithmetic filter with reduced memory size, we reformulate Eq. (30.12) as

$$\hat{u}_n = \sum_{\ell=0}^{L-1}\sum_{w=0}^{W-1} 2^w b_{n-\ell K-k_0}^{(w)}\left[c_{n,\ell K+k_0} + b_{n-\ell K-k_0}^{(w)} \sum_{\substack{k=0\\k\neq k_0}}^{K-1} b_{n-\ell K-k}^{(w)} c_{n,\ell K+k}\right] \tag{30.13}$$

where k_0 can be any element of the set $\{0,\ldots,K-1\}$. In the following, we take $k_0 = 0$. Then the binary symbol $b_{n-\ell K}^{(w)}$ determines whether a selected value is to be added or subtracted. Each table has now 2^{K-1} memory locations, and the filter output is given by

$$\hat{u}_n = \sum_{\ell=0}^{L-1}\sum_{w=0}^{W-1} 2^w b_{n-\ell K}^{(w)} d_n(i_{n,\ell}^{(w)},\ell) \tag{30.14}$$

where $d_n(k,\ell)$, $k = 0,\ldots,2^{K-1}-1$, $\ell = 0,\ldots,L-1$, are the lookup values, and $i_{n,\ell}^{(w)}$, $w = 0,\ldots,W-1$, $\ell = 0,\ldots,L-1$, are the look up indices computed as follows:

$$i_{n,\ell}^{(w)} = \begin{cases} \sum_{k=1}^{K-1} a_{n-\ell K-k}^{(w)} 2^{k-1} & \text{if } a_{n-\ell K}^{(w)} = 1 \\ \sum_{k=1}^{K-1} \bar{a}_{n-\ell K-k}^{(w)} 2^{k-1} & \text{if } a_{n-\ell K}^{(w)} = 0 \end{cases} \tag{30.15}$$

We note that, as long as Eqs. (30.12) and (30.13) hold for some coefficient vector $(c_{n,0},\ldots,c_{n,N-1})$, a distributed-arithmetic filter emulates the operation of a linear transversal filter. For arbitrary values $d_n(k,\ell)$, however, a nonlinear filtering operation results.

The expression of the LMS algorithm to update the values of a distributed-arithmetic echo canceller takes the form

$$\boldsymbol{d}_{n+1} = \boldsymbol{d}_n - \frac{1}{2}\alpha\nabla_{\boldsymbol{d}}\{z_n^2\} = \boldsymbol{d}_n + \alpha z_n \boldsymbol{y}_n \tag{30.16}$$

where $\boldsymbol{d}'_n = [\boldsymbol{d}'_n(0),\ldots,\boldsymbol{d}'_n(L-1)]$, with $\boldsymbol{d}'_n(\ell) = [d_n(0,\ell),\ldots,d_n(2^{K-1}-1,\ell)]$, and $\boldsymbol{y}'_n = [\boldsymbol{y}'_n(0),\ldots,\boldsymbol{y}'_n(L-1)]$, with

$$\boldsymbol{y}'_n(\ell) = \sum_{w=0}^{W-1} 2^w b_{n-\ell K}^{(w)} (\delta_{0,i_{n,\ell}^{(w)}},\ldots,\delta_{2^{K-1}-1,i_{n,\ell}^{(w)}})$$

are $L2^{K-1}\times 1$ vectors and where $\delta_{i,j}$ is the Kronecker delta. We note that at each iteration, only those values that are selected to generate the filter output are updated. The block diagram of an adaptive distributed-arithmetic echo canceller with input symbols from a quaternary alphabet is shown in Fig. 30.6.

The analysis of the mean square error convergence behavior and steady-state performance has been extended to adaptive distributed-arithmetic echo cancellers [1]. The dynamics of the mean square error are, in this case, given by

$$E\{\varepsilon_n^2\} = \varepsilon_0^2\left[1 - \frac{\alpha\sigma_a^2}{2^{K-1}}(2-\alpha L\sigma_a^2)\right]^n + \frac{2\varepsilon_{\min}^2}{2-\alpha L\sigma_a^2} \tag{30.17}$$

The stability condition for the echo canceller is $0 < \alpha < 2/(L\sigma_a^2)$. For a given adaptation gain, echo canceller stability depends on the number of tables and on the variance of the transmitted symbols.

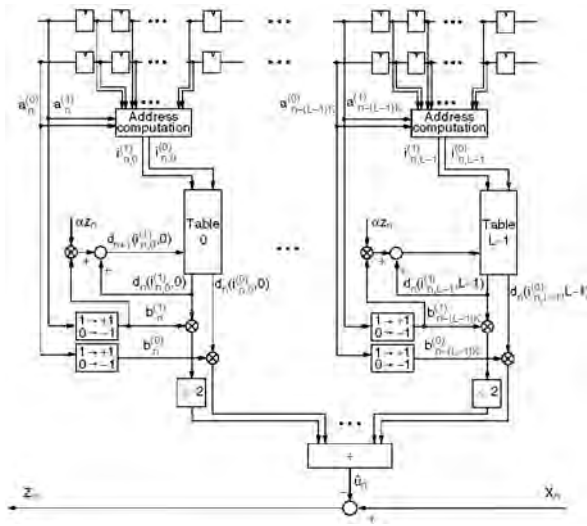


FIGURE 30.6 Block diagram of an adaptive distributed-arithmetic echo canceller.

Therefore, the time span of the echo canceller can be increased without affecting system stability, provided that the number L of tables is kept constant, In that case, however, mean square error convergence will be slower. From Eq. (30.17), we find that the optimum adaptation gain that permits the fastest mean square error convergence at the beginning of the adaptation process is $\alpha_{\text{opt}} = 1/(L\sigma_a^2)$. The time constant of the convergence mode is $\tau_{\text{opt}} = L2^{K-1}$. The smallest achievable time constant is proportional to the total number of values. The realization of a distributed-arithmetic echo canceller can be further simplified by updating at each iteration only the values that are addressed by the most significant bits of the symbols stored in the delay line. The complexity required for adaptation can thus be reduced at the price of a slower rate of convergence.

30.3 Echo Cancellation for Quadrature-Amplitude Modulation (QAM) Systems

Although most of the concepts presented in the preceding sections can be readily extended to echo cancellation for communications systems employing QAM, the case of full-duplex transmission over a voiceband data channel requires a specific discussion. We consider the system model shown in Fig. 30.7. The transmitter generates a sequence $\{a_n\}$ of i.i.d. complex-valued symbols from a two-dimensional constellation $\mathcal{A}$, which are modulated by the carrier $e^{j2\pi f_c nT}$, where T and f_c denote the modulation interval and the carrier frequency, respectively. The discrete-time signal at the output of the transmit Hilbert filter may be regarded as an analytic signal, which is generated at the rate of m/T samples/s, $m > 1$. The real part of the analytic signal is converted into an analog signal by a D/A converter and input to the channel. We note that by transmitting the real part of a complex-valued signal, positive- and negative-frequency components become folded. The image band attenuation of the transmit Hilbert filter thus determines the achievable echo suppression. In fact, the receiver cannot extract aliasing image-band components from desired passband frequency components, and the echo canceller is able to suppress only echo arising from transmitted passband components.

The output of the echo channel is represented as the sum of two contributions. The near-end echo $u^{\text{NE}}(t)$ arises from the impedance mismatch between the hybrid and the transmission line, as in the case of baseband transmission. The far-end echo $u^{\text{FE}}(t)$ represents the contribution due to echos that are generated at intermediate points in the telephone network. These echos are characterized by additional

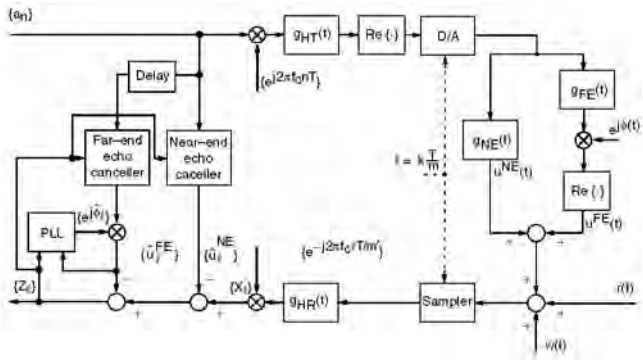


FIGURE 30.7 Configuration of an echo canceller for a QAM transmission system.

impairments, such as jitter and frequency shift, which are accounted for by introducing a carrier-phase rotation of an angle $\phi(t)$ in the model of the far-end echo.

At the receiver, samples of the signal at the channel output are obtained synchronously with the transmitter timing, at the sampling rate of m/T sample/s. The discrete-time received signal is converted to a complex-valued baseband signal $\{x_{nm'+i}, i = 0, \ldots, m' - 1\}$, at the rate of m'/T samples/s, $1 < m' < m$, through filtering by the receive Hilbert filter, decimation, and demodulation. From delayed transmit symbols, estimates of the near- and far-end echo signals after demodulation, $\{\hat{u}^{\mathrm{NE}}_{nm'+i},\ i = 0,\ldots,m' - 1\}$ and $\{\hat{u}^{\mathrm{FE}}_{nm'+i},\ i = 0,\ldots,m' - 1\}$, respectively, are generated using m' interleaved near- and far-end echo cancellers. The cancellation error is given by

$$z_\ell = x_\ell - \hat{u}^{\mathrm{NE}}_\ell - \hat{u}^{\mathrm{FE}}_\ell \tag{30.18}$$

A different model is obtained if echo cancellation is accomplished before demodulation. In this case, two equivalent configurations for the echo canceller may be considered. In one configuration, the modulated symbols are input to the transversal filter, which approximates the passband echo response. Alternatively, the modulator can be placed after the transversal filter, which is then called a baseboard transversal filter [15].

In the considered realization, the estimates of the echo signals after demodulation are given by

$$\hat{u}^{\mathrm{NE}}_{nm'+i} = \sum_{k=0}^{N_{\mathrm{NE}}-1} c^{\mathrm{NE}}_{n,km'+i}\, a_{n-k}, \quad i = 0,\ldots,m' - 1 \tag{30.19}$$

and

$$\hat{u}^{\mathrm{FE}}_{nm'+i} = \left[\sum_{k=0}^{N_{\mathrm{FE}}-1} c^{\mathrm{FE}}_{n,km'+i}\, a_{n-k-D_{\mathrm{FE}}}\right] e^{j\hat{\phi}_{nm'+i}}, \quad i = 0,\ldots,m' - 1 \tag{30.20}$$

where $(c^{\mathrm{NE}}_{n,0},\ldots,c^{\mathrm{NE}}_{n,m'N_{\mathrm{NE}-1}})$ and $(c^{\mathrm{FE}}_{n,0},\ldots,c^{\mathrm{FE}}_{n,m'N_{\mathrm{FE}}-1})$ are the coefficients of the m' interleaved near- and far-end echo cancellers, respectively, $\{\phi_{nm'+i},\ i = 0,\ldots,m' - 1\}$ is the sequence of far-end echo phase estimates, and D_{FE} denotes the bulk delay accounting for the round-trip delay from the transmitter to the point of echo generation. To prevent overlap of the time span of the near-end echo canceller with the time span of the far-end echo canceller, the condition $D_{\mathrm{FE}} > N_{\mathrm{NE}}$ must be satisfied. We also note that,

because of the different nature of near- and far-end echo generation, the time span of the far-end echo canceller needs to be larger than the time span of the near-end canceller, i.e., $N_{\mathrm{FE}} > N_{\mathrm{NE}}$.

Adaptation of the filter coefficients in the near- and far-end echo cancellers by the LMS algorithm leads to

$$\begin{aligned} c^{\mathrm{NE}}_{n+1,km'+i} &= c^{\mathrm{NE}}_{n,km'+i} + \alpha z_{nm'+i}(a_{n-k})^* \\ k &= 0,\ldots,N_{\mathrm{NE}}-1, \quad i = 0,\ldots,m'-1 \end{aligned} \tag{30.21}$$

and

$$\begin{aligned} c^{\mathrm{FE}}_{n+1,km'+i} &= c^{\mathrm{FE}}_{n,km'+i} + \alpha z_{nm'+i}(a_{n-k-D_{\mathrm{FE}}})^* e^{-j\hat{\phi}_{nm'+i}} \\ k &= 0,\ldots,N_{\mathrm{FE}}-1, \quad i = 0,\ldots,m'-1 \end{aligned} \tag{30.22}$$

respectively, where the asterisk denotes complex conjugation.

The far-end echo phase estimate is computed by a second-order phase-lock loop algorithm, where the following stochastic gradient approach is adopted:

$$\begin{cases} \hat{\phi}_{\ell+1} = \hat{\phi}_{\ell} - \frac{1}{2}\gamma_{\mathrm{FE}}\nabla_{\hat{\phi}}|z_\ell|^2 + \Delta\phi_\ell \pmod{2\pi} \\ \Delta\phi_{\ell+1} = \Delta\phi_\ell - \frac{1}{2}\zeta_{\mathrm{FE}}\nabla_{\hat{\phi}}|z_\ell|^2 \end{cases} \tag{30.23}$$

where $\ell = nm' + i$, $i = 0,\ldots,m'-1$, γ_{FE} and ζ_{FE} are step-size parameters, and

$$\nabla_{\hat{\phi}}|z_\ell|^2 = \frac{\partial|z_\ell|^2}{\partial\hat{\phi}_\ell} = -2\mathrm{Im}\{z_\ell(\hat{u}^{\mathrm{FE}}_\ell)^*\} \tag{30.24}$$

We note that algorithm (30.23) requires m' iterations per modulation interval, i.e., we cannot resort to interleaving to reduce the complexity of the computation of the far-end echo phase estimate.

30.4 Echo Cancellation for Orthogonal Frequency Division Multiplexing (OFDM) Systems

Orthogonal frequency division multiplexing (OFDM) is a modulation technique whereby blocks of M symbols are transmitted in parallel over M subchannels by employing M orthogonal subcarriers. We consider a real-valued discrete-time channel impulse response $\{h_i, i = 0,\ldots,L\}$ having length $L + 1 \ll M$. To illustrate the basic principles of OFDM systems, let us consider a noiseless ideal channel with impulse response given by $\{h_i\} = \{\delta_i\}$, where $\{\delta_i\}$ is defined as the discrete-time delta function. Modulation of the complex-valued input symbols at the nth modulation interval, denoted by the vector $\boldsymbol{A}_n = \{A_n(i), i = 0,\ldots,M-1\}$, is performed by an inverse discrete Fourier transform (IDFT), as shown in Fig. 30.8. We assume that M is even, and that each block of symbols satisfies the Hermitian symmetry conditions, i.e., $A_n(0)$ and $A_n(M/2)$ are real valued, and $A_n(i) = A_n^*(M-i)$, $i = 1,\ldots,M/2-1$. Then the signals $\boldsymbol{a}_n = \{a_n(i),$

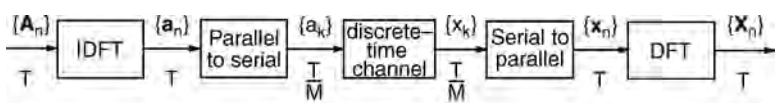


FIGURE 30.8 Block diagram of an OFDM system.

$i = 0,\ldots,M-1\}$ obtained at the output of the IDFT are real valued. After parallel-to-serial conversion, the M signals are sent over the channel at the given transmission rate, M/T, where T denotes the modulation interval. At the output of the channel, the noiseless signals are received without distortion. Serial-to-parallel conversion yields blocks of M elements, with boundaries placed such that each block obtained at the modulator output is also presented at the demodulator input. Then demodulation performed by a discrete Fourier transform (DFT) will reproduce the blocks of M input symbols. The overall input–output relationship is therefore equivalent to that of a bank of M parallel, independent subchannels.

In the general case of a noisy channel with impulse response having length greater than one, M independent subchannels are obtained by a variant of OFDM also known as discrete multitone modulation (DMT) [12]. In a DMT system, modulation by the IDFT is performed at the rate $1/T' = M/(M+L)T < 1/T$. After modulation, each block of M signals is cyclically extended by copying the last L signals in front of the block and converted from parallel to serial. The resulting $L+M$ signals are sent over the channel. At the receiver, blocks of samples with length $L+M$ are taken. Block boundaries are placed such that the last M samples depend only on the elements of one cyclically extended block of signals. The first L samples are discarded, and the vector $\boldsymbol{x}_n$ of the last M samples of the block received at the nth modulation interval can be expressed as

$$\boldsymbol{x}_n = \Gamma_n \boldsymbol{h} + \boldsymbol{w}_n \tag{30.25}$$

where $\boldsymbol{h}$ is the vector of the impulse response extended with $M-L-1$ zeros, $\boldsymbol{w}_n$ is a vector of additive white Gaussian noise samples, and Γ_n is a $M \times M$ circulant matrix given by

$$\Gamma_n = \begin{bmatrix} a_n(0) & a_n(M-1) & \cdots & a_n(1) \\ a_n(1) & a_n(0) & \cdots & a_n(2) \\ \vdots & \vdots & & \vdots \\ a_n(M-1) & a_n(M-2) & \cdots & a_n(0) \end{bmatrix} \tag{30.26}$$

Recalling that $\mathcal{F}_M \Gamma_n \mathcal{F}_M^{-1} = diag(\boldsymbol{A}_n)$, where $\mathcal{F}_M$ is the $M \times M$ DFT matrix defined as $\mathcal{F}_M = [(e^{-\frac{j2\pi}{M}})^{km}]$, k, $m = 0,\ldots,M-1$, and $diag(\boldsymbol{A}_n)$ denotes the diagonal matrix with elements on the diagonal given by $\boldsymbol{A}_n$, we find that the output of the demodulator is given by

$$\boldsymbol{X}_n = diag(\boldsymbol{A}_n)\boldsymbol{H} + \boldsymbol{W}_n \tag{30.27}$$

where $\boldsymbol{H}$ denotes the DFT of the vector $\boldsymbol{h}$, and $\boldsymbol{W}_n$ is a vector of independent Gaussian random variables. Equation (30.27) indicates that the sequence of transmitted symbol vectors can be detected by assuming a bank of M independent subchannels, at the price of a decrease in the data rate by a factor $(M+L)/M$. Note that in practice the computationally more efficient inverse fast Fourier transform and fast Fourier transform are used instead of IDFT and DFT.

We discuss echo cancellation for OFDM with reference to a DMT system [7], as shown in Fig. 30.9. The real-valued discrete-time echo impulse response is $\{h_{E,i}, i = 0,\ldots,N-1\}$, having length $N < M$. We initially assume $N \le L+1$. Futhermore, we assume that the boundaries of the received blocks are placed such that the last M samples of the nth received block are expressed by the vector

$$\boldsymbol{x}_n = \Gamma_n^R \boldsymbol{h} + \Gamma_n \boldsymbol{h}_E + \boldsymbol{w}_n \tag{30.28}$$

where Γ_n^R is the circulant matrix with elements given by the signals from the remote transmitter, and $\boldsymbol{h}_E$ is the vector of the echo impulse response extended with $M-N$ zeros. In the frequency domain, the

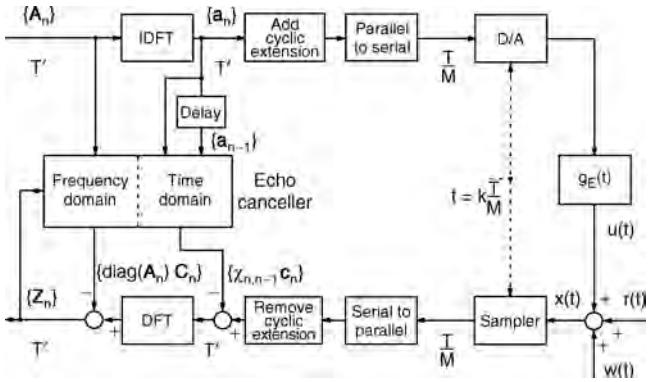


FIGURE 30.9 Configuration of an echo canceller for a DMT transmission system.

echo is expressed as $\boldsymbol{U}_n = diag(\boldsymbol{A}_n)\boldsymbol{H}_E$, where $\boldsymbol{H}_E$ denotes the DFT of the vector $\boldsymbol{h}_E$. In this case, the echo canceller provides an echo estimate that is given by $\hat{\boldsymbol{U}}_n = diag(\boldsymbol{A}_n)\boldsymbol{C}_n$, where $\boldsymbol{C}_n$ denotes the DFT of the vector $\boldsymbol{c}_n$ of the N coefficients of the echo canceller filter extended with $M - N$ zeros. In practice, however, we need to consider the case $N > L + 1$. The expression of the cancellation error is then given by

$$\boldsymbol{z}_n = \boldsymbol{x}_n - \boldsymbol{\Psi}_{n,n-1}\boldsymbol{c}_n \tag{30.29}$$

where the vector of the last M elements of the nth received block is now $\boldsymbol{x}_n = \Gamma_n^R \boldsymbol{h} + \boldsymbol{\Psi}_{n,n-1}\boldsymbol{h}_E + \boldsymbol{w}_n$, and $\boldsymbol{\Psi}_{n,n-1}$ is an $M \times M$ Toeplitz matrix given by

$$\boldsymbol{\Psi}_{n,n-1} = \begin{bmatrix} a_n(0) & a_n(M-1) & \cdots & a_n(M-L) & a_{n-1}(M-1) & \cdots & a_{n-1}(L+1) \\ a_n(1) & a_n(0) & \cdots & a_n(M-L+1) & a_n(M-L) & \cdots & a_{n-1}(L+2) \\ \vdots & & & \mathrm{O} & & & \vdots \\ a_n(M-1) & a_n(M-2) & \cdots & a_n(M-L-1) & a_n(M-L-2) & \cdots & a_n(0) \end{bmatrix} \tag{30.30}$$

In the frequency domain, the cancellation error can be expressed as

$$\boldsymbol{Z}_n = \mathcal{F}_M(\boldsymbol{x}_n - \chi_{n,n-1}\boldsymbol{c}_n) - diag(\boldsymbol{A}_n)\boldsymbol{C}_n \tag{30.31}$$

where $\chi_{n,n-1} = \boldsymbol{\Psi}_{n,n-1} - \Gamma_n$ is an $M \times M$ upper triangular Toeplitz matrix. Equation (30.31) suggests a computationally efficient, two-part echo cancellation technique. First, in the time domain, a short convolution is performed and the result is subtracted from the received signals to compensate for the insufficient length of the cyclic extension. Second, in the frequency domain, cancellation of the residual echo is performed over a set of M independent echo subchannels. Observing that Eq. (30.31) is equivalent to $\boldsymbol{Z}_n = \boldsymbol{X}_n - \tilde{\boldsymbol{\Psi}}_{n,n-1}\boldsymbol{C}_n$, where $\tilde{\boldsymbol{\Psi}}_{n,n-1} = \mathcal{F}_M \tilde{\boldsymbol{\Psi}}_{n,n-1}\mathcal{F}_M^{-1}$, the echo canceller adaptation by the LMS algorithm in the frequency domain takes the form

$$\boldsymbol{C}_{n+1} = \boldsymbol{C}_n + \alpha \tilde{\boldsymbol{\Psi}}^*_{n,n-1}\boldsymbol{Z}_n \tag{30.32}$$

where α is the adaptation gain, and $\tilde{\boldsymbol{\Psi}}^*_{n,n-1}$ denotes the transpose conjugate of $\tilde{\boldsymbol{\Psi}}_{n,n-1}$. We note that, alternatively, echo canceller adaptation may also be performed by the algorithm $\boldsymbol{C}_{n+1} = \boldsymbol{C}_n + \alpha\, diag(\boldsymbol{A}_n^*)\boldsymbol{Z}_n$, which entails a substantially lower computational complexity than the LMS algorithm, at the price of a slower rate of convergence.

In DMT systems, it is essential that the length of the channel impulse response be much less than the number of subchannels, so that the reduction in data rate due to the cyclic extension may be considered negligible. Therefore, time-domain equalization is adopted in practice to shorten the length of the channel impulse response. From Eq. (30.31), however, we observe that transceiver complexity depends on the relative lengths of the echo and of the channel impulse responses. To reduce the length of the cyclic extension as well as the computational complexity of the echo canceller, various methods have been proposed to shorten both the channel and the echo impulse responses jointly [9].

Recently, multitone modulation techniques based on filter banks have been proposed, where the M-branch filters yield very high spectral containment of individual subchannel signals [4]. The filters are frequency-shifted versions of a prototype filter, designed to reduce the interchannel interference to a level that is negligible compared to the level of other noise signals, without having to resort to time-domain equalization. Echo cancellation can be performed in this case entirely in the frequency domain by taking into account, for each subchannel, only the echo generated by transmission in the opposite direction on the same subchannel. The implementation of per-subchannel echo cancellation then requires M adaptive echo cancellers.

30.5 Summary and Conclusions

Digital signal processing techniques for echo cancellation provide large echo attenuation and eliminate the need for additional line interfaces and digital-to-analog and analog-to-digital converters that are required by echo cancellation in the analog signal domain.

The realization of digital echo cancellers in transceivers for high-speed full-duplex data transmission today is possible at a low cost thanks to the advances in VLSI technology. Digital techniques for echo cancellation are also appropriate for near-end crosstalk cancellation in transceivers for transmission over voice-grade cables at rates of several megabits per second for local-area network applications.

In voiceband modems for data transmission over the telephone network, digital techniques for echo cancellation also allow a precise tracking of the carrier phase and frequency shift of far-end echos.

References

1. Cherubini, G., Analysis of the convergence behavior of adaptive distributed-arithmetic echo cancellers, *IEEE Trans. Commun.,* 41(11), 1703–1714, 1993.
2. Cherubini, G., Ölçer, S., and Ungerboeck, G., A quaternaty partial-response class-IV transceiver for 125 Mbit/s data transmission over unshielded twisted-pair cables: principles of operation and VLSI realization, *IEEE J. Sel. Areas Commun.,* 13(9), 1656–1669, 1995.
3. Cherubini, G., Creigh, J., Ölçer, S., Rao, S.K., and Ungerboeck, G., 100BASE-T2: a new standard for 100 Mb/s Ethernet transmission over voice-grade cables, *IEEE Commun. Mag.,* 35(11), 115–122, 1997.
4. Cherubini, G., Cioffi, J.M., Eleftheriou, E., and Ölçer, S., Filter bank modulation techniques for very high-speed digital subscriber lines, *IEEE Commun. Mag.,* 38(5), 98–104, 2000.
5. Duttweiler, D.L., Adaptive filter performance with nonlinearities in the correlation multiplier, *IEEE Trans. Acoust. Speech Signal Processing,* 30(8), 578–586, 1982.
6. Falconer, D.D., Adaptive reference echo-cancellation, *IEEE Trans. Commun.,* 30(9), 2083–2094, 1982.
7. Ho, M., Cioffi, J.M., and Bingham, J.A.C., Discrete multitone echo cancellation, *IEEE Trans. Commun.,* 44(7), 817–825, 1996.
8. Lee, E.A. and Messerschmitt, D.G., *Digital Communication,* 2nd ed., Kluwer Academic Publishers, Boston, MA, 1994.
9. Melsa, P.J.W., Younce, R.C., and Rohrs, C.E., Impulse response shortening for discrete multitone transceivers, *IEEE Trans. Commun.,* 44(12), 1662–1672, 1996.
10. Messerschmitt, D.G., Echo cancellation in speech and data transmission, *IEEE J. Sel. Areas Commun.,* 2(2), 283–297, 1984.

11. Messerschmitt, D.G., Design issues for the ISDN U-Interface transceiver, *IEEE J. Sel. Areas Commun.,* 4(8), 1281–1293, 1986.
12. Ruiz, A., Cioffi, J.M., and Kasturia, S., Discrete multiple tone modulation with coset coding for the spectrally shaped channel, *IEEE Trans. Commun.,* 40(6), 1012–1029, 1992.
13. Smith, M.J., Cowan, C.F.N., and Adams, P.F., Nonlinear echo cancellers based on transpose distributed arithmetic, *IEEE Trans. Circuits and Systems,* 35(1), 6–18, 1988.
14. Ungerboeck, G., Theory on the speed of convergence in adaptive equalizers for digital communication, *IBM J. Res. Develop.,* 16(6), 546–55, 1972.
15. Weinstein, S.B., A passband data-driven echo-canceller for full-duplex transmission on two-wire circuits, *IEEE Trans. Commun.,* 25(7), 654–666, 1977.

III

Networks

31

The Open Systems Interconnections (OSI) Seven-Layer Model

Frederick Halsall
University of Wales, Swansea

31.1 Computer Communications Requirements

Although in many instances computers are used to perform their intended role in a stand-alone mode, in others there is a need to interwork and exchange data with other computers—in financial applications, for example, to carry out fund transfers from one institution computer to another, in travel applications to access the reservation systems belonging to various airlines, and so on. The general requirement in all of these applications is for application programs running in different computers to cooperate to achieve a specific distributed application function. To achieve this, three basic issues must be considered. These are shown in diagrammatic form in Fig. 31.1.

The fundamental requirement in all applications that involve two or more computers is the provision of a suitable data communications facility. This may comprise a **local area network** (LAN), if the computers are distributed around a single site, a **wide area network** (WAN), if the computers are situated at different sites, or an **internetwork**, if multiple interconnected network types are involved. Associated with these different network types is a set of access protocols which enables a communications path between two computers to be established and for data to be transferred across this path. Typically, these protocols differ for the different network types. In addition to these access protocols, the communication subsystem in each computer must provide additional functionality. For example, if the communicating computers are of different types, possibly with different word sizes and character sets, then a means of ensuring the transferred data is interpreted in the same way in each computer must be incorporated. Also, the computers may use different file systems and, hence, functionality to enable application programs, normally referred to as **application processes** (APs), to access these in a standardized way must also be included. All of these issues must be considered when communicating data between two computers.

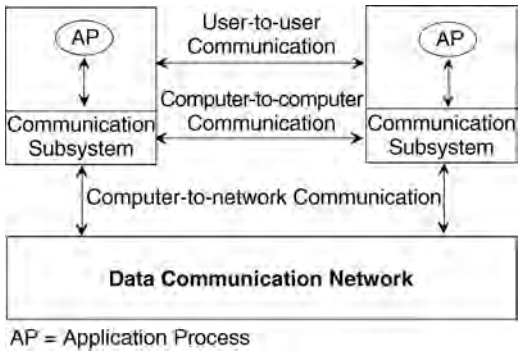


FIGURE 31.1 Computer communication schematic.

31.2 Standards Evolution

Until recently, the standards established for use in the computer industry by the various international bodies were concerned primarily with either the internal operation of a computer or the connection of a local peripheral device. The result was that early hardware and software communication subsystems offered by manufacturers only enabled their own computers, and so-called *plug-compatible systems*, to exchange information. Such systems are known as *closed systems* since computers from other manufacturers cannot exchange information unless they adhere to the (proprietary) standards of a particular manufacturer.

In contrast, the various international bodies concerned with public-carrier telecommunication networks have, for many years, formulated internationally agreed standards for connecting devices to these networks. The *V-series recommendations,* for example, are concerned with the connection of equipment, normally referred to as a *data terminal equipment* (DTE), to a modem connected to the public switched telephone network (PSTN); the *X-series recommendations* for connecting a DTE to a public data network; and the *I-series recommendations* for connecting a DTE to the integrated services digital networks (ISDNs). The recommendations have resulted in compatibility between the equipment from different vendors, enabling a purchaser to select suitable equipment from a range of manufacturers.

Initially, the services provided by most public carriers were concerned primarily with data transmission, and, hence, the associated standards only related to the method of interfacing a device to these networks. More recently, however, the public carriers have started to provide more extensive distributed information services, such as the exchange of electronic messages (*teletex*) and access to public databases (*videotex*). To cater to such services, the standards bodies associated with the telecommunications industry have formulated standards not only for interfacing to such networks but also so-called higher level standards concerned with the format (syntax) and control of the exchange of information (data) between systems. Consequently, the equipment from one manufacturer that adheres to these standards can be interchangeable with equipment from any other manufacturer that complies with the standards. The resulting system is then known as an *open system* or, more completely, as an *open systems interconnection environment* (*OSIE*).

In the mid-1970s, as different types of distributed systems (based on both public and private data networks) started to proliferate, the potential advantages of open systems were acknowledged by the computer industry. As a result, a range of standards started to be introduced. The first was concerned with the overall structure of the complete communication subsystem within each computer. This was produced by the International Standards Organization (ISO) and is known as the *ISO reference model* for **open systems interconnection** (OSI).

The aim of the ISO reference model is to provide a framework for the coordination of standards development and to allow existing and evolving standards activities to be set within a common framework.

The aim is to allow an application process in any computer that supports a particular set of standards to communicate freely with an application process in any other computer that supports the same standards, irrespective of its origin of manufacture.

Some examples of application processes that may wish to communicate in an open way are the following:

- a process (program) executing in a computer and accessing a remote file system
- a process acting as a central file service (server) to a distributed community of (client) processes
- a process in an office workstation (computer) accessing an electronic mail service
- a process acting as an electronic mail server to a distributed community of (client) processes
- a process in a supervisory computer controlling a distributed community of computer-based instruments or robot controllers associated with a process or automated manufacturing plant
- a process in an instrument or robot controller receiving commands and returning results to a supervisory system
- a process in a bank computer that initiates debit and credit operations on a remote system

Open systems interconnection is concerned with the exchange of information between such processes. The aim is to enable application processes to cooperate in carrying out a particular (distributed) information processing task irrespective of the computers on which they are running.

31.3 International Standards Organization Reference Model

A communication subsystem is a complex piece of hardware and software. Early attempts at implementing the software for such subsystems were often based on a single, complex, unstructured program (normally written in assembly language) with many interacting components. The resulting software was difficult to test and often very difficult to modify.

To overcome this problem, the ISO has adopted a layered approach for the reference model. The complete communication subsystem is broken down into a number of layers, each of which performs a well-defined function. Conceptually, these layers can be considered as performing one of two generic functions, network-dependent functions and application-oriented functions. This, in turn, gives rise to three distinct operational environments:

1. The **network environment** is concerned with the protocols and standards relating to the different types of underlying data communication networks.
2. The **OSI environment** embraces the network environment and adds additional application-oriented protocols and standards to allow end systems (computers) to communicate with one another in an open way.
3. The **real systems environment** builds on the OSI environment and is concerned with a manufacturer's own proprietary software and services which have been developed to perform a particular distributed information processing task.

This is shown in diagrammatic form in Fig. 31.2.

Both the network-dependent and application-oriented (network-independent) components of the OSI model are implemented as a number of layers. The boundaries between each layer and the functions performed by each layer have been selected on the basis of experience gained during earlier standardization activity.

Each layer performs a well-defined function in the context of the overall communication subsystem. It operates according to a defined protocol by exchanging messages, both user data and additional control information, with a corresponding peer layer in a remote system. Each layer has a well-defined interface between itself and the layer immediately above and below. Consequently, the implementation of a particular protocol layer is independent of all other layers.

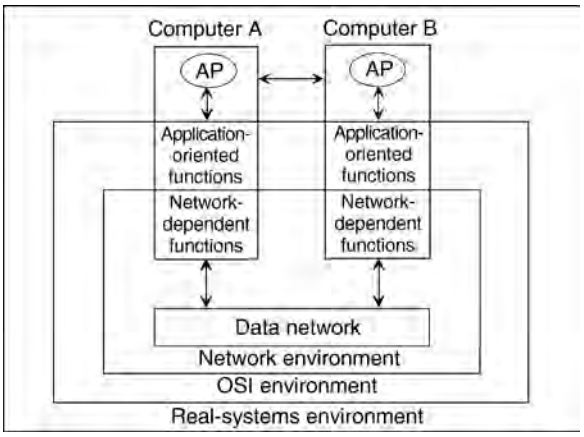


FIGURE 31.2 Operational environments.

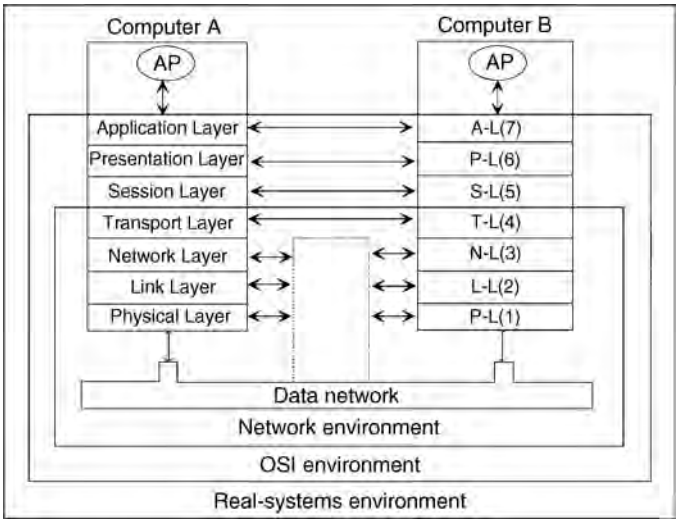


FIGURE 31.3 Overall structure of the ISO reference model.

The logical structure of the ISO reference model is made up of seven protocol layers, as shown in Fig. 31.3. The three lowest layers (1–3) are network dependent and are concerned with the protocols associated with the data communication network being used to link the two communicating computers. In contrast, the three upper layers (5–7) are application oriented and are concerned with the protocols that allow two end user application processes to interact with each other, normally through a range of services offered by the local operating system. The intermediate transport layer (4) masks the upper application-oriented layers from the detailed operation of the lower network-dependent layers. Essentially, it builds on the services provided by the latter to provide the application-oriented layers with a network-independent message interchange service.

The function of each layer is specified formally as a protocol that defines the set of rules and conventions used by the layer to communicate with a similar peer layer in another (remote) system. Each layer provides a defined set of services to the layer immediately above. It also uses the services provided by the layer immediately below it to transport the message units associated with the protocol to the remote peer layer. For example, the transport layer provides a network-independent message transport service to the session layer above it and uses the service provided by the network layer below it to transfer the set of message

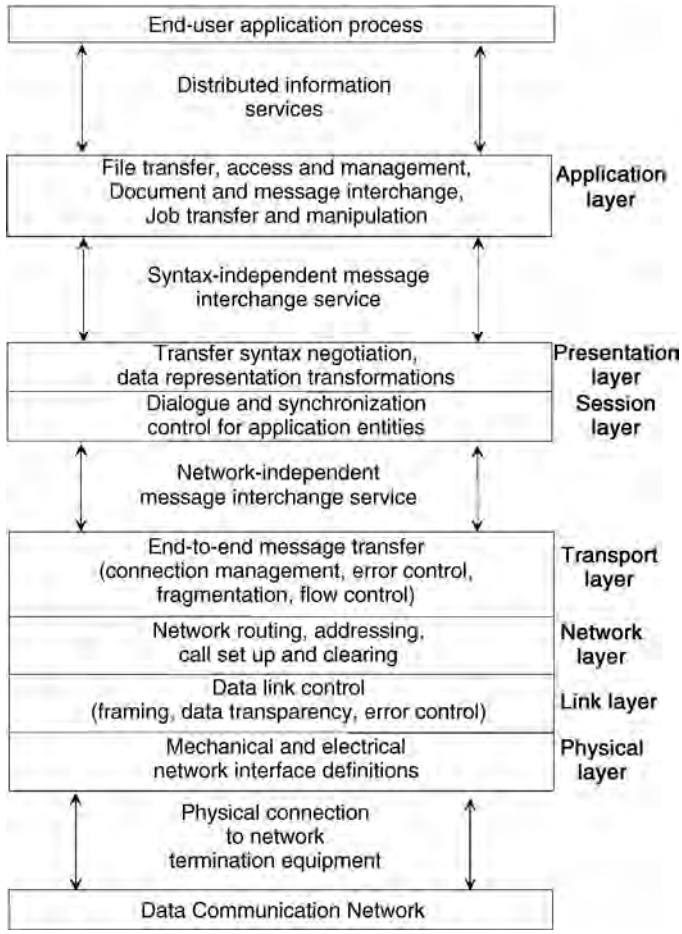


FIGURE 31.4 Protocol layer summary.

units associated with the transport protocol to a peer transport layer in another system. Conceptually, therefore, each layer communicates with a similar peer layer in a remote system according to a defined protocol. However, in practice, the resulting protocol message units of the layer are passed by means of the services provided by the next lower layer. The basic functions of each layer are summarized in Fig. 31.4.

The Application-Oriented Layers

The Application Layer

The application layer provides the user interface, normally an application program/process, to a range of networkwide distributed information services. These include file transfer access and management, as well as general document and message interchange services such as electronic mail. A number of standard protocols are either available or are being developed for these and other types of service.

Access to application services is normally achieved through a defined set of primitives, each with associated parameters, which are supported by the local operating system. The access primitives are the same as other operating system calls (as used for access to, say, a local file system) and result in an appropriate operating system procedure (process) being activated. These operating system procedures use the communication subsystem (software and hardware) as if it is a local device, similar to a disk controller, for example. The detailed operation and implementation of the communication subsystem is thus transparent to the (user) application process. When the application process making the call is

rescheduled (run), one (or more) status parameters are returned, indicating the success (or otherwise) of the network transaction that has been attempted.

In addition to information transfer, the application layer provides such services as:

- identification of the intended communication partner(s) by name or by address
- determination of the current availability of an intended communication partner
- establishment of authority to communicate
- agreement on privacy (encryption) mechanisms
- authentication of an intended communication partner
- selection of the dialogue discipline, including the initiation and release procedures
- agreement on responsibility for error recovery
- identification of constraints on data syntax (character sets, data structures, etc.)

The Presentation Layer

The presentation layer is concerned with the representation (syntax) of data during transfer between two communicating application processes. To achieve true open systems interconnection, a number of common *abstract data syntax* forms have been defined for use by application processes together with associated *transfer* (or *concrete*) *syntaxes*. The presentation layer negotiates and selects the appropriate transfer syntax(es) to be used during a transaction so that the syntax (structure) of the messages being exchanged between two application entities is maintained. Then, if this form of representation is different from the internal abstract form, the presentation entity performs the necessary conversion.

To illustrate the services provided by the presentation layer, consider a telephone conversation between a French speaking person and a Spanish speaking person. Assume each uses an interpreter and that the only language understood by both interpreters is English. Each interpreter must translate from their local language to English, and vice versa. The two correspondents are thus analogous to two application processes with the two interpreters representing presentation layer entities. French and Spanish are the local syntaxes and English is the transfer or concrete syntax. Note that there must be a universally understood language which must be defined to allow the agreed transfer language (syntax) to be negotiated. Also note that the interpreters do not necessarily understand the meaning (semantics) of the conversation.

Another function of the presentation layer is concerned with data security. In some applications, data sent by an application is first encrypted (enciphered) using a key, which is (hopefully) known only by the intended recipient presentation layer. The latter decrypts (deciphers) any received data using the corresponding key before passing it on to the intended recipient.

The Session Layer

The session layer provides the means that enables two application layer protocol entities to organize and synchronize their dialogue and manage their data exchange. It is thus responsible for setting up (and clearing) a communication (dialogue) channel between two communicating application layer protocol entities (presentation layer protocol entities in practice) for the duration of the complete network transaction. A number of optional services are provided, including the following:

- Interaction management: The data exchange associated with a dialogue may be duplex or half-duplex. In the latter case, it provides facilities for controlling the exchange of data (dialogue units) in a synchronized way.
- Synchronization: For lengthy network transactions, the user (through the services provided by the session layer) may choose periodically to establish synchronization points associated with the transfer. Then, should a fault develop during a transaction, the dialogue may be restarted at an agreed (earlier) synchronization point.
- Exception reporting: Nonrecoverable exceptions arising during a transaction can be signaled to the application layer by the session layer.

The Transport Layer

The transport layer acts as the interface between the higher application-oriented layers and the underlying network-dependent protocol layers. It provides the session layer with a message transfer facility that is independent of the underlying network type. By providing the session layer with a defined set of message transfer facilities, the transport layer hides the detailed operation of the underlying network from the session layer.

The transport layer offers a number of **classes of service** which cater to the varying *quality of service* (*QOS*) provided by different types of network.

There are five classes of service ranging from class 0, which provides only the basic functions needed for connection establishment and data transfer, to class 4, which provides full error control and flow control procedures. As an example, class 0 may be selected for use with a packet-switched data network (PSDN), whereas class 4 may be used with a local area network (LAN) providing a best-try service; that is, if errors are detected in a frame, then the frame is simply discarded.

The Network-Dependent Layers

As the lowest three layers of the ISO reference model are network dependent, their detailed operation varies from one network type to another. In general, however, the *network layer* is responsible for establishing and clearing a networkwide connection between two transport layer protocol entities. It includes such facilities as network routing (addressing) and, in some instances, flow control across the computer-to-network interface. In the case of internetworking, it provides various harmonizing functions between the interconnected networks.

The *link layer* builds on the physical connection provided by the particular network to provide the network layer with a reliable information transfer facility. It is thus responsible for such functions as error detection and, in the event of transmission errors, the retransmission of messages. Normally, two types of service are provided:

1. *Connectionless service* treats each information frame as a self-contained entity that is transferred using a best-try approach.
2. *Connection-oriented service* endeavors to provide an error-free information transfer facility.

Finally, the *physical layer* is concerned with the physical and electrical interfaces between the user equipment and the network terminating equipment. It provides the link layer with a means of transmitting a serial bit stream between the two equipments.

31.4 Open System Standards

The ISO reference model has been formulated simply as a template for the structure of a communication subsystem on which standards activities associated with each layer may be based. It is not intended that there should be a single standard protocol associated with each layer. Rather, a set of standards is associated with each layer, each offering different levels of functionality. Then, for a specific open systems interconnection environment, such as that linking numerous computer-based systems in a fully automated manufacturing plant, a selected set of standards is defined for use by all systems in that environment.

The three major international bodies actively producing standards for computer communications are the ISO, the American Institute of Electrical and Electronic Engineers (IEEE), and the International Telegraph and Telephone Consultative Committee (CCITT). Essentially, the ISO and IEEE produce standards for use by computer manufacturers, whereas the CCITT defines standards for connecting equipment to the different types of national and international public networks. As the degree of overlap between the computer and telecommunications industries increases, however, there is an increasing level of cooperation and commonality between the standards produced by these organizations.

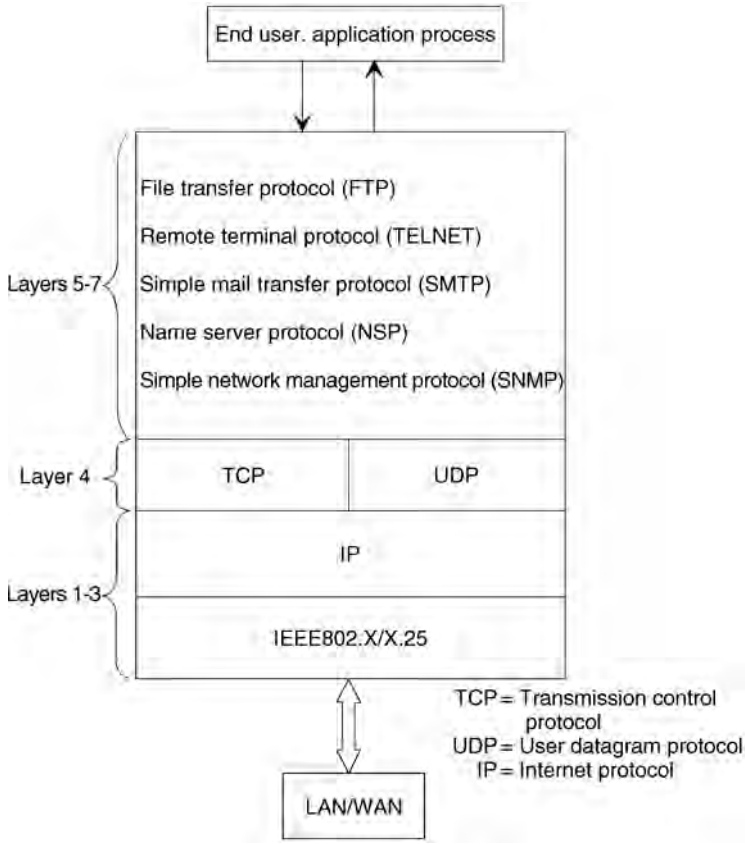


FIGURE 31.5 TCP/IP protocol suite.

In addition, prior to and concurrent with ISO standards activity, the U.S. Department of Defense has for many years funded research into computer communications and networking through its Defense Advanced Research Projects Agency (DARPA). As part of this research, the computer networks associated with a large number of universities and other research establishments were linked to those of DARPA. The resulting internetwork, known as ARPANET, has been extended to incorporate internets developed by other government agencies. The combined internet is now known simply as the Internet.

The protocol suite used with the Internet is known as t**ransmission control protocol/internet protocol (TCP/IP)**. It includes both network-oriented protocols and application support protocols. Because TCP/IP is in widespread use with an existing internet, many of the TCP/IP protocols have been used as the basis for ISO standards. Moreover, since all of the protocol specifications associated with TCP/IP are in the public domain, and hence no licence fees are payable, they have been used extensively by commercial and public authorities for creating open system networking environments. In practice, therefore, there are two major open system (vendor-independent) standards: the TCP/IP protocol suite and those based on the evolving ISO standards.

Figure 31.5 shows some of the standards associated with the TCP/IP protocol suite. As can be seen, since TCP/IP has developed concurrently with the ISO initiative, it does not contain specific protocols relating to all of the seven layers in the OSI model.

Moreover, the specification methodology used for the TCP/IP protocols differs from that used for the ISO standards. Nevertheless, most of the functionality associated with the ISO layers is embedded in the TCP/IP suite. A range of standards has been defined by the ISO/CCITT and a selection of these is shown in Fig. 31.5. Collectively, they enable the administrative authority that is establishing the open system

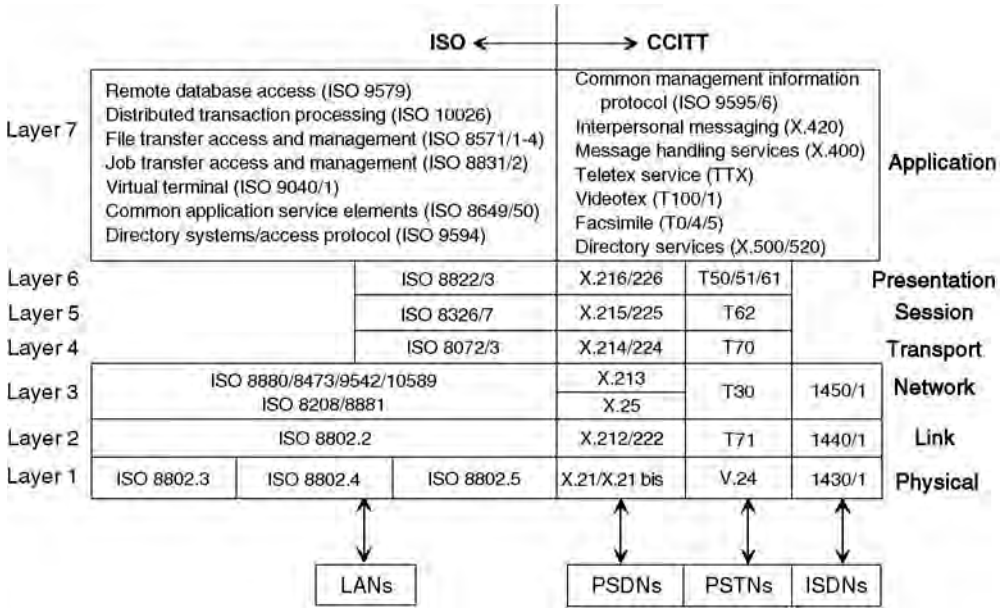


FIGURE 31.6 Standards summary.

environment to select the most suitable set of standards for the application. The resulting protocol suite is known as the *open system interconnection profile*. A number of such profiles have now been defined, including TOP, a protocol set for use in technical and office environments, MAP, for use in manufacturing automation, U.S. and U.K. GOSIP, for use in U.S. and U.K. government projects, respectively, and a similar suite used in Europe known as the CEN functional standards. The latter has been defined by the Standards Promotion and Application Group (SPAG), a group of 12 European companies.

As Fig. 31.6 shows, the lower three layers vary for different network types. CCITT has defined V, X, and I series standards for use with public-carrier networks. The V series is for use with the existing switched telephone network (PSTN), the X series for use with existing switched data networks (PSDN), and the I series for use with the integrated services digital networks (ISDN). Those produced by ISO/IEEE for use with local area networks are known collectively as the 802 (IEEE) or 8802 (ISO) series.

31.5 Summary

This chapter has reviewed the requirements for the communications subsystem in each of a set of interconnected computers that enables them to communicate in an open way to perform various distributed application functions. The philosophy behind the structure of the ISO reference model for open systems interconnection has been presented and the functionality of the seven layers that make up the reference model has been described. Finally, a selection of the ISO/CCITT standards that were defined have been identified.

Defining Terms

Internetwork: A collection of different network types that are linked together.
Local area network (LAN): Network used to link computers that are distributed around a single site.
Open system interconnection (OSI): A standard set of communication protocols that are independent of a specific manufacturer.

Transmission control protocol/Internet protocol (TCP/IP): The standard set of communication protocols that are used by each computer that is connected to the Internet.

Wide area network (WAN): Network used to link computers that are situated at different sites.

Further Information

Further coverage of the material covered in this chapter can be found in:

Halsall, F., *Data Communications, Computer Networks and Open Systems,* 4th ed., Addison Wesley, Reading, MA, 1996.

32

Ethernet Networks

Ramesh R. Rao
University of California

32.1 Overview

The term ethernet describes a collection of hardware, software, and control algorithms that together provide a technique for interconnecting dozens of nodes, spread over hundreds of meters, at aggregate rates ranging from 1 to 100 Mb/s. These nodes are typically computer work stations or peripheral devices that are part of a community of users that frequently exchange files, messages, and other types of data among each other. Shielded coaxial cables and unshielded twisted pairs have been used for the physical interconnection.

The primary hardware element used in an ethernet network is the *network interface card* (NIC). These cards are attached to a computer bus at one end and to a *media attachment unit* (MAU) at the other end via an attachment unit interface (AUI) cable. An MAU is an active device that includes transceivers and other elements matched to the nature of the physical transmission medium. Some vendors consolidate two or more of these elements into a single package, thereby rendering certain elements, such as AUI cables, unnecessary. *Repeaters* are sometimes used to regenerate signals to ensure reliable communication. The NICs perform a number of control functions in software; chief amongst these is the execution of a random access protocol for exchanging packets of information.

Nodes attached to an ethernet network exchange packets of information by broadcasting on a common communication channel. Simultaneous transmission by two or more nodes results in a detectable collision that triggers a collision resolution protocol. This decentralized protocol induces a stochastic rescheduling of the retransmission instants of the colliding nodes. If the overall traffic on the network is below a certain threshold, eventual delivery of all new and previously collided messages is guaranteed. The value of the threshold establishes the maximum aggregate information transfer rate that can be sustained. This threshold is influenced by the span of the network, the length of the packets, latency in collision detection, and the dynamics of the external traffic, as well as the rate at which bits can be transported over the physical medium. The aggregate information transfer rate can never exceed the bit transfer rate on the physical medium, but understanding the impact of the other variables has been the subject of many studies.

Standards bodies stipulate the values of some of these variables. For example, both the 10BASE2 and 1BASE5 versions of the IEEE 802.3 standards restrict packet lengths to values between 512 and 12,144 b. The 10BASE2 version restricts the network span to 925 m and relies on a bit transfer rate of 10 Mb/s over 50-W coaxial cable with a diameter of 5 mm, whereas 1BASE5 restricts the span to 2.5 kms and

relies on a bit transfer rate of 1 Mb/s over unshielded twisted pair. It is not uncommon to observe aggregate information transfer rates of 1.5 Mb/s or less in a 10BASE2 installation.

32.2 Historical Development

Ethernet evolved from a random access protocol that was developed at the University of Hawaii in the 1970s for use in the packet-switched Aloha broadcast radio network. The Aloha network was meant to serve many users that generated infrequent and irregularly spaced bursts of data. Techniques such as synchronous time-division multiplexing that require nodes to access a common channel only during precoordinated, nonoverlapping intervals of time are ill suited to serving many bursty users. This drawback motivated the development of the Aloha random access protocol, in which users access a common channel freely and invoke a collision resolution process only when collisions occur.

Spurred by the success of Aloha, efforts were undertaken to make the protocol more efficient. Early refinements included the use of a common clock to avoid partial collisions (slotted Aloha) as well as the development of hybrid protocols that combined elements of random access with reservations (reservation Aloha). Another series of refinements became possible in the context of limited span local area networks (LANs) that use low-noise communication links, such as coaxial cables, for the transmission medium. In such environments, a node can accurately and quickly sense ongoing transmissions on the cable. This first led to the development of carrier sense multiple access (CSMA) in which nodes withhold transmissions whenever they sense activity on the channel, thereby reducing the occurrence of collisions and improving efficiency. The next refinement for the LAN environment was based on a node's ability to detect a collision in less time than is required for a complete packet transmission. Early collision detection allows the colliding users to abort the transmission of colliding packets, thereby reducing the duration of collisions and further improving efficiency.

The ethernet protocol uses carrier sense multiple access with collision detection (CSMA/CD). The topology of an ethernet network may differ from the Aloha packet radio network. The Aloha packet radio network was configured with a controller node at the center of the network. The member nodes transmitted to the controller on a common inbound frequency band. On a different outbound frequency band, the controller broadcasts to the member nodes all of the packets it received as well as information, from which the occurrence of a collision could be inferred. In contrast, in some types of ethernet networks, nodes attach to a common communication channel and directly monitor it for carrier sensing and collision detection, as well as to transmit and receive packets. There is, thus, no central controller in some ethernet networks.

Ethernet was conceived at Xerox Palo Alto Research Center (PARC) in 1973 and described in a 1976 Association for Computing Machinery (ACM) article by Robert M. Metcalfe and David R. Boggs. The first prototype, also developed at Xerox PARC, operated at a speed of 3 Mb/s. In 1980, the Digital-Intel-Xerox 10 Mb/s ethernet was announced. The non-proprietary nature of the technology contributed to its widespread deployment and use in local area networks. By the mid-1980s, the Institute of Electrical and Electronics Engineers/American National Standards Institute (IEEE/ANSI) standards effort resulted in a set of CSMA/CD standards that were all based on the original ethernet concept.

32.3 Standards

In the terminology of the open systems interconnections (OSI) protocol architecture, the 802.3 series of standards are media access control (MAC) protocols. There exists a separate logical link control (LLC) sublayer, called the 802.2, which is above the MAC layer and, together with the MAC layer, forms a data link layer.

Currently, there are five IEEE/ANSI MAC standards in the 802.3 series. They are designated as 10BASE5, 10BASE2, 10BROAD36, 1BASE5, and 10BASET. A new version called 100BASET is currently under study. The numeric prefix in these designations refers to the speed in megabits per second. The term BASE refers

to baseband transmissions and the term BROAD refers to transmission over a broadband medium. The numeric suffix, when it appears, is related to limits on network span.

All 802.3 MAC standards share certain common features. Chief among them is a common MAC framing format for the exchange of information. This format includes a fixed preamble for synchronization, destination and source addresses for managing frame exchanges, a start of frame delimiter and length of frame field to track the boundaries of the frames, a frame check sequence to detect errors, and a data field and pad field, which have a role related to collision detection, which is explained further in the section on operation. In addition to a common framing format, all of the 802.3 MAC protocols constrain the frame lengths to be between 512 and 12,144 b. A 32-b jam pattern is used to enforce collisions. An interframe gap of 9.6 μs is used with the the 10-Mb/s standard and a gap of 96 μs is used with the 1 Mb/s standard. The parameters that characterize the retransmission backoff algorithm are also the same.

The primary difference between the various 802.3 standards is with regard to the physical communication medium used and consequences thereof.

10BASE5

This was the first of the IEEE/ANSI 802.3 standards to be finalized. A 10-mm coaxial cable with a characteristic impedance of 50 W is the physical medium. The length of any one segment cannot exceed 500 m, but a maximum of five segments can be attached together, through the use of four repeaters, to create a network span of 2.5 km. There can be no more than 100 attachments to a segment, and the distance between successive attachments to this cable must be at points that are multiples of 2.5 m to prevent reflections from adjacent taps from adding in phase. In view of the inflexibility of the 10-mm cable, a separate transceiver cable is used to connect the network interface cards to the coaxial cable. Transmitting nodes detect collisions when voltages in excess of the amount that can be attributed to their transmission are measured. Nontransmitting nodes can detect collisions when voltages in excess of the amount that can be produced by any one node is detected.

10BASE2

10BASE2 networks are based on a coaxial cable that has a characteristic impedance of 50 W and a diameter of 5 mm. The flexibility of the thinner cable makes it possible to dispense with the transceiver cable and directly connect the coaxial cable to the network interface card through the use of a *t*-connector. Furthermore, nodes can be spaced as close as 0.5 m. On the other hand, signals on the thinner cable are less immune to noise and suffer greater degradation. Consequently, segment lengths cannot exceed 185 m and no more than 30 nodes can be attached to a segment. As with 10BASE5, up to five segments can be attached together with repeaters to create a maximum network span of 925 m. The 10BASE2 and 10BASE5 segments can be attached together as long as the less robust 10BASE2 segments are at the periphery of the network.

10BROAD36

The 10BROAD36 networks are based on standard cable TV coaxial cables with a characteristic impedance of 75 W. The 10BROAD36 networks operate in a manner that is closer to the original Aloha network than 10BASE2 or 10BASE5. Nodes on a 10BROAD36 generate differential phaseshift keyed (DPSK) RF signals. A head-end station receives these signals and repeats them on a different frequency band. Thus, sensing and transmissions are done on two different channels just as in the Aloha network. The distance between a node and the head-end station can be as much as 1.8 km for a total network span of 3.6 km. A separate band of frequencies is set aside for the transmission of a collision enforcement signal. Interestingly, in this environment, there is a possibility that when two signals collide, one of the two may be correctly captured by some or even all of the stations. Nonetheless, 10BROAD36 requires that transmitting

stations compare their uplink transmissions with the downlink head-end transmission and broadcast a collision enforcement signal whenever differences are observed.

1BASE5

Also referred to as StarLAN, 1BASE5 networks operate at the lower rate of 1 Mb/s over unshielded twisted pair cables whose diameters range between 0.4 and 0.6 mm. The 1BASE5 networks are physically configured in the shape of a star. Each node is attached to a hub via two separate pairs of unshielded twisted wires: one pair is for transmissions to the hub, and the other pair is for receptions from the hub. Neither of these pairs is shared with any other node. A hub has multiple nodes attached to it and detects collisions by sensing activity on more than one incoming port. It is responsible for broadcasting all correctly received packets on all outgoing links or broadcasting a collision presence signal when one is detected. The 1BASE5 standard does not constrain the number of nodes that can be attached to a hub.

It is also possible to cascade hubs in a five-layer hierarchy with one header hub and a number of intermediate hubs. The node-to-hub distance and the inter-hub distances are limited to 250 m and, therefore the largest network span that can be created is limited to 2.5 kms. In view of the lower speed of 1BASE5 networks, they can not be easily integrated with the other IEEE 802.3 networks.

10BASET

The 10BASET network is similar to the older 1BASE5 network in that both use unshielded twisted pairs, but 10BASET networks can sustain data rates of 10 Mb/s. This is accomplished partly by limiting individual segment lengths to 100 m. Instead of the hierarchical hubs used in 1BASE5, 10BASET uses multiport repeater sets, which are connected together in exactly the same way as nodes are connected to a repeater set. Two nodes on a 10BASET network can be separated by up to four repeater sets and up to five segments. Of the five segments, no more than three can be coaxial cable-based 10BASE2 or 10BASE5 segments; the remaining segments can be either unshielded twisted pair, point-to-point links of length less than 100 m each, or fiber optic point-to-point links of length less than 500 m each. The network span that can be supported depends on the mix of repeater sets and segments used.

100BASET

Currently, there is an effort to create a 100-Mb/s standard in the 802.3 series. The 100 BASET network will use multiple pairs of unshielded twisted pairs, a more efficient code for encoding the data stream, and a slightly faster clock rate to generate the tenfold increase in bit transfer rates.

32.4 Operation

The core algorithm used in all of the 802.3 series of MAC standards and the original ethernet is essentially the same. The MAC layer receives LLC data and encapsulates into MAC frames. Of particular interest in the MAC frame is the pad field, whose role is related to collision detection and will be explained. Prior to transmitting frames, the MAC layer senses the physical channel. If transmission activity is detected, the node continues to monitor until a break is detected, at which time the node transmits its frame. In spite of the prior channel sensing performed, it is possible that the frame will collide with one or more other frames. Such collisions may occur either because two or more nodes concurrently initiate transmissions following the end of the previous transmission or because propagation delays prevent a node from sensing the transmission initiated by another node at an earlier instant.

Colliding users detect collisions through physical mechanisms that depend on the nature of the attachment to the underlying communication medium. As described in the section on standards, in

10BASE2 and 10BASE5 networks, transmitting nodes recognize electrical currents in excess of what they injected into the coaxial cable. In 1BASE5 and 10BASET networks, nodes recognize activity on both the transmitter and receiver ports, and in 10BROAD36 networks, nodes spot differences in the transmitted and received waveforms. All of these collision sensing methods rely on a node's knowledge of its own transmission activity. Since this information is not available to all of the other nodes, the disposition of some partially received MAC frames may be in doubt. Therefore, to enforce a common view of the channel outcome, colliding users transmit a jam pattern as soon as they sense a collision.

Thus, the full duration of a collision is the sum of the maximum collision detect time and the jam pattern. The first of these two components, the time to detect a collision, is primarily determined by the time it takes for signals to propagate between any two nodes on the cable. By constraining the span of a network, standards such as the IEEE 802.3 limit the maximum collision detect time. The second component, the jam pattern length, is also constrained by standards. Thus, the longest period of collision activity, which can be thought of as an invalid frame, is limited. The pad field of the MAC frame can then be used to ensure that valid frames are larger than the longest invalid frame that may occur as a result of collisions. This provides a simple mechanism by which nodes that were not involved in a collision can also sense the collision.

In order to resolve collisions, ethernet nodes retransmit after a randomly chosen delay. Further more, the probability distribution of the retransmission delay of a given packet is updated each time the packet experiences a collision. The choice of the initial probability distribution of the retransmission delay and the dynamics of its updates determines whether or not the total population of new and retransmitted users can transmit successfully. Thus, the retransmission algorithm lies at the heart of the ethernet protocol.

The IEEE 802.3 standard specifies the use of a truncated binary exponential backoff algorithm. In this algorithm, a packet that has collided n times is withheld for a duration that is proportional to a quantity that is uniformly distributed between 0 and $2^m - 1$, where $m = \min(10, n)$. The constant of proportionality is chosen to be large enough to ensure that nodes that choose to withhold for different durations do not collide, regardless of their location on the cable. Furthermore, when the number of times a packet has collided exceeds a user-defined threshold, which is commonly chosen to be 16, it is rejected by the MAC layer and has to be resubmitted at a later time by a higher layer.

As a consequence of the collision resolution process, the order in which ethernet serves packets depends on the traffic intensity. During periods of very light traffic, service is first come first served. During periods of heavy contention, some colliding users will defer their transmissions into the future even as some later arrivals transmit sooner. Because of this, packet delivery in ethernet may occur out of sequence and has to be rectified at the receiver.

The lack of first come first served service and the stochastic nature of the rescheduling algorithm are sometimes considered to be weaknesses of the ethernet network. On the other hand, the ability to use unshielded twisted pair wires for 10-Mb/s communication, the large installed base of users, and the promise of even faster communication at rates of 100 Mb/s over multiple pairs of unshielded twisted pair wires makes ethernet networks very attractive.

References

Abramson, N., The Aloha system—another alternative for computer communications, in *Proceedings of the Fall Joint Computer Conference, AFIPS Conference,* Vol. 37, 281–285, 1970.

Metcalfe, R.M. and Boggs, D.R., Ethernet: distributed packet switching for local computer networks, *Commun. ACM* (reprinted in *The Ethernet Source Book*), 3–12, 1976.

Shotwell, R., Ed., *The Ethernet Source Book,* North Holland, New York, 1985.

Stallings, W., *Handbook of Computer-Communications Standards,* 2nd ed., H.W. Sams, Carmel, IN, 1989–1990.

Walrand, J., *Communication Networks: A First Course,* Aksen Associates, Irwin, Homewood, IL, 1991.

Further Information

The Ethernet Source Book (1985) is a compilation of papers on various aspects of ethernet including the first published paper on ethernet entitled "Ethernet: Distributed Packet Switching for Local Computer Networks" by R.M. Metcalfe and D.R. Boggs. *The Ethernet Source Book* includes an introduction by Robert M. Metcalfe who, along with David Boggs, Chuck Thacker, and Tat Lampson, holds the patent on "Multipoint Data Communications System with Collision Detection." The *Handbook of Computer Communication Standards*, Vol. 2, by William Stallings, describes most of the details of the IEEE 802.3 standards. N. Abramson described the ALOHA system in an article entitled "The ALOHA System—Another Alternative for Computer Communications." *Communication Networks—A First Course*, by Jean Walrand, is an accessible introductory book on all aspects of networks.

33

Fiber Distributed Data Interface and Its Use for Time-Critical Applications*

Biao Chen
University of Texas

Nicholas Malcolm
Hewlett-Packard (Canada) Ltd.

Wei Zhao
Texas A&M University

33.1 Introduction

Fiber distributed data interface (FDDI) is an American National Standards Institute (ANSI) standard for a 100-Mb/s fiber optic token ring network. High-transmission speed and a bounded access time [Sevcik and Johnson, 1987] make FDDI very suitable for supporting real-time applications. Furthermore, FDDI employs a dual-loop architecture, which enhances its reliability; in the event of a fault on one loop, the other loop can be used for transmission. Several new civil and military networks have adopted FDDI as the backbone network.

In order to achieve efficiency, a high-speed network such as an FDDI network requires simple (i.e., low overhead) protocols. In a token ring network, the simplest protocol is the *token passing* protocol. In this protocol, a node transmits its message whenever it receives the token. After it completes its transmission, the node passes on the token to the next neighboring node. Although it has the least overhead, the token passing protocol cannot bound the time between two consecutive visits of the token to a node (called the *token rotation time*), which makes it incapable of guaranteeing message delay requirements. The *timed token protocol*, proposed by Grow [1982], overcomes this problem. This protocol has been

*The work described in this chapter was supported in part by the National Science Foundation under Grant NCR-9210583, by Office of Naval Research under Grant N00014-95-J-0238, and by Texas A&M University under an Engineering Excellence Grant. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing official positions or policies of NSF, ONR, University of Texas at Dallas, Hewlett-Packard (Canada) Ltd., or Texas A&M University.

adopted as a standard for the FDDI networks. The idea behind the timed token protocol is to control the token rotation time. As a result, FDDI is able to support real-time applications.

In the rest of this chapter, we will first introduce the architecture of FDDI networks and discuss its fault management capabilities. We then address how to use FDDI to support time-critical applications.

33.2 Architecture and Fault Management

FDDI is a token ring network. The basic architecture of an FDDI network consists of nodes connected by two counter-rotating loops, as illustrated in Fig. 33.1(a). A node in an FDDI network can be a single attachment station (SAS), a dual attachment station (DAS), a single attachment concentrator (SAC), or a dual attachment concentrator (DAC). Whereas stations constitute the sources and destinations for user frames, the concentrators provide attachments to other stations. The single attachment stations and concentrators are so called because they connect to only one of the two loops. The two fiber loops of an FDDI network are usually enclosed in a single cable. These two loops will be collectively referred to as the FDDI *trunk ring*.

The FDDI standards have been influenced by the need to provide a certain degree of built-in fault tolerance. The station management (SMT) layer of the FDDI protocol deals with initialization/control, monitoring, and fault isolation/recovery in FDDI networks [Jain, 1994]. In ring networks such as FDDI, faults are broadly classified into two categories: node faults and link faults. The FDDI standard specifies explicit mechanisms and procedures for detection of and recovery from both kinds of faults. These fault management capabilities of FDDI provide a foundation for capabilities of FDDI enabled fault-tolerant operation. We present here a brief sketch of FDDI's fault management capabilities. For a comprehensive discussion, see ANSI [1992].

To deal with node faults, each node in FDDI is equipped with an optical bypass mechanism. Using this mechanism, a faulty node can be isolated from the ring, thus letting the network recover from a node fault. Link faults are handled by exploiting the dual-loop architecture of FDDI. Many FDDI networks use one loop as the primary loop and the other as a backup. In the event of a fault on the primary loop, the backup loop is used for transmission [Fig. 33.1(b)].

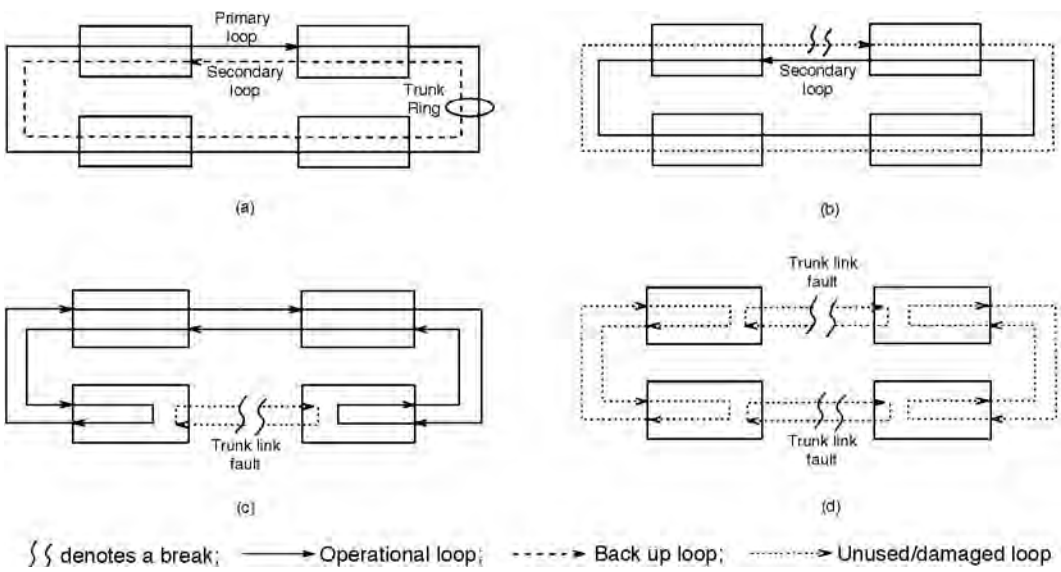


FIGURE 33.1 Architecture and fault management: (a) FDDI trunk ring consists of primary and secondary loops; (b) a fault on the primary loop, secondary loop used for transmission; (c) single trunk link fault, primary and secondary loops wrap up to form a new operational loop; and (d) two trunk link faults segment the network, no loop connects all the nodes.

Because of the proximity of the two fiber loops, a link fault on one loop is quite likely to be accompanied by a fault on the second loop as well, at approximately the same physical location. This is particularly true of faults caused by destructive forces. Such link faults, with both loops damaged, may be termed as *trunk link faults*.

An FDDI network can recover from a single trunk link fault using the so called *wrap-up* operation. This operation is illustrated in Fig. 33.1(c). The wrap-up operation consists of connecting the primary loop to the secondary loop. Once a link fault is detected, the two nodes on either side of the fault perform the wrap-up operation. This process isolates the fault and restores a single closed loop.

The fault detection and the wrap-up operation is performed by a sequence of steps defined by FDDI's link fault management. Once the network is initialized, each node continuously executes the link fault management procedure. The flow chart of the link fault management procedure is shown in Fig. 33.2. Figure 33.2 only presents the gist of the fault management procedure. For details, the reader is referred

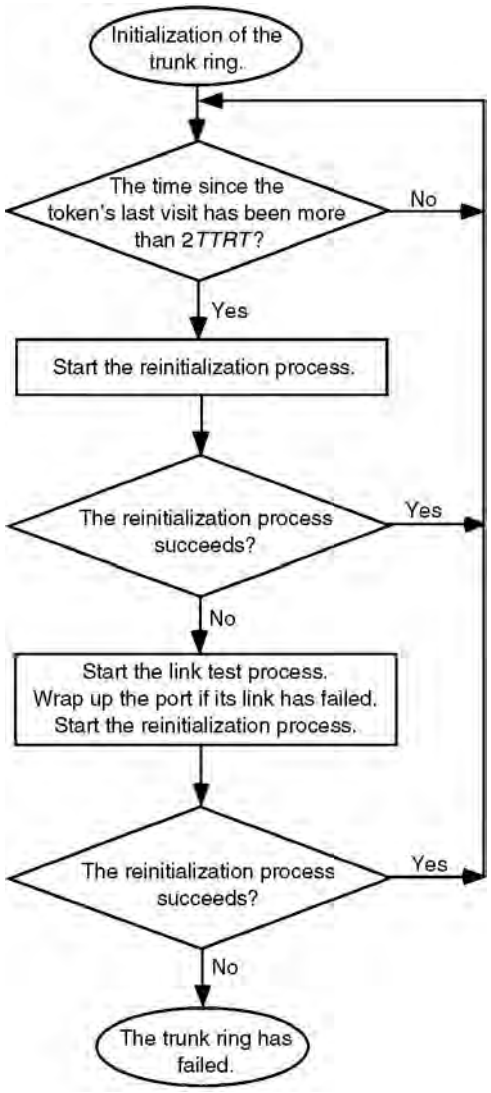


FIGURE 33.2 Flow chart of FDDI link fault management procedure.

to ANSI [1992]. The major steps of this procedure are as follows:

- An FDDI node detects a fault by measuring the time elapsed since the previous token visit to that node. At network initialization time, a constant value is chosen for a protocol parameter called the target token rotation time (*TTRT*). The FDDI protocol ensures that under fault-free operating conditions, the duration between two consecutive token visits to any node is not more than 2*TTRT*. Hence, a node initiates the fault tracing procedure when the elapsed time since the previous token arrival at the node exceeds 2*TTRT*.
- The node first tries to reinitialize the network (using claim frames). If the fault was temporary, the network will recover. Several nodes may initiate the reinitialization procedure asynchronously. However, the procedure converges within *TTRT* time (which is usually on the order of a few milliseconds). The details of this procedure are beyond the scope of this chapter. The reader is referred to ANSI [1992].
- If the reinitialization procedure times out, then the node suspects a permanent fault in the trunk ring. The nodes then execute a series of further steps in order to locate the ring's fault domain. The nodes within the fault domain perform link tests to check the status of the links on all their ports. If a link is faulty, the ring is wrapped up at the corresponding port. The node then tries to reinitialize the wrapped up ring. Once again, note that these operations are carried out concurrently and asynchronously by all the nodes in the trunk ring. That is, any of them may initiate the series of fault recovery steps, but their actions will eventually converge to locate and recover from the fault(s).
- In the event that the last step has also failed, the network is considered unrecoverable. It is possible that operator intervention is required to solve the problem.

With the basic fault management mechanism just described, an FDDI network may not be able to tolerate two or more faults. This is illustrated in Fig. 33.1(d), where two trunk link faults leave the network disconnected. Hence, the basic FDDI architecture needs to be augmented to meet the more stringent fault tolerance requirements of mission-critical applications. In addition, the bandwidth of a single FDDI ring may not be sufficient for some high-bandwidth applications. FDDI-based reconfigurable networks (FBRNs) provide increased fault tolerance and transmission bandwidth as compared to standard FDDI networks. See Agrawal, Kamat, and Zhao [1994]; Chen, Kamart, and Zhao [1995]; and Kamat, Agrawal, and Zhao [1994] for a discussion of this subject.

33.3 The Protocol and Its Timing Properties

FDDI uses the timed token protocol [Grow, 1982]. The timed token protocol is a token passing protocol in which each node gets a guaranteed share of the network bandwidth. Messages in the timed token protocol are segregated into two separate classes: the *synchronous* class and the *asynchronous* class [Grow, 1982]. Synchronous messages are given a guaranteed share of the network bandwidth and are used for real-time communication.

The idea behind the timed token protocol is to control the token rotation time. During network initialization, a protocol parameter called the target token rotation time (*TTRT*) is determined, which indicates the expected token rotation time. Each node is assigned a portion of the *TTRT*, known as its *synchronous bandwidth* (H_i), which is the maximum time a node is permitted to transmit synchronous messages every time it receives the token. When a node receives the token, it transmits its synchronous messages, if any, for a time no more than its allocated synchronous bandwidth. It can then transmit its asynchronous messages only if the time elapsed since the previous token departure from the same node is less than the value of *TTRT*, that is, only if the token arrives earlier than expected.

In order to use the timed token protocol for real-time communications, that is, to guarantee that the deadlines of synchronous messages are met, parameters such as the synchronous bandwidth, the target

token rotation time, and the buffer size must be chosen carefully.

- The synchronous bandwidth is the most critical parameter in determining whether message deadlines will be met. If the synchronous bandwidth allocated to a node is too small, then the node may not have enough network access time to transmit messages before their deadlines. Conversely, large synchronous bandwidths can result in a long token rotation time, which can also cause message deadlines to be missed.
- Proper selection of *TTRT* is also important. Let τ be the token walk time around the network. The proportion of time taken due to token walking is given by $\tau/TTRT$. The maximum network utilization available to user applications is then $1 - \tau/TTRT$ [Ulm, 1982]. A smaller *TTRT* results in less available utilization and limits the network capacity. On the other hand, if *TTRT* is too large, the token may not arrive at a node soon enough to meet message deadlines.
- Each node has buffers for outgoing and incoming synchronous messages. The sizes of these buffers also affect the real-time performance of the network. A buffer that is too small can result in messages being lost due to buffer overflow. A buffer that is too large is wasteful of memory.

In Section 33.4, we will discuss parameter selection for real-time applications. Before that, we need to define network and message models and investigate protocol properties.

The Network and the Message Models

The network contains n nodes arranged in a ring. Message transmission is controlled by the timed token protocol, and the network is assumed to operate without any faults. Outgoing messages at a node are assumed to be queued in first-in-first-out (FIFO) order.

Recall that the token walk time is denoted by τ, which includes the node-to-node delay and the token transmission time. Here τ is the portion of *TTRT* that is not available for message transmission. Let α be the ratio of τ to *TTRT*, that is, $\alpha = \tau/TTRT$. Then α represents the proportion of time that is not available for transmitting messages.

There are n streams of synchronous messages, $S_1,\ldots,S_n$, with stream S_i originating at node i. Each synchronous message stream S_i may be characterized as $S_i = (C_i, P_i, D_i)$:

- C_i is the maximum amount of time required to transmit a message in the stream. This includes the time to transmit both the payload data and the message headers.
- P_i is the interarrival period between messages in the synchronous message stream. Let the first message in stream S_i arrive at node i at time $t_{i,ji}$. The jth message in stream S_i will arrive at node i at time $t_{i,j} = t_{i,j} + (j - 1)\, P_i$, where $j \geq 1$.
- D_i is the *relative deadline* of messages in the stream. The relative deadline is the maximum amount of time that may elapse between a message arrival and completion of its transmission. Thus, the transmission of the jth message in stream S_i, which arrives at $t_{i,j}$, must be completed by $t_{i,j} + D_i$, which is the *absolute deadline* of the message. To simplify the discussion, the terms relative and absolute will be omitted in the remainder of the chapter when the meaning is clear from the context.

Throughout this chapter, we make no assumptions regarding the destination of synchronous message streams. Several streams may be sent to one node. Alternatively, multicasting may occur in which messages from one stream are sent to several nodes.

If the parameters (C_i, P_i, D_i) are not completely deterministic, then their worst-case values must be used. For example, if the period varies between 80 and 100 ms, a period of 80 ms must be assumed. Asynchronous messages that have time constraints can have their deadlines guaranteed by using pseudosynchronous streams. For each source of time-constrained asynchronous messages, a corresponding synchronous message stream is created. The time-constrained asynchronous messages are then promoted to the synchronous class and sent as synchronous messages in the corresponding synchronous message stream.

Each synchronous message stream places a certain load on the system. We need a measure for the load. We define the *effective utilization* U_i of stream S_i as follows:

$$U_i = \frac{C_i}{\min(P_i, D_i)} \tag{33.1}$$

Because a message of length C_i arrives every P_i time, $U_i = C_i/P_i$ is usually regarded as the load presented by stream S_i. An exception occurs when $D_i < P_i$. A message with such a small deadline must be sent relatively urgently, even if the period is very large. Thus, $U_i = C_i/D_i$ is used to reflect the load in this case.

The total effective utilization U of a synchronous message set can now be defined as

$$U = \sum_{i=1}^{n} U_i \tag{33.2}$$

The total effective utilization is a measure of the demands placed on the system by the entire synchronous message set. In order to meet message deadlines, it is necessary that $U \leq 1$.

Constraints

The timed token protocol requires several parameters to be set. The synchronous bandwidths, the target token rotation time, and the buffer sizes parameters are crucial in guaranteeing the deadlines of synchronous messages. Any choice of these parameters must satisfy the following constraints.

The Protocol Constraint

This constraint states that the total bandwidth allocated to synchronous messages must be less than the available network bandwidth, that is,

$$\sum_{i=1}^{n} H_i \leq TTRT - \tau \tag{33.3}$$

This constraint is necessary to ensure stable operation of the timed token protocol.

The Deadline Constraint

This constraint simply states that every synchronous message must be transmitted before its (absolute) deadline. Formally, let $s_{i,j}$ be the time that the transmission of the jth message in stream S_i is completed. The deadline constraint requires that for $i = 1, \ldots, n$ and $j = 1, 2, \ldots,$

$$S_{i,j} \leq t_{i,j} + D_i \tag{33.4}$$

where $t_{i,j}$ is the arrival time and D_i is the (relative) deadline. Note that in this inequality, $t_{i,j}$ and D_i are given by the application, but $s_{i,j}$ depends on the synchronous bandwidth allocation and the choice of *TTRT*.

The Buffer Constraint

This constraint states that the size of the buffers at each node must be sufficient to hold the maximum number of outgoing or incoming synchronous messages that could be queued at the node. This constraint is necessary to ensure that messages are not lost due to buffer overflow.

Timing Properties

According to Eq. (33.4), the deadline constraint of message M is satisfied if and only if the minimum synchronous bandwidth available to M within its deadline is bigger than or equal to its message size. That is, for a given *TTRT*, whether or not deadline constraints can be satisfied is solely determined by

bandwidth allocation. In order to allocate appropriate synchronous bandwidth to a message stream, we need to know the worst-case available synchronous bandwidth for a stream within any time period.

Johnson and Sevcik [1987] showed that the maximum amount of time that may pass between two consecutive token arrivals at a node can approach $2TTRT$. This bound holds regardless of the behavior of asynchronous messages in the network. To satisfy the deadline constraint, it is necessary for a node to have at least one opportunity to send each synchronous message before the message deadline expires. Therefore, in order for the deadline constraint to be satisfied, it is necessary that for $i = 1, \ldots, n$,

$$D_i \geq 2TTRT \tag{33.5}$$

It is important to note that Eq. (33.5) is only a necessary but not a sufficient condition for the deadline constraint to be satisfied. For all message deadlines to be met, it is also crucial to choose the synchronous bandwidths H_i appropriately.

Further studies on timing properties have been performed in Chen, Li, and Zhao [1994] and the results can be summarized by the following theorem.

Theorem 33.1. Let $X_i(t, t+I, \boldsymbol{H})$ be the minimum total transmission time available for node i to transmit its synchronous message within the time interval $(t, t+I)$ under bandwidth allocation $\boldsymbol{H} = (H_1, H_2, \ldots, H_n)^T$ then

$$= \begin{cases} \left\lfloor \dfrac{I}{TTRT} - 1 \right\rfloor \cdot H_i + \max\left(0, \min\left(r_i - \tau - \displaystyle\sum_{j=1,\ldots,n, j\neq i} H_j, H_i\right)\right) & \text{if } I \geq TTRT \\ 0 & \text{otherwise} \end{cases} \tag{33.6}$$

where $r_i = I - \left\lfloor \frac{I}{TTRT} \right\rfloor \cdot TTRT$.

Using Theorem 33.1, deadline constraint (33.4) is satisfied if and only if, for $i = 1, \ldots, n$,

$$X_i(t, t + D_i, H) \geq C_i \tag{33.7}$$

33.4 Parameter Selection for Real-Time Applications

To support real-time applications on FDDI, we have to properly set the following three types of parameters: synchronous bandwidth, *TTRT*, and buffer size. We address this parameter selection problem in this section.

Synchronous Bandwidth Allocation

In FDDI, synchronous bandwidths are assigned by a synchronous bandwidth allocation scheme. This subsection examines synchronous bandwidth allocation schemes and discusses how to evaluate their effectiveness.

A Classification of Allocation Schemes

Synchronous bandwidth allocation schemes may be divided into two classes: local allocation schemes and global allocation schemes. These schemes differ in the type of information they may use. A local synchronous bandwidth allocation scheme can only use information available locally to node i in allocating H_i. Locally available information at node i includes the parameters of stream S_i (i.e., C_i, P_i, and D_i). *TTRT* and τ are also locally available at node i because these values are known to all nodes. On the other hand, a global synchronous bandwidth allocation scheme can use global information in its allocation of synchronous bandwidth to a node. Global information includes both local information and information regarding the parameters of synchronous message streams originating at other nodes.

A local scheme is preferable from a network management perspective. If the parameters of stream S_i change, then only the synchronous bandwidth H_i of node i needs to be recalculated. The synchronous bandwidths at other nodes do not need to be changed because they were calculated independently of S_i. This makes a local scheme flexible and suited for use in dynamic environments where synchronous message streams are dynamically initiated or terminated.

In a global scheme, if the parameters of S_i change, it may be necessary to recompute the synchronous bandwidths for all nodes. Therefore, a global scheme is not well suited for a dynamic environment. In addition, the extra information employed by a global scheme may cause it to handle more traffic than a local scheme. However, it is known that local schemes can perform *very closely* to the optimal synchronous bandwidth allocation scheme when message deadlines are equal to message periods. Consequently, given the previously demonstrated good performance of local schemes and their desirable network management properties, we concentrate on local synchronous bandwidth allocation schemes in this chapter.

A Local Allocation Scheme

Several local synchronous bandwidth allocation schemes have been proposed for both the case of $D_i = P_i$ [Agrawal, Chen, and Zhao, 1993] and the case of $D_i \neq P_i$ [Malcolm and Zhao, 1993; Zheng and Shin, 1993]. These schemes all have similar worst-case performance. Here we will consider the scheme proposed in Malcolm and Zhao [1993]. With this scheme, the synchronous bandwidth for node i is allocated according to the following formula:

$$H_i = \frac{U_i D_i}{\left\lfloor \frac{D_i}{TTRT} - 1 \right\rfloor} \tag{33.8}$$

Intuitively, this scheme follows the flow conservation principle. Between the arrival of a message and its absolute deadline, which is D_i time later, node i will have at least $\left\lfloor \frac{D_i}{TTRT} - 1 \right\rfloor H_i$ of transmission time available for synchronous messages by Eq. (33.6). This transmission time is available regardless of the number of asynchronous messages in the network. During the D_i time, $U_i D_i$ can loosely be regarded as the load on node i. Thus, the synchronous bandwidth in Eq. (33.8) is just sufficient to handle the load on node i between the arrival of a message and its deadline.

The scheme defined in Eq. (33.8) is a simplified version of those in Malcolm and Zhao [1993] and Zheng and Shin [1993]. In the rest of this chapter, we assume this scheme is used because it is simple, intuitive, and well understood. Another reason for concentrating on this scheme is that it has been adopted for use with the SAFENET standard, and thus will be used in distributed real-time systems in the future.

Schedulability Testing

We now consider schedulability testing. A message set is schedulable if the deadlines of its synchronous messages can be satisfied. This can be determined by referring to the tests that both the protocol and deadline constraints [Eqs. (33.3) and (33.4) or Eq. (33.7)] are satisfied.

Testing to see if the protocol constraint is satisfied is very straightforward. But the test of deadline constraints is more complicated and requires more information. This test can be greatly simplified if the bandwidths are allocated according to Eq. (33.8). It was shown that if this scheme is used, the protocol constraint condition defined in Eq. (33.3) implies the deadline constraint condition Eq. (33.4). Testing the protocol constraint alone is sufficient to ensure that both constraints are satisfied. This is a big advantage of using the allocation scheme defined in Eq. (33.8).

A second method of schedulability testing is to use the *worst-case achievable utilization* criteria. This criteria has been widely used in real-time systems. For a synchronous bandwidth allocation scheme, the worst-case achievable utilization U^* defines an upper bound on the effective utilization of a message set: if the effective utilization is no more than the upper bound, both the protocol and the deadline constraints

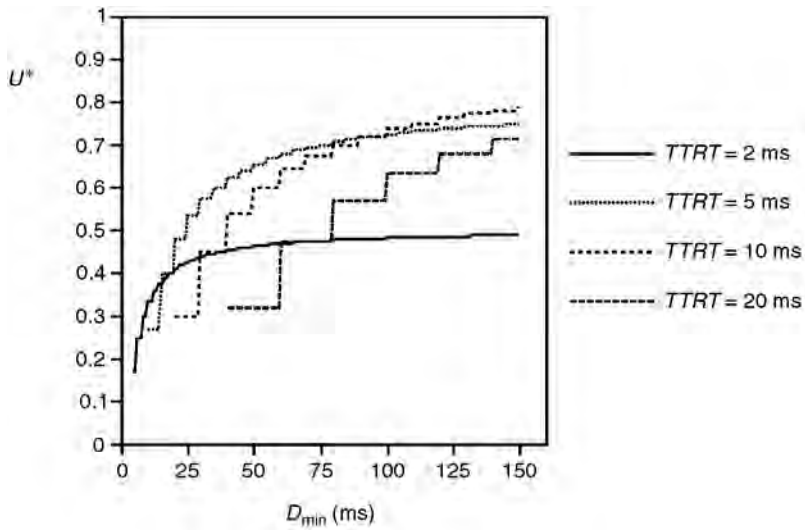


FIGURE 33.3 U^* vs. $D_{\min}$.

are always satisfied. The worst-case achievable utilization U^* for the scheme defined in Eq. (33.8) is

$$U^* = \frac{\left\lfloor \frac{D_{\min}}{TTRT} \right\rfloor - 1}{\left\lfloor \frac{D_{\min}}{TTRT} \right\rfloor + 1}(1 - \alpha) \tag{33.9}$$

where $TTRT$ is the target token rotation time and $D_{\min}$ is the minimum deadline. For any message set, if its effective utilization Eq. (33.2) is less than U^*, both the protocol and deadline constraints are guaranteed to be satisfied.

We would like to emphasize that, in practice, using this criteria can simplify network management considerably. The parameters of a synchronous message set can be freely modified while still maintaining schedulability, provided that the effective utilization remains less than U^*.

Let us examine the impact of $TTRT$ and $D_{\min}$ on the worst-case achievable utilization given in Eq. (33.9). Figure 33.3 shows the worst-case achievable utilization vs. $D_{\min}$ for several different values of $TTRT$. This figure was obtained by plotting Eq. (33.9) with τ taken to be 1 ms (a typical value for an FDDI network). Several observations can be made from Fig. 33.3 and formula (33.8).

1. For a fixed value of $TTRT$, the worst-case achievable utilization increases as $D_{\min}$ increases. From Eq. (33.9) it can be shown that when $D_{\min}$ approaches infinity, U^* approaches $(1 - \alpha) = (1 - \tau/TTRT)$. That is, as the deadlines become larger, the worst-case achievable utilization approaches the available utilization of the network.
2. In Agrawal, Chen, and Zhao [1993], it was shown that for a system in which all relative deadlines are equal to the corresponding message periods ($D_i = P_i$), a worst-case achievable utilization of $\frac{1}{3}(1 - \alpha)$ can be achieved. That result can be seen as a special case of Eq. (33.9): if $D_{\min} = 2TTRT$, we have $\left\lfloor \frac{D_{\min}}{TTRT} \right\rfloor = 2$ and $U^* = \frac{1}{3}(1 - \alpha)$.
3. $TTRT$ clearly has an impact on the worst-case achievable utilization. From Fig. 33.3, we see that when $D_{\min} = 50$ ms, the case of $TTRT = 5$ ms gives a higher worst-case achievable utilization than the other plotted values of $TTRT$. When $D_{\min} = 125$ ms, the case of $TTRT = 10$ ms gives a higher U^* than the other plotted values of $TTRT$. This observation provides motivation for maximizing U^* by properly selecting $TTRT$ once $D_{\min}$ is given.

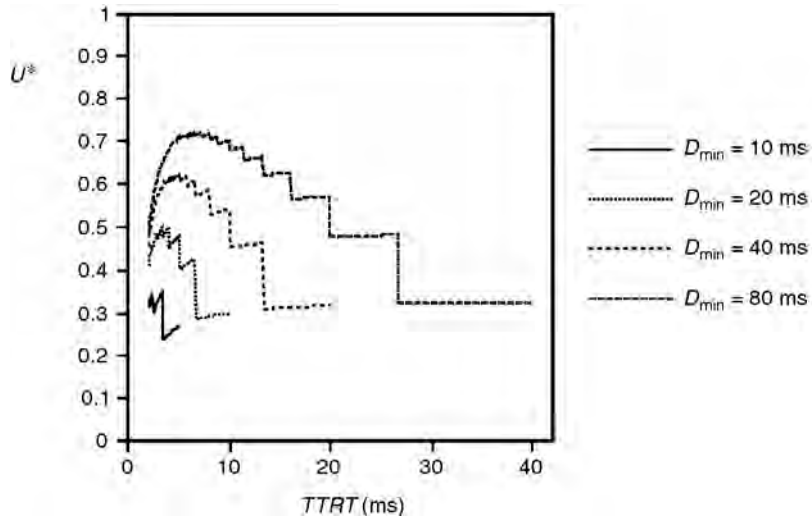


FIGURE 33.4 U^* vs. *TTRT*.

Selection of Target Token Rotation Time

Recall that *TTRT*, the target token rotation time, determines the expected value of the token rotation time. In contrast to the synchronous bandwidths of individual nodes, *TTRT* is common to all of the nodes and should be kept constant during run time. As we observed in the last section, *TTRT* has an impact on the worst-case achievable utilization. Thus, we would like to choose *TTRT* in an optimal fashion so that the worst-case achievable utilization U^* is maximized. This will increase the amount of real-time traffic that can be supported by the network.

The optimal value of *TTRT* has been derived in Malcolm and Zhao [1993] and is given by

$$TTRT = \frac{D_{\min}}{\left\lceil \frac{-3+\sqrt{9+\frac{8D_{\min}}{\tau}}}{2} \right\rceil} \tag{33.10}$$

The impact of an appropriate selection of *TTRT* is further evident from Fig. 33.4, which uses Eq. (33.10) to show U^* vs. *TTRT* for several different values of $D_{\min}$. As with Fig. 33.3, τ is taken to be 1 ms. From Fig. 33.4, the following observations can be made:

1. The curves in Fig. 33.4 verify the prediction of the optimal *TTRT* value given by Eq. (33.10). For example, consider the case of $D_{\min} = 40$ ms. By Eq. (33.10), the optimal value of *TTRT* is 5 ms. The curve clearly indicates that at $TTRT = 5$ ms, the worst-case achievable utilization is maximized. Similar observations can be made for the other cases.
2. As indicated in Eq. (33.10), the optimal *TTRT* is a function of $D_{\min}$. This coincides with the expectations from the observations of Fig. 33.3. A general trend is that as $D_{\min}$ increases, the optimal *TTRT* increases. For example, the optimal values of *TTRT* are approximately 2.5 ms for $D_{\min} = 10$ ms, 4 ms for $D_{\min} = 20$ ms, 5 ms for $D_{\min} = 40$ ms, and 6.67 ms for $D_{\min} = 80$ ms.
3. The choice of *TTRT* has a large effect on the worst-case achievable utilization U^*. Consider the case of $D_{\min} = 40$ ms shown in Fig. 33.4. If *TTRT* is too small (say, $TTRT = 2$ ms), U^* can be as low as 45%. If *TTRT* is too large (say, $TTRT = 15$ ms), U^* can be as low as 31%.

However, when the optimal value of *TTRT* is used (i.e., $TTRT = 5$ ms), U^* is 62%. This is an improvement of 17% and 31%, respectively, over the previous two cases.

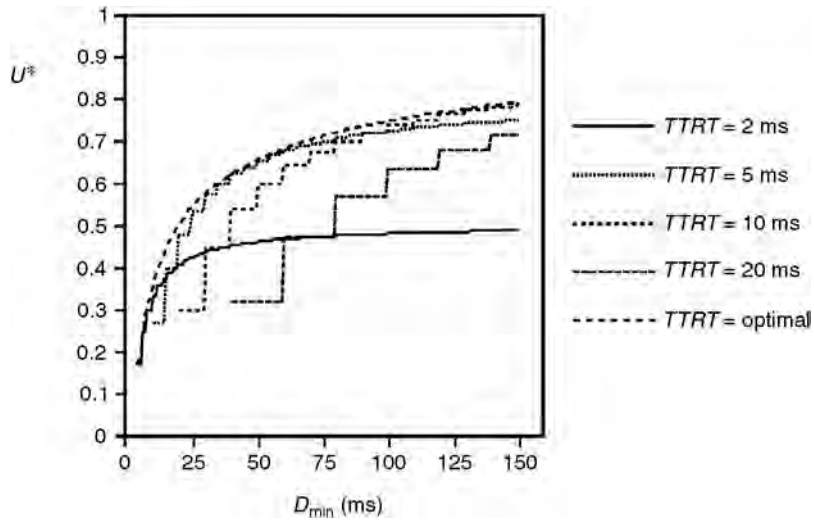


FIGURE 33.5 U^* vs. D_{min} with optimal *TTRT*.

The effect of choosing an optimal value for *TTRT* is also shown in Fig. 33.5. Figure 33.5 shows the worst-case achievable utilization vs. D_{min} when an optimal value of *TTRT* is used. For ease of comparison, the earlier results of Fig. 33.3 are also shown in Fig. 33.5.

Buffer Requirements

Messages in a network can be lost if there is insufficient buffer space at either the sending or the receiving node. To avoid such message loss, we need to study the buffer space requirements.

There are two buffers for synchronous messages on each node. One buffer is for messages waiting to be transmitted to other nodes and is called the *send buffer*. The other buffer is for messages that have been received from other nodes and are waiting to be processed by the host. This buffer is called the *receive buffer*.

We consider the send buffer first.

- Case of $D_i \le P_i$. In this case, a message must be transmitted within the period in which it arrives. At most, one message will be waiting. Hence, the send buffer need only accommodate one message.
- Case of $D_i > P_i$. In this case, a message may wait as long as D_i time without violating its deadline constraint. At node *i*, messages from stream S_i arrive every P_i time, requesting transmission. During the D_i time following the arrival of a message, there will be further $\left\lfloor \frac{D_i}{P_i} \right\rfloor$ message arrivals. Thus, the send buffer may potentially have to accommodate $\left\lfloor \frac{D_i}{P_i} \right\rfloor + 1$ messages. When the deadline is very large (tens or hundreds times larger than the period, for example, which can occur in voice transmission), this might become a problem in practice due to excessive buffer requirements.

However, when the synchronous bandwidths are allocated as in Eq. (33.8), the impact of D_i on required buffer size is limited: it can be shown that if Eq. (33.8) is used, the waiting time of a message from stream S_i is bounded by $\min(D_i, P_i + 2TTRT)$ [Malcolm, 1994]. Thus, the maximum number of messages from S_i that could be queued at node *i* is no more than $\left\lceil \frac{\min(D_i, P_i + 2TTRT)}{P_i} \right\rceil$. This allows us to reduce the required size of the send buffer. Let BS_i denote the send buffer size at node *i*. The send buffer will never overflow if

$$BS_i \ge \left\lceil \frac{\min(D_i, P_i + 2TTRT)}{P_i} \right\rceil B_i \tag{33.11}$$

where B_i is the number of bytes in a message of stream S_i.

An interesting observation is that if $D_i \geq P_i + 2TTRT$, then

$$BS_i \geq \left\lceil \frac{P_i + 2TTRT}{P_i} \right\rceil B_i \tag{33.12}$$

That is, the send buffer requirements for stream S_i are independent of the deadline D_i. As mentioned earlier, one would expect that increasing the deadline of a stream could result in increased buffer requirements. When the scheme in Eq. (33.8) is used, however, the buffer requirements are not affected by increasing the deadline once the deadline reaches certain point.

Now let us consider the receive buffer size. Suppose that node j is the destination node of messages from nodes $j_1, \ldots, j_k$. This means that messages from streams $S_{j1}, \ldots, S_{jk}$ are being sent to node j. The size of the receive buffer at node j depends not only on the message traffic from these streams but also on the speed at which the host of node j is able to process incoming messages from other nodes. The host at node j has to be able to keep pace with the incoming messages, otherwise the receive buffer at node j can overflow. When a message from stream S_{ji} arrives at node j, we assume that the host at node j can process the message within P_{ji} time.

Assuming that the host is fast enough, then it can be shown that the number of messages from S_i that could be queued at the destination node is bounded by $\left\lceil \frac{\min(D_i, P_i + 2TTRT)}{P_i} \right\rceil + 1$. With this bound, we can derive the space requirements for the receive buffer at node j. Let BR_j denote the receive buffer size at node j. The receive buffer will never overflow if

$$BR_j \geq \sum_{i=1}^{k} \left\lceil \frac{\min(D_{ji}, P_{ji} + 2TTRT)}{P_{ji}} + 1 \right\rceil B_{ji} \tag{33.13}$$

As in the case of the send buffer, we can observe from Eq. (33.13) that if the deadlines are large, the required size of the receive buffer will not grow as the deadlines increase.

33.5 Final Remarks

In this chapter, we introduced the architecture of FDDI and discussed its fault-tolerant capability. We presented a methodology for the use of FDDI networks for real-time applications. In particular, we considered methods of selecting the network parameters and of schedulability testing. The parameter selection methods are compatible with current standards. The synchronous bandwidth allocation method has the advantage of only using information local to a node. This means that modifications in the characteristics of synchronous message streams, or the creation of new synchronous message streams, can be handled locally without reinitializing the network. Schedulability tests that determine whether the time constraints of messages will be met were presented. They are simple to implement and computationally efficient.

The materials presented in this chapter complement much of the published work on FDDI. Since the early 1980s, extensive research has been done on the timed token protocol and its use in FDDI networks [Albert and Jayasumana, 1994; Jain, 1994; Mills, 1995; Shah and Ramakrishnan, 1994]. The papers by Ross [1989], Iyer and Joshi [1985], and others [Mccool, 1988; Southard, 1988; Stallings, 1987] provided comprehensive discussions on the timed token protocol and its use in FDDI. Ulm [1982] discussed the protocol performance with respect to parameters such as the channel capacity, the network cable length, and the number of stations. Dykeman and Bux [1988] developed a procedure for estimating the maximum total throughput of asynchronous messages when using single and multiple asynchronous priority levels. The analysis done by Pang and Tobagi [1989] gives insight into the relationship between the bandwidth allocated to each class of traffic and the timing parameters. Valenzo, Montuschi, and Ciminiera [1990] concentrated on the asynchronous throughput and the average token rotation time when the asynchronous

traffic is heavy. The performance of the timed token ring depends on both the network load and the system parameters. A study of FDDI by Jain [1991] suggests that a value of 8 ms for *TTRT* is desirable as it can achieve 80% utilization on all configurations and results in around 100 ms maximum access delay on typical rings. Further studies have been carried out by Sankar and Yang [1989] to consider the influence of the target token rotation time on the performance of varying FDDI ring configurations.

References

ANSI. 1992. FDDI station management protocol (SMT). ANSI Standard X3T9.5/84-89, X3T9/92-067, Aug.

Agrawal, G., Chen, B., and Zhao, W. 1993. Local synchronous capacity allocation schemes for guaranteeing message deadlines with the timed token protocol. In *Proc. IEEE Infocom'93,* 186–193.

Agrawal, G., Kamat, S., and Zhao, W. 1994. Architectural support for FDDI-based reconfigurable networks. In *Workshop on Architectures for Real-Time Applications (WARTA).*

Albert, B., and Jayasumana, A.P. 1994. *FDDI and FDDI-II Architecture, Protocols, and Performance,* Artech House, Norwood, MA.

Chen, B., Li, H., and Zhao, W. 1994. Some timing properties of timed token medium access control protocol. In *Proceedings of International Conference on Communication Technology,* 1416–1419, June.

Chen, B., Kamart, S., and Zhao, W. 1995. Fault-tolerant real-time communication in fddi-based networks. *Proc. IEEE Real-Time Systems Symposium,* 141–151, Dec.

Dykeman, D. and Bux, W. 1988. Analysis and tuning of the FDDI media access control protocol. *IEEE J. Selec. Areas Commun.,* 6(6).

Grow, R.M. 1982. A timed token protocol for local area networks. In *Proc. Electro/ 82, Token Access Protocols,* May.

Iyer, V. and Joshi, S.P. 1985. FDDI's 100 M-bps protocol improves on 802.5 Spec's 4-M-bps limit. *Elec. Design News,* (May 2):151–160.

Jain, R. 1991. Performance analysis of FDDI token ring networks: effect of parameters and guidelines for setting TTRT. *IEEE Lett.*, (May): 16–22.

Jain, R. 1994. *FDDI Handbook—High-Speed Networking Using Fiber and Other Media,* Addison Wesley.

Kamat, S., Agrawal, G., and Zhao, W. 1994. On available bandwidth in FDDI-based reconfigurable networks. In *Proc. IEEE Infocom'94.*

Malcolm, N. and Zhao, W. 1993. Guaranteeing synchronous messages with arbitrary deadline constraints in an FDDI network. In *Proc. IEEE Conference on Local Computer Networks,* (Sept.): 186–195.

Malcolm, N. 1994. Hard real-time communication in high speed networks. Ph.D. Thesis, Dept. of Computer Science, Texas A&M University, College Station.

Mccool, J. 1988. FDDI—getting to the inside of the ring. *Data Commun.,* (March): 185–192.

Mills, A. 1995. *Understanding FDDI,* Prentice–Hall, Englewood Cliffs, NJ.

Pang, J. and Tobagi, F.A. 1989. Throughput analysis of a timer controlled token passing protocol under heavy load. *IEEE Trans. Commun.,* 37(7):694–702.

Ross, F.E. 1989. An overview of FDDI: the fiber distributed data interface. *IEEE J. Selec. Areas Commun.,* 7(Sept.):1043–1051.

Sankar, R. and Yang, Y.Y. 1989. Performance analysis of FDDI, In *Proceedings of the IEEE Conference on Local Computer Networks,* Minneapolis MN, Oct. 10–12, 328–332.

Sevcik, K.C. and Johnson, M.J. 1987. Cycle time properties of the FDDI token ring protocol. *IEEE Trans. Software Eng.,* SE-13(3):376–385.

Shah, A. and Ramakrishnan, G. 1994. *FDDI: A High Speed Network,* Prentice–Hall, Englewood Cliffs, NJ.

Southard, R. 1988. Fibre optics: A winning technology for LANs. *Electronics* (Feb.): 111–114.

Stallings, W. 1987. *Computer Communication Standards, Vol 2: Local Area Network Standards,* Howard W. Sams.

Ulm, J.N. 1982. A timed token ring local area network and its performance characteristics. In *Proceedings of the IEEE Conference on Local Computer Networks,* Feb., 50–56.

Valenzano, A., Montuschi, P., and Ciminiera, L. 1990. Some properties of timed token medium access protocols. *IEEE Trans. Software Eng.,* 16(8).

Zheng, Q. and Shin, K.G. 1993. Synchronous bandwidth allocation in FDDI networks. In *Proc. ACM Multimedia'93,* Aug., 31–38.

34

Broadband Local Area Networks

Joseph A. Bannister
The Aerospace Corporation

34.1 Introduction

The local area network (LAN) became a permanent fixture of the computer world during the 1980s, having achieved penetration primarily in the form of the Institute of Electrical and Electronics Engineers (IEEE) 802.3 ethernet* LAN. Thereafter, other types of LANs, such as the IEEE 802.5 token ring and the American National Standards Institute (ANSI) X3T9.5 fiber distributed data interface (FDDI), also established themselves as lesser—though potent—players. These LANs now form the dominant communications infrastructures that connect computers and related devices within individual enterprises. They provide the transmission substrate for common services such as file transfer, network file services, network window service, electronic mail, hypertext transfer, and others.

The installed base of ethernet, token ring, and FDDI LANs consists fundamentally of shared-media broadcast networks in which each node's transmission may be heard by all other nodes on the LAN. In any such **shared-media** LAN, a **media-access control protocol** must be provided to coordinate the different nodes' transmissions, lest they interfere adversely with one another. The most serious drawback of shared-media LANs is that, as more nodes are added, each node receives a proportionally smaller share of the media's bandwidth. The bandwidth of the media is, therefore, a limiting resource. This inability to scale up to meet increased demands for bandwidth makes shared-media LANs inherently unsuitable for new classes of communications services, such as the transmission of high-resolution image and video data. Furthermore, the nature of their media-access control protocols, which can introduce substantial and variable delays into transmissions, often makes shared-media LANs unable to meet the tight **latency** and **jitter** requirements of real-time video and audio transmission.

To address the problems of bandwidth scalability and real-time transmission, a new type of LAN, called the **broadband local area network (BLAN)**, has been steadily taking root. BLANs are designed to accommodate increases in traffic demand by modular expansion. They are also capable of dedicating a set amount

*Originally, ethernet referred to a specific product offering, whereas the IEEE 802.3 LAN was formally known as the carrier sense multiple access with collision detection LAN. The two LANs differ minutely, and henceforth we consider them to be the same.

of bandwidth to a requesting application, guaranteeing that the latency of the application's transmissions is bounded and unvarying, if required. The BLAN will thus carry integrated traffic, comprising data, voice, and video. In this respect, the BLAN is a local manifestation of the **broadband integrated services digital network** (**BISDN**), which is intended for deployment in the global telecommunications system. The common ground of the BLAN and the BISDN is their reliance on **asynchronous transfer mode** (**ATM**). Special considerations, however, distinguish the BLAN from the BISDN, including different cost structures, the need to support different services, and different approaches to network management and security.

This chapter will describe the common approaches to BLANs, including discussion of the major requirements and the hardware and software technologies essential to their operation. The central role of ATM networks in BLANs as well as other less well-known network technologies, will be covered. The protocols used in BLANs themselves will also be discussed. Finally, some of the new applications being enabled by BLANs will be presented.

34.2 User Requirements

BLAN users expect to run network-based applications that require services not provided by conventional LANs. These requirements include bandwidth scalability, integrated support for different traffic classes, multicasting, error recovery, manageability, and security.

To meet the increased demand for bandwidth that accompanies the addition of new nodes and the replacement of old nodes by more-capable equipment, bandwidth scalability in shared-media LANs has been achieved by segmenting users into distinct **collision domains**, which are interconnected by bridges and routers so that intrasegment traffic is filtered out before leaving its originating domain, and intersegment traffic is forwarded to its destination. This approach is, however, not truly scalable, and the resulting web of LANs becomes fundamentally unmanageable. To overcome this difficulty, packet switches replace the collision domains, and each node is connected to its switch by a dedicated link. Growth of the network is then accomplished by adding switches or selectively replacing the interfaces between switches and nodes with higher speed interfaces.

Network traffic is traditionally divided into three classes: **isochronous, synchronous**, and **asynchronous** traffic. Isochronous traffic flows at a constant rate and has a strict timing relationship between any consecutively transmitted units of data. An example is periodically sampled fixed-size data blocks such as one would encounter in the transmission of a video frame buffer every 1/30 of a second. Synchronous traffic flows at a variable rate and has a loose timing relationship between consecutively transmitted units of data. An example is the transmission of a compressed video frame buffer every 1/30 of a second. Asynchronous traffic flows at a variable rate but has no definite timing relationship among its data units. An example is file transfer. Although all conventional LANs provide support for asynchronous traffic, few support synchronous traffic. Support for isochronous traffic is not common in conventional LANs. Currently, only BLANs support all three classes of traffic.

Users who desire to run video applications will use a BLAN that supports **multicasting**, a generalization of broadcasting in which a sender may transmit the same message to a number of different receivers without sending multiple copies. Easily supported in shared-media LANs, this capability must be designed in to switch-based networks. Multicasting is not only used in disseminating video broadcasts, but has become entrenched into the way that shared-media LANs operate, for example, in registering addresses. Thus, it must be supported in BLANs to accommodate legacy protocols.

As components of a complex system—such as a BLAN—fail, the system should continue to provide service with a minimum of interruption. Having assumed the responsibility for transporting nearly all of the traffic in an enterprise, the BLAN of necessity must be robust enough to provide continuous service even in the face of failures. Given that a BLAN can fail in many different ways, it is essential that error-recovery procedures are integrated into its design. The BLAN must detect transmission errors and notify the system administrator of them. Failures of resource (e.g., a switch or link) must also be recognized and bypassed. Since failures or oversubscription can cause congestion, there must be the means to recognize and clear congestion from the network.

Owned and operated by a single enterprise, the BLAN usually falls under the responsibility of a specialized department charged with its management. It is often necessary to track all resources within the BLAN so that the network manager knows the precise disposition of all equipment at any given time. Each resource also must be monitored continuously to assess its level of performance. Resources also are controlled and configured by the network manager. For the convenience of the network manager, these functions are normally concentrated in a network-management platform. The dynamically changing state of all resources within the BLAN is referred to as the **management information base** (MIB), which may be viewed as a readable and writeable distributed database through which the network is monitored and controlled. User information may be extracted from the MIB in order to produce usage and billing information.

Although the BLAN resides within the relatively friendly confines of an enterprise, external and internal attacks on the BLAN's computers are a threat that cannot be ignored. External attacks on a BLAN connected to a larger public network are effectively controlled by use of a network firewall, which bars unauthorized outside access to the enterprise's computers. Because sensitive information might have to be denied to internal users, it is also necessary to provide for privacy within some BLANs. Techniques for encrypting stored or transmitted data and for providing access controls are then required. To enhance security further, it also could be necessary to provide intrusion-detection software, which passively monitors BLAN resources for internal and external attacks.

34.3 BLAN Technologies

The spread of BLANs has been promoted by technological advances on several fronts. Fast packet switching has been one of the enabling technologies. So, too, has been high-speed transmission technology. Other important software technologies in BLANs are routing, signalling, and network management.

Packet switching networks are composed of switching nodes connected by links. The switches forward input packets to the appropriate output links after reading the packets' header information and possibly buffering packets temporarily to resolve contention for common output links. The move to packet switching LANs was motivated by the realization that shared-media LANs do not scale, because the total throughput of all nodes can never exceed the bandwidth of the media. Similarly, latency and jitter grow as nodes are added to a shared-media LAN because the media-access control protocol tends to perform like a single-server queueing system, in which the mean and variance of delay increase as the server's workload increases. Switching plays a key role in BLANs because a **mesh topology**, which consists of switches connected by links, is necessary to overcome the bandwidth bottleneck of the shared-media LAN, as illustrated in Fig. 34.1. In the shared-media LAN only one node at a time can transmit

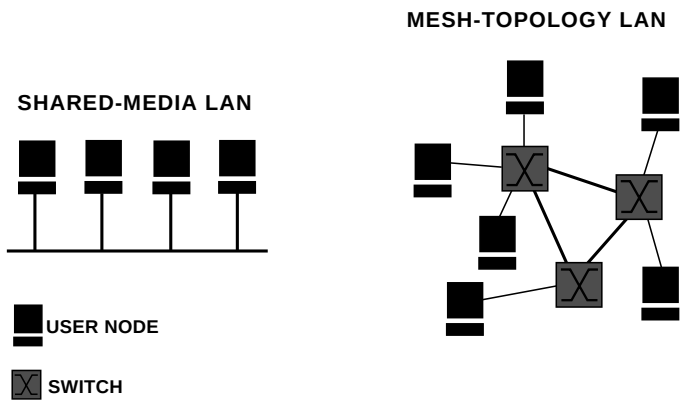


FIGURE 34.1 The mesh-topology LAN provides higher bandwidth and supports more nodes than the shared-media LAN.

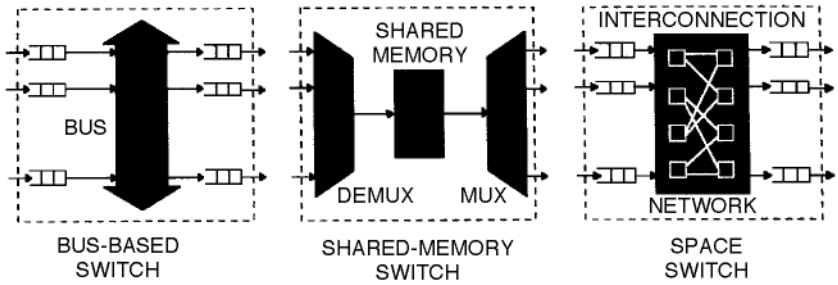


FIGURE 34.2 Packet switch architectures.

information, but in the mesh-topology LAN several nodes can transmit information simultaneously. The result is a dramatic increase in the throughput of the mesh-topology LAN, compared to the shared-media LAN. If one can build a packet switch with a sufficient number of input/output ports, then a mesh network may be easily expanded by adding new nodes, switches, and links without significantly reducing the amount of bandwidth that is available to other nodes. It is also true that latency and jitter in a mesh topology may be kept low, since the capacity of the network is an increasing function of the number of nodes.

Paced by advances in integrated circuits, switch technology has progressed rapidly. Packet switching fabrics based on high-speed backplane buses, shared memory, or multistage interconnection networks (MINs), such as delta, omega, shuffle-exchange, or Clos networks [Tobagi, 1990], can be constructed to switch pre-established connections with a throughput of several gigabits per second. Examples of each of these switch architectures are shown in Fig. 34.2. To increase switch throughput, buses may be replicated and shared memories interleaved. Although backplane-bus fabrics cannot switch more than a few dozen input and output ports, because the electrical limits on bus length restrict the number of modules that may be attached, two-by-two switching elements can be combined in very deep configurations to implement MINs with arbitrary numbers of input and output ports. Switch latency can be reduced by the use of cut-through routing, which permits the head of a packet to emerge from the switch before its tail has even entered the switch. This is especially beneficial for long packets, but the switch design is more complex and special facilities must often be added to prevent deadlock. Bus-based switching fabrics support multicasting easily. MIN-based switching fabrics have also been implemented to support multicasting by forking copies of a packet at various stages or using a separate copy network [Turner, 1987]. Today's nonblocking, output-queued MIN-based fabrics can achieve very low latency and high throughput [Hluchyj and Karol, 1988].

High-performance microprocessors—increasingly with reduced instruction set computer architectures—often form an integral part of switch and interface architectures. These provide necessary decision-making and housekeeping functions such as routing-table maintenance and call processing. They also may be used for such functions as segmenting messages into smaller transmission units at the sender and then reassembling the message at the receiver. Such segmentation and reassembly functions are sometimes performed by special-purpose hardware under the control of the microprocessor. High-speed random-access and first-in-first-out memories are also critical for storage and packet buffering.

Examples of chipsets that support high-speed switching are found in Denzel, Engbersen, and Iliadis [1995] and Collivignarelli et al. [1994]. Supporting chipsets are generally implemented as complementary metal-oxide-semiconductor devices. As higher switching and transmission speeds are needed, however, it is expected that supporting chipsets will be implemented as gallium arsenide or bipolar devices, which operate in a much faster regime.

Hand in hand with fast packet switching technology comes high-speed transmission technology. The enabling technology in this area has been optical fiber technology, which can achieve link speeds of several gigabits per second. The principal transmission technology for BLANs is multimode optical fibers (MOFs). Using infrared signals with a 1.3-μm wavelength, BLANs can achieve data rates of more than

100 Mb/s over distances of up to 2000 m. Transmitters are usually light-emitting diodes, receivers are photodiodes, and fibers are 125-μm cladding/62.5-μm core multimode silica fibers. To transmit over greater distances or at higher data rates, 100-μm cladding/9-μm core single-mode silica fibers are used with 1.3- or 1.5-μm wavelength laser-diode transmitters. Because LANs are of limited geographical reach, single-mode optical fibers can carry information at rates of several gigabits per second without the need for optical amplification or regeneration. Optical fiber BLAN links typically operate with direct direction and return-to-zero, 4B/5B, or 8B/10B encoding (nB/mB denotes a scheme in which n data bits are encoded as m transmission bits). Error-detecting or -correcting codes are used sparingly because the transmission quality is normally very high.

The relatively high cost of optical systems and the installed base of metallic-wire systems make the use of inexpensive metallic-wire media very desirable. Most conventional LANs now use category-3 twisted-pair 24-gauge insulated wires, deployed in a manner similar to telephone wires. By upgrading to category-5 wires or using multiple category-3 wires, it is possible to transmit data at rates up to 155 Mb/s. BLANs are beginning to use category-5 twisted-pair wires for transmission at data rates greater than 100 Mb/s. This is especially attractive for linking office or laboratory equipment to telecommunications closets, which are designed to be situated no more than 100 m away from their users. Switches located inside distinct telecommunications closets are then usually connected to each other by means of optical fibers, which enjoy a greater geographical span and lower error rates than electronic links. To reduce electromagnetic interference, it is sometimes necessary to scramble high-speed data traveling over twisted-pair wires. Higher data rates can be achieved more easily by shielding wires to reduce crosstalk and electromagnetic emission. However, unshielded twisted-pair wires are desirable for BLANs because they are compatible with the type of wire used to connect desktop telephones to telecommunications closet equipment.

LANs, being complex systems, require substantial software to operate properly. This is even more true of BLANs, which require more sophisticated control to support all features of conventional LANs, as well as traffic integration. Common software technologies for BLANs include routing, signalling, and network management. Routing implies the ability to discover paths to a destination, signalling the ability to reserve resources along the best path, and network management implies the ability to monitor and control the state of those resources.

Wireless and mobile technologies are on the verge of being introduced into BLANs, but prototypes have yet to appear.

34.4 ATM BLANs

ATM occupies the dominant position among BLANs. Originally conceived as a method for transferring information over the public telecommunications network, ATM was quickly accepted as a way to provide high bandwidth within the premises of a private enterprise. Vendors arose to provide ATM products specialized for LAN use. This includes inexpensive switches, computer-interface cards, and software.

A factor that militates against the spread of ATM BLANs is the incomplete state of ATM standardization. ATM standards are officially under the aegis of the International Telecommunications Union—Telecommunications Standardization Sector (ITU-T). The overall framework of BISDN falls within the ITU's purview. The ATM Forum, a large consortium of organizations whose purpose is to accelerate the acceptance of ATM networking by producing agreements among its members on how to implement specific standards, augments the contributions of the ITU-T. The ATM Forum has defined many of the physical- and link-layer standards for ATM, including 100/140-Mb/s transparent asynchronous transceiver interface (TAXI) and 25-Mb/s Desktop 25 link protocols, which operate on MOFs and twisted pairs, respectively. The Internet Engineering Task Force (IETF) also plays a role in ATM BLANs by its consideration of how to integrate ATM into the Internet protocol suite. It produced the standard for running the classical internet protocol (IP) over ATM.

ATM is a **connection-oriented protocol**, in which a caller requests to exchange data with a recipient. The network is responsible for finding the route and reserving any resources needed to complete the exchange.

Once the call has been established, each hop of the call is associated with a virtual path identifier/virtual circuit identifier (VPI/VCI). The VPI/VCI has only local significance at switches or interfaces and is used to map an incoming cell (the 53-octet [an octet equals 8 bits] fixed-length ATM packet) to its output port. The connection guarantees a specific **quality of service** (QOS), which specifies a contract between the network and the user that the network will satisfy the bandwidth, delay, and jitter requirements of the user provided that the user's traffic meets the constraints of the contract. Several QOS classes may be specified:

- Constant bit rate (CBR), which supports isochronous traffic
- Variable bit rate/real time (VBR/RT), which supports synchronous traffic
- Variable bit rate/nonreal time (VBR/NRT), which supports asynchronous traffic that requires a minimum bandwidth guarantee
- Available bit rate (ABR), which supports asynchronous traffic with no QOS guarantees (best effort) but increases and reduces bandwidth according to competing demands of other connections
- Unspecified bit rate (UBR), which supports asynchronous traffic but possibly also drops cells

ABR and UBR correspond most closely to the QOS provided by existing LANs. CBR will be needed to carry uncompressed video, whereas VBRRT will be used to carry compressed video.

An important consideration is to be able to coexist with existing software packages. For IP-based networks, this largely means supporting applications that have been written using the UNIX sockets application program interface. This type of BLAN uses the classical-IP-over-ATM approach defined by the IETF. For other networks, such as those based on Novell Netware or AppleTalk protocols, the answer is to use the LAN emulation (LANE) standard developed by the ATM forum. Both of these approaches attempt to hide from the application the fact that its underlying network is a connection-oriented ATM BLAN rather than a connectionless ethernet or token ring shared-media LAN. One of the central difficulties in achieving this is that ATM must mimic the broadcasting that is used by the older LANs.

ATM BLANs may use a variety of media. Currently, the most popular media are optical fibers, but unshielded twisted pair (UTP) is fast becoming the media of choice for ATM desktop connection. MOFs frequently use the synchronous digital hierarchy (SDH) to carry synchronous transport signal (STS) level-3 concatenated (STS-3c) frames, which consist of 2430 octets sent every 125 μs for an overall transmission rate of 155.5 Mb/s. In the U.S., SDH is known as synchronous optical network (SONET) and SDH-3c is known as optical carrier level 3 concatenated (OC-3c). The maximum reach of a segment must be less than 2000 m. STS-3c may also be used with UTP, but the maximum reach of a segment is only 100 m.

Many ATM BLAN devices were based on the TAXI standard, a 100-Mb/s MOF interface with 4B/5B encoding that was made popular by the FDDI standard. The appeal of TAXI for ATM reflected the fact that several TAXI chipsets were available at the time of initial ATM BLAN introduction. Also used were 140-Mb/s TAXI devices. The TAXI devices, however, have rapidly been supplanted by SDH devices because of their more universal appeal.

A host of other physical interfaces have been defined for ATM BLANs. These include the Fibre Channel—compatible 155.5-Mb/s MOF and shielded twisted-pair interfaces with 8B/l0B encoding. The STS-1/2 interface—also known as Desktop 25—is a 25.9-Mb/s category-3 UTP interface. An STS-1 UTP interface operates at 51.8 Mb/s. None of these interfaces is widely used in BLANs at this time.

Two principal types of ATM BLAN interfaces are the **private user-network interface** (PUNI) and the **private network–node interface** (PNNI), where private designates that the interfaces do not impinge on public networks. The PUNI is an interface between user equipment and BLAN switches, whereas the PNNI is an interface between BLAN switches. Each type of interface has a protocol associated with it to implement signalling, routing, connection admission control, traffic shaping, traffic policing, and congestion control. These protocols use reserved VCIs to communicate among peers.

Because of their popularity, a great deal of effort has gone into developing special protocols for ATM BLANs. As most networks use IP, the classical-IP-over-ATM protocol was designed to permit connectionless IP to establish a new ATM virtual connection or bind to an existing one by means of an address

resolution protocol similar to that used in shared-media LANs. To support multicasting, the ATM forum has also defined leaf-initiated join protocols that allow point-to-multipoint multicast trees to be grown from the destination leaves to the source root; this is important for LANE operations, which rely heavily on multicasting. Security is also an important concern for BLAN administrators, and the U.S. government-sponsored Fastlane device allows key-agile end-to-end encryption and decryption of cells on a virtual-connection-by-virtual-connection basis [Stevenson, 1995].

34.5 Other BLANs

Although ATM offers the most capable and widely deployed approach to implementing BLANs, several other approaches are available. These approaches are IEEE 802.3 ethernet, IEEE 802.9 isochronous ethernet, and IEEE 802.12 100VG-AnyLAN.

Given ethernet's popularity, it is not surprising that there should be ethernet-based BLANs. Ethernet originally was designed to use the carrier sense multiple access with collision detection protocol, which allows a node to share a LAN segment with other nodes by transmitting only when the medium is inactive and randomly deferring transmission in the event of a collision (simultaneous transmission by two or more nodes). The inefficiency and nondeterminism that accompany this protocol can be partially overcome by partitioning the LAN so that each node resides on its own dedicated segment. The simplest ethernet BLAN consists of 10-Mb/s IEEE 802.3 10BaseT dedicated links connecting nodes to switching hubs. Using ethernet in this way, one can implement a BLAN that delivers 10 Mb/s of bandwidth directly to each user if the switch has adequate packet buffers and the user interface can keep up. Although they provide no mechanisms for supporting specific levels of QOS, such switched ethernets are able to support high and scalable throughput. Nevertheless, the need of some applications for more than 10 Mb/s of bandwidth justifies the same switched-ethernet approach but with 100-Mb/s IEEE 802.3 links. The IEEE 802.3 100-Mb/s ethernet (sometimes called fast ethernet) is virtually identical to the 10-Mb/s standard, except that the data rate has been increased tenfold and almost all time constants have been decimated. User nodes are attached to switching or nonswitching hubs (the latter of which merely acts as a shared medium) either by UTPs or MOFs. The 100BaseTX option uses two category-5 UTPs with 4B/5B encoding, the 100BaseT4 option uses four category-3 UTPs with 8B/6T encoding (which encodes 8 b as six ternary digits), and the 100BaseFX option uses two MOFs with 4B/5B encoding. Moreover, all of these options (including 10BaseT) may be mixed in a single switching hub. Fast ethernet is rapidly gaining wide acceptance, and several vendors supply ethernet and fast ethernet BLAN equipment.

A drawback of using switched ethernet for a BLAN is that there is no explicit support for isochronous traffic. Isochronous ethernet (isoethernet), as defined by IEEE 802.9, overcomes those drawbacks. Isoethernet maintains compatibility with the 10BaseT wiring infrastructure and interoperates with 10BaseT at the hub level. Isoethernet multiplexes onto a category-3 UTP a 10-Mb/s ethernet and a 6.14-Mb/s ISDN stream. Using 4B/5B encoding, these streams are mapped into a 125-μs frame, which dedicates specific quartet (a quartet equals 4 bit) positions of the frame to ethernet or ISDN traffic. The ethernet subframe is subject to the usual collision-backoff-retransmit cycle, but the ISDN subframe can be allocated to several CBR connections, which have a fundamental frequency of 8000 Hz. Its signalling protocol is similar to ATM's. The isoethernet BLAN has not yet been widely deployed.

The IEEE 802.12 100VG-AnyLAN standard is designed to be run over voice-grade (hence, VG) wiring infrastructure. It uses four category-3 bidirectional UTPs to connect user nodes to repeaters, which may be cascaded in a tree configuration. Repeaters poll user nodes for transmission requests, which are granted in a round-robin discipline. Transmissions are not restricted to any single protocol (hence, AnyLAN) and can, therefore, support traffic integration. Transmission requests have either a low or high priority, with high-priority requests pre-empting those of lower priority. The 100VG-AnyLAN BLAN has not been widely deployed.

The complete replacement of a legacy LAN being impractical, it is not uncommon to build a BLAN with a mix of technologies. A backbone ATM network is frequently the point of access for 802.3 LANs,

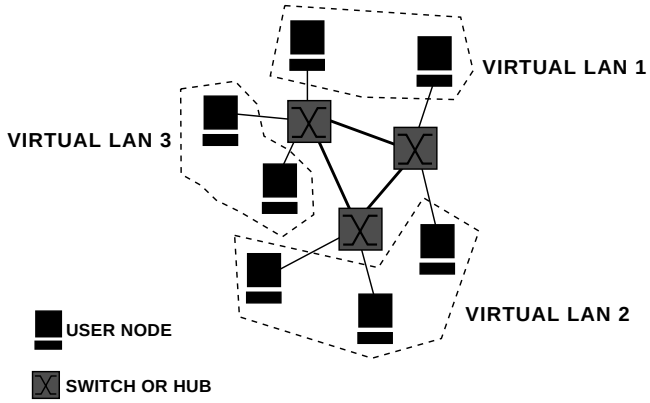


FIGURE 34.3 Virtual LANs.

which can provide from 10 to 100 Mb/s bandwidth to the desktop. User nodes that have higher bandwidth requirements or need a specific QOS may be attached directly to the ATM backbone. The attachment equipment in the backbone—called an edge device—provides ATM and other BLAN interfaces. Working together, these edge devices have the capability to form virtual LANs (VLANs) of nodes logically joined regardless of their physical locations, as shown in Fig. 34.3. Each VLAN forms a unique broadcast domain, in which broadcasts of a node in the VLAN are forwarded to all other members of the VLAN but prevented from propagating to nonmembers. In principle, VLANs should simplify BLAN administration and enhance security and performance. The edge devices and the configurations of VLANs that they support may be managed through MIB variables specially designed to manipulate VLANs. VLAN creation by address, by attachment port, or by several other attributes is currently supported.

34.6 New Applications

Conventional LANs are used principally for a narrow range of client–server applications. BLANs open the door to new computing applications.

Multimedia computing, which processes data, video, image, and audio, is the driving application area for BLANs. The local distribution of video to computer users requires the bandwidth and QOS of a BLAN. With the use of corporate video servers growing, it is essential to have a BLAN that can transport several simultaneous video streams. Such servers are commonly used in training and film production. Other applications include videoconferencing, with workstation-based frame grabbers, cameras, microphones, and electronic whiteboards.

The success of hypermedia browsers, especially those based on the hypertext transport protocol (HTTP), has given rise to the phenomenon of intranetworking, in which enterprises use HTTP to share data among members. As HTTP browsers embrace more functions such as audio, animation, video, and image transfer, the need for higher bandwidth LANs grows. Given the unparalleled traffic growth on the Internet that resulted from HTTP usage, it is reasonable to expect an intranet to experience similar growth. Deployment of BLANs within these intranets thus becomes imperative.

Today, most enterprise LANs are collections of segmented collision domains interconnected by bridges and hubs. A cost-effective strategy for expanding a LAN is to use an ATM BLAN as the backbone. In this manner, existing LANs can be attached to the ATM backbone and power users can also be attached as the need arises. The use of VLAN technology and edge devices just described makes it relatively simple to link existing LANs through an ATM BLAN.

Coupled with high-performance workstations, BLANs make it possible to approximate the power of special-purpose supercomputers for a small fraction of the cost. Needing high throughput and low latency between communicating processes, parallel–distributed computations of hundreds of processes can tax

conventional LANs. The use of networks of workstations (NOWs) to perform tasks formerly delegated to expensive supercomputers demands the power of a BLAN. Experimental NOWs based on ATM BLANs are proving the effectiveness of this approach.

34.7 Conclusion

Although still in its infancy, the BLAN is gaining wide acceptance for upgrading legacy LANs and enabling new applications that require a range of QOS. A wide variety of technologies is used to implement the BLAN, including high-speed integrated circuits, electronic and optical links, and high-level protocol software. By providing guaranteed bandwidth on demand, BLANs enable the integration of different classes of traffic in a single network. This, in turn, makes possible a host of new applications, such as video servers, image servers, hypermedia browsing, and parallel-distributed computing.

ATM has taken the early lead in the race to become the leading BLAN technology. Strong contenders, however, are fast ethernet, isoethernet, and 100BaseVG-AnyLAN. Moreover, these different technologies are often mixed together in edge switches to realize hybridized BLANs.

Although much of the hardware technology behind the BLAN is relatively mature, the complex software-based protocols needed to support broadband applications are only partially complete. Advances in signalling, QOS management, resource reservation, and internetworking must continue before BLANs become permanently established.

Defining Terms

Asynchronous traffic: Traffic with no particular timing relationship between successive transmission units.

Asynchronous transfer mode: A broadband network technology based on fast cell switching.

Broadband integrated services digital network: The cell-relay based technology upon which the next-generation telecommunications infrastructure is to be based.

Broadband local area network: A high-speed LAN that carries integrated traffic.

Collision domains: A group of nodes that share a single broadcast medium.

Connection-oriented protocol: A protocol that must establish a logical association between two communicating peers before exchanging information and then release the association after the exchange is complete.

Desktop 25: A low-speed (25-Mb/s) link protocol for use in desktop ATM networks.

Isochronous traffic: Constant-rate traffic in which successive transmisson units have a strictly fixed timing relationship.

Jitter: The amount of variation in the time between successive arrivals of transmission units of a data stream.

Latency: The delay from transmission of a packet by the source until its reception by the destination.

Management information base: The collection of a network's parameters monitored and controlled by a network management system.

Media-access control protocol: The link-layer protocol in a shared-media LAN that guarantees fair and orderly access to the media by nodes with packets to transmit.

Mesh topology: An arrangement of packet switches connected by dedicated links.

Multicasting: Transmitting a packet within a network so that it will be received by a specified group of nodes.

Quality of service: The bandwidth, latency, and error characteristics provided by a network to a group of users.

Shared-media topology: Refers to a network in which any transmission is heard by all nodes.

Synchronous traffic: Variable-rate traffic in which successive transmission must be delivered within a specific deadline.

Transparent asynchronous transceiver interface: A 100-Mb/s link protocol originally developed for FDDI and later adopted for use in ATM networks.

Virtual path/circuit identifier: A pair of designators in an ATM cell which uniquely identifies within a switch the logical connection to which the cell belongs.

References

Collivignarelli, M., Daniele, A., De Nicola, P., Licciardi, L., Turolla, M., and Zappalorto, A. 1994. A complete set of VLSI circuits for ATM switching. In *Proc. IEEE GLOBECOM'94,* 134–138. Nov., Inst. of Electrical and Electronics Engineers, New York.

Denzel, W.E., Engbersen, A.P.J., and Iliadis, I. 1995. A flexible shared-buffer switch for ATM at Gb/s rates. *Comput. Networks ISDN Syst.,* 27(4):611–634.

Hluchyj, M.G. and Karol, M.J. 1988. Queueing in high-performance packet switching. *IEEE J. Select. Areas Commun.,* 6(9):1587–1597.

Stevenson, D., Hillery, N., and Byrd, G. 1995. Secure communications in ATM networks. *Commun. ACM,* 38(2):45–52.

Tobagi, F.A. 1990. Fast packet switch architectures for broadband integrated services digital networks. *Proc. IEEE,* 78(1):133–167.

Turner, J.S. 1987. Design of a broadcast packet switching network. *IEEE Trans. Commun.,* 36(6):734–743.

Further Information

Two leading textbooks on ATM networks and fast ethernet are:

- M. Händel, N. Huber, and S. Schröder, *ATM Networks: Concepts, Protocols, Applications,* 2nd ed., Addison–Wesley, Reading, MA, 1994.
- W. Johnson, *Fast Ethernet: Dawn of a New Network,* Prentice–Hall, Englewood Cliffs, NJ, 1995.

Several conferences and workshops cover topics in BLANs, including:

- IEEE INFOCOM (Conference on Computer Communications)
- ACM SIGCOMM (Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications)
- IEEE Conference on Local Computer Networks
- IEEE LAN/MANWorkshop

The important journals that cover BLANs are:

- *IEEE/ACM Transactions on Networking*
- *IEEE Network Magazine*
- *Computer Communications Review*
- *Computer Networks and ISDN Systems*
- *Internetworking Research and Experience*

Three prominent organizations that create and maintain BLAN standards are:

- ATM Forum
- IETF
- IEEE Project 802

35

Multiple Access Methods for Communications Networks

Izhak Rubin
University of California

35.1 Introduction

Modern computer communications networks, particularly local area networks (LANs) and metropolitan area networks (MANs), employ multiple access communications methods to share their communications resources. A **multiple-access** communications channel is a network system whose communications media are shared among distributed stations (terminals, computers, users). The stations are distributed in the sense that there exists no relatively low-cost and low-delay mechanism for a single controlling station to gain access to the status of all stations. If such a mechanism exists, the resulting sharing mechanism is identified as a **multiplexing system**.

The procedure used to share a multiple access communications medium is the **multiple access algorithm**. The latter provides for the control, coordination, and supervision of the sharing of the system's communications resources among the distributed stations, which transport information across the underlying multiple access communications network system.

In the following, we present a categorization of the various medium access control (MAC) methods employed for the sharing of multiple access communications channel systems. We demonstrate these schemes by considering applications to many different classes of computer communication networks, including wired and wireless local and metropolitan area networks, satellite communications networks, and local area optical communications networks.

35.2 Features of Medium Access Control Systems

Typically employed topologies for shared-medium communications networks are shown in Fig. 35.1 and are characterized as follows [Rubin and Baker, 1990].

A *star* topology, under which each station can directly access a single central node, is shown in Fig. 35.1(a). A switching star network results when the star node provides store and forward buffering and switching functions, whereas a broadcast star network involves the employment of the star node as an unbuffered repeater, which reflects all of the incoming signals into all outgoing links. A wired star configuration is shown in Fig. 35.1(a1). In Fig. 35.1(a2), we show a wireless cell which employs radio channels for the mobile terminals to communicate with a central base station. The terminals use a multiple access algorithm

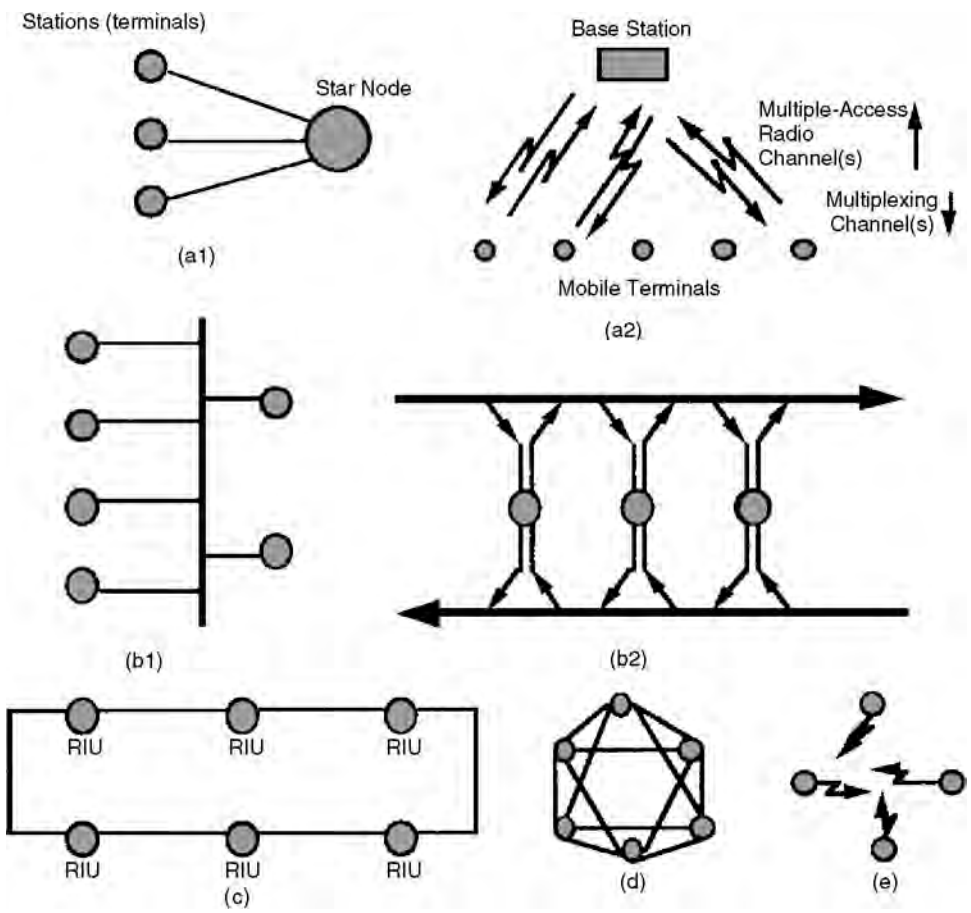


FIGURE 35.1 Multiple-access network topologies: (a) star, (b) bus, (c) ring, (d) mesh, and (e) broadcast radio net.

to gain access to the shared (mobile to base station, reverse access) radio channel(s), whereas the base-station employs a multiplexing scheme for the transmission of its messages to the mobile terminals across the (base-station to terminals, forward access) communications channels.

A *bus* topology, under which stations are connected to a bus backbone channel, is shown in Fig. 35.1(b). Under a logical bus topology, each station's message transmissions are broadcast to all network stations. Under a physical bus implementation, the station is passively connected to the bus, so that a station's failure does not interfere with the operation of the bus network. The station is then able to sense and receive all transmissions that cross its interface with the bus; however, it is not able to strip information off the bus or to properly overwrite information passed along the bus; it can transmit its own messages across the bus when assigned to do so by the employed MAC protocol.

It is noted that when fiber links are used, due to the unidirectional nature of the channel, to achieve full station-to-station connectivity, a bus network implementation necessitates the use of two buses. As depicted in Fig. 35.1(b2), a common approach is to employ two separate counterdirectional buses. Stations can then access each fiber bus through the use of a proper read (sense) tap followed by a write tap.

A *ring* topology, under which the stations are connected by point-to-point links in, typically, a closed-loop topology, is shown in Fig. 35.1(c). In a physical ring network implementation, each station connects to the ring through an active interface, so that transmissions across the ring pass through and are delayed in the register of the ring interface units (RIUs) they traverse. To increase network reliability, such implementations include special circuits for ensuring a rapid elimination of a failed RIU, to prevent such a failure from leading to a prolonged overall network failure. Because of the use of an active interface, the station is now able to strip characters or messages it receives off the medium, as well as to overwrite onto symbols and messages transmitted across the medium, when the latter pass its interface.

An active interface into the medium further enables the station to amplify the signals it passes along, thus leading to considerable reduction of the insertion loss. This is of particular importance in interfacing a fiber optic medium, whereby a passive interface causes a distinct insertion loss, thus leading to a significant limitation in the number of stations that can be passively connected to a fiber bus, without the incorporation of optical amplifiers.

For fiber optic links, electrical amplification and MAC processing at the active interface involve double conversion: optical-to-electronic of the received signal and electronic-to-optical of the amplified transmitted signal. As a result, the station's access burst rate needs to be selected so that it is compatible with the electronics processing rate at the station interface, so that no rate mismatch exists at the electronics/optics interfaces between the potentially very high transmission burst rate across the optical channel and the limited processing rate capacity of the electronically based very large-scale integrated (VLSI) access processor at the station's medium interface unit.

A *mesh* topology is one under which the stations are connected by point-to-point links in a more diverse and spatially distributed (mesh) topology [Fig. 35.1(d)]. To traverse a mesh topology, switching nodes are required. Through the use of cross-connect switches, multiple embedded multiple-access subnetworks are constructed. These subnetworks can be dynamically reconfigured (by adjusting the cross-connect matrices at the switches) to adapt to variations in network loading conditions and in interconnection patterns and to allow the support of private virtual networks. Such an architecture, as demonstrated by the SMARTNet optical network, is presented in Section. 35.7.

A *broadcast* multiaccess radio net, as shown in Fig. 35.1(e) provides, by the physical features of the radio propagation across the shared medium, a direct link between any pair of nodes (assuming the terrain topography to induce no blockages in the node-to-node line-of-sight clearances). A packet transmitted by a single station (node) will be received directly by all other stations. Note that this broadcast property is physically achieved also by the wireline two-way bus network systems shown in Figs. 35.1(bl) and 35.1(b2). In turn, the other network topologies shown in Fig. 35.1 require signal repeating or retransmission protocols to be used to endow them with the broadcast property.

In considering general high-speed local and metropolitan area networks, the following frame distribution topologies can be distinguished for basic implementations: broadcast (logical bus) topology and mesh (switching) topology.

Broadcast (Logical Bus) Topology

Under such a distribution method, station messages are routed through the use of a broadcasting method. Since each message frame (the MAC level protocol data unit) contains addressing information, it is copied automatically from the medium by the intended destination station or stations, so that MAC routing is automatically achieved. Typically, bus, ring, or broadcast star topologies are used to simplify the characterization of the broadcasting path.

The communications link is set up as a bus (for a physical bus implementation), or as a point-to-point link (for a physical ring implementation). The corresponding station interface unit (SIU) acts as a passive or active MAC repeater. A passive MAC node does not interfere with ongoing transmissions along the bus, while being able to copy the messages transmitted across the bus system.

An active MAC station interface unit operates in one out of two possible modes: (1) *repeat mode*, whereby it performs as a repeater, serving to repeat the frame it receives, and (2) *nonrepeat mode*, under which the SIU is not repeating the information it receives from the medium. In the latter case, under a logical bus configuration, the SIU is also typically set to be in a *stripping mode*, stripping from the medium the information it receives. During this time, the SIU is able to transmit messages across the medium, provided it gains permission to do so from the underlying MAC protocol. Note that if the SIU is provided with a store-and-forward capability, it can store all of the information it receives while being in a nonrepeat mode and retransmit this information, if so desired, at a subsequent time.

A logical-bus topology can also be associated with an active interface of the station onto the fiber bus/buses. This is, for example, the case for the Institute of Electrical and Electronics Engineers (IEEE) 802.6 distributed queue dual bus (DQDB) implementation [Rubin and Baker, 1990]. Under DQDB, the station can overwrite each passing bus bit through optical to electrical conversion, an electrical OR write operation, and an electrical to optical reconversion. The active interface of the station can provide it with the capacity to also strip bits, and thus message frames, off the bus.

MAC procedures for logical bus configurations can also be characterized by the following features relating to the method used for removing message frames and the constraints imposed upon the number of simultaneous transmissions carried along the medium. As to the latter feature, we differentiate between the following implementations.

Frame Transmission Initiations in Relation to the Number of Simultaneous Transmissions Along the Logical Bus

Single Message Frame Transmission Across the Medium

A single MAC-frame transmission is permitted across the medium at any time instant; thus, no more than one station can initiate a transmission onto the medium at any given time instant, and no other station can initiate a transmission until this later transmission is removed from the medium.

This is the technique employed by the IEEE 802.5 token ring LAN under a late token release mode, whereby a station holding the token does not release it until it receives its own frame (following the latter's full circulation around the ring). This simplifies the operation of the protocol and provides the station with the ability to review its acknowledged message frame prior to releasing the token, thus enabling it to immediately retransmit its MAC message frame if the latter is determined to have not been properly received by the destination.

The latter token ring LAN uses twisted-pair or coaxial media at transmission rates of 4–16 Mb/s. At 10 Mb/s, the transmission time of a 1000-b frame is equal to 100 μs, which is longer than the typical propagation time across the medium when the overall LAN length is shorter than 20 km, considering a propagation rate of about 5 μs/km.

In turn, when considering a 100-Mb/s LAN or MAN fiber-optic based system, such as the fiber distribution data interface (FDDI) token ring or the DQDB reservation bus systems, which can span longer distances of around 100 km, the corresponding frame transmission time and networkwide propagation delay are equal to 10 and 500 μs, respectively. For a corresponding MAN system, which operates

at a channel rate of 1 Gb/s, the propagation delay of 500 μs is much larger than the frame transmission time of 1 μs. Thus, under such high-transmission rate conditions, each message transmission occupies only a small physical length of the logical bus network medium. Therefore, it is not efficient, from message delay and channel bandwidth MAC utilization considerations, to provide for only a single transmission at a time across the logical bus medium. The following mode of operation is thus preferred. [The following mode, single transmission initiation, can also be used for the token-ring network under the early token release option, which can lead to performance improvements at the higher data rate levels.]

Multiple Simultaneous Frame Transmissions Along the Logical Bus Medium

While permitting multiple frame transmissions across the medium, we can consider two different procedures as it relates to whether a single or multiple simultaneous message frame transmission initiations are allowed.

A *single transmission initiation* is allowed at any instant of time. When a fiber optic-based token ring MAC scheme, such as FDDI, is used, the station releases the token immediately following the transmission of the station's message, rather than waiting to fully receive its own message prior to releasing the token. In this manner, multiple simultaneous transmissions can take place across the bus, allowing for a better utilization of the bus spatial-bandwidth resources.

A slotted access scheme is often used for logical bus linear topologies. Under such a scheme, a bus controller is responsible for generating successive time slots within recurring time cycles. A slot propagates along the unidirectional fiber bus as an idle slot until captured by a busy station; the slot is then designated as busy and is used for carrying the inserted message segment. A busy station senses the medium and captures an available idle slot, which it then uses to transmit its own message segment. For fairness reasons, a busy station may be allowed to transmit only a single segment (or a limited maximum number of segments) during each time cycle. In this manner, such a MAC scheme strives to schedule busy station accesses onto the medium such that one will follow the other as soon as possible, so that a message train is generated efficiently utilizing the medium bandwidth and space (length) dimensions. However, note that no more than a single station is allowed at any time to initiate transmissions onto the unidirectional medium; transmission initiations occur in an order that matches the positional location of the stations along the fiber. As a result, the throughput capacity of such networks is limited by the shared medium's data rate.

Multiple simultaneous frame transmission initiations are included in the MAC protocol. In using such a MAC scheduling feature, multiple stations are permitted to initiate frame transmissions at the same time, accessing the medium at sufficiently distant physical locations, so that multiple transmissions can propagate simultaneously in time along the space dimension of the shared medium. The underlying MAC algorithm needs to ensure that these simultaneous transmissions do not cause any frame overlaps (collisions).

In a ring system, such a procedure can involve the proper use of multiple tokens or ring buffer insertions at each station's interface. In a logical bus operation, when a *slotted* channel structure is employed, such an operation can be implemented through the designation of slots for use by a proper station or group of stations, in accordance with various system requests and interconnectivity conditions. For example, in the DQDB MAN, stations indicate their requests for channel slots so that propagating idle slots along each fiber bus can be identified with a proper station to which they are oriented; in this manner, multiple simultaneous frame transmission initiations and on-going message propagations can take place. Similarly, when time-division multiple access (TDMA) and demand-assigned circuit or packet-switched TDMA MAC schemes are used.

Further enhancements in bandwidth and space utilization can be achieved by incorporating in the MAC scheme an appropriate message frame removal method, as indicated in the following.

Frame Removal Method

When considering a logical bus network with active SIUs, the removal of frames from the logical bus system can be carried out in accordance with the following methods.

Source Removal

Considering loop topologies, under a source removal method the source station is responsible for the removal of its own transmitted frames. This is, as previously noted, the scheme employed by the IEEE 802.5 and fiber-based FDDI token ring systems. Such a MAC feature permits the source station (following the transmission of a frame) to receive an immediate acknowledgment from its destination station (which is appended as a frame trailer), or to identify immediate no-response when the destination station is not operative.

Destination Removal

Under such a scheme, a station, upon identifying a passing frame destined for itself, removes this frame from the medium. Under such a removal policy, a frame is not broadcast across the overall length of medium but just occupies a space segment of the medium that spans the distance between the source and destination stations. Such a method can lead to a more complex MAC protocol and management scheme. Improvement in delay and throughput levels can, however, be realized through spatial reuse, particularly when a noticeable fraction of the system traffic flows among stations that are closely located with respect to each other's position across the network medium.

To apply such a scheme to a token ring system, multiple tokens are allowed and are properly distributed across the ring. When two closely located stations are communicating, other tokens can be used by other stations, located away from the occupied segment(s), to initiate their own nonoverlapping communications paths. Concurrent transmissions are also attained by the use of a slotted access scheme or through the use of a buffer insertion ring architecture, as illustrated later.

When such a scheme is applied to a bus system with actively connected SIUs, which is controlled by a *slotted* access scheme, the destination station is responsible for stripping the information contained in the slot destined to itself from the bus and for releasing the corresponding slot for potential use by a downstream station.

Removal by Supervisory Nodes

It can be beneficial to employ special supervisory nodes, located across the medium, to remove frames from the medium. In this manner, the frame destination removal process can be implemented in only supervisory nodes, relieving regular nodes of this task.

Using such frame removal supervisor nodes, the system interconnectivity patterns can be divided (statically or dynamically) into a number of modes. Under an extensively divisive mode, the supervisory stations allow only communications between stations that are located between two neighboring supervisory stations to take place. Under a less divisive connectivity mode, the system is divided into longer disjoint communications segments. Under a full broadcast mode, each frame is broadcast to all network stations. Depending on the network traffic pattern, time cycles can be defined such that a specific mode of operation is invoked during each time cycle.

If such supervisory stations (or any station) are operated as store-and-forward buffered switching units, then clearly the distribution processes across the medium segments can be isolated, and inter-segment message frames would then be delivered across the logical bus system in a *multiple-hop* (multi-retransmission) fashion. This is the role played by buffer-insertion ring architectures for the specialized ring topology and, in general, by the mesh switching architectures discussed in the following.

Mesh (Switching) Topology

Under a mesh (switching) topology, the network topology can be configured as an arbitrary mesh graph. The network nodes provide buffering and switching services. Communications channels are set up as point-to-point links interconnecting the network nodes. Messages are routed through the network through the use of specially developed routing algorithms.

Depending on whether the switching node performs store-and-forward or cross-connect switching functions, we distinguish between the following architectures.

Store-and-Forward Mesh Switching Architecture

Under a store-and-forward mesh switching architecture, the nodal switches operate in a store-and-forward fashion as packet switches. Thus, each packet received at an incoming port of the switch is examined, and based on its destination address it is queued, switched, and forwarded on the appropriate outgoing port. Each point-to-point channel in such a mesh network needs to be efficiently shared among the multitude of messages and connections that are scheduled to traverse it. A statistical multiplexing scheme is selected (and implemented at each switch output module) for dynamically sharing the internodal links.

Cross-Connect Mesh Switching Architecture

In this case, each switching node operates as a cross-connect switch. The latter serves as a circuit (or virtual circuit) switch, which transfers the messages belonging to an established circuit (or virtual circuit) from their incoming line and time slot(s) [or logical connection groups for cross-connect virtual path switches used by asynchronous transfer mode (ATM) networks] to their outgoing line and time slot(s) (or logical group connections). The cross-connect matrix used by the node to implement this switching function is either preset and kept constant, or it can be readjusted periodically, as traffic characteristics vary, or even dynamically in response to the setup and establishment of end-to-end connections. (See also Section 35.7 for such an optical network identified as SMARTNet.)

Hybrid Mesh Switching Architecture

The nodal switch can integrate fixed assigned and statistical operations in multiplexing traffic across the mesh topology links. For example, in supporting an integrated circuit switched and packet-switched implementation, a time cycle (time frame) is typically defined, during which a number of slots are allocated to the supported circuits that use this channel, while the remainder of the cycle slots are allocated for the transmission of the packets waiting in the buffers feeding this channel. Frequently, priority-based disciplines must be employed in statistically multiplexing the buffered packets across the packet-switched portion of the shared link, so that packets belonging to different service classes can be guaranteed their desired quality of service, as it relates to their characteristic delay and throughput requirements. For example, voice packets must subscribe to strict end-to-end time delay and delay jitter limits, whereas video packets induce high throughput support requirements. Asynchronous transfer mode network structures, which employ combined virtual path and virtual circuit switches, serve as another example.

Hybrid Logical Bus and Buffered Switching Topologies

High-speed multigigabit communications networking architectures that cover wider areas often need to combine broadcast (logical bus) and mesh (switching) architectures to yield an efficient, reliable, and responsive integrated-services fiber-based network.

In considering logical-bus topologies with active SIUs, we previously noted that certain nodes can be designated to act as store-and-forward processors and to also serve as frame removal supervisory nodes. Thus, such nodes actually operate as MAC bridge gateways. These gateways serve to isolate the segments interconnecting them, each segment operating as an independent logical bus network. These gateways act to filter MAC bridge interconnections between these segments. The individual segments can operate efficiently when they serve a local community of stations, noting that each segment spans a shorter distance, thus reducing the effect of the end-to-end propagation delay on message performance. To ensure their timely distribution, it can be effective to grant access priority, in each segment, to the intersegment packets that traverse this segment.

Layered Protocols and the Medium Access Control Sublayer

In relation to the open systems interconnect (OSI) reference model, the data link layer of multiple access (such as typical LAN and MAN) networks is subdivided into the MAC lower sublayer and the logical link control (LLC) upper sublayer. Services provided by the MAC layer allow the local protocol entity to exchange MAC message frames (which are the MAC sublayer protocol data units) with remote MAC entities.

In considering typical LAN and MAN systems, we note that the MAC sublayer provides the following services:

1. MAC sublayer services are provided to the higher layer (such as the LLC sublayer for LAN and MAN systems). LLC service data units (SDUs) are submitted to the MAC sublayer for transmission through proper multiple-access medium sharing. In turn, MAC protocol data units (PDUs) received from the medium and destined to the LLC are transferred to the LLC sublayer as proper LLC-SDUs. The underlying LLC SDUs include source and destination addresses, the data itself, and service class and quality of service parameters.
2. MAC sublayer services are similarly provided to directly connected isochronous and nonisochronous (connection-oriented circuit and packet-switched) channel users (CUs) allowing a local CU entity to exchange CU-data units with peer CU entities. These are connection-oriented services, whereby after an initial connection set-up, the channel user is able to directly access the communications channel through the proper mediation of the MAC sublayer. The CU generates and receives its data units through the MAC sublayer over an existing connection on an isochronous or nonisochronous basis. Such CU-SDUs contain the data itself and possibly service quality parameters; no addressing information is needed since an established connection is involved.

Corresponding services are provided for connectionless flows.

3. MAC sublayer services are provided to the local MAC station management entity via the local MAC layer management interface.

Examples of services include: the opening and closing of an isochronous or nonisochronous connection; its profile, features, and the physical medium it should be transmitted on; and the establishment and disestablishment of the binding between the channel user and the connection endpoint identifier.

The MAC sublayer requires services from the physical layer that provide for the physical transmission and reception of information bits. Thus, in submitting information (MAC frames) to the physical layer, the MAC sublayer implements the medium access control algorithm, which provides its clients (the LLC or other higher layer messages) access to the shared medium. In receiving information from the physical layer, the MAC sublayer uses its implemented access control algorithm to select the MAC frames destined to itself and then to provide them to the higher layer protocol entities. MAC layer addresses are used by the MAC layer entity to identify the destination(s) of a MAC frame.

Figure 35.2 shows the associated layers as implemented by various local and metropolitan area networks, as they relate to the OSI data link and physical layers.

35.3 Categorization of Medium Access Control Procedures

A multitude of access control procedures are used by stations to share a multiaccess communications channel or network system. We provide in this section a classification of MAC schemes.

Medium Access Control Dimensions

The multiple-access communications medium resource can be shared among the network stations through the allocation of a number of key resource dimensions. The assignment space considered is the (T, F, C, S) = {time, frequency, code, space} space. The allocation of access to the shared medium is provided to stations' messages and sessions/calls through the assignment of slots, segments, and cells of the time, frequency, code, and space dimensions.

1. *Time-division scheduling.* Stations share a prescribed channel frequency band by having their transmissions scheduled to take place at different segments of time. Typically, only a single message can be transmitted successfully across the designated channel at any instant of time.
2. *Frequency- and wavelength-division allocation.* The bandwidth of the communications channel is divided into multiple disjoint frequency bands (and wavelength channels for an optical channel)

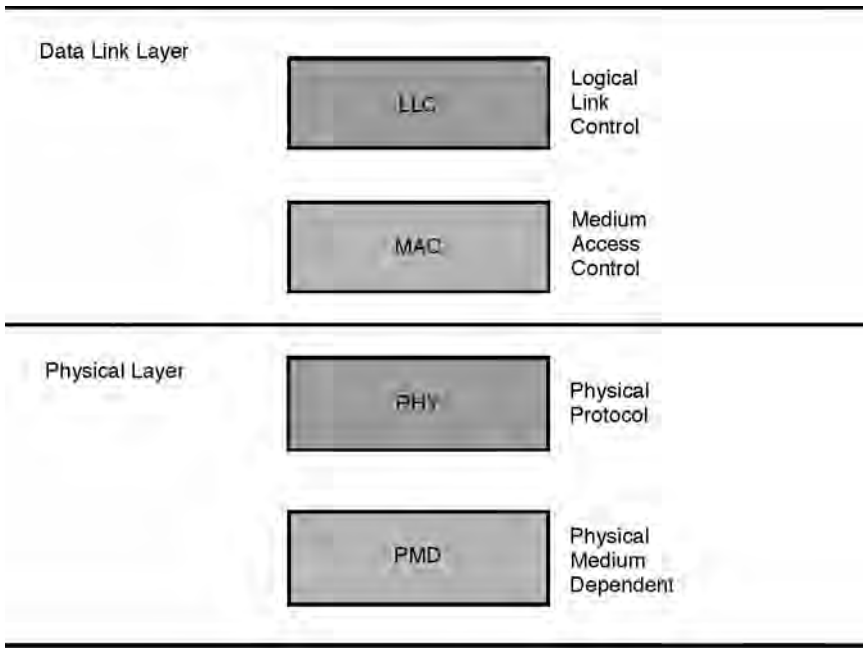


FIGURE 35.2 The MAC sublayer and its interfacing layers across the OSI data link and physical layers.

so that a station can be allocated a group, consisting of one or more frequency/wavelength bands, for use in accessing the medium. Multiple time-simultaneous transmissions can take place across the channel, whereby each message transmission occupies a distinct frequency/wavelength channel.

3. *Code-division multiple access* (*CDMA*). Each station's message is properly encoded so that multiple messages can be transmitted simultaneously in time in a successful manner using a single frequency band of the shared communications channel, so that each message transmission is correctly received by its destined station. Typically, orthogonal (or nearly orthogonal) pseudonoise sequences are used to randomly spread segments of a message over a wide frequency band (frequency hopping method) or to time correlate the message bit stream (direct sequencing method). A message can be encoded by an address-based key sequence that is associated with the identity of the source, the destination, the call/connection, or their proper combination. A wider frequency band is occupied by such code-divided signals. In return, a common frequency band can be used by network stations to successfully carry, simultaneously in time, multiple message transmissions.
4. *Space-division multiple access*. Communications channels are shared along their space dimension. For example, this involves the sharing of groups of physically distinct links or multiple space segments located across a single high-speed logical bus network.

Considering the structure of the access control procedure from the dimensional allocation point of view, note the similarity between the space-division and frequency/wavelength-division methods in that they induce a *channel selection* algorithm that provides an allocation of distinct frequency/wavelength or physical channels to a user or to a group of users. In turn, once a user has been assigned such a frequency/wavelength band, the sharing of this band can be controlled in accordance with a time-division and/or a code-division MAC method. Thus, under a combined use of these dimensions, the medium access control algorithm serves to schedule the transmission of a message by an active station by specifying the selected channel (in a frequency/wavelength-division or space-division manner) and subsequently specifying the time slot(s) and/or multiple-access codes to be used, in accordance with the employed time-division and/or code-division methods.

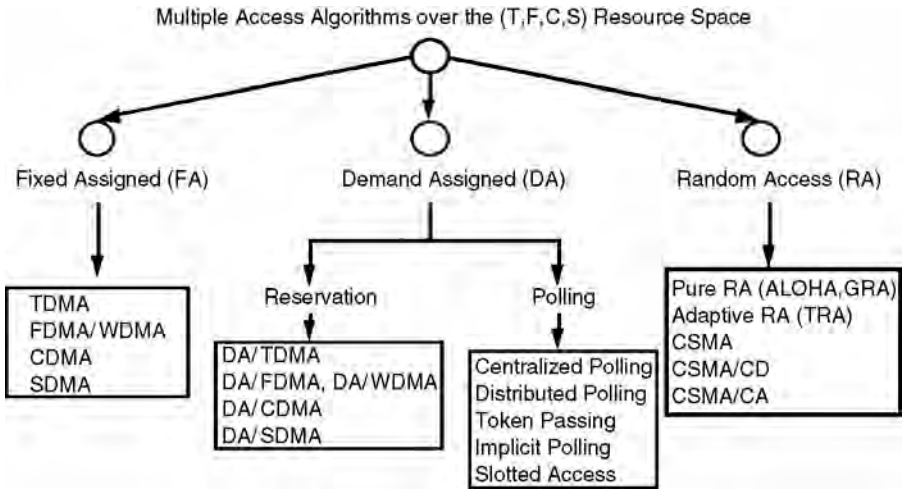


FIGURE 35.3 Classes of medium access control schemes with illustrative access procedures noted for each class.

Medium Access Control Categories

In Fig. 35.3, we show our categorization of medium access control procedures over the previously defined (T, F, C, S) assignment space. Three classes of access control policies are identified: **fixed assignment** (FA), **demand assignment** (DA), and **random access** (RA). Within each class, we identify the signalling (SIG) and control component and the information transmission (IT) method used. Note that within each class, circuit-switching as well as packet-switching mechanisms (under connectionless and/or connection-oriented modes) can be used, in isolation or in an integrated fashion.

Under a *fixed-assignment* scheme, a station is dedicated, over the (T, F, C, S) space, a communications channel resource, which it can permanently use for accessing the channel. Corresponding access control procedures thus include **time-division multiple access (TDMA), frequency-division multiple access (FDMA), wavelength-division multiple access (WDMA), code-division multiple access (CDMA),** and **space-division multiple access (SDMA)** schemes. A corresponding signalling/control procedure is implemented to ensure that station transmission boundaries are well recognized by the participating stations along the respective (T, F, C, S) dimensions.

Medium resources (along each dimension) can be assigned on a fixed basis to specified sessions/connections, to a station, or to a group of stations. When allocated to a connection/call, such schemes provide the basis for the establishment of isochronous circuits and can lead to efficient sharing of the medium-bandwidth resources for the support of many voice, video, high-speed data, and real-time session connections, when a steady traffic stream of information is generated. Effective channel sharing can also result when such channel resources are dedicated for the exclusive use by a station or a group of stations, provided the latter generate steady traffic streams which can efficiently utilize the dedicated resources at an acceptable quality of service level.

For example, under a TDMA procedure, a station is allocated a fixed number of slots during each frame. Under a packet-switched TDMA (PS-TDMA) operation, the station is accessing the shared medium by multiplexing its packets into the allocated slots. In turn, under a circuit-switched TDMA (CS-TDMA) operation, the station uses its allocated time slots to establish circuits. A circuit can consist of a fixed number of the station's time slots allocated in each time frame (e.g., a single slot per frame). Connections are assigned by the station's available circuits at the requested rate. The messages generated by a connection are then transmitted in the time slots belonging to the circuit allocated to this connection.

Under a *random-access* scheme, stations contend for accessing the communications channel in accordance with an algorithm that can lead to time-simultaneous (overlapping, colliding) transmissions by several stations across the same frequency band, causing at times the retransmission of certain packets.

Under fixed random-access schemes, a station transmits its message frame across the channel at a random time, in a slotted or unslotted fashion, without coordinating its access with other stations. A station that detects (through channel sensing, or through nonreceipt of a positive acknowledgment from its destination) its transmission to collide with another transmission, will retransmit its message after a properly computed random retransmission delay (whose duration may depend on the estimated channel state of congestion). Under an adaptive channel-sensing random-access algorithm, a ready station first senses the channel to gain certain channel state information and then uses this information to schedule its message for transmission in accordance with the underlying protocol, again without undertaking full access coordination with all other network stations. Various fixed (such as ALOHA [Abramson, 1973] and group random access [Rubin, 1978]) and adaptive channel-sensing random-access algorithms such as carrier sense multiple access (CSMA) [Kleinrock and Tobagi, 1975], carrier sense multiple access/collision detection (CSMA/CD), dynamic group random access (DGRA) [Rubin, 1983], tree-random-access collision resolution-based [Capetanakis, 1979] and others [Sachs, 1988; Maxemchuk, 1982; Stallings, 1993] can be invoked.

In general, random-access schemes are not well suited as information transmission methods for governing the access of information messages onto a very high-speed fiber communications channel, unless provisions are made to considerably limit the extent of message overlaps due to collisions or to limit the use of this access technique to certain low-throughput messaging classes.

A random-access procedure is typically employed for supporting packet-switching services. At higher normalized throughput levels, high packet delay variances can occur. A proper flow control regulation mechanism must be employed to guarantee maximum delay limits for packets associated with isochronous and real-time based services.

Nonchannel-sensing random-access procedures yield low channel utilization levels. In turn, the efficiency of a channel-sensing random-access procedure critically depends upon the timeliness of the sensed channel occupancy information and, thus, upon the overhead durations associated with sensing the channel. For high-speed logical-bus systems, when channel-sensing type random-access schemes are used, performance degradation is caused by the very high ratio of the channel propagation delay to the packet transmission time. An active station is required to obtain the channel state prior to initiating a transmission. However, since a station's transmission must propagate across the total logical-bus medium before its existence can be sensed by all network stations, and since stations make uncoordinated access decisions under a random-access scheme, these decisions can be based on stale channel state information; as a result, this can lead to the occurrence of excessive message collisions.

The degrading effects of the propagation delay to message transmission time ratio can be somewhat reduced by the following approaches.

Limit the overall propagation delay by dividing the overall network into multiple tiers or segments, as illustrated by the double-tier LAN and MAN architecture presented in Rubin and Tsai [1987]. The network stations are divided into groups. A shared backbone medium can be used to provide interconnection between groups (on a broadcast global basis using repeaters as gateways, or on a store-and-forward bridged/routing basis) using a noncontention (nonrandom-access)-based MAC algorithm. A random-access MAC strategy can be used for controlling the access of messages across the shared group (segment) medium, which now spans a much shorter distance and thus involves considerably reduced propagation delays. Similarly, such a network configuration is realized by interconnecting LAN (such as ethernet) segments by routers, bridges, or switches.

While the propagation delay depends on the conducting medium and is independent of the transmission speed, the packet transmission time is reduced as the transmission speed is increased. The MAC message frame transmission time can be increased by using FDMA/WDMA or SDMA methods to replace the single-band high-bandwidth medium with a group of shared multiple channels, each of lower bandwidth [Chlamtac and Ganz, 1988]. Clearly, this will provide an improvement only if the underlying services can be efficiently supported over such a reduced bandwidth channel when a random-access (or other MAC) scheme is used.

Random-access schemes can serve, as part of other MAC architectures, as effective procedures for transmitting signalling packets for demand-assigned schemes. Random-access schemes can also be used in conjunction with CDMA and spread spectrum multiple access (SSMA) procedures over high-speed channels.

Random-access protocols are further presented and discussed in Section 35.5.

Under a *demand-assigned* scheme, a signalling procedure is implemented to allow certain network entities to be informed about the transmission and networking needs and demands of the network stations. Once these network entities are informed, a centralized or distributed algorithm is used to allocate to the demanding stations communications resource segments over the (T, F, C, S) assignment space.

The specific methods used in establishing the signalling channel and in controlling the access of individual demanding stations onto this channel characterize the features of the established demand-assignment scheme.

The signalling channel can be established as an out-of-band or in-band fixed channel to which communications resources are dedicated, as is the case for many DA/TDMA, DA/FDMA, or DA/WDMA schemes. In turn, a more dynamic and adaptive procedure can be used to announce the establishment of a signalling channel at certain appropriate times, as is the case for the many polling, implicit polling, slotted, and packet-train-type access methodologies.

The schemes used to provide for the access of stations onto the signalling channel can be divided into two categories: **polling** and **reservation** procedures.

Under a reservation scheme, it is up to the individual station to generate a reservation packet and transmit it across the signalling (reservation) channel to inform the network about its needs for communications resources. The station can identify its requirements as they occur or in advance, in accordance with the type of service required (isochronous vs. asynchronous, for example) and the precedence level of its messages and sessions/connections. Reservation and assignment messages are often transmitted over an in-band or out-of-band multiple-access channel so that a MAC procedure must be implemented to control the access of these MAC-signalling messages.

Under a polling procedure, it is the responsibility of the network system to query the network stations so that it can find out their transmission needs, currently or in advance. Polling and polling-response messages can be transmitted over an in-band or out-of-band multiple-access channel, so that a MAC procedure must be implemented to control the access of these MAC-signalling messages.

Under an implicit polling procedure, stations are granted access into the medium (one at a time or in a concurrent noninterfering fashion) when certain network state conditions occur; such conditions can be deduced directly by each station without the need to require the station to capture an explicit polling message prior to access initiation. Such schemes are illustrated by slotted bus systems whereby the arrival of an idle slot at a station serves as an implicit polling message, and by buffer-insertion schemes whereby a station is permitted to access the ring if its ring buffer is empty. Additional constraints are typically imposed on the station regarding utilizing these implicit polling messages for access to the channel to ensure a fair allocation of communications network bandwidth resources. Such constraints often guarantee that a station can transmit a quota of packets within a properly defined cycle.

A multitude of access policies can be adopted in implementing integrated services polling and reservation algorithms. Note that polling techniques are used in implementing the IEEE 802.4 token bus and IEEE 802.5 token ring LANs, the FDDI fiber LAN, and many other fiber bus and ring network systems. Reservation techniques are employed by the IEEE 802.6 DQDB metropolitan area network (whereby a slotted positional priority implicit polling access procedure is used to provide access to reservation bits [Newman and Hullett, 1986]), by demand-assigned TDMA, FDMA, CDMA, and WDMA systems, and many other multiple-access network systems. Integrated reservation and polling schemes are also used. See the following section for further discussion of polling and implicit polling methods.

35.4 Polling-Based Multiple Access Networks

In this section, we describe a number of multiple-access networks whose medium access control architectures are based on polling procedures.

Token-Ring Local Area Network

The token-ring network is a local area network whose physical, MAC, and link layer protocol structures are based on the IEEE 802.5 Standard. The ring interface units are connected by point-to-point directional links, which operate at data rates of 4 Mb/s or 16Mb/s. A distributed polling mechanism, known as a **token passing** protocol, is used as the medium access control scheme. A polling packet, known as a token, is used to poll the stations (through their RIUs). When a station receives the token (buffering it for a certain number of bit times to allow the station to identify the packet as a token message and to react to its reception), it either passes the token along to its downstream neighboring station (when it has no message waiting for transmission across the ring network) or it seizes the token. A captured token is thus removed from the ring so that only a single station can hold a token at a time. The station holding the token is then allowed to transmit its ready messages. The dwell time of this station on the ring is limited through the setting of the station's token holding timer. When the latter expires, or when the station finishes transmitting its frames, whichever occurs first, the station is required to generate a new token and pass it to its downstream neighboring station. The transmitted frames are broadcasted to all stations on the ring: they fully circulate the ring and are removed from the ring by their source station. Such a procedure is also known as a source removal mechanism. The token-ring protocol permits a message priority-based access control operation. For this purpose, the token contains a priority field. A token of a certain priority level can be captured by only those stations which wish to transmit across the medium a message of a priority level equal to or higher than the token's priority level. Furthermore, stations can make a reservation for the issue of a token of a higher priority level by setting a priority request tag in a reservation field contained in circulating frames (such tags can also be marked in the reservation field of a busy token, which is a token that has been captured but not yet removed from the ring, and tagged as busy by the station which has seized it). Upon its release of a new idle token, the releasing station marks the new token at a priority level which is the highest of all of those levels included in the recently received reservations. At a later time, when no station is interested in using a token of such a high-priority level, this station is responsible for downgrading the priority of the token. In setting priority-based timeout levels for the token holding timer, different timeout values can be selected for different priority levels.

The *throughput capacity* attainable by a polling scheme, including a token passing distributed polling mechanism such as the token ring LAN, is dependent on the *walk time* parameter. The system *walk time* is equal to the total time it takes for the polling message (the token) to circulate around the network (the ring) when no station is active. For a token ring network, the walk time is thus calculated as the sum of the transmission times, propagation delay, and ring buffer delays incurred by the token in circulating across an idle ring. Clearly, the throughput inefficiency of the network operation is proportional to the ratio between the walk time and the overall time occupied by message transmissions during a single rotation of the token around the ring. The network's maximum achievable throughput level, also known as the network's throughput capacity, is denoted as $L(C)$ (b/s); the normalized throughput capacity index is set to $s(C) = L(C)/R$, where R (b/s) denotes the data rate of each network link, noting that $0 \leq s(C) \leq 1$. Thus, if $S(C) = 0.8$, the network permits a maximum throughput level that is equal to 80% of the link's data rate. To assess the network's throughput capacity level, we assume the network to be highly loaded so that all stations have frames ready for transmission across the ring. In this case, each cycle (token circulation around the ring) has an average length of $E(C) = NK(F/R) + W$, where N denotes the number of ring stations (RIUs), K denotes the maximum number of frames that a station is allowed to transmit during a single visit of the token, F is the frame's average length (in bits), so that F/R represents the average frame transmission time across a ring's link, whereas W denotes the average duration (in seconds) of the token's walk time. The network's normalized throughput capacity index is thus given by $s(C) = NK(F/R)/E(C)$. Clearly, higher throughput levels are attained as the walk time durations (W) are reduced. Note that $W = R(p)L + (T/R) + N(M/R)$, where $R(p)$ denotes the propagation delay per unit-distance across the medium (typically equal to 5 μs/km for wired links); L (km) is the distance spanned by the ring network so that $R(p)L$ represents the overall propagation delay around the ring; T is the token's length, so that T/R represents the token's transmission time; M denotes the number of bit time

delays incurred at the RIU's interface buffer, so that $N(M/R)$ expresses the overall delay incurred by a frame in being delayed at each interface around the ring.

The delay-throughput performance behavior exhibited by polling systems, such as the distributed polling token-passing scheme of the token-ring LAN, follows the behavior of a single server queueing system in which the server dynamically moves from one station to the next, staying (for a limited time) only at those stations that have messages requiring service (i.e., transmission across the shared medium). This is a stable operation for which message delays increase with the overall loading on the network, as long as the latter is lower than the throughput capacity level $L(C)$. As the loading approaches the latter capacity level, message delays rapidly increase and buffer overflows can be incurred.

The Fiber Data Distribution Interface Network

The fiber data distribution interface network also employs a token passing access control algorithm. It also uses a ring topology. Two counter-rotating fiber optic rings are used (one of which is in a standby mode) so that, upon the failure of a fiber segment, the other ring is employed to provide for a closed-loop topology. The communications link operates at a data rate of 100 Mb/s. The network can span a looped distance of up to 200 km.

As previously described for the token ring network, on the receipt of a token, an idle station passes it to its downstream station after an interface delay. If the station is busy, it will capture the token and transmit its frames until its token holding timer times out. As for the token-ring network, a source removal mechanism is used so that the transmitted frames are broadcasted to all stations on the ring; they fully circulate the ring and are removed from the ring by their source station. The timeout mechanism and the priority support procedure used by FDDI are different than those employed by the token ring network and are described subsequently. When a station terminates its dwell time on the medium, it immediately releases the token. This is identified as an *early token release* operation. For lower speed token-ring implementations, a late token release operation can be selected. Under the latter, the token is released by the station only after it has received all of its transmitted messages and removed them from the ring. Such an operation leads to throughput performance degradations when the network uses higher speed links since then the station holds the token for an extra time, which includes as a component the ring's propagation delay. The latter can be long relative to the frame's transmission time. As a result, higher speed token-ring networks are generally set to operate in the early token release mode.

The FDDI MAC scheme distinguishes between two key FDDI service types: synchronous and asynchronous. Up to eight priority levels can be selected for asynchronous services. To date, most commonly employed FDDI adapter cards implement only the asynchronous priority 1 service; some also provide a synchronous priority service.

The access of frames onto the medium is controlled by a *timed token rotation* (TTR) protocol. Each station continuously records the time elapsed since it has last received the token (denoted as the token rotation time). An initialization procedure is used to select a target token rotation time (TTRT) through a bidding process whereby each station bids for a token rotation time (TRT) and the minimum such time is selected. As previously noted, two classes of service are defined: synchronous, under which a station can capture the token whenever it has synchronous frames to transmit, and asynchronous, which permits a station to capture a token only if the current TRT is lower than the established TTRT. To support multiple priority levels for asynchronous frames, additional time thresholds are defined for each priority level. In this manner, a message of a certain priority level is allowed to be transmitted by its station, when the latter captures the token, only if the time difference between the time this station has already used (at this ring access) for transmitting higher priority messages and the time since the token last visited this station is higher than the corresponding time threshold associated with the underlying message priority level. This priority-based access protocol is similar to the one used for the IEEE 802.4 token bus LAN system.

Using this procedure, stations can request and establish guaranteed bandwidth and response time for synchronous frames. A guaranteed maximum cycle latency-based response time is established for the

ring, since the arrival time between two successive tokens at a station can be shown to not exceed the value of 2 × TTRT.

As a polling scheme, the performance of the FDDI network is limited by the ring walk (W) time. The ring throughput is thus proportional to 1 - W/TTRT. While lower TTRT values (such as 4–8 ms) yield lower guaranteed cycle response times (token intervisit times lower than 8–16 ms), higher TTRT values need to be selected to provide for better bandwidth utilization under higher load conditions. The ring latency varies from a small value of 0.081 ms for 50 stations, 10-km LAN, to a value of 0.808 ms for 500 stations, 100-km LAN. Using a TTRT value of 50 ms for a LAN that supports 75 stations and 30 km of fiber, and having a ring latency W = 0.25 ms, a maximum utilization of 99.5% can be achieved [Ross, 1986].

To provide messages their desired delay performance behavior across the FDDI ring network, it is important to calibrate the FDDI network so that acceptable levels of queueing delays are incurred at the stations' access queues for each service class [Shah et al., 1992]. This can be achieved by the proper selection of the network's MAC parameters, such as the TTRT level, the timeout threshold levels when multipriority asynchronous services are used, and the station synchronous bandwidth allocation level when a station's FDDI adapter card is set also to provide a synchronous FDDI service. The latter service is effective in guaranteeing quality of service support to real-time streams, such as voice, compressed video, sensor data, and high-priority critical message processes, which require strictly limited network delay jitter levels [Shah et al., 1992]. Note that when a token is received by a station that provides an FDDI synchronous service, the station is permitted to transmit its frames that receive such a service (for a limited time, which is equal to a guaranteed fraction of the TTRT) independently of the currently measured TRT. When no such messages are queued at the station at the arrival of the token, the token immediately starts to serve messages which receive asynchronous service, so that the network's bandwidth is dynamically shared among all classes of service.

The delay-throughput performance features of the FDDI networks follow the characteristics of a distributed polling scheme already discussed. The FDDI performance results reported in Shah et al. [1992] were obtained by using the PLANYST™ tool developed by the IRI Corporation. This tool has been also used to calibrate the parameters of FDDI networks through the use of its expert-based analytical routines.

Implicit Polling Schemes

Under an implicit polling multiple-access mechanism, the network stations monitor the shared medium and are then granted access to it by identifying proper status conditions or tags. To illustrate such structures, we consider the slotted channel, register-insertion, positional-priority, and collision-avoidance access protocols.

Under a *slotted channel access protocol*, the shared medium link(s) are time shared through the generation of time slots. Each time slot contains an header which identifies it as busy or idle. In addition, time slots may also be reserved to connections, so that a circuit-switched mode can be integrated with the packet-switched multiple-access mode described in the following. To regulate station access rates, to assign time circuits, and to achieve a fair allocation of channel resources, the time slots are normally grouped into recurring time frames (cycles). A ready station, with packets to transmit across the medium, monitors the medium. When this station identifies an idle slot, which it is allowed to capture, it marks it to be in a busy state and inserts a single segment into this slot. Clearly, in transmitting a packet, a station must break it into multiple segments, whereby the maximum length of a segment is selected such that it fits into a single slot. The packet segments must then be assembled into the original packet at the destination station.

To be able to insert packets into moving slots, the station must actively interface the medium by inserting an active buffer into the channel. A common configuration for such a network is the slotted ring topology. A slotted access protocol can also be used in sharing a linear logical bus topology with active station interfaces. The later configuration is used by the distributed queue dual-bus MAN defined

by the IEEE 802.6 standard. The latter uses a fiber optic-based dual-bus configuration so that each station is connected to two counterdirectional buses.

To regulate the maximum level of bandwidth allocated to each station, in accordance with the class of service provided to the station, and to control the fair allocation of network resources among stations that receive the same class of service, the access algorithm can limit (statically or dynamically) the number of slots that can be captured by a station during each frame. For the DQDB network, a reservation subchannel is established for stations to insert reservation tags requesting for slots to be used for the transmission of their packets. A station is allowed to capture an idle slot which passes its interface only if it has satisfied all of the earlier requests it has received signifying slot reservations made by other stations. The DQDB network also integrates a circuit-switched mode through the establishment of isochronous circuits as part of the call setup procedure. A frame header is used to identify the slots that belong to dedicated time circuits.

Under a *register-insertion* configuration, each station's medium interface card includes a register (buffer), which is actively inserted into the medium. Each packet is again broken down into segments. A station is permitted to insert its segment(s) into the medium when its register contains no in-transit segments (thus deferring the transmission of its own segments until no in-transit segments are passing by) or when the gap between in-transit segments resident in its register is sufficiently long. In-transit packets arriving at the station's interface when the station is in the process of transmitting its own segment are delayed in the register. To avoid register overflows, its size is set to be equal to at least the maximum segment length. The IBM Metaring/Orbit [Cidon and Ofek, 1989] network is an example of such a network system which employs a ring topology as well as the destination removal spatial reuse features to be presented.

At higher speeds, to further increase the throughput efficiency of the shared medium network, a *destination removal* mechanism is used. This leads to *spatial reuse* since different space segments of the network's medium can be used simultaneously in time by different source–destination pairs. For example, for a spatial-reuse slotted channel network (such as slotted ring or bus-based topologies), once a segment has reached its destination station, the latter marks the slot as idle, allowing it to be reused by subsequent stations it visits. Similarly, the use of an actively inserted buffer (as performed by the register-insertion ring network) allows for operational isolation of the network links, providing spatial reuse features.

To assess the increase in throughput achieved by spatial reuse ring networks, assume the traffic matrix to be uniform (so that the same traffic loading level is assumed between any source–destination pair of stations). Also assume that the system employs two counter-rotating rings, so that a station transmits its segment across the ring that offers the shortest path to the destination. Clearly, the maximum path length is equal to half the ring length, while the average path length is equal to one-quarter of the ring length. Hence, an average of four source–destination station pairs simultaneously communicate across the dual ring network. As a result, the normalized throughput capacity achieved by the spatial reuse dual ring network is equal to 400% (across each one of the rings), as compared with a utilization capacity of 100% (per ring) realized when a destination removal mechanism is used. Hence, such spatial reuse methods lead to substantial throughput gains, particularly when the network links operate at high speeds. They are thus especially important when used in ultra-high-speed optical networks, as noted in Section 35.7.

Positional-Priority and Collision Avoidance Schemes

Under a hub-polling (centralized polling) positional-priority scheme, a central station (such as a computer controller) polls the network stations (such as terminals) in an order which is dictated by their physical position in the network (such as their location on a ring network with respect to the central station), or by following a service order table [Baker and Rubin, 1987]. Stations located in a higher position in the ordering list are granted higher opportunities for access to the network's shared medium. For example, considering a ring network, the polling cycle starts with the controller polling station 1

(the one allocated highest access priority). Subsequently, station 2 is polled. If the latter is found to be idle, station 1 is polled again.

In a similar manner, a terminal priority-based distributed implicit polling system is implemented. Following a frame start tag, station 1 is allowed to transmit its packet across the shared medium. A ready station 2 must wait a single slot of duration T (which is sufficiently long to allow station 1 to start transmitting and for its transmission to propagate throughout the network so that all stations monitoring the shared channel can determine that this station is in the process of transmission) before it can determine whether it is allowed to transmit its packet (or segment). If this slot is monitored by station 2 to be idle, it can immediately transmit its packet (segment). Similarly, station i must wait for $(i - 1)$ idle slots (giving a chance to the $i - 1$ higher priority terminals to initiate their transmissions) following the end of a previous transmission (or following a frame start tag when the channel has been idle) before it can transmit its packet across the shared communications channel.

Such a multiple-access mechanism is also known as a collision avoidance scheme. It has been implemented as part of a MAC protocol for high-speed back-end LANs (which support a relatively small number of nodes) such as HyperChannel. It has also been implemented by wireless packet radio net works, such as the TACFIRE field artillery military nets. In assessing the delay-throughput behavior of such an implicit polling distributed control mechanism, we note that the throughput efficiency of the scheme depends critically on the monitoring slot duration T, whereas the message delay behavior depends on the terminal's priority level [Rubin and Baker, 1986]. At lower loading levels, when the number of network nodes is not too large, acceptable message delays may be incurred by all terminals. In turn, at higher loading levels, only higher priority terminals will manage to attain timely access to the network while lower priority ones will be effectively blocked from entering the shared medium.

Probing Schemes

When a network consists of a large number of terminals, each generating traffic in a low duty cycle manner, the polling process can become highly inefficient in that it will require relatively high bandwidth and will occupy the channel with unproductive polling message transmissions for long periods of time. This is induced by the need to poll a large number of stations when only a few of them will actually have a packet to transmit.

A probing [Hayes, 1984] scheme can then be employed to increase the efficiency of the polling process. For this purpose, rather then polling individual stations, groups of stations are polled. The responding individual stations are then identified through the use of a collision resolution algorithm. For example, the following *tree-random-access* [Capetanakis, 1979; Hayes, 1984] algorithm can be employed. Following an idle state, the first selected polling group consists of all of the net stations. A group polling message is then broadcasted to all stations. All stations belonging to this group which have a packet to transmit then respond. If multiple stations respond, a collision will occur. This will be recognized by the controller, which will subsequently subdivide the latest group into two equal subgroups. The process will proceed with the transmission of a subgroup polling message. Using such a binary search type algorithm, all currently active stations are eventually identified. At this point, the probing phase has been completed and the service phase is initiated. All stations which have been determined to be active are then allocated medium resources for the transmission of their messages and streams.

Note that this procedure is similar to a reservation scheme in that the channel use temporally alternates between a signalling period (which is used to identify user requests for network support) and a service period (during which the requesting stations are provided channel resources for the transmission of their messages). Under a reservation scheme, the stations themselves initiate the transmission of their requests during the signalling periods. For this purpose, the stations may use a random-access algorithm or other access methods. When the number of network stations is not too large, dedicated minislots for the transmission of reservation tags can be used.

35.5 Random-Access Protocols

In this section, we describe a number of random-access protocols which are commonly used by many wireline and wireless networks. Random-access protocols are used for networks which require a distributed control multiple-access scheme, avoiding the need for a controller station which distributes (statically or dynamically) medium resources to active network stations. This results in a survivable operation, which avoids the need for investment of significant resources into the establishment and operation of a signalling subnetwork. This is of particular interest when the shared communications medium supports a relatively large number of terminals, operating each at a low duty cycle. When a station has just one or a few packets, it needs to transmit in a timely manner (at low delay levels) on infrequent occasions; it is not effective to allocate to the station a fixed resource of the network (as performed by a fixed assignment scheme). It is also not effective to go through a signalling procedure to identify the station's communications needs prior to allocating it a resource for the transport of its few packets (as performed by demand-assigned polling and reservation schemes). As a consequence, for many network systems, particularly for wireless networks, it is effective to use random-access techniques for the transport of infrequently generated station packets or, when active connections must be sustained, to use random-access procedures for the multiaccess transport of signalling packets.

The key differences among the random-access methods described subsequently are reflected by the method used in performing shared medium status monitoring.

When stations use full-duplex radios and the network is characterized by a broadcast communications medium (so that every transmission propagates to all stations), each station receives the transmissions generated by the transmitting station, including the transmitting station itself. Stations can then rapidly assess whether their own transmission is successful (through data comparison or energy detection). This is the situation for many local area network implementations. In turn, when stations are equipped with half-duplex transceivers (so that they need to turn around their radios to transition between a reception mode and a transmission mode) and/or when a fully broadcast channel is not available (as is the case for mobile radio nets for which topographical conditions lead to the masking of certain stations, lacking line-of-sight connections between certain pairs of stations), the transmitting station cannot automatically determine the status of its transmission. The station must then rely on the receipt of a positive acknowledgment packet from the destination station.

ALOHA Multiple Access

Under the ALOHA random-access method, the network stations do not monitor the status of the shared communications channel. When a ready station receives a packet, it transmits it across the channel at any time (under an unslotted ALOHA scheme) or at the start of time slot (under a slotted ALOHA algorithm, where the length of a slot is set equal to the maximum MAC frame length). If two or more stations transmit packets (frames) at the same time (or at overlapping times), the corresponding receivers will not usually be able to correctly receive the involved packets, resulting in a destructive collision. (Under communications channel capture conditions, the stronger signal may capture the receiver and may be received correctly, while the weaker signals may be rejected.) When a station has determined that its transmitted packet has collided, it then schedules this packet for retransmission after a random retransmission delay. The latter delay can be selected at random from an interval whose length is dynamically determined based on the estimated level of congestion existing across the shared communications channel. Under a binary exponential backoff algorithm, each station adapts this retransmission interval on its own by doubling its span each time its packet experiences an additional collision.

The throughput capacity of an unslotted ALOHA algorithm is equal to $s(C) = 1/(2e) = 18.4\%$, whereas that of a slotted ALOHA scheme is equal to $s(C) = 1/e = 36.8\%$. The remainder of the shared channel's used bandwidth is occupied by original and retransmitted colliding packets. In effect, to reduce the packet delay level and the delay variance (jitter), the loading on the medium must be reduced significantly below

the throughput capacity level. Hence, the throughput efficiency of ALOHA channels is normally much lower than that attainable by fixed-assigned and demand-assigned methods. The random-access network system is, however, more robust to station failures and is much simpler to implement, not requiring the use of complex signalling subnetworks.

The ALHOA shared communications channel exhibits a bistable system behavior. Two distinctly different local equilibrium points of operation are noted. Under sufficiently low-loading levels the system state resides at the first point, yielding acceptable delay-throughput behavior. In turn, under high-loading levels, the system can transition to operate around the second point. Loading fluctuations around this point can lead to very high packet delays and diminishing throughput levels. Thus, under high-loading bursts, the system can experience a high level of collisions, which in turn lead to further retransmissions and collisions, causing the system to produce very few successful transmissions. To correct this unstable behavior of the random-access multiaccess channel, flow control mechanisms must be used. The latter regulate admission of new packets into the shared medium at times during which the network is congested. Of course, this in turn induces an increase in the packet blocking probability or in the delay of the packet at its station buffer.

Carrier Sense Multiple-Access

Under the CSMA random-access method, the network stations monitor the status of the shared communications channel to determine if the channel is busy (carrying one or more transmissions) or is idle. A station must listen to the channel before it schedules its packet for transmission across the channel.

If a ready station senses the channel to be busy, it will avoid transmitting its packet. It then either (under the nonpersistent CSMA algorithm) takes a random delay, after which it will remonitor the channel, or (under the 1-persistent CSMA algorithm) it will keep persisting on monitoring the channel until it becomes idle.

Once the channel is sensed to be idle, the station proceeds to transmit its packet. If this station is the only one transmitting its packet at this time, a successful transmission results. Otherwise, the packet transmission normally results in a destructive collision. Once the station has determined that its packet has collided, it schedules its packet for retransmission after a random retransmission delay.

Many local and regional area packet-radio multiaccess networks supporting stationary and mobile stations, including those using half-duplex radio transceivers, have been designed to use a CSMA protocol. The performance efficiency of CSMA networks is determined by the *acquisition delay index* $a = t(a)/T(P)$, where $T(P)$ denotes the average packet transmission time and $t(a)$ denotes the system's acquisition time delay. The latter is defined as the time elapsed from the instant that the ready station initiates its transmission (following the termination of the last activity on the channel) to the instant that the packet's transmission has propagated to all network stations so that the latter can sense a transmitting station and thus avoid initiating their own transmissions. The acquisition delay $t(a)$ includes as components the network's end-to-end propagation delay, the radio turn-around time, radio detection times of channel busy-to-idle transitions, radio attack times (times to build up the radio's output power), various packet preamble times, and other components. As a result, for half-duplex radios, the network's acquisition delay $t(a)$ may assume a relatively large value. The efficiency of the operation is, however, determined by the factor a, which is given as the ratio of $t(a)$ and the packet transmission time $T(P)$, since once the station has acquired the channel and is the only one currently active on the channel [after a period of length $t(a)$], it can proceed with the uninterrupted transmission of its full packet [for a period of duration $T(P)$].

Clearly, this can be a more efficient mechanism for packet radio networks which operate at lower channel transmission rates. As the transmission rate increases, $T(P)$ decreases while the index a increases so that the delay-throughput efficiency of the operation rapidly degrades. A CSMA network will attain good delay-throughput performance behavior for acquisition delay index levels lower than about 0.05. For index levels around or higher than 0.2, the CSMA network exhibits a throughput level that is lower than that obtained by a slotted (or even unslotted) ALOHA multiaccess net. Under such conditions, the

channel sensing mechanism is relatively ineffective since the window of vulnerability for collisions $[t(a)]$ is now relatively too long. It is thus highly inefficient to use a CSMA mechanism for higher data rate channels as well as for channels which induce relatively long-propagation delays (such as a satellite communication network) or for systems which include other mechanisms that contribute to an increase in the value of the acquisition delay index.

As for the ALOHA scheme, a CSMA network exhibits a bistable behavior. Thus, under loading bursts, the channel can enter a mode under which the number of collisions is excessive so that further loading of the channel results in diminishing throughput levels and higher packet delays. A flow control-based mechanism can be used to stabilize the behavior of the CSMA dynamic network system.

CSMA/CD Local Area Networks

A local area network that operates by using a CSMA/CD access control algorithm incorporates into the CSMA multiple-access scheme previously described the capability to perform collision detection (CD). The station's access module uses a full-duplex radio and appends to its CSMA-based sensing mechanism a CD operation. Once the ready station has determined the channel to be idle, it proceeds to transmit its packet across the shared channel while at the same time it is listening to the channel to determine whether its transmission has resulted in a collision. In the latter case, once the station has determined its packet to be involved in a process of collision, the station will immediately abort transmission. In this manner, colliding stations will occupy the channel with their colliding transmissions only for a limited period of time, the collision detection time $T(\mathrm{CD})$. Clearly, if $T(\mathrm{CD}) < T(P)$, where $T(P)$ denotes the transmission time of a packet (MAC frame) across the medium, the CSMA/CD operation leads to improved delay-throughput performance over that of a CSMA operation.

The ethernet LAN developed by Xerox Corporation is a bus-based network operating at a channel data rate of 10 Mb/s and using a 1-persistent CSMA/CD medium access control algorithm. The physical, MAC, and link layers of such a CSMA/CD network are defined by the IEEE 802.3 standard (and a corresponding ISO standard). Ethernet nets can employ different media types: twisted pair, coaxial cable, fiber optic line, as well as radio links (in the case of an ethernet-based wireless LAN system). The configuration of the ethernet LAN is that of a logical bus so that a frame transmission by a station propagates across the bus medium to all other stations. In turn, the physical layout of the medium can assume a bus or a star topology. Under the latter configuration, all stations are connected to a central hub node, at which point the access lines are connected by a reflecting repeater module so that a transmission received from an access line is repeated into all other lines. Typical ethernet layouts are limited in their geographical span to an overall distance of about 500–2000 m.

The delay-throughput efficiency of the CSMA/CD network operation is determined by the acquisition time delay $t(a)$ and index a, as for the CSMA network. In addition, the network efficiency also depends on the CD time $T(\mathrm{CD})$. It is noted that $T(\mathrm{CD})$ can be as long as twice the propagation delay across the overall span of the bus medium, plus the time required by the station to establish the occurrence of a collision. To ensure that a collision is reliably detected, the power levels of received packets must be sufficiently high. This also serves as a factor limiting the length of the bus and of the distance at which repeaters must be placed. As a result, the network stations are attached to several ethernet segments, each assuming a sufficiently short span. The different ethernet segments are interconnected by gateways or routers. The latter act as store-and-forward switches which serve to isolate the CSMA/CD multiaccess operation of each segment.

As for the random-access mechanisms previously discussed, burst loads applied to the CSMA/CD network can cause large delay-throughput degradations. In supporting application streams which require limited packet delay jitters, it is thus required to prevent the bus from being excessively loaded. Many implementations thus plan the net's offered traffic loading levels to be no higher than 40% (or 4 Mb/s). Also note that, as for other random access schemes, large loading variations can induce an unstable behavior. A flow control mechanism must then be employed to regulate the maximum loading level of the network.

When shorter bus spans are used, or when shorter access link distances are employed in a star configuration, it is possible to operate this network at higher channel data rates. Under such conditions, an ethernet operation at a data rate of 100 Mb/s (or higher) has been implemented.

35.6 Multiple-Access Schemes for Wireless Networks

Under a cellular wireless network architecture, the geographical area is divided into cells. Each cell contains a central base station. The mobile terminals communicate with the base station controlling the cell in which they reside. The terminals use the reverse traffic and signalling channel(s) to transmit their messages to the base station. The base station multiplexes the messages it wishes to send to the cell's mobile terminals across the forward traffic and signalling channels.

First-generation cellular wireless networks are designed to carry voice connections employing a circuit switching method. Analog communications signals are used to transport the voice information. The underlying signalling subnetwork is used to carry the connection setup, termination, and handover signalling messages. The voice circuits are allocated through the use of a reservation based demand-assigned/FDMA scheme. A ready mobile uses an allocated signalling channel for its handset to transmit to its cell's base station a request message for the establishment of a voice circuit. If a frequency channel is available for the allocation of such a circuit (traffic channel), the base station will make this allocation by signalling the requesting handset.

Second-generation cellular wireless networks use digital communications channels. A circuit-switching method is still employed, with the primary service providing for the accommodation of voice connections. Circuits are formed across the shared radio medium in each cell by using either a TDMA-access control scheme (through the periodic allocation of time slots to form an established circuit, as performed by the European GSM and the U.S. IS-54 standards) or by employing a CDMA procedure (through the allocation of a code sequence to a connection's circuit, as carried out by the IS-95 standard). A signalling subnetwork is established. Reverse signalling channels (RSCs) are multiple-access channels which are used by the mobile terminals to transmit to their cells' base station their channel-request packets (for a mobile originating call), as well as for the transmission of paging response packets (by those mobiles which terminate calls). Forward signalling channels (FSCs) are multiplexing channels configured for the transmission of packets from the base station to the mobiles. Such packets include channel-allocation messages (which are sent in response to received channel request packets) and paging packets (which are broadcasted in the underlying location area in which the destination mobile may reside).

For the reverse signalling channels, a random access algorithm such as the ALOHA scheme is frequently employed. For TDMA systems, time circuit(s) are allocated for the random access transmission of signalling packets. For CDMA systems, codes are allocated for signalling channels; each code channel is time shared through the use of a random access scheme (employing, for example, a slotted or unslotted ALOHA multiple-access algorithm).

Paging and channel allocation packets are multiplexed (in a time-shared fashion) across the forward signalling channels. To reduce battery consumption at the mobile terminals, a slotted mode operation can be invoked. In this case, the cell mobiles are divided into groups, and paging messages destined to a mobile belonging to a certain group are transmitted within that group's allocated channels (time slots) [Rubin and Choi, 1996]. In this manner, an idle mobile handset needs to be activated for the purpose of listening to its FSC only during its group's allocated slots.

Under the IEEE 802.11 protocol, terminals access the base station by sharing the radio channel through the use of a CSMA/CA (collision avoidance) scheme.

Third-generation cellular wireless networks are planned to provide for the support of both voice and data services. These networks employ packet-switching principles. Many of the multiple-access schemes described in the previous sections can be employed to provide for the sharing of the mobile-to-base-station multiple-access reverse communications channels. In particular, random-access polling (or implicit polling) and reservation protocols can be effectively implemented. For example, the following versions of a reservation method have been investigated. Under the packet reservation multiple access

(PRMA) [Goodman et al., 1989] or random-access burst reservation procedure, a random-access mechanism is used for the mobile to reserve a time circuit for the transmission of a burst (which consists of a number of consecutively generated packets). For voice bursts, a random-access algorithm is used to govern the transmission of the first packet of the burst across the shared channel (by randomly selecting an unused slot, noting that, in each frame, the base station notifies all terminals as to which slots are unoccupied). If this packet's transmission is successful, its terminal keeps the captured slot's position in each subsequent frame until the termination of the burst's activity. Otherwise, the voice packet is discarded and the next voice packet is transmitted in the same manner. In selecting the parameters and capacity levels of such a multiaccess network, it is necessary to ensure that connections receive acceptable throughput levels and that the packet discard probability is not higher than a prescribed level ensuring an acceptable voice quality performance [Rubin and Shambayati, 1995].

For connection-oriented packet-switching network implementations which also provide data packet transport, including wireless ATM networks, it is necessary to avoid frequent occurrences of packet discards to reduce the rate of packet retransmissions and to lower the chances for packet reordering procedures needed to guarantee the ordered delivery of packets. Reverse and forward signalling channels are established to set up the virtual-circuit connection. Channel resources can then be allocated to established connections in accordance with the statistical features of such connections. For example, real-time connections can be accommodated through the periodic allocation of time slots (or frequency/code resources), whereas burst sources can be supported by the allocation of such resources only for the limited duration of a burst activity. A mechanism must then be employed to identify the start and end times of burst activities. For example, the signalling channels can be used by the mobiles to signal the start of burst activity, whereas the in-band channel is used to detect the end of activity (directly and/or through the use of tagging marks).

35.7 Multiple Access Methods for Spatial-Reuse Ultra-High-Speed Optical Communications Networks

As we have seen above, token-ring and FDDI LANs use token-passing methods to access a shared medium. Furthermore, these networks use a source removal procedure so that each station is responsible for removing its own transmitted frames from the ring. In this manner, each transmitted frame circulates the ring and is then removed by its source station. As a result, the throughput capacity of such networks is limited to a value which is not higher than the ring's channel data rate.

In turn, as observed in Section 35.4 (in connection with implicit polling-based multiple-access networks), the bandwidth utilization of shared medium networks, with particular applications to local and metropolitan area networks operating at high speed using generally fiber optic links, can be significantly upgraded through the employment of spatial-reuse methods. For example, consider a ring network which consists of two counter-rotating unidirectional fiber optic ring topologies. Each terminal (station) is connected to both rings through ring interface units. The bandwidth resources of the rings are shared among the active stations. Assume now that a destination removal method is employed. Thus, when a frame reaches its destination node (RIU), it is removed from the ring by the latter. The network communications resources occupied by this frame (such as a time slot) can then be made available to source nodes located downstream to the removing station (as well as to the removing station itself).

A source station which has a frame to transmit selects the ring that offers the shortest path to the destination node. In this manner, the length of a path will be no longer than 1/2 the ring's length. If we assume a uniform traffic matrix (so that traffic flows between the network terminal nodes are symmetrically distributed), the average path length will be equal to 1/4 the ring's length. As a result, using such a destination removal technique, we conclude that an average of four source-destination flows will be carried simultaneously in time by each one of the two rings. Therefore, the throughput capacity of such a spatial reuse network can reach a level which is equal to 400% the ring's channel data rate for each one of the two rings.

Spatial reuse ring networks can be used to carry traffic on a circuit-switched or packet-switched (connectionless or connection-oriented) basis. For packet-switched network implementations, the following two distributed implicit polling multiple-access methods have been used. Under a slotted ring operation, time slots are generated and circulated across the ring. Each time slot contains a header which identifies it as either empty or busy. A station which senses an empty time slot can mark it as busy and insert a segment (a MAC frame which, for example, can carry an ATM cell) into the slot. The destination station will remove the segment from its slot and will designate the slot as idle so that it can be reused by itself or other stations. Under a buffer-insertion ring multiple-access scheme, each station inserts a buffer into the ring. A busy station will defer the transmission of its packet to packets currently being received from an upstream station. In turn, when its buffer is detected to be empty, the station will insert its segment into the ring. The buffer capacity is set equal to a maximum size packet so that a packet received by a station while it is in a process of transmission of its own packet can be stored in the station's inserted buffer. Each packet is removed from the ring by its destination node.

As previously noted, the DQDB MAN is a slotted dual-bus network which employs active station interfaces. To achieve spatial bandwidth reuse gains of the bus' bandwidth resources, frame removal stations can be positioned across the bus. These stations serve to remove frames which have already reached their destinations so that their slots can be reused by downstream stations. Note that such stations must use an inserted buffer of sufficient capacity to permit reading of the frame's destination node(s) so that they can determine whether a received frame has already reached its destination(s).

In designing optical communications networks, it is desirable to avoid the need for executing store-and-forward switching and operations at the network nodes. Such operations are undesirable due to the large differences existing between the optical communications rates across the fiber optic links and the electronic processing speeds at the nodes, as well as due to the difficulty involved in performing intelligent buffering and switching operations in the optical domain. Networks that avoid such operations (at least to a certain extent) are known as *all-optical networks*.

Systems commonly used for the implementation of a shared medium local area all-optical network employ a star topology. At the center of the star, an optical coupler is used. The coupler serves to repeat the frame transmissions received across any of its incoming fiber links to all of its outgoing fiber links. The coupler can be operated in a passive transmissive mode or in an active mode (using optical amplification to compensate for power losses incurred due to the distribution of the received optical signal across all output links). Each station is then connected by a fiber link to the star coupler so that the stations share a multiple-access communications channel. A multiple-access algorithm must then be employed to control the sharing of this channel among all active stations. Typical schemes employed for this purpose use random-access, polling, and reservation methods. Furthermore, multiple wavelengths can be employed so that a WDMA component can be integrated into the multiple-access scheme. The wavelengths can be statically assigned or dynamically allocated. To illustrate the latter case, consider a reservation-based DA/WDMA procedure. Under a connection-oriented operation, wavelengths are assigned to source-destination nodes involved in a connection. Both transmitter and receiver based wavelength assignment configurations can be used: the transmitter's or receiver's (or both) operating wavelengths can be selected dynamically to accommodate a configured connection.

Because of its broadcast feature, the star-configured optical network does not offer any spatial or wavelength reuse advantages. In turn, the meshed-ring scalable multimedia adaptable meshed-ring terabit (SMARTNet) optical network introduced by Rubin [1994] capitalizes on the integrated use of spatial-reuse and wavelength-reuse methods. In the following, we illustrate the architecture of the SMARTNet configuration. Two counter-rotating fibers make up the ring periphery. The stations access the network through their RIUs, which are actively connected to the peripheral fiber rings. The fiber link resources are shared through the use of a wavelength-division scheme. Multiple wavelengths are used, and each station has access to all wavelengths (or to a limited number of them). The rings are divided into segments. Each segment provides access to multiple stations through their connected RIUs. Each segment is connected to its neighboring segments through the use of wavelength cross-connect routers. Such a router switches messages received at multiple wavelengths from its incoming ports to its outgoing ports.

No store-and-forward switching operations are performed. The switching matrix is preset (on a static or adjustable basis) so that messages arriving across an incoming link at a certain wavelength are always immediately switched to a prescribed outgoing link (at the same wavelength, although wavelength translations can also take place). The wavelength routers are also connected (to certain other wavelength routers) by chord links (each consisting of a pair of counterdirectional fiber links). The routers and their chord links form the chord graph topology.

For each configuration of the chord graph, for each assignment of wavelengths, and for each switching table configuration of the routers, the SMARTNet topology can be divided into wavelength graphs. Each wavelength graph represents a subnetwork topology and is associated with a specific wavelength; all stations accessing this subnetwork (which is set here to be a ring or a bus topology) can communicate to each other across this subnetwork by using the associated wavelength. Within each subnetwork, destination-removal spatial-reuse techniques are used. Among the subnetworks, wavelengths can be reused by a number of subnetworks which do not share links.

As previously discussed for the spatial-reuse ring networks, a SMARTNet system can be used to support connection-oriented circuit-switching and packet-switching (including ATM) architectures as well as connectionless packet-switching networking modes. For packet-switching networks, the links (with each wavelength channel) can be time shared by using slotted or buffer-insertion implicit polling schemes or through the use of other multiple-access methods.

As a network which internally employs cross-connect switches, no random delays or packet discards are incurred within the network. Such a performance behavior is advantageous for the effective support of multimedia services, including voice and video streams (which require low network delay jitter levels) and data flows (which prefer no network packet discards, avoiding the need for packet and segment retransmission and reordering operations). An integrated circuit and packet-switched operation is also readily implemented through the allocation of time, wavelength, and subnetwork channels. Such allocations can also be used to support and isolate virtual subnetworks and their associated user communities.

The SMARTNet configuration illustrated in Fig. 35.4 consists of six wavelength routers; each router has a nodal degree of 4, connected to four incoming and outgoing pairs of counterdirectional links. The chord topology is such that every router is connected to its second-neighboring router (rather than to its neighboring router). Analysis shows the network to require the use of a minimum of five wavelengths [Rubin and Hua, 1995]. Five different groups of wavelength graphs are noted. Each group includes three wavelength graph subnetworks (rings) which reuse the same wavelength. Hence, a wavelength reuse factor of 3 is attained. Each ring subnetwork offers a spatial reuse gain of 2.4 (rather than 4, due to the nonuniform

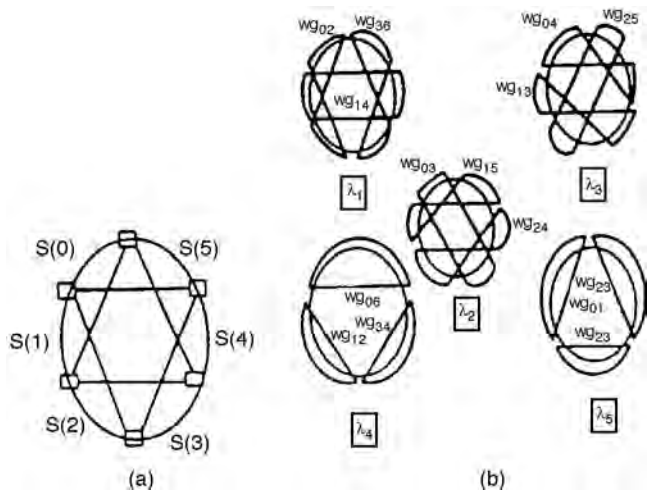


FIGURE 35.4 SMARTNet all-optical network: (a) the network layout; nodes represent wavelength routers; (b) group of wavelength graphs. Five groups are defined, and each group contains three wavelength graphs.

distribution of the traffic across the wavelength graph). Hence, this SMARTNet configuration provides a throughput capacity level of $3 \times 2.4 = 720\%$ of the data rate across each wavelength channel, multiplied by the number of wavelengths used (equal to at least five, and can be selected to be any integral multiple of five to yield the same throughput efficiency factor) and by the number of parallel fiber links employed (two for the basic configuration previously discussed).

It can be shown that an upper bound on the throughput gain achieved by a SMARTNet configuration, which employs wavelength routers of degree four, is equal to 800%. Hence, the illustrated configuration, which requires only six wavelength routers and a minimum of only five wavelengths, yields a throughput gain efficiency which reaches 90% of the theoretical upper bound. Clearly, higher throughput gains can be obtained when the router nodes are permitted to support higher nodal degrees (thus increasing the number of ports sustained by each router). Furthermore, it can be shown that such a network structure offers a throughput performance level which is effectively equal to that obtained when store-and-forward switching nodes are employed for such network configurations.

Defining Terms

Carrier sense multiple access (CSMA): A random access scheme that requires each station to listen to (sense) the shared medium prior to the initiation of a packet transmission; if a station determines the medium to be busy, it will not transmit its packet.

Carrier sense multiple access with collision avoidance (CSMA/CA): A CSMA procedure under which each station uses the sensed channel state and an *a priori* agreed upon scheduling policy to determine whether it is permitted, at the underling instant of time, to transmit its packet across the shared medium; the scheduling function can be selected so that collision events are reduced or avoided.

Carrier sense multiple access with collision detection (CSMA/CD): A CSMA procedure under which each station is also equipped with a mechanism that enables it to determine if its ongoing packet transmission is in the process of colliding with other packets; when a station detects its transmission to collide, it immediately acts to stop the remainder of its transmission.

Code-division multiple access (CDMA): A multiple access procedure under which different messages are encoded by using different (orthogonal) code words while sharing the same time and frequency resources.

Demand assignment multiple access (DAMA): A class of multiple access schemes that allocate dynamically (rather than statically) a resource of the shared medium to an active station; polling and reservation policies are protocols used by DAMA systems; reservation-based TDMA and FDMA systems are also known as demand assigned (DA)/TDMA and DA/FDMA systems.

Fixed assignment multiple access scheme: A multiple access procedure that allocates to each station a resource of the communications medium on a permanent dedicated basis.

Frequency-division multiple access (FDMA): A multiple access method under which different stations are allocated different frequency bands for the transmission of their messages.

Multiple access algorithm: A procedure for sharing a communications resource (such as a communications medium) among distributed end users (stations).

Multiplexing scheme: A system that allows multiple colocated stations (or message streams) to share a communications medium.

Polling scheme: A multiple access method under which the system controller polls the network stations to determine their information transmission (or medium resource) requirements; active stations are then allowed to use the shared medium for a prescribed period of time.

Random access algorithm: A multiple access procedure that allows active stations to transmit their packets across the shared medium without first coordinating their selected access time (or other selected medium's resource) among themselves; this can result in multiple stations contending for access to the medium at overlapping time periods (or other overlapping resource sets), which may lead to destructive collisions.

Reservation scheme: A multiple access method under which a station uses a signalling channel to request a resource of the shared medium when it becomes active; also known as demand assigned scheme.

Time-division multiple access (TDMA): A multiple access scheme under which a station is assigned dedicated distinct time slots within (specified or recurring) time frames for the purpose of transmitting its messages; different stations are allocated different time periods for the transmission of their messages.

Token passing scheme: A polling multiple access procedure that employs a distributed control access mechanism; a token (query) message is passed among the network stations; a station that captures the token is permitted to transmit its messages across the medium for a specified period of time.

References

Abramson, N. 1973. The ALOHA system-another alternative for computer communications. In *Computer Communications Networks,* Eds. N. Abramson and F. Kuo, Prentice–Hall, Englewood Cliffs, NJ.

Baker, J.E. and Rubin, I. 1987. Polling with a general service order table. *IEEE Trans. Commun.,* 35(3):283–288.

Capetanakis, J.I. 1979. Tree algorithm for packet broadcast channels. *IEEE Trans. Inf. Theory,* 25(5):505–515.

Chlamtac, I. and Ganz, A. 1988. Frequency-time controlled multichannel networks for high-speed communication. *IEEE Trans. Commun.,* 36(April):430–440.

Cidon, I. and Ofek, Y. 1989. METARING—a ring with fairness and spatial reuse. Research Report, IBM T.J. Watson Research Center.

Goodman, D.J., Valenzuela, R.A., Gayliard, K.T., and Ramamurthi, B. 1989. Packet reservation multiple access for local wireless communications. *IEEE Trans. Commun.,* 37:885–890.

Hayes, J.F. 1984. *Modeling and Analysis of Computer Communications Networks,* Plenum Press, New York.

Kleinrock, L. and Tobagi, F.A. 1975. Packet switching in radio channels, Part I—Carrier sense multiple access models and their throughput-delay characteristics. *IEEE Trans. Commun.,* 23(Dec.): 1417–1433.

Maxemchuk, N.F. 1982. A variation on CSMA/CD that yields movable TDM slots in integrated voice/data local networks. *Bell Syst. Tech. J.,* 61(7).

Newman, R.M. and Hullett, J.L. 1986. Distributed queueing: a fast and efficient packet access protocol for QPSX. In *Proc. ICCC'86,* Munich.

Ross, F.E. 1986. FDDI-A tutorial. *IEEE Commun. Mag.,* 24(5):10–17.

Rubin, I. 1978. Group random-access disciplines for multi-access broadcast channels. *IEEE Trans. Inf. Theory,* Sept.

Rubin, I. 1983. Synchronous and carrier sense asynchronous dynamic group random access schemes for multiple-access communications. *IEEE Trans. Commun.,* 31(9):1063–1077.

Rubin, I. 1994. SMARTNet: a scalable multi-channel adaptable ring terabit network. Gigabit Network Workshop. *IEEE INFOCOM'94,* Toronto, Canada, June.

Rubin, I. and Baker, J.E. 1986. Performance analysis for a terminal priority contentionless access algorithm for multiple access communications. *IEEE Trans. Commun.,* 34(6):569–575.

Rubin, I. and Baker, J.E. 1990. Medium access control for high speed local and metropolitan area networks. *Proc. IEEE,* Special Issue on High Speed Networks, 78(1):168–203.

Rubin, I. and Choi, C.W. 1996. Delay analysis for forward signalling channels in wireless cellular networks. In *Proc. IEEE INFOCOM'96 Conf.,* San Francisco, CA, March.

Rubin, I. and Hua, H.K. 1995. SMARTNet: an all-optical wavelength-division meshed-ring packet switching network. In *Proc. IEEE GLOBECOM'95 Conf.,* Singapore, Nov.

Rubin, I. and Shambayati, S. 1995. Performance evaluation of a reservation access scheme for packetized wireless systems with call control and hand-off loading. *Wireless Networks J.,* 1(II):147–160.

Rubin, I. and Tsai, Z. 1987. Performance of double-tier access control schemes using a polling backbone for metropolitan and interconnected communication networks. *IEEE Trans. Spec. Topics Commun.,* Dec.

Sachs, S.R. 1988. Alternative local area network access protocols. *IEEE Commun. Mag.,* 26(3):25–45.
Shah, A., Staddon, D., Rubin, I., and Ratkovic, A. 1992. Multimedia over FDDI. In *Proc. IEEE Local Comput. Networks Conf.,* Minneapolis, MN.
Stallings, W. 1993. *Networking Standards: A Guide to OSI, ISDN, LAN, and MAN Standards,* Addison-Wesley, Reading, MA.

Further Information

For performance analysis of multiple access local area networks, see *Performance Analysis of Local Computer Networks*, by J.J. Hammond and P.J.P. O'Reilly, Addison-Wesley, 1986. For further information on related networking standards for local and metropolitan area networks, consult *Networking Standards: A Guide to OSI, ISDN, LAN and MAN Standards,* by W. Stallings, Addison-Wesley, 1993.

36

Routing and Flow Control

Rene L. Cruz
University of California

36.1 Introduction

There are a large number of issues relating to routing and flow control, and we cannot hope to discuss the subject comprehensively in this short chapter. Instead, we outline generic concepts and highlight some important issues from a high-level perspective. We limit our discussion to the context of **packet switched networks**.

Routing and flow control are concerned with how information packets are transported from their source to their destination. Specifically, routing is concerned with *where* packets travel, that is, the paths traveled by packets through the network. Flow control is concerned with *when* packets travel, that is, the timing and rate of packet transmissions.

The behavior of routing and flow control can be characterized by a number of interrelated performance measures. These performance measures can be categorized into two groups: user metrics and network metrics. User metrics characterize performance from the perspective of a single user. Common user metrics are throughput (rate of information transfer), average delay, maximum delay, delay jitter (variations in delay), and packet loss probability. Network metrics characterize performance from the perspective of the network as a whole and all users of it. Common network metrics are network throughput, number of users that can be supported, average delay (averaged over all users), and buffer requirements. Fairness issues may also arise from the perspective of the network.

The goals of routing and flow control are to balance these different performance measures. Optimization of user metrics is often at the expense of network metrics, and vice versa. For example, if the number of users allowed to access the network is too large, the quality of service seen by some users might fall below acceptable levels.

In reality, it is often even difficult to optimize a single performance measure due to the uncertainty of demand placed on the network. Statistical models can be used, but the appropriate parameters of the models are often unknown. Even with an accurate model of demand, the problem of predicting performance for a given routing or flow control strategy is often intractable. For this reason, routing and flow control algorithms are typically based on several simplifying assumptions, approximations, and heuristics.

Before we discuss routing and flow control in more detail, we begin by classifying different modes of operation common in packet switched networks.

36.2 Connection-Oriented and Connectionless Protocols, Services, and Networks

In a connection-oriented protocol or service between two network terminals, the two terminals interact with one another prior to the transport of data, to set up a connection between them. The interaction prior to the exchange of data is required, for example, to identify the terminals to each other and to define the rules governing the exchange of data. For example, users of the telephone network engage in a connection-oriented protocol in the sense that a phone number must first be dialed, and, typically, the parties involved must identify themselves to each other before the essential conversation can begin. The transport control protocol (TCP) used on the Internet is an example of a connection-oriented protocol. Examples of connection-oriented services are rlogin and telnet, which can be built upon TCP. Connection-oriented protocols and services typically involve the exchange of data in both directions. In addition, connection-oriented protocols are frequently used to provide reliable transport of data packets between network endpoints, whereby the network endpoints continue to interact during the data exchange phase in order to manage retransmission of lost or errored packets.

In connectionless protocols or services between two network terminals, no interaction between the terminals takes place prior to the exchange of data. For example, when a letter is mailed over the postal network, a connectionless protocol is used in the sense that the recipient of the letter need not interact with the sender ahead of time. An example of a connectionless protocol is the current Internet protocol (IP) used to transport packets on the Internet. Electronic mail provides a connectionless service, in the same sense that the postal network provides a connectionless service. Because of the lack of interaction between terminals engaging in a connectionless protocol, retransmission of lost or errored packets is not possible, and, hence, connectionless protocols are not considered intrinsically reliable. On the other hand, connectionless protocols are often appropriate in environments where lost or errored packets are sufficiently rare.

A connection-oriented network is a network specifically designed to support connection-oriented protocols. When a connection request is made to a connection-oriented network, the network establishes a path between the source and destination terminal. The switching nodes are then configured to route packets from the connection along the established path. Each switching node is aware of the connections that pass through it and may reserve resources (e.g., buffers and bandwidth) on a per-connection basis. A packet switched network that is connection oriented is also referred to as a **virtual circuit switched network**. Asynchronous transfer mode (ATM) networks are an example of virtual circuit switched networks.

A connectionless network is a network designed to support connectionless protocols. Each node of a connectionless network typically knows how to route packets to each possible destination since there is no prior coordination between terminals before they communicate. Connectionless networks are sometimes referred to as **datagram networks**.

The terminology as defined is not standard, and indeed there does not appear to be widespread agreement on terminology. We have made a distinction between a connection-oriented protocol and a connection-oriented network, and a distinction connectionless protocol and a connectionless network. The reason for these distinctions is because of the many ways in which protocols and networks can be layered. In particular, connection-oriented protocols can be built upon connectionless networks and protocols, and connectionless protocols can be built upon connection-oriented networks and protocols. For example, TCP is in widespread use on end hosts in the Internet, but the Internet itself is inherently a connectionless network which utilizes IP. Another example that has received attention recently is in using an ATM network (connection oriented) as a lower layer transport mechanism to support the IP protocol between Internet nodes. This may be considered somewhat odd, since IP will, in turn, often support the connection-oriented TCP.

Packet switched networks can be further classified as lossy or lossless according to whether or not network nodes will sometimes be forced to drop (destroy) packets without forwarding them to their

intended destination. Connectionless networks are inherently lossy, since the lack of coordination of data sources creates the possibility that buffers inside the network will overflow. Connection-oriented networks can be designed to be lossless. On the other hand, connection-oriented networks can also be lossy, depending on if and how flow control is exercised.

36.3 Routing in Datagram Networks

We will first confine our discussion on routing to the context of connectionless networks. A common approach to routing in connectionless networks is for each node in the network to have enough intelligence to route packets to every possible destination in the network.

A common approach to route computation is *shortest path routing*. Each link in the network is assigned a weight, and the length of a given path is defined as the sum of the link weights along the path. The weight of a link can reflect the desirability of using that link and might be based on reliability or queueing delay estimates on the link, for example. However, another common choice for link weights is unity, and shortest paths in this case correspond to minimum hop paths.

We now briefly review some popular algorithms for computing shortest paths. Suppose there are n nodes in the network, labeled $1, 2, \ldots, n$. The label (i, j) is assigned to a link that connects node i to node j. The links are assumed to be oriented, that is, link (i, j) is distinct from (j, i). We assume that if there is a link (i, j) in the network, then there is also a link (j, i) in the network. We say that node i is a neighbor of node j if the link (i, j) exists. Let $N(i)$ be the set of nodes for which node i is a neighbor. Let w_{ij} denote the weight assigned to link (i, j), and define $w_{ij} = +\infty$ if node i is not a neighbor of node j. We assume that $w_{ij} > 0$ for all i and j. Formally, a path from node i_1, to node i_k is a sequence of nodes $(i_1, i_2, \ldots, i_k)$. The length of such a path is the sum of the weights of links between successive nodes, and the path is said to have $k - 1$ hops. We assume that the network is *strongly connected*, that is, there exists a path between all ordered pairs of nodes with finite length.

Consider the problem of finding shortest (i.e., minimum length) paths from each node to a given destination node. Since each node in the network is a potential destination, this problem will be solved separately for each destination. Without loss of generality, assume that the destination node is node 1. The distance from a node i to the destination node is defined to be the length of the shortest path from node i to the destination node and is denoted as D_i. Define $D_1 = 0$.

Consider a shortest path from a node i to the destination: $(i, j, k, \ldots, 1)$. Note that the subpath that begins at node j, $(j, k, \ldots, 1)$, must be a shortest path from node j to the destination, for otherwise there exists a path from node i to the destination that is shorter than $(i, j, k, \ldots, 1)$. Thus, this subpath has length D_j, and $D_i = w_{ij} + D_j$. Furthermore, since D_i is the length of a shortest path, we have $D_i \le w_{ik} + D_k$ for any k. Thus, we have $D_1 = 0$ and

$$D_i = \min\{w_{ij} + D_j \ : \ 1 \le j \le n\}, \quad 1 < i \le n$$

This is known as *Bellman's equation* and forms the basis of an optimization technique known as dynamic programming. Let $\boldsymbol{D} = (D_1, D_2, \ldots, D_n)$ denote the vector of shortest path lengths, and let $\boldsymbol{f}$ denote the vector valued mapping in Bellman's equation. With these definitions, Bellman's equation says that $\boldsymbol{D}$ is a fixed point of the mapping $\boldsymbol{f}$, that is, $\boldsymbol{D} = \boldsymbol{f}(\boldsymbol{D})$. It can be shown that under our assumption of positive link weights, there is only one solution $\boldsymbol{x}$ to the equation $\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x})$, that is, $\boldsymbol{D}$ is the unique solution to Bellman's equation.

The *Bellman–Ford algorithm* for shortest paths finds the solution to Bellman's equation iteratively. Specifically, let $\boldsymbol{D}^0 = (0, \infty, \infty, \ldots, \infty)$. In the first iteration of the Bellman–Ford algorithm, $\boldsymbol{D}^1 = \boldsymbol{f}(\boldsymbol{D}^0)$ is computed. In the mth iteration, $\boldsymbol{D}^m = \boldsymbol{f}(\boldsymbol{D}^{m-1})$ is computed. It can be shown by induction on m that $\boldsymbol{D}^m$ is the vector of shortest path lengths, where the paths are constrained to have at most m hops. Since link weights are positive, any shortest path visits a node at most once, and, hence, any shortest path has at most $n - 1$ hops. Thus, it follows that the Bellman–Ford algorithm converges within $n - 1$ iterations. In other words, we have $\boldsymbol{f}(\boldsymbol{D}^{n-1}) = \boldsymbol{D}^{n-1} = D$. The shortest paths can be generated from $\boldsymbol{D}$ by examining the argument of the minimum in Bellman's equation. Specifically, for each i, let $\text{Next}(i) = j$, where j is an

argument of the minimum in the equation for D_i, that is, $D_i = w_{ij} + D_j$. With this definition, Next(i) is clearly the first node along a shortest path from node i to the destination. The worst case time complexity of the Bellman–Ford algorithm is easily seen to be $O(n^3)$.

The Bellman–Ford shortest path algorithm is sometimes called a *distance vector* routing algorithm. The algorithm lends itself easily to a distributed, asynchronous implementation that forms the basis for many commonly used routing algorithms. In such an implementation, each node i is responsible for computing the distance from itself to the destination D_i. Each node initially has an estimate of D_i. These initial estimates may be arbitrary, except the destination itself always knows the correct distance from itself to itself, that is, node 1 knows that $D_1 = 0$. Each node i sends, or advertises, to its neighbors the estimate of D_i that it currently has. This advertisement is either done periodically or whenever the estimate changes. Whenever a node receives an estimate from a neighbor, it updates its own estimate in accordance with Bellman's equation. Specifically, if $D_i(t)$ is the estimate of D_i at time t and if $D_j(t - \tau_j)$ is the estimate of D_j from node j that was most recently advertised to node i, then $D_i(t) = \min\{w_{ij} + D_j(t - \tau_j): j \in N(i)\}$. Note that this minimum can be taken over only nodes j which are neighbors of node i, since $w_{ij} = \infty$ if node i is not a neighbor of node j. It can be shown that each node's estimate converges to the actual distance to the destination under essentially arbitrary assumptions regarding propagation delays of advertisements and delays associated with computing estimates [Bertsekas and Gallager, 1992].

In the preceding discussion, we assumed that the link weights were constant. In practice, however, the link weights may change with time. For example, this may be as a result of link failures or changing traffic characteristics on the links. If the link weights change slowly enough, the distributed asynchronous Bellman–Ford algorithm will track the shortest paths correctly as the link weights change. However, if the link weights change fast enough so that the algorithm does not have a chance to converge, problems may arise. For example, if link weights change quickly in response to changing traffic characteristics, it may be possible for packets to travel around in loops indefinitely. In such a scenario, the link weights oscillate as the traffic pattern on the links change due to oscillating estimates of shortest paths [Bertsekas and Gallager, 1992].

Another popular approach to finding shortest paths is for each node to independently compute the shortest paths. Each node i broadcasts the link weight w_{ij} for all outgoing links (i, j) to all other nodes. Each node then has complete information about the network and can compute shortest paths independently. Networks that use this approach are sometimes said to use *link state* routing. Each node could use the Bellman–Ford algorithm as described to compute shortest paths. However, with link state routing, it is common to use *Dijkstra's algorithm* for shortest paths, described next.

Dijkstra's algorithm for finding shortest paths first finds the closest node to the destination node (i.e., the node i other than the destination node with the smallest value of D_i), then the second closest node, then the third closest node, and so on. Let us assume again without loss of generality that node 1 is the destination node. Dijkstra's algorithm is summarized as follows:

Initialization: $m \leftarrow 1$, $P_m \leftarrow \{1\}$, $D_1 \leftarrow 0$.
For all i not in P_m:
$X_i \leftarrow w_{i1} + D_1$
Next(i) $\leftarrow 1$
Iteration:
$i^* \leftarrow$ argmin $\{X_i : i$ not in $P_m\}$
$P_{m+1} \leftarrow P_m \cup \{i^*\}$
$D_{i^*} \leftarrow X_{i^*}$
For all k not in P_{m+1}:
    If $X_k > w_{ki^*} + D_{i^*}$ then
        $X_k \leftarrow w_{ki^*} + D_{i^*}$
        Next(k) $\leftarrow i^*$
$m \leftarrow m + 1$
If $m = n$, then terminate, Otherwise, execute iteration again.

The set P_m consists of the m closest nodes to the destination, including the destination. In other words, for each m we have $|P_m| = m$, and $D_i \geq D_j$ if j belongs to P_m and i does not belong to P_m. To see why Dijkstra's algorithm works, assume after executing the iteration m times that P_{m+1} is the set of $m + 1$ closest nodes, that $X_k = \min\{w_{kj} + D_j : j \in P_{m+1}\}$, and that D_j has been computed correctly for all $j \in P_{m+1}$. Clearly this is true for $m = 0$, and we now show by induction that it is true for all m. Consider the $(m + 1)$st iteration. Suppose that node i^* is a node that could be added to P_{m+1} to yield P_{m+2} with the desired property and that node k is an arbitrary node that cannot be added to P_{m+1} to yield P_{m+2} with the desired property. Note that $D_k > D_{i^*}$. By the induction hypothesis and the fact that all link weights are positive, it thus follows from Bellman's equation that $D_{i^*} = X_{i^*}$. Furthermore, we have $X_k \geq D_k > D_{i^*} = X_{i^*}$. Thus, in the $(m+1)$st iteration, node i^* is chosen correctly to form P_{m+2} and $D_{i^*} = X_{i^*}$. Furthermore, we see that X_k is updated correctly to preserve the induction hypothesis.

It is seen that the time complexity of Dijkstra's algorithm is $O(n^2)$, which is less than the worst case-time complexity of $O(n^3)$ for the Bellman–Ford algorithm. The description of Dijkstra's algorithm can also easily be adapted to address the computation of shortest paths from a given source node to all possible destinations, which makes it well suited for link state routing.

Deflection routing [Maxemchuk, 1987] has been proposed for high-speed networks. With deflection routing, each node attempts to route packets along a shortest path. If a link becomes congested, however, a node will intentionally send a packet to a node which is *not* along a shortest path. Such a packet is said to be *deflected*. If the number of links incoming to a node is equal to the number of outgoing links, and the capacity of all links is the same, packet buffering can be essentially eliminated with deflection routing. It is possible for packets to circulate in loops with deflection routing. With a suitable priority structure, however, delay can be bounded. Deflection routing has also been proposed within high-speed packet switches [Cruz and Tsai, 1996].

In another approach to routing, sometimes called *optimal routing*, the rate of traffic flow between each possible source–destination pair is estimated. Suppose that $r(w)$ is the estimated rate (bits per second) of traffic flow for a given source–destination pair w. A number of possible paths for routing traffic between each source–destination pair are identified, say path 1 through P_w for source–destination pair w. The total flow $r(w)$ is allocated among all of the P_w paths according to a vector $(f_1, f_2, \ldots, f_{P_w})$ such that $r(w) = f_1 + f_2 + \cdots + f_{P_w}$, and such that f_p flows on path p. Typically, bifurcation is allowed, so that f_p may be positive for more than one path p. The path flow allocation for all source–destination pairs is determined simultaneously in order to minimize some cost function. Typically, the cost function is of the form $\Sigma\, C_e(f_e)$, where the sum is taken over all links e in the network and f_e denotes the total flow over all paths which cross link e. A common choice for $C_e(f_e)$ is $C_e(f_e) = f_e/(C - f_e)$, where C is the capacity in bits per second of link e. This would correspond to the average delay on link e if the output queue for link e were an M/M/1 queue. Nonlinear optimization techniques can be used to optimize the path flow allocations with respect to the cost function. The disadvantage of the optimal routing approach is that it is difficult to estimate the rates $r(w)$ *a priori*. In addition, the choice of the cost function to minimize is somewhat arbitrary.

Typically, in connectionless networks, the network nodes will determine where to route a packet so that it gets forwarded to its destination. In such an environment, the source places the address of the destination in a **header field** of each packet. Upon receiving a packet, a network node examines the destination address and uses this to index into a lookup table to determine the outgoing link that the packet should be forwarded on.

In contrast, *source routing* describes the situation where the route that a packet travels is determined by the source. In this case, the source explicitly specifies what route the packet should take within a header field of the packet. Typically, this specification consists of the list of node addresses along the path that the packet should travel. This eliminates the need for large lookup tables within each network node. For example, link state routing can be used, where the sources compute shortest paths.

36.4 Routing in Virtual Circuit Switched Networks

In connection-oriented networks, sources make connection requests to the network, transmit data to the destinations through the connection, and then terminate the connection. Since connection requests are distributed in time, a routing decision is typically made for each connection request as it is made.

In other words, the sequence of connection requests that will be made is unknown, and a routing decision for a connection is not made until the connection request occurs.

Shortest path routing is still a viable approach in connection-oriented networks. The network nodes calculate shortest paths to all possible destinations. When a connection request is made, the connection request is forwarded along a shortest path to the destination, and the switching nodes along the path can determine if they have the resources (e.g., bandwidth, buffers) to support the connection request. If all network nodes determine that they can support the request, the network accepts the connection, link weights are updated to reflect the newly formed connection, and data for the connection is forwarded along the established path. If a network node along the shortest path is not able to support the connection request, the network would then reject the connection request.

It is possible for the network to attempt to route a connection request along another path if the first attempted path fails. However, this should be done carefully, since routing a connection on a path with many hops may use up bandwidth that could be better used to support other connection requests. **Trunk reservation** techniques may be used to advantage in this regard.

Alternatively, the sources might be responsible for specifying a desired path in a connection request, which would be a form of source routing. Obviously in this case, the sources would need knowledge of the network topology in order to make intelligent choices for the routes they request.

In connection-oriented networks, the source need not place the destination address into each packet. Rather, it could put a *virtual circuit identifier* (VCI) in each packet, which labels the connection that the packet belongs to. This can result in a savings of bandwidth if there are a large number of potential destinations and a comparatively smaller number of connections passing through each link. A given connection can have a different VCI assigned on every link along the path assigned to it. When a connection is set up through a node, the node assigns a VCI to the connection to be used on the outgoing link, which may be different from the VCI used for the connection on the incoming link. The node then establishes a binding between the VCI used on the incoming link, the VCI assigned on the outgoing link, as well as the appropriate output link itself. Upon receiving a packet, the network node reads the VCI used on the incoming link and uses this as an index into a lookup table. In the lookup table, the node determines the appropriate outgoing link to forward the packet on, as well as the new VCI to use. This is sometimes referred to as VCI translation.

36.5 Hierarchical Routing

In a network with a very large number of nodes, lookup tables for routing can get very large. It is possible to reduce the size of the lookup tables, as well as the complexity of route computation, by using a hierarchical routing strategy. One approach to a hierarchical routing strategy is as follows. A subset of the network nodes are identified as *backbone* nodes. A link connecting two backbones is called a backbone link. The set of backbone nodes and backbone links is called the backbone network. The nonbackbone nodes are partitioned into clusters of nodes called domains. Each backbone node belongs to a distinct domain and is called the *parent* of all nodes in the domain. The nodes within a domain are typically geographically close to one another. Routes between pairs of nodes in the same domain are constrained not to go outside the domain. Routes between pairs of nodes in different domains are constrained to be a concatenation of a path completely within the source domain, a path completely within the backbone network, and a path completely within the destination domain. Routes are computed at two levels, the backbone level and the domain level. Route computations at the backbone level need only information about the backbone network and can proceed independently of route computations at the domain level. Conversely, route computations at the domain level need only information about the specific domain for which routes are computed. Because of this decoupling, the problem of route computation and packet forwarding can be considerably simplified in large networks.

With a hierarchical routing approach it is often convenient to have a hierarchical addressing strategy as well. For example, for the hierarchical routing strategy just described, a node address could be grouped

into two parts, the domain address and the address of the node within the domain. Packet forwarding mechanisms at the backbone level need only look at the domain part of the node address, and packet forwarding mechanisms at the domain level can ignore the domain part of the node address. Of course, it is possible to have hierarchical routing strategies with more than two levels as described.

36.6 Flow Control in Datagram Networks

Since sources in a connectionless network do not interact with the network prior to the transmission of data, it is possible that the sources may overload the network, causing buffers to overflow and packets to be lost. The sources may become aware that their packets have been lost and, as a result, retransmit the lost packets. In turn, it is possible that the retransmission of packets will cause more buffers to overflow, causing more retransmissions, and so on; the network may be caught in a traffic jam where the network throughput decreases to an unacceptable level. It is also possible that buffers may overflow at the destinations if the sources send data too fast. The objective of flow control in connectionless networks is to throttle the sources, but only as necessary, so that rate of packet loss due to buffer overflow is at an acceptably low level.

A number of schemes have been proposed for flow control in connectionless networks. One technique that has been proposed is that network nodes that experience congestion (e.g., buffer overflow) should send the network nodes causing the congestion special control packets, which, in effect, request that the offending nodes decrease the rate at which packets are injected into the network. This has been implemented in the Internet, and the control packets are called *source quench* packets. A practical problem with this approach is the lack of accepted standards that specify under what conditions source quench packets are sent and exactly how network nodes should react to receiving a source quench packet.

We now focus our discussion on a common flow control technique used on connectionless networks, whereby the network carries traffic predominately from connection-oriented protocols, and *window flow control* is exercised within the connection-oriented protocols.

Window flow control is a technique where a source may have to wait for acknowledgements of packets previously sent before sending new packets. Specifically, a *window size W* is assigned to a source, and the source can have at most *W* unacknowledged packets outstanding. Typically, window flow control is exercised end-to-end, and the acknowledgements are sent back to the source by the destination as soon as it is ready to receive another packet. By delaying acknowledgements, the destination can slow down the source to prevent buffer overflow within the destination.

Another motivation behind window flow control is that if the network becomes congested, packet delays will increase. This will slow down the delivery of acknowledgements back to the source, which will cause the source to decrease its rate of transmissions. The window size *W* is typically chosen large enough so that in the absence of congestion, the source will not be slowed down by waiting for acknowledgements to arrive.

Although it is possible to prevent buffer overflow by choosing the window sizes small enough and appropriately managing the creation of connections, this is typically not how datagram networks operate. In many cases, packet loss due to buffer overflow is possible, and window flow control is combined with an **automatic repeat request (ARQ) protocol** to manage the retransmission of lost packets. In particular, the source will wait only a limited time for an acknowledgement to arrive. After this amount of time passes without an acknowledgement arriving, the source will timeout and assume that the packet has been lost, causing one or more packets to be retransmitted. The amount of time that the source waits is called the *timeout value*.

As already mentioned, retransmissions should be carefully controlled to prevent the network throughput from becoming unacceptably low. We now provide an explanation of how this is commonly handled within the current Internet, specifically within the TCP protocol. The explanation is oversimplified, but it should give a rough idea of how congestion control mechanisms have been built into the TCP protocol. This is the primary method of flow control within the current Internet.

The first idea [Jain, 1986] is that a timeout indicates congestion, and sources should wait for the congestion to clear and not add to it. In the TCP protocol [Jacobson, 1988], each timeout causes the timeout value to be doubled. This is called *exponential backoff*, and is similar to what happens on ethernet networks. In practice, so that the timeout values do not get unacceptably large, exponential backoff is stopped when the timeout value reaches a suitable threshold. After acknowledgements begin to arrive at the source again, the source can reset the timeout value appropriately.

In addition to exponential backoff, the TCP protocol dynamically adjusts the window size in response to congestion. In particular, a congestion window W' is defined. Initially, W' is set to the nominal window size W, which is appropriate for noncongestion conditions. After each timeout, W' is reduced by a factor of two, until $W' = 1$. This is called *multiplicative decrease*.

The window size used by the TCP protocol is at most W', and could be less. The idea is that after a period of congestion, a large number of sources could still retransmit a large number of packets and cause another period of congestion. To prevent this, each time a timeout occurs, the window size is reset to one. Hence, a source is then allowed to send one packet. If the source receives an acknowledgement of this packet, it increases the window size by one, and, hence, is then allowed to send two packets. In general, the source will increase the window size by one for each acknowledgement it receives, until the window size reaches the congestion window size W'. This is called *slow start* and is also used for new connections. After the window size reaches W', the window size is increased by one for every roundtrip delay, assuming no timeouts occur, until the original window size appropriate for noncongestion conditions is reached. This has been called *slower start*.

36.7 Flow Control in Virtual Circuit Switched Networks

In connection-oriented networks, flow control, in a broad sense, can be exercised at two different levels. At a coarse level, when a connection request is made, the network may assess whether it has the resources to support the requested connection. If the network decides that it cannot support the requested connection, the request is blocked (rejected). This is commonly called *admission control*. A connection request may carry information such as the bandwidth of the connection that is required, which the network can use to assess its ability to carry the connection. Some sources do not formally require any specified bandwidth from the network, but would like as high a bandwidth as possible. Such sources are not acutely sensitive to delay; as an example, a file transfer might be associated with such a source. These types of sources have been recently named *available bit rate* (ABR) sources. If all sources are ABR, it may be unnecessary for the network to employ admission control.

At a finer level, when a connection is made, flow control may be exercised at the packet level. In particular, the rate as well as the burstiness of sources may be controlled by the network. We focus our discussion on flow control entirely at the packet level. Proposals have also been made to exercise flow control at an intermediate level, on groups of packets called bursts, but we do not discuss that here.

We first discuss flow control for ABR sources. Given that the bandwidth allocated to an ABR source is variable, the issue of fairness arises. There are many possible ways to define fairness. One popular definition of fairness is *max–min fairness*, which is defined as follows. Each link in the network has a given maximum total bandwidth that is allocated to ABR sources using that link. Each ABR source has an associated path through the network that is used to deliver data from the source to the associated destination. An allocation of bandwidths to the ABR sources is called max–min fair if the minimum bandwidth that is allocated is as large as possible. Given that constraint, the next lowest bandwidth that is allocated is as large as possible, and so on.

A simple algorithm to compute the max–min fair allocation of bandwidths to the ABR sources is as follows. A link is said to be saturated if the sum of the bandwidths of all ABR sources using that link equals the total bandwidth available to ABR sources on that link. Each ABR source is initially assigned a bandwidth of zero. The bandwidth of all ABR sources is increased equally until a link becomes saturated. The bandwidth allocated to each ABR source that crosses a saturated link is then frozen. The bandwidth of all remaining ABR sources are then increased equally until another link becomes saturated, and the

bandwidth of each ABR source that crosses a newly saturated link is then frozen. The algorithm repeats this process until the bandwidth of each ABR source is frozen.

One proposed way to impose flow control on ABR sources is to use window flow control for each ABR connection at the hop level. That is, each ABR connection is subject to a window flow control algorithm for each hop along the path of the connection. Thus, each ABR connection is assigned a sequence of window sizes which specify the window size to be used at each hop along the path for the connection. Each node along the path of a connection reserves buffer space to hold a number of packets equal to the window size for the link incoming to that node. Packet loss is avoided by sending acknowledgements only after a packet has been forwarded to the next node along the path. This scheme results in backpressure that propagates toward the source in case of congestion. For example, if congestion occurs on a link, it inhibits acknowledgements from being sent back on the upstream link, causing the buffer space allocated to the connection at the node feeding the upstream link to fill. This will cause that node to delay sending acknowledgements to the node upstream to that node, and so forth. If congestion persists long enough, the source will eventually be inhibited from transmitting packets. It has been shown that if each node transmits buffered packets from ABR connections that pass through it on a round robin basis, that if window sizes are sufficiently large, and that if each source always has data to send, then the resulting bandwidths that the ABR sources receive are max–min fair [Hahne, 1991]. This scheme is also known as *credit-based flow control* [Kung and Morris, 1995], and algorithms have been proposed for adaptively changing the window sizes in response to demand.

Another way to achieve max–min fairness (or fairness with respect to any criterion for allocated bandwidths) is to explicitly compute the bandwidth allocations and enforce the allocations at the entry points of the network. Explicit enforcement of bandwidth allocations is known as *rate-based flow control*.

Bandwidth is defined in a time-average sense, and this needs to be specified precisely in order to enforce a bandwidth allocation. Suppose packet sizes are fixed. One way to precisely define conformance to a bandwidth allocation of ρ packets per unit time is as follows. In any interval of time of duration x, a source can transmit at most $\rho x + \sigma$ packets. The packet stream departing the source is then said to be (σ, ρ) *smooth*. The parameter σ is a constant positive integer and is a measure of the potential burstiness of the source. Over short intervals of time, the bandwidth of the source can be larger than ρ, but over sufficiently large (depending on the value σ) intervals of time, the bandwidth is, essentially, at most ρ. The source or network can ensure conformance to this by using *leaky bucket flow control* [Turner, 1986], defined as follows.

A source is not allowed to transmit a packet unless it has a permit, and each transmitted packet by the source consumes one permit. As long as the number of permits is less than σ, the source receives new permits once every $(1/\rho)$ units of time. Thus, a source can accumulate up to σ permits. A packet that arrives when there are no permits can either be buffered by the source or discarded. Each ABR source could be assigned the parameters (σ, ρ), where ρ is the allocated bandwidth of the source, and σ represents the allocated burstiness of the source, in some sense.

Non-ABR sources require a certain minimum bandwidth from the network for satisfactory operation and may also require that latency be bounded. Examples are multimedia sources, interactive database sources, distributed computing applications, and real-time applications. Leaky bucket flow control can also be applied to non-ABR sources. The parameter ρ specifies the maximum average bandwidth of the source, and the parameter σ measures the potential burstiness of the source, in some sense. If leaky bucket rate control is applied to all of the sources in the network, delay and buffering requirements can be controlled, and packet loss can be reduced or eliminated. Some recent analyses of this issue are presented in Cruz [1995, 1991a, 1991b]. We now briefly describe one important idea in these analyses.

A packet stream may become more bursty as it passes through network nodes. For example, suppose a packet stream arriving to a node is (σ, ρ) smooth, and that each packet has a delay in the node of at most d seconds. Consider the packet stream as it departs the node. Over any interval of length x, say, $[t, t+x]$, the packets that depart must have arrived to the switch in the interval $[t-d, t+x]$ since the delay is bounded by d. The number of these packets is bounded by $\rho(x + d) + \sigma = \rho x + (\sigma + \rho d)$, since the packet stream is (σ, ρ) smooth as it enters the node. Hence, the packet stream as it departs the node is

$(\sigma + \rho d, \rho)$ smooth and is potentially more bursty than the packet stream as it entered the node. Thus, as a packet stream passes through network nodes, it can get more bursty at each hop, and the buffering requirements will increase at each hop.

We note that buffer requirements at intermediate nodes can be reduced by using rate control (with buffering) to govern the transmission of packets *within* the network [Cruz, 1991b; Golestani, 1991]. This limits the burstiness of packet flow within the network, at the expense of some increase in average delay. It is also possible to reduce delay jitter by using this approach.

Defining Terms

ARQ protocol: An automatic repeat request (ARQ) protocol ensures reliable transport of packets between two points in a network. It involves labeling packets with sequence numbers and management of retransmission of lost or errored packets. It can be exercised on a link or an end-to-end basis.

Datagram network: A packet switched network that treats each packet as a separate entity. Each packet, or datagram, is forwarded to its destination independently. The data sources do not coordinate with the network or the destinations prior to sending a datagram. Datagrams may be lost in the network, without notification of the source.

Header field: A packet is generally divided into a header and a payload. The payload carries the data being transported, whereas the header contains control information for routing, sequence numbers for ARQ, flow control, error detection, etc. The header is generally divided into fields, and a header field contains information specific to a type of control function.

Packet switched network: A network where the fundamental unit of information is a packet. To achieve economy, the communication links of a packet switched network are statistically shared. Buffering is required at network nodes to resolve packet contentions for communication links.

Trunk reservation: A technique used in telephone networks whereby some portion of a communications link is reserved for circuits (trunks) that make most efficient use of the link. A link may reject a circuit request because of trunk reservation so that other circuits, which make better use of the link, may later be accepted.

Virtual circuit switched network: A packet switched network that organizes data packets according to the virtual circuit that they belong to. Packets from a given virtual circuit are routed along a fixed path, and each node along the path may reserve resources for the virtual circuit. A virtual circuit must be established before data can be forwarded along it.

References

Bertsekas, D. and Gallager, R. 1992. *Data Networks,* 2nd ed., Prentice–Hall, Englewood Cliffs, NJ.

Comer, D. 1991. *Internetworking with TCP/IP,* Vol. 1: *Principles, Protocols, and Architecture,* 2nd ed., Prentice–Hall, Englewood Cliffs, NJ.

Cruz, R.L. 1995. Quality of service guarantees in virtual circuit switched networks. *IEEE J. Selec. Areas Commum.,* 13(6):1048–1056.

Cruz, R.L. 1991a. A calculus for network delay, Part I: network elements in isolation. *IEEE Trans. Inf. Th.,* 37(1):114–131.

Cruz, R.L. 1991b. A calculus for network delay, Part II: network analysis. *IEEE Trans. Inf. Th.,* 37(1):132–141.

Cruz, R.L. and Tsai, J.T. 1996. COD: alternative architectures for high speed packet switching. *IEEE/ACM Trans. Netwk.,* 4(1):11–21.

Golestani, S.J. 1991. Congestion-free communication in high speed packet networks. *IEEE Trans. Commun.,* 39(12):1802–1812.

Hahne, E. 1991. Round-robin scheduling for max-min fairness in data networks. *IEEE J. Selec. Areas Commun.,* 9(7):1024–1039.

Jacobson, V. 1988. Congestion avoidance and control. In *Proc. ACM SIGCOMM'88,* Stanford, CA, 314–329.

Jain, R. 1986. A timeout-based congestion control scheme for window flow-controlled networks. *IEEE J. Selec. Areas Commun.,* 4(7):1162–1167.

Kung, H.T. and Morris, R. 1995. Credit-based flow control for ATM networks. *IEEE Netwk. Mag.,* 9(2):40–48.

Maxemchuk, N. 1987. Routing in the Manhattan street network. *IEEE Trans. Commun.,* 35(5):503–512.

Steenstrup, M., Ed. 1995. *Routing in Communication Networks,* Prentice–Hall, Englewood Cliffs, NJ.

Turner, J.S. 1986. New directions in communications. *IEEE Commun. Mag.,* 24(10):8–15.

Further Information

Routing in Communication Networks, edited by Martha E. Steenstrup, is a recent survey of the field of routing, including contributed chapters on routing in circuit switched networks, packet switched networks, high-speed networks, and mobile networks.

Internetworking with *TCP*, Vol. 1, by Douglas E. Comer, provides an introduction to protocols used on the Internet and provides pointers to information available on-line through the Internet.

Proceedings of the INFOCOM Conference are published annually by the Computer Society of the IEEE. These proceedings document some of the latest developments in communication networks.

The journal *IEEE/ACM Transactions on Networking* reports advances in the field of networks. For subscription information contact: IEEE Service Center, 445 Hoes Lane, P.O. Box 1331, Piscataway NJ, 088855-1331. Phone (800) 678-IEEE.

37

Transport Layer

A. Udaya Shankar
University of Maryland

37.1 Introduction

The transport layer of a computer network is situated above the network layer and provides reliable communication service between any two users on the network. The users of the **transport service** are e-mail, remote login, file transfer, Web browsers, etc. Within the transport layer are **transport entities**, one for each user. The entities execute distributed algorithms, referred to as **transport protocols**, that make use of the message-passing service of the network layer to offer the transport service. The most well-known transport protocols today are the Internet Transmission Control Protocol (TCP), the original transport protocol, and the ISO TP4.

Historically, transport protocol design has been driven by the need to operate correctly in spite of unreliable network service and failure-prone networks and hosts. The channels provided by the network layer between any two entities can lose, duplicate, and reorder messages in transit. Entities and channels can fail and recover. An entity failure is fail-stop, that is, a failed entity performs no actions and retains no state information except for stable storage. A channel failure means that the probability of message delivery becomes negligible; that is, even with retransmissions, a message is not delivered within a specified time. These fault-tolerance considerations are still valid in today's internetworks.

To a first approximation, we can equate transport service to reliable data transfer between any two users. But reliable data transfer requires resources at the entities, such as buffers and processes for retransmitting data, reassembling data, etc. These resources typically cannot be maintained across failures. Furthermore, maintaining the resources continuously for every pair of users would be prohibitively inefficient because only a very small fraction of user pairs in a network exchange data with any regularity, especially in a large network such as the Internet.

Therefore, a transport service has two parts: **connection management** and **data transfer**. Data transfer provides for the reliable exchange of data between connected users. Connection management provides for the establishment and termination of connections between users. Users can open and close connections to other users and can accept or reject incoming connection requests. Resources are acquired when a user enters a connection and released when the user leaves the connection. An incoming connection request is rejected if the user has failed or its entity does not have adequate resources for new connections.

A key concern of transport protocols is to ensure that a connection is not infiltrated by old messages that may remain in the network from previous terminated connections. The standard techniques are to use the **3-way handshake** mechanism for connection management and the **sliding window** mechanism

for data transfer within a connection. These mechanisms use **cyclic sequence numbers** to identify the connection attempts of a user and the data blocks within a connection. The protocol must ensure that received cyclic sequence numbers are correctly interpreted, and this invariably requires the network to enforce a **maximum message lifetime**.

In the following section, we describe the desired transport service. We then describe a simplified transport protocol that illustrates the core of TCP and ISO TP4. We conclude with a brief mention of performance issues and *minimum-latency* transport protocols.

Unless explicitly mentioned, we consider a **client–server** architecture. That is, the users of the transport layer are partitioned into clients and servers, and clients initiate connections to servers. Every user has a unique identity (id).

37.2 Transport Service

This section explains the correctness properties desired of the transport layer. We first define the notion of **incarnations**. An incarnation of a client is started whenever the client requests a connection to any server. An incarnation of a server is started whenever the server accepts a (potentially new) connection request from any client. Every incarnation is assigned an **incarnation number** when it starts; the incarnation is uniquely distinguished by its incarnation number and entity id.

Once an incarnation x of an entity c is started in an attempt to connect to an entity s, it has one of two possible futures. The first possibility is that at some point x becomes open and acquires an incarnation number y of some incarnation of s—we refer to this as "x becomes open to incarnation y of s"; at some later point, x becomes closed. The second possibility is that x becomes closed without ever becoming open. This can happen to a client incarnation either because its connection request was rejected by the server or because of failure (in the server, the client, or the channels). It can happen to a server incarnation either because of failure or because it was started in response to a connection request that later turns out to be a duplicate request from some old, now closed, incarnation.

Because of failures, it is also possible that an incarnation x of c becomes open to incarnation y of s but y becomes closed without becoming open. This is referred to as a *half-open* connection.

A **connection** is an association between two open incarnations. Formally, a connection exists between incarnation x of entity c and incarnation y of entity s if y has become open to x and x has become open to y. The following properties are desired of connection management:

- Consistent connections: If an incarnation x of entity c becomes open to an incarnation y of entity s, then incarnation y is either open to x or will become open to x unless there are failures.
- Consistent data transfer: If an incarnation x of entity c becomes open to an incarnation y of entity s, then x accepts received data only if sent by y.
- Progress: If an incarnation x of a client requests a connection to a server, then a connection is established between x and an incarnation of the server within some specified time, provided the server does not reject x's request and neither client, server, nor channels fail within that time.
- Terminating handshakes: An entity cannot stay indefinitely in a state (or set of states) where it is repeatedly sending messages expecting a response that never arrives. (Such *infinite chatter* is worse than deadlock because in addition to not making progress, the protocol is consuming precious network resources.)

Given a connection between incarnations x and y, the following properties are desired of the data transfer between x and y:

- In-sequence delivery: Data blocks are received at entity $y(x)$ in the same order as they were sent by $x(y)$.
- Progress: A data block sent by $x(y)$ is received at $y(x)$ within some specified time, provided the connection is not terminated (either intentionally or due to failures) within that time.

37.3 Data-Transfer Protocol

This section describes the sliding-window method for achieving reliable flow controlled data transfer, assuming that users are always connected and correctly initialized. Later we add connection management to this protocol.

Consider two entities *c* and *s* connected by unreliable network channels. The user at *c* produces data blocks to be delivered to the user at *s*. Because the channels can lose messages, every data block has to be retransmitted until it is acknowledged. For throughput reasons, entity *c* should be able to have several data blocks outstanding, that is, sent but not acknowledged. Similarly, entity *s* should be able to buffer data blocks received out-of-sequence (due to retransmissions or network reordering).

Let us number the data blocks produced by user *c* with successively increasing sequence numbers, starting from 0. The sliding window mechanism maintains two windows of sequence numbers, one at each entity.

Entity *c* maintains the following variables:

- *na*: {0, 1,… }; initially 0. This is the number of data blocks sent and acknowledged.
- *ns*: {0, 1,…}; initially 0. This is the number of data blocks produced by the local user.
- *sw*: {0, 1,…,*SW*}; initially 0. This is the maximum number of outstanding data blocks that the entity can buffer.

Data blocks *na* to *ns* - 1 are outstanding and must be buffered. The entity can accept an additional *sw* - *ns* + *na* data blocks from its user, that is, data blocks *ns* to *na* + *sw* - 1. The sequence numbers *na* to *na* + *sw* - 1 constitute the **send window**, and *sw* is its size.

Entity *s* maintains the following variables:

- *nr*: {0, 1,…}; initially 0. This is the number of data blocks delivered to the user.
- *rw*: {0, 1,…,*RW*}; initially 0. This is the maximum number of data blocks that the entity can buffer.

Data blocks 0 to *nr* - 1 have been delivered to the user in sequence. A subset, perhaps empty, of the data blocks *nr* to *nr* + *rw* - 1 has been received and the blocks are buffered. The sequence numbers *nr* to *nr* + *rw* - 1 constitute the **receive window**, and *rw* is its size.

The easiest way for an entity to identify a received data block or acknowledgement is for the message to include the sequence number of the concerned data block. But such a sequence number field would grow without bound. Instead, the protocol uses cyclic sequence numbers; that is, *j* mod *N*, for some *N*, instead of the *unbounded* sequence number *j*.

To ensure that a received cyclic sequence number is correctly interpreted, it is necessary for the network to enforce a maximum message lifetime, that is, no message older than the lifetime remains in the network. It then suffices if *N* satisfies

$$N \geq SW + RW + \frac{L}{\delta}$$

where *SW* and *RW* are the maximum sizes of the send and receive windows, *L* is the maximum message lifetime, and δ is the minimum time between successive data block productions.

Flow control is another issue in data transfer, that is, entity *c* should not send data faster than the network or entity *s* can handle. By dynamically varying the send window size, the sliding window mechanism can also achieve flow control. In particular, consumer-directed flow control works as follows: entity *s* dynamically varies its receive window size to reflect local resource constraints and regularly requests entity *c* to set its send window size to the receive window size. Note that in this case, the previous condition on *N* reduces to $N \geq 2RW + L/\delta$.

We finish this section with a specification of the sliding window protocol. The data messages of the protocol have the form (D, *sid, rid, data, cj*), where *sid* is the sender's id, *rid* is the intended receiver's id, *data* is a data block, and *cj* is its cyclic sequence number.

```
Entity c

Accept datablock from user (data)
 ec: ns-na < sw
 ac: buffer data as datablock ns;
     ns:= ns+1

Send datablock (j)        //also resends
 ec: 0 ≤ j ≤ ns-na-1
 ac: Send (D, c, s, datablock j, j mod N)

Reception of (ACK, s, c, cj, w)
 ac: tmp := (cj-na) mod N;
     if 1 ≤ tmp ≤ ns-na then
        delete datablocks na to na+tmp
                        from send window;
        na := na+tmp;
        sw := w
     else if tmp = 0 then sw :=  max(sw, w)
     // else tmp > ns-na; do nothing
```

```
Entity s

Expand receive window
 ec: rw < RW
 ac: rw := rw+1

Deliver datablock to user (data)
 ec: datablock nr in receive window
 ac: data := datablock nr;
     nr := nr+1;
     rw := rw-1

Send acknowledgement
 ec: true       // also resends
 ac: Send (ACK, s, c, nr mod N, rw)

Reception of (D, c, s, data, cj)
 ac: tmp := (cj-nr) mod N;
     if 0 ≤ tmp < rw then
         buffer data as datablock nr+tmp;
     // else tmp≥rw; do nothing
```

FIGURE 37.1 Events of sliding window protocol.

The acknowledgment (ack) messages of the protocol have the form (ACK, *sid*, *rid*, *cj*, *w*), where *sid* and *rid* are as defined previously, *cj* is a cyclic sequence number, and *w* is a window size. When the message is sent, *cj* and *w* are set to the values of *nr* mod *N* and *rw*, respectively. Thus, the message indicates the data block next expected in sequence. Because it acknowledges all earlier data blocks, it is referred to as a *cumulative ack*.

The events of the producer and consumer entities are shown in Fig. 37.1. There are two types of events. A *nonreceive* event has an enabling condition, denoted *ec*, and an action, denoted *ac*; the action can be executed whenever the event is enabled. A receive event for a message has only an action; it is executed whenever the message is received.

37.4 Connection-Management Protocol

This section describes a connection-management protocol. Traditional transport protocols, including TCP and TP4, identify successive incarnations by increasing, though not necessarily successively, incarnation numbers from some modulo-*N* space. Every entity uses a counter or a real-time clock to generate incarnation numbers for local incarnations.

Another feature of traditional transport protocols is that an entity stores a remote incarnation's number only while it is connected to the remote incarnation. This necessitates a 3-way handshake for connection establishment. A client that wants to connect to a server sends a connection request with its incarnation number, say, *x*. When the server receives this, it responds by sending a response containing *x* and a server incarnation number, say, *y*. When the client receives the response, it becomes open to *y* and responds by sending an ack containing *x* and *y*. The server becomes open when it receives the ack. The server could not become open when it received the connection request containing only *x* because it may have been a duplicate from a previous, now terminated, connection.

A 2-way handshake suffices for connection closing. An open entity sends a disconnect request that is acknowledged by the other entity. A 2-way handshake also suffices for connection rejection. It is obvious that a server may have to reject a connection request of a client. What is not so obvious is that a client may have to reject a "connection request" of a server. Specifically, if a server receives an old connection request from a terminated incarnation of the client, the server attempts to complete the second stage of the 3-way handshake. In this case, the client has to reject the server.

The unreliable channels imply that a k-way handshake has the following structure: In every stage except the last, a message is sent repeatedly until the message of the next stage is received. The message of the last stage is sent only in response, otherwise the handshake would never terminate.

It is convenient to think of the protocol as a distributed system that is driven by user requests. Each user request causes the associated entity to initiate a 2- or 3-way handshake with the other entity. At each stage of the handshake, one entity learns something about the other entity and may issue an appropriate indication to its local user. At the end of the handshake, the protocol has served the user request. The protocol's behavior can be complex because two handshakes can be executing concurrently, with one of them conveying information that is relevant to the other.

There is an intricate relationship between the modulo-N space of the incarnation numbers and the handshaking algorithms, much more so than in the case of data transfer, since the latter assumes correctly initialized users. To achieve correct interpretation of received cyclic incarnation numbers, it is necessary to have bounds on message lifetime, incarnation lifetime, wait duration, and recovery duration. Under the reasonable assumption that the incarnation lifetime dominates the wait and recovery durations, it is sufficient and necessary to have

$$N \geq \frac{4L + I}{\alpha}$$

where L is the maximum message lifetime, I is the maximum incarnation lifetime, and α is the minimum time between successive incarnation creations at an entity. Most references in the literature incorrectly assume that $N \geq 2L/\alpha$ is sufficient.

This bound may not be satisfiable for exceedingly long-lived incarnations, say, of the order of days. In that case, if we assume that the probability of two successive connections having identical modulo-N client and server incarnation numbers is negligible (it is approximately $1/N^2$ under reasonable assumptions of incarnation lifetimes), then the following bound, which does not depend on I, suffices:

$$N \geq \frac{4L}{\alpha}$$

We now give a specification of the connection-management protocol. A client entity maintains the following variables for each server s:

- *status*[s]: {CLOSED, OPENING, OPEN, CLOSING}; initially CLOSED. This is the status of the client's relationship with server s: CLOSED if client has no incarnation involved with s. OPENING means client has an incarnation requesting a connection with s. OPEN means client has an incarnation open to s. CLOSING means client has an incarnation closing a connection with s.
- *lin*[s]: {NIL, 0, 1,…}; initially NIL. This is the local incarnation number: NIL if *status*[s] equals CLOSED; otherwise identifies client incarnation involved with server s.
- *din*[s]: {NIL, 0, 1,…}; initially NIL. This is the distant incarnation number: NIL if *status*[s] equals CLOSED or OPENING; otherwise identifies the incarnation of server s with which the client incarnation is involved.

A server entity maintains the following state variables for each client c:

- *status*[c]: {CLOSED, OPENING, OPEN}; initially CLOSED. This is status of server's relationship with client c: CLOSED iff server has no incarnation involved with c. OPENING means server has an incarnation accepting a connection request from c. OPEN means server has an incarnation open to c.
- *lin*[c]: {NIL, 0, 1,…}; initially NIL. This is local incarnation number: NIL if *status*[c] equals CLOSED; otherwise identifies server incarnation involved with client c.
- *din*[c]: {NIL, 0, 1,…}; initially NIL. This is the distant incarnation number. NIL if *status*[c] equals CLOSED; otherwise identifies the incarnation of client c with which the server incarnation is involved.

The messages of the protocol have the form (M, *sid, rid, sin, rin*), where M is the type of the message, *sid* is the sender's id, *rid* is the intended receiver's id, *sin* is the sender's incarnation number, and *rin* is the intended receiver's incarnation number. In some messages, *sin* or *rin* may be absent.

Each message is either a *primary* message or a *secondary* message. A primary message is sent repeatedly until a response is received or the maximum wait duration has elapsed. A secondary message is sent only in response to the reception of a primary message. Note that the response to a primary message may be another primary message, as in a 3-way handshake.

The messages sent by clients are as follows:

- (CR, *sid, rid, sin*): Connection request; sent when opening; primary message.
- (CRRACK, *sid, rid, sin, rin*): Acknowledgement to connection request reply (CRR); secondary message.
- (DR, *sid, rid, sin, rin*): Disconnect request; sent when closing; primary message.
- (REJ, *sid, rid, rin*): Reject response to a connection request reply that is received when closed. The *sin* of the received CRR is used as the value of *rin*; secondary message.

The messages sent by servers are as follows:

- (CRR, *sid, rid, sin, rin*): Reply to connection request in 3-way handshake; sent when opening; primary message.
- (DRACK, *sid, rid, sin, rin*): Response to disconnect request; secondary message.
- (REJ, *sid, rid, rin*): Reject response to a CR received when closed. The *sin* of the received message is used as the value of *rin*, secondary message.

The events of the client and server entities are shown in Figs. 37.2 and 37.3, assuming unbounded incarnation numbers.

37.5 Transport Protocols

A transport protocol between a client entity *c* and a server entity *s* consists of a connection-management protocol augmented with two data-transfer protocols, one for data transfer from *c* to *s* and another for data transfer from *s* to *c*. At each entity, the data-transfer protocol is initialized each time the entity becomes open and its events are executed only while open. The data-transfer messages are augmented by incarnation number fields, which are used by receiving entities to filter out data-transfer messages of old connections.

We illustrate with the protocols of the previous sections. Start with the connection-management protocol of Section 37.4 between *c* and *s*. Add two sliding window protocols, one from *c* to *s* and one from *s* to *c*, as follows:

- At each entity, introduce variables *ns, na, sw,* and the send window buffers for the outgoing data transfer and *nr, rw*, and the receive window buffers for the incoming data transfer. These data-transfer variables are initialized whenever the entity becomes open. Whenever the entity becomes closed, these variables are deallocated.
- Modify the client as follows. Add $status[s]$ = OPEN to the enabling condition of every data-transfer event. Add *sin* and *rin* fields to the sliding window protocol messages. When a data-transfer message is sent, *sin* is set to the local incarnation number $lin[s]$ and *rin* is set to the remote incarnation number $din[s]$. When a data-transfer message is received, first test for $status[s]$ = OPEN, $sin = din[s]$ and $rin = lin[s]$. If the test fails, ignore the message, otherwise, process the message as in the sliding window protocol specification.
- Modify the server similarly.

There are various ways to extend and integrate the data-transfer protocols. The messages of the two data-transfer protocols can be combined. For example, the data messages sent by an entity can have

```
Client entity c: events concerning server s

Connect Request(s)
 ec: status[s] := CLOSED
 ac: status[s] := OPENING ; lin[s] := new incarnation number

Disconnect Request(s)
 ec: status[s] = OPEN
 ac: status[s] := CLOSING

Abort(s)
 ec: status[s] ≠ CLOSED & "response timeout"
 ac: status[s] := CLOSED ; lin[s] := NIL ; din[s] := NIL

SendCR(s)
 ec: status[s] = OPENING
 ac: Send ( CR, c, s, lin[s] )

SendDR(s)
 ec: status[s] = CLOSING
 ac: Send ( DR, c, s, lin[s], din[s] )

Receive (CRR, s, c, sin, rin )
 ac: if status[s] = OPENING & rin = lin[s] then
       status[s] := OPEN ; din[s] := sin ;
       Send (CRRACK, c, s, lin[s], din[s])

     else if status[s] = OPEN & rin = lin[s] & sin = din[s] then
       // duplicate CRR
       Send (CRRACK, c, s, lin[s], din[s])

     else if status[s] = OPEN & rin = lin[s] & sin>din[s] then
       // server crashed, recovered, responding to old CR
       Send ( REJ, c, s, sin ) ; status[s] := CLOSED ;
       din[s] := NIL ; lin[s] := NIL

     else if status[s] is CLOSED or CLOSING then Send ( REJ, c, s, sin )

Receive ( REJ, s, c, rin )
 ac: if status[s] is OPENING or CLOSING & rin = lin[s] then
       status[s] := CLOSED ; din[s] :=NIL ; lin[s] :=NIL
     //else status[s] is OPEN or CLOSED; do nothing

Receive (  DRACK, s, c, sin, rin )
 ac: if status[s] = CLOSING & rin = lin[s] & sin = din[s] then
       status[s] := CLOSED ; din[s] := NIL; lin[s] := NIL
     // else status[s] is OPENING or OPEN or CLOSED; do nothing
```

FIGURE 37.2 Client events of connection management protocol.

additional fields to *piggy-back* acknowledgement information for incoming data, that is, fields for *nr* and *rw*. This is done in TCP and ISO TP4.

The preceding protocol uses cumulative acknowledgments. We can also use *reject* messages to indicate a gap in the received data. Rejects allow the data source to retransmit missing data sooner than cumulative acks. The protocol can also use *selective* acknowledgements to indicate correctly received out-of-sequence data. This allows the data source to retransmit only what is needed, rather than everything outstanding. Selective acks and rejects are not used in TCP and ISO TP4, although there are studies indicating that they can improve performance significantly.

```
Server entity s: events concerning client c

Abort(c)
 ec: status[c] ≠ CLOSED & "response timeout"
 ac: status[c] := CLOSED ; lin[c] := NIL ; din[c] := NIL

Send CRR(c)
 ec: status[s] = OPENING
 ac: Send ( CRR, s, c, lin[c], din[c] )

Receive ( CR, c, s, sin )
 ac: if status[c] = CLOSED & "rejecting connections" then
        Send ( REJ, s, c, sin ) ;
     else if status[c] = CLOSED & "accepting connections" then
        lin[c] := new incarnation number ;
        status[c] := OPENING ; din[c] := sin

     else if status[c] = OPENING & sin>din[c] then
        // previous din[c] value was from some old CR
        din[c] := sin

     else if status[c] = OPEN & sin>din[c] then
        // client crashed, reconnecting
        if "willing to reopen" then
           lin[c] := new incarnation number ;
           din[c] := sin ; status[c] := OPENING
        else status[c] := CLOSED ; lin[c] := NIL ; din[c] := NIL
      // else status[c] = OPEN & sin ≤ din[c]; do nothing

Receive ( CRRACK, c, s, sin,, rin, )
 ac: if status[c] = OPENING & sin = din[c] & rin = lin[c] then
        status[c] := OPEN
      // else status[c] is OPEN or CLOSED; do nothing

Receive ( DR, c, s, sin, rin )
 ac: if status[c] = OPEN & sin = din[c] & rin = lin[c] then
        Send ( DRACK, s, c, lin[c], din[c] );
        status[c] := CLOSED ; lin[c] := NIL ; din[c] := NIL

    else if status[c] = CLOSED then Send ( DRACK, s, c, rin, sin ) ;
    // else status[c] = OPENING; do nothing

Receive ( REJ, c, s, rin )
 ac: if status[c] = OPENING & rin = lin[c] then
        status[c] := CLOSED ; lin[c] := NIL ; din[c] := NIL
    // else status[c] is OPEN or CLOSED; do nothing
```

FIGURE 37.3 Server events of connection management protocol.

The preceding protocol uses fixed-size data blocks. An alternative is to send variable-sized data blocks. This can be done by augmenting the data messages with a field indicating the size of the data block. TCP does this with an octet, or byte, as the basic unit. A similar modification would be needed for selective acks.

The connection-management protocol can be extended and integrated in several ways.

The preceding protocol allows either user to close a connection at any point, without waiting for data transfer to be completed. An alternative is so-called *graceful closing*, where a user can close only its outgoing data transfer. The user must continue to handle incoming data until the remote user issues a close also. TCP has graceful closing, whereas ISO TP4 does not. It is a simple matter to add graceful closing to a protocol that does not have it.

It is possible to merge connection establishment, data transfer, and connection termination. The connection request can contain data, which would be delivered after the server becomes open. The connection request can also indicate that after the data is delivered, the connection is to be closed. TCP allows this.

TCP uses a single 32-b cyclic sequence number space to identify both incarnations and data blocks. When an incarnation is created at an entity, an **initial sequence number** is chosen and assigned to the incarnation. Successive new messages sent by the incarnation, whether of connection management or data transfer, occupy increasing sequence numbers starting from this initial sequence number.

TCP messages integrate both data transfer and connection management. Every TCP message has fields indicating the sequence number of the message, the next sequence number expected, the data segment (if any), the segment length, and receive window size. A connection-management message that requires an acknowledgement is considered to use up a sequence number, say, n. This means that the next new connection-management message sent by the entity, whether or not it requires an acknowledgement, would have the sequence number $n + 1$. The remote entity can acknowledge a connection-management message by sending a message of any type with its next expected sequence number field equal to $n + 1$.

The TCP messages SYN, SYN-ACK, ACK, FIN, FIN-ACK, and RST correspond, respectively, to the messages CR, CRR, CRRACK, DR, DRACK, and REJ of our protocol.

TCP provides **balanced opening**, a service that is outside the client–server paradigm. Here, if two entities request connections to each other at the same time, a *single* connection is established. In the client–server architecture, each entity would be considered to have a client-half and a server-half, and two connections would be established. In fact, TCP's algorithm for balanced opening is flawed: it may result in valid connection requests being rejected and invalid connection requests leading to connections. Fortunately, no application seems to use TCP's balanced-opening service. ISO TP4 does not have balanced opening.

37.6 Conclusions

This chapter has described the service expected of a transport layer and a distributed algorithm, or protocol, that achieves this service. The protocol faithfully illustrates the inner workings of TCP and ISO TP4. We conclude by outlining two currently active research areas in transport protocols.

One area is that of minimum-latency transport protocols. The delay in connection establishment incurred by the 3-way handshake is unacceptable for many transaction-oriented applications such as remote procedure calls (RPCs). Note that although transaction data can be sent with a connection request, the server cannot process the transaction until it confirms that this is a new request. This has motivated the development of transport protocols where the server can determine the *newness* of a connection request as soon as it is received, thereby achieving connection establishment with a 2-way handshake.

To achieve this, the server has to retain information about clients even when it is not connected to them. Consider a 3-way handshake between client incarnation x and server incarnation y. If the server had remembered the incarnation number, say, z, that the client had previously used when it connected to the server, then the server could determine that the connection request with x was new (because x would be greater than z). In that case, the server could have become open at once, resulting in a 2-way handshake connection establishment. A server cannot be expected to indefinitely remember the last incarnation number of every client to which it was connected, due to the enormous number of clients in a typical internetwork. However, a caching scheme is feasible, and several have been proposed, culminating recently in a proposed modification to TCP.

An alternative to caching is to use timer-based mechanisms. Here also, a server is required to maintain information on each client it has served for a duration comparable to that in cache-based mechanisms (the major component in both is the message lifetime). In most timer-based protocols, if a client's entry is removed before the specified duration, for example, due to a crash or memory limitation, then the server can incorrectly accept old connection requests of that client. Simple connection management protocol (SCMP) is an exception: assuming synchronized clocks maintains correctness, but it may reject

new connections for a period of time depending on clock skews and other parameters. In any case, timer-based approaches do not have a backup 3-way handshake.

Another area of current research concerns the performance of transport protocols. As mentioned earlier, transport protocol design has been historically driven by the need to operate correctly in spite of unreliable network service and failure-prone networks and hosts. This has resulted in the premise that transport protocols should have minimal knowledge of the network state. For example, message roundtrip times are the only knowledge that TCP has of the current network state.

In recent years, Internet traffic has grown both in quantity and variety, resulting in increasing utilization, delays, and often congestion. This has prompted the development of retransmission and windowing policies to reduce network congestion, mechanisms to reduce processing time for transport protocol messages, etc. However, this performance work is still governed by the premise, motivated by fault tolerance, that a transport protocol should have minimal knowledge of the network state.

It is not clear whether this premise would be valid for the high-speed networks of the near and far future. Current implementations of TCP use a very conservative flow control scheme: the source reduces its send window by a multiplicative factor whenever roundtrip times indicate network congestion, and increases the send window by an additive factor in the absence of such indication. Although this is very robust, it tends to underutilize the network resources. It should be noted that most protocol proposals for high-speed networks make use of some form of resource reservation to offer varying quality-of-service guarantees. It is not clear how this can be incorporated within the current TCP framework, or even whether it should be.

Defining Terms

3-way handshake: A handshake between two entities involving three messages, say $m1$, $m2$, $m3$. The initiating entity repeatedly sends $m1$. The other entity responds by repeatedly sending $m2$. The initiating entity responds by sending a single $m3$ for each $m2$ received.

Balanced-opening: Where two entities requesting connections to each other at the same time establish a single connection.

Client–server architecture: A paradigm where the users of the transport layer consist of clients and servers, connections are always between clients and servers, and connections are always initiated by clients.

Connection: An association between two open incarnations.

Connection management: That portion of the transport service and protocol responsible for the establishment and termination of connections between users.

Cyclic sequence numbers: The modulo-N sequence numbers used for identifying incarnations and data blocks.

Data transfer: That portion of the transport service and protocol responsible for the reliable exchange of data between connected users.

Incarnation: An instance of an entity created for a connection attempt.

Incarnation numbers: Sequence numbers for identifying incarnations.

Initial sequence numbers: The incarnation numbers of TCP.

Maximum message lifetime: The maximum duration for which a message can reside in the network.

Receive window: The data blocks that can be buffered at the consuming entity.

Send window: The data blocks that can be outstanding at the producing entity.

Sliding-window mechanism: A method for sending data reliably over unreliable channels, using cyclic sequence numbers to identify data, a send window at the data producer, and a receive window at the data consumer.

Transport entities: The processes within the transport layer.

Transport protocols: The distributed algorithm executed by the transport entities communicating over the network service.

Transport service: The service expected of the transport layer.

Further Information

The specification of TCP can be found in:

Postel, J. 1981. *Transmission Control Protocol: DARPA Internet Program Protocol Specification*. Request for Comment RFC-793, STD-007, Network Information Center, Sept.

The specification of ISO TP4 can be found in:

International Standards Organization 1989. *Information Processing System–Open Systems Interconnection–Transport Protocol Specification Addendum 2: Class four operation over connectionless network service*. International Standard ISO 8073/DAD 2, ISO, April.

The full specification and analysis of the data-transfer protocol in Section 37.3, including selective acks and rejects, can be found in:

Shankar, A.U. 1989. Verified data transfer protocols with variable flow control. *ACM Trans. Comput. Syst.,* 7(3):281–316.

The full specification and analysis of the connection management protocol in Section 37.4, as well as caching-based protocols and references to timer-based protocols, can be found in:

Shankar, A.U. and Lee, D. 1995. Minimum latency transport protocols with modulo-*N* incarnation numbers. *IEEE/ACM Trans. Networking,* 3(3):255–268.

Composing data transfer and connection management is explained in:

Shankar, A.U. 1991. Modular design principles for protocols with an application to the transport layer. *Proc. IEEE,* 79(12):1687–1709.

Balanced opening and the flaws of TCP are described in:

Murphy, S.L. and Shankar, A.U. 1991. Connection management for the transport layer: service specification and protocol verification. *IEEE Trans. Commun.,* 39(12):1762–1775.

38

Gigabit Networks*

Jonathan M. Smith
University of Pennsylvania

38.1 Introduction

This chapter summarizes what we have learned in the past decade of research into extremely high throughput networks. Such networks are colloquially referred to as *gigabit networks* in reference to the billion bit per second throughput regime they now operate in. The engineering challenges are in delivering this bandwidth *end-to-end* from hosts, through access networks and fast transmission systems, to receiving hosts. Hosts range through PCs, high-performance engineering workstations, to parallel supercomputing systems.

First, some history: high-throughput fiber optic networks, prototype high-speed packet switches, and high-performance workstations were all available in the late 1980s. A major engineering challenge was integrating these elements into a computer networking system capable of high application-to-application throughput. As a result of a proposal from D. Farber and R. Kahn (the Kahn/Farber Initiative) the U.S. Government [National Science Foundation (NSF) and DARPA] funded sets of collaborators in five gigabit testbeds. [Computer Staff, 1990]. These testbeds were responsible for investigating different issues, such as applications, metropolitan-area networks (MANs) vs. local-area networks (LANs), and technologies such as high performance parallel interface (HIPPI), asynchronous transfer mode (ATM), and synchronous optical network (SONET) [Partridge, 1993]. The Aurora gigabit testbed linked University of Pennsylvania (Penn), Bellcore, IBM, and MIT, with gigabit transmission facilities provided by collaborators Bell Atlantic, MCI, and Nynex (both Bell Atlantic and Nynex are part of Verizon), and was charged with exploring technologies for gigabit networking, whereas the other four testbeds were applications-focused and, hence, used off-the-shelf technologies. The results of Aurora work underpin most of today's high-speed network infrastructures.

Clark et al. [1993] set out the research goals and plans for the testbed and outlined major research directions addressed by the testbed. Aurora uniquely addressed the issues of providing switched infrastructure and high end-to-end throughput between workstation class machines. In contrast to supercomputers, these machines were, and are, the basis for most computing today.

Figures 38.1 and 38.2 provide illustrations of the Aurora geography and logical topologies, respectively. Sections 38.2–38.4 describe network host interface architectures that can operate in gigabit ranges,

*Support for research reported by University of Pennsylvania came from Bellcore (which is now called Telcordia Technologies) (through Project Dawn), IBM, Hewlett-Packard, National Science Foundation and Defense Advanced Research Projects Agency (DARPA) through Corporation for National Research Initiavtive (CNRI) under cooperative agreement NCR-89-19038, and the National Science Foundation under CDA-92-14924.

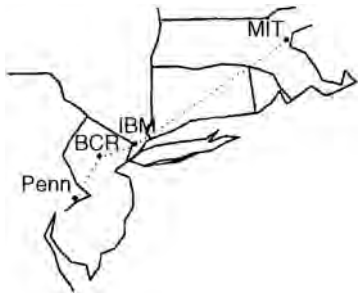


FIGURE 38.1 Aurora geography.

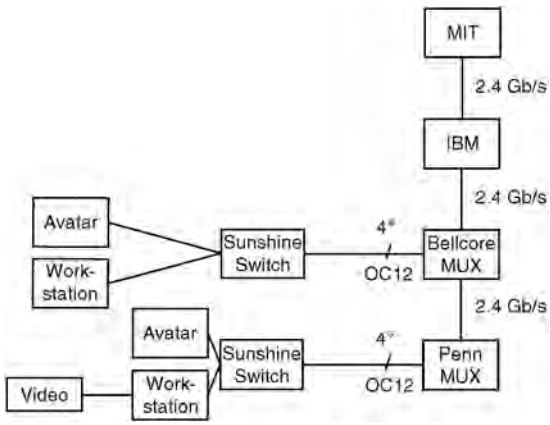


FIGURE 38.2 Partial Aurora logical topology.

multimedia aspects of gigabit networks, and the World Wide Web as an applications programming interface for gigabit (and beyond) networks. Section 38.5 summarizes our state of knowledge. Table 38.1 provides a compact summary of some major Aurora milestones.

38.2 Host Interfacing

Asynchronous transfer mode (ATM) networks were an attempt to use the same data units end-to-end, from host to host through the network. These units were small packets called *cells*, which were fixed in size (48 byte payloads and 5 byte headers). Aurora work showed that efficient, low-cost host/computer interfaces to ATM networks could be built and incorporated into a hardware/software architecture for workstation-class machines. This was believed to be problematic due to the nature of small, fixed-size ATM cells and their mismatch with workstation memory architectures. Penn designed [Traw and Smith, 1991] and implemented an ATM host interface for the IBM RISC System/6000 workstation, which connects to the machine's micro channel bus. It translated variable-size application data units into streams of fixed-size ATM cells using dedicated segmentation-and-reassembly logic. The novel application of a content-addressable memory, hardware-implemented linked-list manager, and the reassembly pipeline structure allowed use of a low clock speed and, hence, low-cost technologies. The cellification and decellification logic have measured performance which could support data rates of 600–700 Mb/s [Traw, 1995].

What this research illustrated was that hosts could achieve substantial throughputs by appropriate choice of protocol parameters (for example, by using large TCP/IP window sizes) over high performance wide-area networks. The ATM adapters carried out the job of converting between the variable-sized packet

TABLE 38.1 Aurora Gigabit Testbed: Selected Milestones

Date	Milestone
5/6/93	2.4 Gb/s OC-48 SONET backbone operational Penn <=> Bellcore
5/7/93	End-to-end data between workstations at Penn and Bellcore, interoperating Penn and Bellcore ATM host interfaces
5/19/93	Sunshine switches ATM cells between IBM RS/6000 at Penn and IBM RS/6000 at Bellcore
6/7/93	Penn and Bellcore ATM interfaces interoperate through Sunshine
6/8/93	End-to-end video over ATM from Penn workstation w/Penn video card to Bellcore workstation display
10/26/93	2nd Sunshine operational, at Penn
11/12/93	Full-motion A/V teleconference over packet transfer mode (PTM)/SONET, Penn <=> IBM
12/31/93	25 Mb/s TCP/IP over Aurora switched loopback
2/25/94	Cheap Video ATM appliance running over Aurora
3/15/94	Telementoring interactive distance learning over Aurora Penn <=> Bellcore using Cheap Video NTSC/ATM
3/30/94	70 Mb/s TCP/IP over Aurora between RS/6000s
4/17/94	MNFS/AIX solving differential heat equations over Aurora
4/21/94	Avatar, w/audio VC operational Penn <=> Bellcore, and IBM PVS IBM <=> over PlaNET, and VuNet Penn <=> MIT
5/6/94	Avatar in operation Penn <=> MIT
12/31/94	Link to IBM and MIT taken out of operation
7/5/95	HP PA-RISC/Afterburner ATM Link Adapter achieves 144 Mb/s TCP/IP
8/22/95	ATM Link Adapter achieves 215 + Mb/s TCP/IP

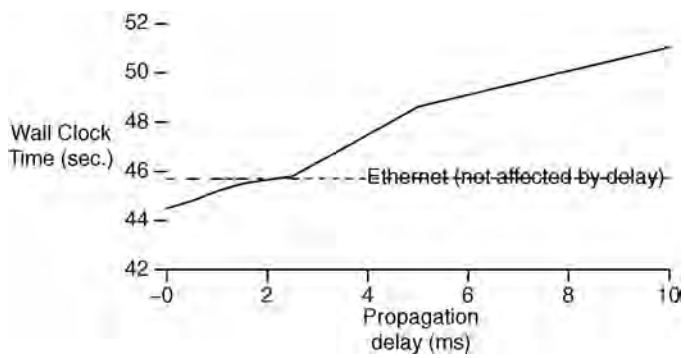


FIGURE 38.3 Performance of distributed heat equation solution, DSM/ATM.

abstraction used by the hosts and applications and the cell abstraction used by the network infrastructure. Today's deployed infrastructures use 100 megabit or gigabit ethernet [Seifert, 1999] local area networks as interconnects rather than ATM switches, and, thus, the size conversions that take place are between the application's preferred data unit size and the approximately 1500 byte MTU resulting from limitations on the size of ethernet frames. Such conversions, as they are on larger data units, are less frequent than with ATM networks and so can be performed by host software.

A continuing concern with advanced applications such as medical imaging, e-commerce and commercial teleconferencing is privacy. Privacy transformations have traditionally been rather slow due to the *bit-complexity* of the substitution—and confusion—introducing operations. Augmentation of network host interfaces with cryptographic hardware to achieve high performance [Smith, Traw, and Farber, 1992] has been designed; this work was based on earlier observations by Broscius and Smith [1991] describing the use of parallelism to achieve high performance in an implementation of the National Bureau of Standards (NBS)

data encryption standard. Smith, Traw, and Farber [1992] describe the history and motivation for the architecture and explain how to insert a high-performance cryptographic chip (for example, the VLSI Technologies VM007 data encryption standard (DES) chip, which operates at 192 Mb/s) into the ATM host interface architecture. The resulting system is able to operate at full network speed while providing a per-cell (agile) per virtual channel identifier (VCI) rekeying; both the performance and the operation are transparent to the host computer while providing much greater key control than possible with link encryption approaches. A current implementation using the new Advanced Encryption Standard (AES) based on the Rijndael cipher may be able to operate up to 10 Gb/s. The difficulty of getting data to and from the encrypting hardware through a bus remains—one attractive solution (which may or may not be economically feasible) is the use of a *cryptographic coprocessor* attached directly to the CPU.

A key principle in interface design is to leave all but the most bit-intensive tasks to software. This is one of the virtues of the ethernet design and is among the many reasons for its longevity. A number of researchers in the early 1990s (in particular Van Jacobson) characterized this approach as WITLESS, for *workstation interface that's low-cost efficient and stupid*. This approach was adopted by Traw and Smith [1991] and used no microcontrollers in a simple all-hardware adapter, with careful separation of functions between hardware and software implementation. Traw and Smith [1993] report on the implementation of the ATM host interface and its support software. The architecture is presented in detail, and design decisions are evaluated. Later work [Smith and Traw, 1993] focused attention on the adapter to application path through software, and some of the key design decisions embedded in the software are examined. Of particular interest are the system performance measures where the adaptor operates with a significant application workload present.

The initial software subsystem provided an application programmer interface roughly equivalent to a raw Internet protocol (IP) socket and was able to achieve more than 90% of the hardware subsystem's performance, thus driving an optical carrier level 3 (OC-3) at its full 155 Mb/s rate. Key innovations were the reduction of data copying (through use of virtual memory (VM) support; this direction was later followed by others, including the University of Arizona team [Druschel, Peterson, and Davie, 1994] designing software for the Osiris [Davie, 1993] interface developed at Bellcore by Bruce Davie) and the partitioning of functions between hardware and software. As can be seen from Table 38.2 [Druschel et al., 1993] this reduction in data copying was necessitated by the memory bandwidth limitations of early-1990s workstations. The bottleneck on the IBM RS/6000 was initially the workstation's system bus to I/O bus interconnect [Traw and Smith, 1993], however, improvements to the I/O subsystem architecture moved the bottleneck to the physical link. For the Hewlett-Packard (HP) PA-RISC implementation [Traw, 1995] designed to demonstrate scaling of the host interface architecture to higher speeds, the bottleneck was the bus attachment (for this environment, the SGC graphics bus served as the attachment point).

The HP PA-RISC/afterburner ATM link adapter held a record for highest reported transport control protocol (TCP)/IP/ATM performance of 215 + Mb/s for almost one year. This performance was measured between two HP PA-RISC 755s, connected by a 320 Mb/s SONET-compliant null modem, using the *netperf* test program. The best performance was achieved using a 32-kilobyte socket buffer size and 256-kilobyte packets.

The lessons from this research have made their way into common use in all modern systems software. As Table 38.2 illustrates, the memory bandwidths of the second generation of RISC workstations, with

TABLE 38.2 Workstation Memory Bandwidths, Early 1990s

	Memory (Mb/s, peak)	CPU/Memory (Mb/s, sustained)		
		Copy	Read	Write
IBM RS/6000 340	2133	405(0.19)	605(0.30)	590(0.28)
Sun SS 10/30	2300	220(0.10)	350(0.15)	330(0.14)
HP 9000/720	1600	160(0.10)	450(0.28)	315(0.20)
Dec 5000/200	800	100(0.13)	100(0.13)	570(0.71)

Source: Tabulation by Druschel, P., Abbott, M.B., Pagels, M.A., and Peterson, L.L., 1993. Network subsystem design. *IEEE Network* special issue, 7(4):8–17.

the exception of the RS/6000, were poor. The designers relied on effective caches to maintain high instruction processing rates. While the architectures were effective on scientific workloads, they were far less effective when the data was moved to and from memory as they are in moving to and from the network. If anything, in the first decade of the 21st century, the ratio between CPU performance and memory throughput has gotten worse rather than better, and thus the need for reducing memory copies has increased. Pragmatically, if a modern host is achieving any approximation to a gigabit of throughput through a Gb/s ethernet adapter, it is copying the data only once.

38.3 Multimedia Services

Multimedia services for a networking infrastructure capable of gigabit throughputs must be designed with scale, endpoint heterogeneity, and application requirements as the key driving elements. For example, an integrated multimedia architecture with which applications define which data are to be bundled together for transport and select which data are unbundled from received packages can be far more receiver-driven in nature. Later work in the Internet community took this approach by "layering" different qualities of media delivery, so that receivers can select the quality of video which they choose to reproduce. This allows sources to choose the degree of resource allocation they wish to provide; receivers choose which elements of the package they wish to produce. Although potentially wasteful of bandwidth, the massive reduction in the multiplicity of customized channels allows sources to service a far greater number of endpoints and receivers to accommodate endpoint resources by reproducing what they are capable of. The scaling advantage of this approach is that much of the complexity of customization is moved closest to the point where customization is required, namely, the endpoint.

Multimedia appliances [Udani, 1992] were developed in anticipation of readily available bandwidth, an approach which is now again revived with the model of "always-on" Webcams. For example, this led to the all-hardware National Television Systems Committee (NTSC)/ATM Avatar ATM [Marcus and Traw, 1995] card, which supports NTSC video and CD-quality audio, and is the first example of an ATM appliance, with a parts cost (in 1993) of under $300. Many of these cards were fabricated. They were used for distance teaching when connecting experimental "Video Windows," for collaborative work between researchers at Penn and Bellcore, and for teleconferencing between Penn and Massachusetts Institute of Technology (which had its own extensive research effort in ubiquitous ATM hardware, the ViewStation, modeled after the Desk Area Network at Cambridge University).

Multimedia provision relies on the development of operating system abstractions which can support high-speed applications. These abstractions use the hardware and low-level operating system support of the operating system to provide an application-centered model of quality of service (QoS) requirements. Novel technical means had to be developed for bandwidth allocation implementations, e.g., support from the operating system scheduling mechanism for true end-to-end service delivery. Nahrstedt [1995] identified the software support services needed to provide QoS guarantees to advanced applications that control the characteristics of their networking system and adapt within parameterized limits. These services form a kernel, or a least common subset of services, required to support advanced applications.

A logical relationship between applications-specified QoS [Nahrstedt and Smith, 1996], as well as operating system policy and mechanism, and network-provided QoS was developed. An example challenge is the kinematic data stream directing a robotic arm, which can tolerate neither packet drops nor packet delays—unlike video or audio, which can tolerate drops but not delays. The approach used is a bidirectional translator (like a compiler/uncompiler pair for a computer language) which resides between the network service interface and the applications' service primitives. This can dynamically change QoS as application requirements or network capabilities change, allowing better use of network capacity, which can be mapped more closely to application's current needs than if a worst-case requirement is maintained. The implementation [Nahrstedt, 1995] outlined the complete requirements for such a strategy, including communication primitives and data necessary for translation between network and application. For example, an application request to zoom and refocus a camera on the busiest part of a scene will certainly require peer-to-peer communication between the application and the

camera management entity. The network may need to understand the implications for interframe compression schemes and required bandwidth allocations. The translation method renegotiates QoS as necessary.

These ideas were described in Nahrstedt and Smith [1995], which describes a mechanism to provide bidirectional negotiation of quality-of-service parameters between applications and the other elements of a workstation participant in advanced network applications. The scheme is modeled on a *broker*, a traditional mechanism for carrying on back-and-forth negotiations while filtering implementation details irrelevant to the negotiation. The QoS broker reflects both the dynamics of service demands for complex applications and the treatment of both applications and service kernels as first-class participants in the negotiation process. The QoS broker was implemented in the context of a system for robotic teleoperation. The broker was implemented and evaluated as part of a complete end-to-end architecture Nahrstedt [1995].

Gigabit multimedia is much desired by the applications community. Applications include high-quality (e.g., HDTV) movies, interactive games, and ultra-high-quality conferencing.

38.4 The World-Wide Web

One of the major questions addressed in the gigabit testbeds was what application programming interface would be used. The connection-oriented model of telephony and TCP/IP was looking limited, especially in the face of long latencies and clashes between the connection-oriented model and the object-access paradigm (memory and files) used in computer systems. Farber [1995] had proposed distributed shared memory (DSM) as a technology solution for integrating computation and communications more closely. At about the same time that the gigabit testbed research was underway, a new paradigm for network resource access was being developed, the World Wide Web. This model relied on a hypertext model—one where documents are named and may link to other named documents. The naming paradigm, such as http://my. location.edu/~me/myfile, incorporates both the object name and its location in its name, allowing non-local objects to be referenced and fetched. In this sense, then, the Web can be seen as a very loosely coupled distributed shared memory. Thus, the Web satisfies some interesting properties when viewed as a memory. First, it can be searched—this is exemplified by various Web-crawling programs, such as http://www.google.com, which make the Web into the equivalent of a large-scale content addressable memory. Second, the objects referenced are often visual. As Shaffer's thesis [Shaffer 1996] showed, the benefits of high bandwidth accrue directly to applications requiring high throughput by reducing the perceived latency to the application (such as a web browser) for remotely referencing the object. This advantage will continue to accrue as Web caches are used to reduce the effects of other latencies such as disk accesses and round-trip delays and will be amplified as protocols such as parallel HTTP and persistent HTTP are used to overcome some of the delay inherent in TCP/IP's slow-start behavior.

The Web is a superb application for end-to-end gigabit networks, as the availability of bandwidth makes more complex content and even the multimedia services of the previous section available to users. Fortunately, the workstations, LANs, and core networks have been evolving rapidly. Less fortunately, deployment of broadband to the home, where most people access the Web, has been far too slow.

38.5 Conclusions

> *There is an old network saying: Bandwidth problems can be cured with money. Latency problems are harder because the speed of light is fixed—you can't bribe God.*

—D. Clark [Patterson and Hennessy, 1996]

The Aurora gigabit testbed research had a fundamental influence on the design and development of network technology. The host interface work answered concerns about the viability of ATM segmentation and reassembly (SAR) and made the performance concerns expressed when the testbeds were begun nonissues. Host interface work from Aurora has influenced all of today's commercial products, both ATM and ethernet [Chu, 1996].

Operating systems research at Penn and later work at the University of Arizona showed how to reduce copying between host interfaces and applications through careful management of DMA, buffer pools, and process data access [Chu, 1996]. Operating systems can deliver gigabit range throughputs to applications with appropriate restructuring and rethinking of copying and protection boundary crossing.

The still unanswered questions revolve around the discovery and evaluation of mechanisms that deliver a practical reduction in latency to applications. These include better cache management algorithms for shared objects as well as techniques for look-ahead referencing, or prefetching. Another important area is the reduction in operating system induced costs in network latency. The multimedia work in the gigabit networking research community [Steinmetz and Nahrstedt, 1995] has had impact on the operating systems and robotics communities. It also pointed out some issues to be avoided in adapter design, such as head-of-line blocking observed in serial DMA of large media objects.

The Web remains the most exciting frontier for high-speed networking, as user demand for bandwidth appears insatiable. The increasing availability of home broadband, in the form of digital subscriber line and ever-faster cable modem bandwidths, and eventually fiber-to-the-home, will make gigabit networking for all a reality, a path illuminated by the early work in the Aurora gigabit testbed.

References

Alexander, D.S., Traw, C.B.S., and Smith, J.M. 1994. Embedding high speed ATM in UNIX IP. In *USENIX High-Speed Networking Symposium,* Oakland, CA, Aug., 1–3, 119–121.

Broscius, A.G. and Smith, J.M. 1991. Exploiting parallelism in hardware implementation of the DES. In *Proceedings, CRYPTO 1991 Conference,* Ed. J. Feigenbaum, Santa Barbara, CA, Aug., 367–376.

Chu, J. 1996. Zero-copy TCP in Solaris. In *Proceedings USENIX 1996 Annual Technical Conference,* San Diego, CA, Jan. 22–26, 253–264.

Clark, D.D. Davie, B.S., Farber, D.J., Gopal, I.S., Kadaba, B.K., Sincoskie, W.D., Smith, J.M., and Tennenhouse, D.L. 1993. The AURORA gigabit testbed. *Computer Networks and ISDN Systems,* 25(6):599–621.

Computer Staff, 1990. Gigabit network testbeds. *IEEE Computer,* 23(9):77–80.

Davie, B.S. 1993. The architecture and implementation of a high speed host interface. *IEEE J. Selec. Areas Common.,* Special issue on high-speed computer/network interfaces, 11(2):228–239.

Druschel, P., Abbott, M.B., Pagels, M.A., and Peterson, L.L. 1993. Network subsystem design. *IEEE Network,* Special issue: End-system support for high-speed networks (breaking through the network I/O bottleneck), 7(4):8–17.

Druschel, P., Peterson, L.L., and Davie, B.S. 1994. Experiences with a high-speed network adapter: A software perspective. In *Proceedings 1994 SIGCOMM Conference,* London, UK, Aug. 31–Sept. 2, 2–13.

Farber, D.J. 1995. The Convergence of Computers and Communications, Part 2, ACM SIGCOMM Award Lecture, Aug. 30.

Marcus, W.S. 1996. An experimental device for multimedia experimentation. *IEEE/ACM Trans. Networking.*

Marcus, W.S. and Traw, C.B.S. 1995. AVATAR: ATM video/audio transmit and receive. Distributed Systems Laboratory, University of Pennsylvania Tech. Rep., March.

Minnich, R.G. 1993. Mether-*NFS*: a modified NFS which supports virtual shared memory. In *Experiences with Distributed and Multiprocessor Systems (SEDMS IV),* USENIX Assoc., San Diego, CA, 89–107.

Nahrstedt, K. 1995. An Architecture for Provision of End-to-End QoS Guarantees, Ph.D. thesis, Tech. Rep. CIS Dept., University of Pennsylvania.

Nahrstedt, K. and Smith, J.M. 1995. The QoS broker. *IEEE Multimedia Mag.,* 2(l):53–67.

Nahrstedt, K. and Smith, J.M. 1996. Design, implementation and experiences of the OMEGA end-point architecture. *IEEE J. Selec. Areas Commun.,* Special issue on multimedia systems.

Partridge, C. 1993. *Gigabit Networking,* Addison-Wesley, Reading, MA.

Patterson, D.A. and Hennessy, J.L. 1996. *Computer Architecture: A Quantitative Approach,* 2nd ed., Morgan Kaufmann, San Francisco, CA.

Seifert, R. 1999. *Gigabit Ethernet,* Addison-Wesley, Reading, MA.

Shaffer, J.H. 1996. The Effects of High-Bandwidth Networks on Wide-Area Distributed Systems, Ph.D. thesis, CIS Dept., University of Pennsylvania.

Smith, J.M. and Traw, C.B.S. 1993. Giving applications access of Gb/s networking. *IEEE Network,* Special issue: End-system support for high-speed networks (breaking through the network I/O bottleneck), 7(4):44–52.

Smith, J.M., Traw, C.B.S., and Farber, D.J. 1992. Cryptographic support for a gigabit network. In *Proceedings INET'92,* Kobe, Japan, Inaugural Conference of the Internet Society, June 15–18, 229–237.

Steinmetz, R. and Nahrstedt, K. 1995. *Multimedia: Computing, Communications, and Applications,* Prentice-Hall, Englewood Cliffs, NJ.

Traw, C.B.S. 1995. Applying Architectural Parallelism in High Performance Network Subsystems, Ph.D. thesis, CIS Dept., University of Pennsylvania.

Traw, C.B.S. and Smith, J.M. 1991. A high-performance host interface for ATM networks. In *Proceedings SIGCOMM 1991,* Zurich, Switzerland, Sept. 4–6, 317–325.

Traw, C.B.S. and Smith, J.M. 1993. Hardware/software organization of a high-performance ATM host interface. *IEEE J. Selec. Areas Commun.,* Special issue on high speed computer/network interfaces 11(2):240–253.

Traw, C.B.S. and Smith, J.M. 1995. Striping within the network subsystem. *IEEE Network* (July/Aug.): 22–32.

Udani, S.K. 1992. Architectural Considerations in the Design of Video Capture Hardware, M.S.E thesis (EE), School of Engineering and Applied Sciences, University of Pennsylvania, April.

39

Local Area Networks

Thomas G. Robertazzi
SUNY at Stony Brook

39.1 Introduction

Local area networks (LANs) are computer networks operating over a small area such as a single department in a company. Typically, a single LAN will support a small number of users (say, less than 25). Most LANs provide communications between terminals, PCs, workstations, and computers. Local area networks were first created in the late 1970s and early 1980s. Today, LANs may carry a variety of traffic, such as voice, video, and data.

In general, the elements of a computer or local area network must follow compatible rules of operation together to function effectively. These rules of operation are known as protocols. A variety of LANs, operating under different protocols, are available today. These are described below.

39.2 Local Area Networks (LANs)

There are a variety of local network architectures that have been commercially produced to date including ethernets, token rings, wireless LANs, and others. A number of these have been standardized in the IEEE 802 series standards.

Ethernet

Ethernet is a very popular LAN technology that has several versions. The original IEEE 802.3 standard deals with a network architecture and protocol first constructed at Xerox in the 1970s and termed "ethernet." All stations in the original ethernet protocol are connected, through interfaces, to a coaxial cable that is run through the ceiling or floor near each user's computer equipment.

The coaxial cable essentially acts as a private radio channel for the users. An interesting protocol called Carrier Sense Multiple Access with Collision Detection (CSMA/CD) is used in such a network. Each station constantly monitors the cable and can detect when it is idle (no user transmitting), when one user is transmitting (successfully), or when more than one user is simultaneously transmitting (resulting in an unsuccessful collision on the channel). The cable basically acts as a broadcast bus. Any station can transmit on the cable if the station detects it to be idle. Once a station transmits, other stations will not interrupt the transmission. As there is no central control in the network, occasionally two or more stations may attempt to transmit at about the same time. The transmissions will overlap and be unintelligible (collision).

The transmitting stations will detect such a situation and each will retransmit at a randomly chosen later time. Ethernet operates at 10 Mbps.

Computer capabilities have increased significantly since the introduction of standard 802.3. To keep pace, a 100 Mbps version of Ethernet was standardized in 1995. Furthermore, gigabit ethernet became available in the late 1990s. All versions of ethernet today are generally wired together using a star type-wiring pattern with hubs at the center. Hubs are boxes that provide media sharing or switching capabilities for interconnecting local computers. Moreover, all versions of ethernet today support the use of twisted pair or fiber optics for wiring.

The 100 Mbps version of ethernet is known as fast ethernet (IEEE 802.3u). To maintain channel utilization at a similar rate as 802.3, in spite of the ten-fold speed increase, the maximum segment length is ten times smaller for 802.3u compared to 802.3. A combination of ternary signaling (i.e., three logic levels) and parallel wiring is used to achieve a fast ethernet version 100 Base-T4 signaling speed only 25% higher (at 25 MHz) than that for 802.3. The 100-Base-TX version uses twisted pair wires and handles full duplex 100 Mbps transmission with the use of 4B5B coding. The 4B5B coding provides sufficient logic transitions, in spite of long message runs of 1s or 0s, for digital receivers to lock on to for synchronization.

The gigabit (1000 Mbps) version of ethernet is known as gigabit ethernet. To maintain channel utilization at a similar value as 802.3 and 802.3u in spite of the speed increase, the minimum frame length has been increased to 512 bytes from 512 bits. Gigabit standard IEEE 802.3z allows the use of fiber optic transmission while IEEE 802.ab allows the use of twisted pair wiring. Older PCs may not be able to keep up with gigabit ethernet, and performance may actually degrade.

There is work on developing a 10 Gbps version of ethernet. This is a more difficult task than creating some of the previous variants as it is not possible to reuse certain technology from other networks, as was done in creating fast ethernet and gigabit ethernet. The forthcoming 10 Gbps ethernet standard is IEEE 802.3ae and should be finalized in 2002. Pre-standards products appeared in the market in late 2000/early 2001 and standards compliant products after the standards ratification. One version of the standard allows single mode fiber optic transport for 40+ km. Another means of 10 Gbps ethernet transport is on SONET OC-192 links.

There have been efforts to create a multimedia ethernet known as Iso-ethernet (IEEE 802.9). By using 4B5B (80% efficiency) instead of 802.3s Manchester encoding (50% efficiency), one can use 802.3 wiring with a 60% boost in speed. Specifically, an Iso-ethernet has a speed of 16.384 Mbps. This can be allocated either mostly to multimedia traffic (all isochronous mode) or split between 10 Mbps of regular ethernet packet service and 6.144 Mbps of multimedia traffic (multiservice mode).

Ethernet, in all its variety, is the most popular LAN to date. For instance, a 1999 Sage Research survey of 700 US organizations indicated that half the organizations planned to buy fast ethernet by the end of 1999.

Token Ring

Token rings are covered by the IEEE 802.5 standard. Token ring LANs were developed by IBM in the early 1980s. Logically, stations are arranged in a circle with point-to-point links between neighbors. Transmissions flow in only one direction (clockwise or counter-clockwise). A message transmitted is relayed over the point-to-point links to the receiving station and then forwarded around the rest of the ring and back to the sender to serve as an acknowledgement.

Only a station possessing a single digital codeword known as a “token” may transmit. When a station is finished transmitting, it passes the token to its downstream neighbor. Thus, there are no collisions in a token ring and utilization can approach 100% under heavy loads.

Because of the use of point-to-point links, token rings can use various transmission media such as twisted pair wire or fiber optic cables. The transmission speed of an 802.5 token ring is generally 4 or 16 Mbps. A standard for a 100 Mbps variant of token ring is 802.5t and a standard for a gigabit variant is 802.5v. Token rings are often wired in star configurations for ease of installation.

Token Bus

Token bus operation is standardized in the IEEE 802.4 standard. A token bus uses a coaxial cable along with the token concept to produce a LAN with improved throughput compared to the 802.3 protocol. That is, stations pass a token from one to another to determine which station currently has permission to transmit. Also, in a token bus (and in a token ring), response times can be deterministically bounded. This is important in factory automation where commands to machines must be received by set times. By way of comparison, response times in an ethernet-like network can only be probabilistically defined. This is one reason General Motor's Manufacturing Automation Protocol made use of the token bus. Token buses can operate at 1,5, and 10 Mbps.

Wireless LANs

Wireless LANs use a common radio channel to provide LAN connectivity without any physical wiring. Protocols for wireless LANs include the IEEE 802.11 standard. One might consider the use of a CSMA/CD protocol in this environment. However, it is not possible to implement the collision detection (CD) part of CSMA/CD in a radio environment. Therefore, a modified protocol known as CSMA/CA is used. CSMA/CA stands for Carrier Sense Multiple Access with Collision Avoidance. In this variant on CSMA, a station requests permission from a station it wishes to transmit to.

An industrial, science, and medical (ISM) band (at 2.4 GHz) is used for 802.11 transmission. The two architectures envisioned are tying stations to a local access point (i.e., base station) connected to a backbone or direct "ad hoc" station to station communication. Spread spectrum technology is used in 802.11. Transmission speed is in the low megabit range. The 802.11 standard was finalized in 1997.

A newer standard, IEEE 802.11b, allows data transfer at up to 11Mbps. It uses WEP (wired equivalency privacy) encryption for security.

There are also other wireless products such as bluetooth (802.15). Bluetooth establishes a wireless link between devices such as laptops and printers.

An alternative wireless LAN technology uses infrared light as the transmission medium. There are both direct infrared systems (range: 1-3 miles) and non-direct systems (bounced off walls or ceilings). For small areas, data rates are consistent with those of existing ethernet and token ring networks with 100 Mbps systems possible.

IEEE 802.12

As mentioned, in going from a speed of 10 Mbps to 100 Mbps in creating fast ethernet, it was necessary to reduce the maximum LAN segment size by a factor of ten in order to maintain channel utilization at a reasonable level. Some people considered this an undesirable compromise. A group of parties split off from the 802.3 standardization effort and sought to develop a completely new protocol, known as 802.12. This protocol operates at 100 Mbps. The standard makes use of hubs called repeaters. A station wishing to transmit sends a request to a hub. The repeater arbitrates among requests using a round robin algorithm. That is, each station requesting access is given some access in turn. Trees of repeaters are possible. Either twisted pair or fiber optics can be used. There is a priority structure in 802.12. This protocol can also be adapted to higher speeds.

ATM LANs and Backbones

ATM stands for asynchronous transfer mode. This is a packet switching technology utilizing relatively short, fixed length packets to provide networking services. ATM was originally seen as a way to develop the next generation wide area telephone network using packet switching, rather than the more conventional circuit switching technology. It was envisioned as a way to transport video, voice, and data in an integrated network. A short packet size was chosen to meet several requirements including minimizing real time queueing delay.

While progress on the original goal of using ATM technology in wide area telephone networks proceeded slowly because of the complexity of the challenge and the large investments involved, a number of smaller companies introduced ATM local area network and backbone products using much the same technology. Backbones are high speed networks used to interconnect LANs.

An ATM network consists of one or more switches. There are several possibilities for the internal architecture of the switch. A low cost switch may essentially be a computer bus in a box. More sophisticated switches may use switching fabrics. These are VLSI implementations of patterned networks of simple switching elements, sometimes referred to as space division switching. A great deal of effort has gone into producing cost-effective ATM switches. It should be pointed out that many of the issues that are difficult for wide area network ATM (i.e. traffic policing, billing) are more tractable in the private ATM LAN/backbone environment.

39.3 The Future

The future is likely to see an increase in data rates. This will spur the development of faster switching nodes through the use of parallel processing and VLSI implementation. Protocols will have to be simplified to increase processor throughput. Higher speed computer networks in general, and LANs in particular, will continue to proliferate throughout the world, making possible ubiquitous communications between any two points on the globe.

Defining Terms

Area networks: LAN, within single building; MAN, metropolitan sized region; WAN, national/international region.

Coaxial cable: A shielded cable that conducts electrical signals and is used in older bus type local area networks.

Fiber-optic cable: A glass fiber cable that conducts light signals. Fiber optic cables can carry data at very high speeds and are immune to electrical interference.

SONET: Synchronous optical network. A widely used fiber optic standard.

Twisted pair wire: Two thin wires that are twisted together. Wires with different numbers of twists per unit length are available that provide different amounts of rejection to electromagnetic interference.

The following are the IEEE 802 series of standards related to local area networks.

IEEE 802.3: Ethernet (CSMA/CD bus) protocol standard.
IEEE 802.3u: Fast ethernet (100 Mbps).
IEEE 802.3z: Gigabit ethernet (fiber).
IEEE 802.3ab: Gigabit ethernet (twisted pair).
IEEE 802.3ae: Ten gigabit ethernet.
IEEE 802.4: Token bus standard.
IEEE 802.5: Token ring standard.
IEEE 802.5t: Token ring (100 Mbps).
IEEE 802.5v: Token ring (gigabit).
IEEE 802.9: Iso-ethernet.
IEEE 802.11: CSMA/CA wireless LAN standard.
IEEE 802.11b: Wireless LAN (11 Mbps).
IEEE 802.12: Demand priority protocol.
IEEE 802.15: Bluetooth.

References

Comer, D., *Computer Networks and Internets,* 2nd ed., Prentice Hall, Upper Saddle River, NJ, 1999.

Crow, B.P., Widjaja, I., Kim, J.G. and Sakai, P.T., IEEE 802.11 wireless local area networks, *IEEE Commun. Mag.*, 116, 1997.

Halsall, F., *Data Communications, Computer Networks and Open Systems,* 4th ed., Addison-Wesley, Reading, MA., 1996.
Molle, M. and Watson, G., 100Base-T/IEEE802.12/packet switching, *IEEE Commun. Mag.,* 64, 1996.
Peterson, L. and Davie, B., *Computer Networks: A Systems Approach,* Morgan Kaufman, San Francisco, CA, 2000.
Ross, F.E. and Vaman, D.R., IsoEthernet: an integrated services LAN, *IEEE Commun. Mag.,* 74, 1996.
Tanenbaum, A., *Computer Networks,* 3rd ed., Prentice Hall, Upper Saddle River, NJ, 1996.
Walrand, J. and Varaiya, P., *High-Performance Communication Networks,* 2nd ed., Morgan Kaufman, San Francisco, CA, 2000.

Further Information

Tutorial articles on LANs appear in *IEEE Communications Magazine* and *IEEE Network Magazine.* Technical articles on LANs appear in *IEEE Transactions on Networking, IEEE Transactions on Communications, IEEE Journal on Selected Areas in Communications*, and the journal *Wireless Networks*.

40

Asynchronous Time Division Switching

Achille Pattavina
Politecnico di Milano

40.1 Introduction

In the last decades, separate communication networks have been deployed to support specific sets of services. For example, voice communication services (and, transparently, some low-speed data services) are supported by circuit switched networks, whereas packet switched networks have been specifically designed for low-to-medium speed data services. During the 1980s, a worldwide research effort was undertaken to show the feasibility of packet switches capable of supporting narrowband services, such as voice and low-to-medium speed communications, together with broadband services, such as high-speed data communications and those services typically based on video applications.

The challenge of the forthcoming broadband-Integrated Services Digital Network (B-ISDN), as envisioned by the International Telecommunication Union—Telecommunication Standardization Sector (ITU-TSS), is to deploy a unique transport network based on *asynchronous time-division* (ATD) multiplexing and switching that provides a B-ISDN interface flexible enough to support all of today's services (voice and data) as well as future narrowband (for data applications) and broadband (typically for video applications) services to be defined. With ATD, the bandwidth is not preallocated in the transmission and switching equipment of the broadband network so as to fully exploit by statistical multiplexing the available capacity of communication resources. ITU-TSS has defined a standard for the transfer mode based on ATD switching called the **asynchronous transfer mode** (**ATM**) [ITU, 1993].

This chapter is intended to present a brief review of the basic ATM concepts, as well as of the main architectural and performance issues of ATM switching. The basic principles of the asynchronous transfer mode are described in Section 40.2, whereas the model of an ATM switch is presented in Section 40.3 by discussing its main functions. Three main classes of ATM switching fabrics are identified therein based on the characteristics of their *interconnection network* (IN), and the basic characteristics and properties for each class are discussed in the following sections. ATM switches whose IN is blocking with minimum

depth and arbitrary depth are dealt with in Sections 40.4 and 40.5, respectively, whereas architectures with nonblocking IN are reported on in Section 40.6.

40.2 The ATM Standard

Time-division multiplexing (TDM) today is the most common technique used to transport on the same transmission medium tens or hundreds of channels, each carrying an independent digital flow of information, typically digitized voice signals. A preventive per-channel allocation of the transmission capacity, in terms of a set of slots within a periodic frame, is done with TDM that time interleaves the supported channels onto the same transmission medium. Therefore, a rigid allocation of the transmission resource is obtained, in which the available bandwidth is fully used only if all of the channels are active at any time; that is, all of the channel slots carry user information. Therefore, TDM is well suited to support communication services with a high activity rate, as in the case of voice services, since other services whose information sources are active only for a small percentage of time, typically data services, would waste transmission bandwidth.

The asynchronous time division (ATD) technique is intended to overcome this problem by enabling the sharing of transmission and switching resources by several channels without any preventive bandwidth allocation to the single users. Therefore, by means of a suitable storage capability, information from the single channels can be statistically multiplexed onto the same communication resource, thus avoiding resource wastage when the source activity is low. Asynchronous time-division multiplexing (ATDM) requires, however, each piece of information to be accompanied by the owner information, which is no longer given by the position of the information within a frame, as in the case of TDM. Switching ATDM channels requires the availability of ad-hoc packet switching fabrics designed to switch enormous amount of information when compared to the switching fabrics of current narrowband packet switched networks.

The B-ISDN envisioned by ITU-TSS is expected to support a heterogeneous set of narrowband and broadband services by sharing as much as possible the functionalities provided by a unique underlying transmission medium. ATM is the ATD-based transport mechanism to be adopted in the lower layers of the protocol architecture for this purpose [ITU, 1993]. Two distinctive features characterize an ATM network: (1) The user information is transferred through the network in small fixed-size units, called **ATM cells**, each 53 bytes long, divided into a *payload* (48 bytes) for the user information and a *header* (5 bytes) for control data. (2) It is a *connection-oriented* network, that is, cells are transferred onto previously set-up virtual links identified by a label carried in the cell header. Both *virtual connections* (VC) and *virtual paths* (VP) are defined in the B-ISDN. A logical connection between two end users consists of a series of $n + 1$ virtual channels if n switching nodes are crossed. A virtual path is a bundle of virtual channels. Since a virtual channel is labeled by means of a hierarchical key virtual path identified/virtual channel identifier (VPI/VCI), a switching fabric can operate either a full VC switching or just a VP switching.

The B-ISDN protocol architecture includes three layers that, from the bottom up, are referred to as the *physical layer, ATM layer*, and *ATM adaptation layer* (AAL) [ITU, 1991]. The task of the physical layer is to provide a transmission capability for the transfer of ATM cells. Its functionalities include some basic tasks, such as timing, cell delineation, cell header verification, etc. Since the standard interface has the minimum rate of about 150 Mb/s, even ATM switching nodes of medium size with, say, 128 input and output links must be capable of carrying loads of tens of gigabits per second. The ATM layer includes the VCI/VPI translation, the header generation and extraction, and the multiplexing/demultiplexing of cells with different VCI/VPI onto the same physical layer connection. The purpose of the ATM adaptation layer is to add different sets of functionalities to the ATM layer so as to differentiate the kinds of services provided to the higher layers. Four service classes are defined to support connection-oriented and connectionless services. These services range from circuit emulation and variable bit rate video to low-speed data services. Even if the quality-of-service parameters to be guaranteed for each class are still to be

defined, the B-ISDN network should be able to provide different traffic performance to each service class in terms of packet loss and delay figures.

40.3 Switch Model

Research in ATM switching has been developed worldwide for several years showing the feasibility of ATM switching fabrics. However, a unique taxonomy of ATM switching architectures is very hard to find since different keys used in different orders can be used to classify ATM switches. Very briefly, we can say that most of the ATM switch proposals rely on the adoption for the interconnection network, which is the switch core, of multistage arrangements of very simple switching elements (SEs), each using the **cell self-routing** concept. This technique consists of allowing each SE to switch (route) autonomously the received cell(s) only using a self-routing label preceding the cell. Even if other kinds of switching architectures can be used as well (e.g., shared memory or shared medium units), multistage INs provide the processing power required to carry the above-mentioned load foreseen for small-to-large-size ATM switches. The interconnection network, which is the ATM switch core, operates by transferring in a time window, called the *slot*, the packets, which are received aligned by the IN from the switch inlets to the requested switch outlets.

Rather than using technological features of the switching architecture, the main key to classify the switch proposals can be found in the functional relationship setup between inlets and outlets by the switch. Multistage INs can be basically classified as **blocking** or **nonblocking**. In the former case, different input/output (I/O) paths within the IN can share some interstage links; thus, the control of packet loss events requires the adoption of additional techniques, such as the packet storage capability in the SEs (**minimum-depth** INs) or deflection routing (**arbitrary-depth** INs). In the latter case, different I/O paths are available so that SEs do not need internal buffers and are much simpler to implement (a few tens of gates per SE). Nevertheless, these INs require more stages than blocking INs. A further distinctive feature of ATM switches is where the buffers holding cells are placed; three configurations are distinguished with reference to each single SE or to the whole IN, that is, **input queueing, output queueing**, and **shared queueing**.

We will refer here only to the cell switching of an ATM node by discussing the operations related to the transfer of cells from the inputs to the outputs of the switch. Thus, all of the functionalities relevant to the setup and teardown of the virtual connections through the switch are intentionally disregarded. The general model of an $N \times N$ switch is shown in Fig. 40.1. The reference switch includes *N input port controllers* (IPC), *N output port controllers* (OPC), and an interconnection network (IN), where the IN is capable of switching up to K cells to the same OPC in one slot. The IN is said to have a speedup K if $K > 1$ and no speedup if $K = 1$, since an internal bit rate higher than the external rate (or an equivalent space-division technique)

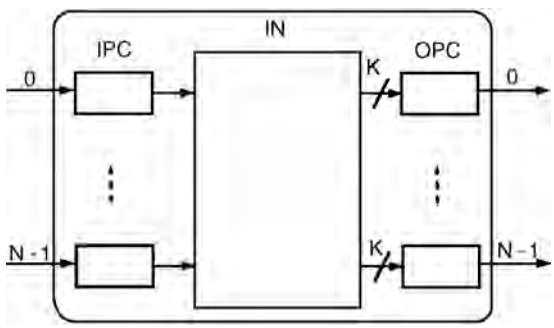


FIGURE 40.1 Model of ATM switch.

is required to allow the transfer of more than one cell to the same OPC. The IN is usually a multistage arrangement of very simple SEs, typically 2×2, which can be either provided with input/output/shared buffers (*SE queueing*), or unbuffered (*IN queueing*). In this last case, input and output queueing, whenever adopted, take place at IPC and OPC, respectively, whereas shared queueing is accomplished by means of additional hardware associated to the IN. Depending on the type of interconnection network and the queueing adopted, three classes of ATM switches will be distinguished:

- *Blocking multistage IN with minimum depth*. SE queueing and no speedup are adopted; the interconnection network is blocking and makes available only one path per I/O pair.
- *Blocking multistage IN with arbitrary depth*. IN queueing (typically, output queueing) and a speedup K are adopted; the interconnection network is blocking but makes available more than one path per I/O pair.
- *Nonblocking IN*. IN queueing and a speedup K are adopted.

In general, two types of conflicts characterize the switching operation in the interconnection network in each slot, the **internal conflicts** and the **external conflicts**. The former occur when two I/O paths compete for the same internal resource, that is, the same interstage link in a multistage arrangement, whereas the latter take place when more than K packets are switched in the same slot to the same OPC. An interconnection network $N \times N$ with speedup K, $K \leq N$, is said to be nonblocking if it guarantees absence of internal conflicts for any arbitrary switching configuration free from external conflicts for the given **network speedup** value K. That is, a nonblocking IN is able to transfer to the OPCs up to N packets per slot, in which at most K of them address the same switch output. Note that the adoption of output queues either in an SE or in the IN is strictly related to a full exploitation of the switch speedup; in fact, a structure with $K = 1$ does not require output queues, since the output port is able to transmit downstream one packet per slot. Whenever queues are placed in different elements of the ATM switch, it is assumed that there is no backpressure between queues, so that packets can be lost in any queue due to buffer saturation.

The main functions of the port controllers are:

- Rate matching between the input/output channel rate and the switching fabric rate.
- Aligning cells for switching (IPC) and transmission (OPC) purposes (it requires a temporary buffer of one cell).
- Processing the cell received (IPC) according to the supported protocol functionalities at the ATM layer; a mandatory task is the routing (switching) function, that is, the allocation of a switch output and a new VPI/VCI to each cell, based on the VCI/VPI carried by the header of the received cell.
- Attaching (IPC) and stripping (OPC) self-routing labels to each cell.
- With IN queueing, storing (IPC) the packets to be transmitted and probing the availability of an I/O path through the IN to the addressed output if input queueing is adopted; queueing (OPC) the packets at the switch output if output queueing is adopted.

The traffic performance of the nonblocking ATM switches will be described by referring to an offered uniform random traffic, with $p(0 < p \leq 1)$ denoting the probability that a packet is received at a switch input in a slot. All packet arrival events at IPCs are mutually independent, and each switch outlet is selected with the same probability by the cells. Typically, three parameters are used to describe the switching fabric performance, all referred to as steady-state conditions: the *switch throughput* $\rho(0 < \rho \leq 1)$ (the traffic carried by the switch) expressed as a utilization factor of its output links; the *average packet delay* $T(T \geq 1)$, that is, average number of slots it takes for a packet received at a switch input to cross the network and thus be transmitted downstream by the addressed switch output, and the *packet loss probability* $\pi(0 < \pi \leq 1)$, defined as the probability that a packet received at a switch input is lost due to buffer overflow. The maximum throughput ρ_{max}, also referred to as *switch capacity*, indicates the load carried by the switch for an offered load $p = 1$.

Needless to say, our dream is a switching architecture with minimum complexity, capacity very close to 1, average packet delay less than a few slots, and a packet loss probability as low as desired, for example, less than 10^{-9}. The target loss performance is usually the most difficult task to achieve in an ATM switch.

40.4 ATM Switch with Blocking Multistage IN and Minimum Depth

The class of ATM switch characterized by a blocking multistage IN with minimum depth is based on the adoption of a *banyan network* as IN. Banyan networks are multistage arrangements of very simple 2×2 switching elements without internal speedup with interstage connection patterns such that only one path exists between any inlet and outlet of the network and different I/O paths can share one or more interstage links. SEs in a banyan network are organized in $n = \log_2 N$ stages, each comprising $N/2$ SEs. Figure 40.2 represents the *reverse baseline* network [Wu and Feng, 1980]. An SE with high and low inlets (and outlets) labeled 0 and 1, respectively, can assume only two states, *straight* giving the I/O paths 0-0 and 1-1 in the SE, and *cross* giving the I/O paths 0-1 and 1-0.

The availability of only one I/O path per inlet/outlet pair with interstage links shared by several I/O paths causes the banyan network blocking due to the internal conflicts. In spite of its high blocking probability, the key property of banyan networks that suggests their adoption in ATM switches is their cell self-routing capability: an ATM cell preceded by an address label, the *routing tag*, is given an I/O path through the network in a distributed fashion by the network itself. For a given topology, this path is uniquely determined by the inlet address and by the routing tag, whose bits are used, one per stage, by the switching elements along the paths to route the cell to the requested outlet. The specific bit to be used, whose value 0(1) means that the cell requires the high (low) SE outlet, depends on the banyan network topology. An example of self-routing in a reverse baseline network is shown in Fig. 40.2, where a cell received on inlet 4 addresses outlet 9.

As is clear from the preceding description, the operations of the SEs in the network are mutually independent, so that the processing capability of each stage in an $N \times N$ switch is $N/2$ times the processing

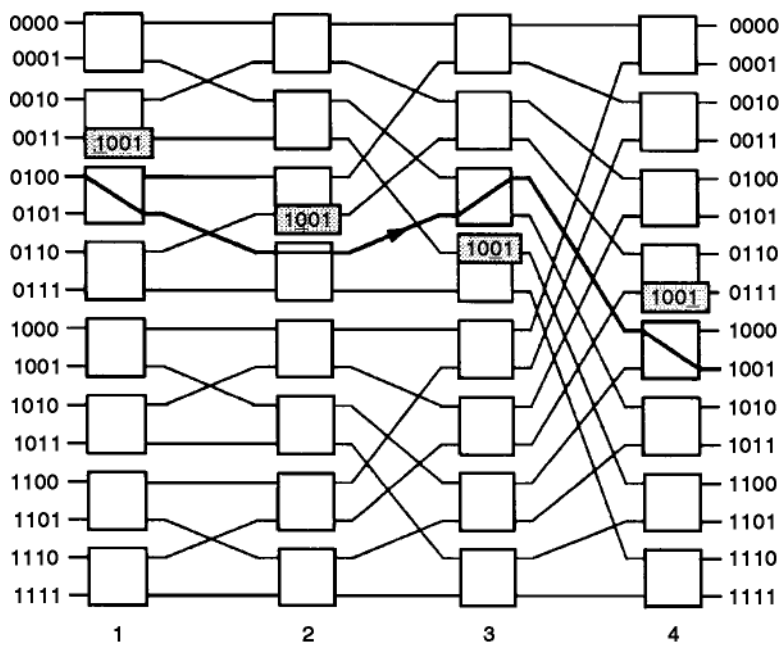


FIGURE 40.2 Reverse baseline network with self-routing example.

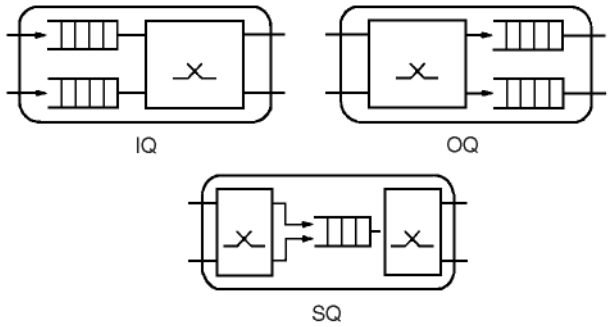


FIGURE 40.3 Structure of a 2 × 2 SE with IQ, OQ, and SQ.

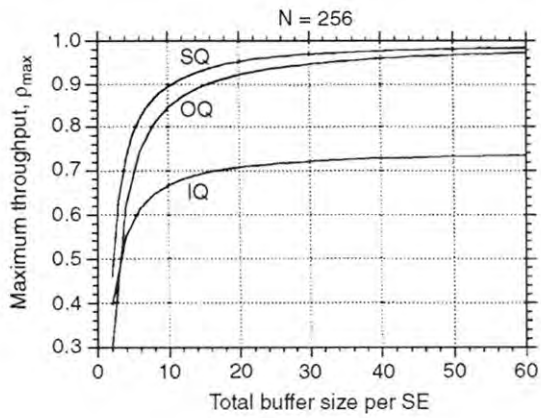


FIGURE 40.4 Throughput performance of banyan networks.

capability of one SE. Thus, a very high parallelism is attained in packet processing within the IN of an ATM switch by relying on space-division techniques.

As already mentioned, interstage links are shared by several I/O paths that cause packet loss if two packets require the same link in a slot. Provision of a queueing capability in the SE (SE queueing) is therefore mandatory to be able to control the packet loss performance of the switch. In general, three configurations are considered: input queueing (IQ), output queueing (OQ), and shared queueing (SQ), where a buffer is associated with each SE inlet, each SE outlet, or shared by all the SE inlets and outlets, respectively. Therefore, the SE routing is operated after (before) the packet storage with IQ (OQ), both before and after the storage with SQ (see Fig. 40.3). Note that these three solutions have an increasing hardware complexity: each buffer is read and written at most once per slot with IQ, read once and written twice per slot with OQ, and read twice and written twice per slot with SQ.

The maximum throughput of these networks increases with the total buffer size B (Fig. 40.4) from a value around $\rho_{max} = 0.4 - 0.5$ for all of the cases to a value close to $\rho_{max} = 0.75$ for input queueing and $\rho_{max} = 1.0$ for output and shared queueing. These values are the asymptotic throughputs expected, since a very large buffer in each SE implies that the throughput degradation is only due to internal conflicts. By definition, these conflicts do not occur with OQ SEs and they give the asymptotic throughput of a single 2 × 2 SE with IQ and OQ SEs. Given a total SE buffer size, SQ gives the best cell loss performance; OQ is significantly better than IQ unless a very low offered load is considered, as shown in Fig. 40.5 for a total SE buffer of 16 ATM cells for a varying offered load p.

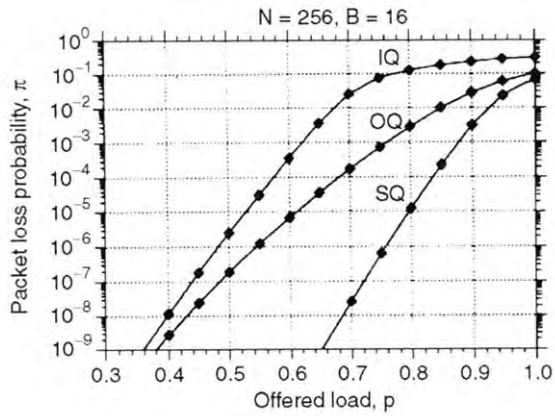


FIGURE 40.5 Loss performance of banyan networks.

40.5 ATM Switch with Blocking Multistage IN and Arbitrary Depth

Another class of ATM switches has been developed using the same concept of cell self-routing characterizing a banyan network but now based on the use of unbuffered SEs. The basic idea behind this type of switching fabric is that packet loss events that would occur because of multiple packets requiring the same interstage link are avoided by deflecting packets onto unrequested SE output links. Therefore, the packet loss performance is controlled by providing several paths between any inlet and outlet of the switch, which is generally accomplished by arranging a given number of self-routing stages cascaded one to another. Since each OPC of the switching fabric transmits at most one packet, queueing is still necessary when the interconnection network switches more than one packet in a slot to the same outlet interface (the IN has an internal speedup K). In this class of switch architectures, output queueing is obviously accomplished by providing each OPC of the switch with a queue.

The interconnection network of an $N \times N$ switch is a cascade of switching stages, each including $N/2$ unbuffered SEs, whose interstage link patterns depend on the specific ATM switch architecture. Unlike all other ATM switch classes, cells cross a variable number of stages before entering the addressed output queue, since direct connections to the output queues, through the SE local outlets, are also available from intermediate stages. Also, the routing strategy operated by the switching block depends on the specific architecture. However, the general switching rule of this kind of architecture is to route the packets onto the local outlet addressed by the cell as early as possible. Apparently, those cells that do not reach this outlet at the last switching stage are lost.

The cell output address, as well as any other control information needed during the routing operation, is transmitted through the network in front of the cell itself as the cell routing tag. A feature common to all ATM switch architectures based on deflection routing is the packet self-routing principle, meaning that each packet carries all of the information needed in the switching element it is crossing to be properly routed toward the network outlet it is addressing. Compared to the simplest case of self-routing operated in a banyan network, where the routing is based on the analysis of a single bit of the output address, here the routing function can be a little more complicated and varies according to the specific architecture considered. Now, however, depending on the specific network architecture, the cell self-routing requires the processing of more than one single bit; in some cases, the overall cell address must be processed in order to determine the path through the network for the cell.

Only two specific ATM switch architectures based on deflection routing will be described here to better clarify the principles of their internal operations, the *shuffleout* switch [Decina et al., 1991] and the *tandem banyan* switch [Tobagi et al., 1991], although other architectures have also been proposed.

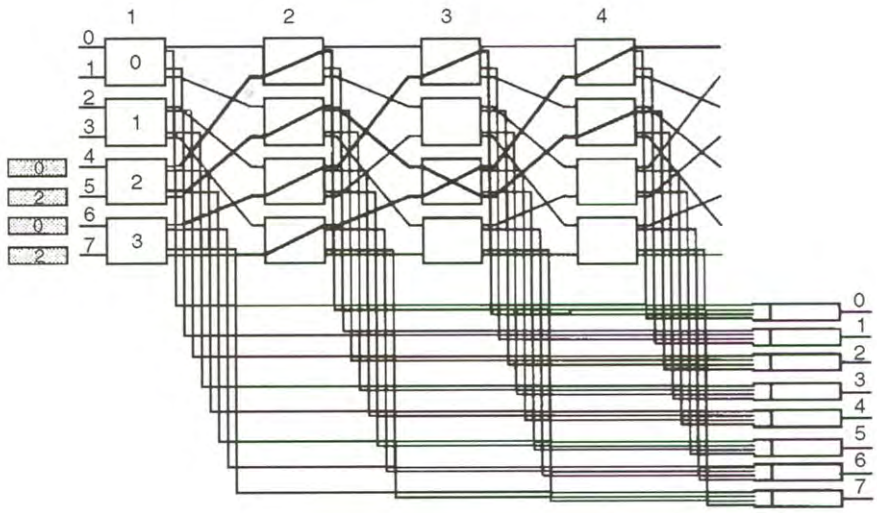


FIGURE 40.6 Routing example in shuffleout.

Each of the K switching stages in the shuffleout switch includes $N/2$ switching elements of size 2×4, and the interstage connection pattern is a shuffle pattern. Thus, the network includes $N/2$ rows of SEs, numbered $0 - N/2 - 1$, each having K SEs. An SE is connected to the previous stage by its two inlets and to the next stage by its two interstage outlets; all of the SEs in row $i(0 \le i \le N/2 - 1)$ have access to the output queues interfacing the outlets $2i$ and $2i + 1$ by means of the *local outlets*, thus accomplishing a speedup K. The shuffleout switch is shown in Fig. 40.6 for $N = 8$ and $K = 4$. The routing tag in the shuffleout switch is just the network output address.

The distributed routing algorithm adopted in the shuffleout interconnection network is jointly based on the *shortest path* and *deflection routing* principles. Therefore, an SE attempts to route the received cells along its outlet belonging to the minimum I/O path length to the required destination. The output distance d of a cell from the switching element it is crossing to the required outlet is defined as the minimum number of stages to be crossed by the cell in order to enter an SE interfacing the addressed output queue. After reading the cell output address, the SE can easily compute the cell output distance, whose value ranges from 0 to $\log_2 N - 1$ because of the shuffle interstage pattern.

When two cells require the same SE outlet (either local or interstage), only one can be correctly switched, whereas the other must be transmitted to a nonrequested interstage outlet, due to the memoryless structure of the SE. Thus, conflicts are resolved by the SE applying the deflection routing principle: if the conflicting cells have different output distances, then the closest one is routed to its required outlet, whereas the other is deflected to the other interstage link. If the cells have the same output distance, a random choice is carried out. If the conflict occurs for a local outlet, the loser packet is deflected onto an interstage outlet that is randomly selected.

An example of packet routing is shown in Fig. 40.6 for $N = 8$. In the first stage, both the SEs 2 and 3 receive two cells requiring the remote switch outlets 0 and 2, so that a conflict occurs for the their upper interstage link. In both cases, the contending cells have the same distance ($d = 1$ in SE 2 and $d = 2$ in SE 3) and the random winner selection results in the deflection of the cells received on the low inlets. The two winner cells enter the output queue 0 at stages 2 and 3, whereas the other two cells addressing outlet 2 contend again at stage 3; the cell received at the low inlet now is the winner. This cell enters the output queue at the following stage, whereas the cell that has been deflected twice cannot reach the proper output queue within the fourth stage of the network and, thus, is lost.

Each output queue, which operates on a first-in first-out FIFO basis, is fed by K lines, one from each stage, so that up to K packets can be concurrently received in each slot. Since K can range up to several tens depending on the network parameter and performance target, it is necessary to limit the maximum

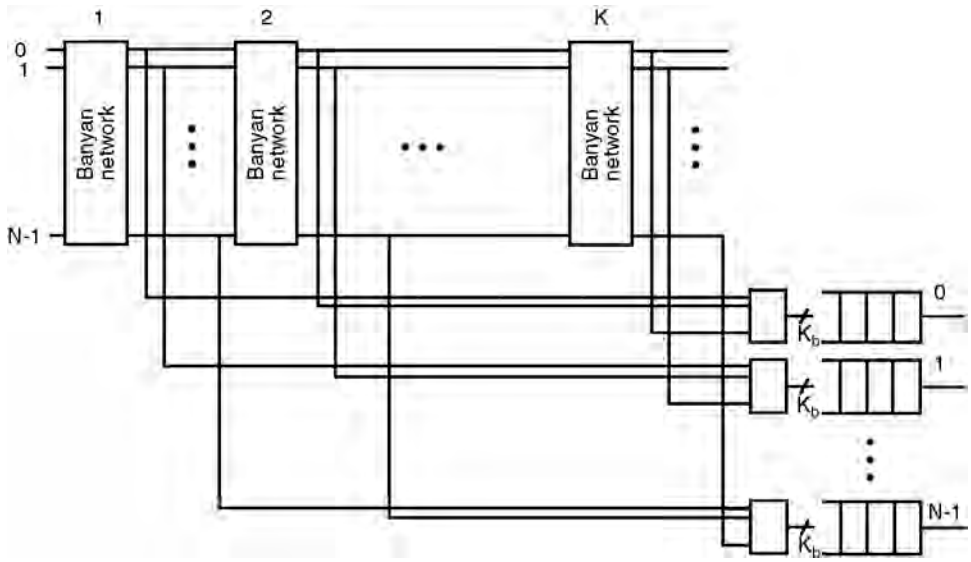


FIGURE 40.7 Architecture of the tandem banyan switch.

number of packets entering the queue in the same slot. Therefore, a *concentrator* with size $K \times C$ is generally equipped in each output queue interface so that up to C packets can enter the queue concurrently. The number C of outputs from the concentrator and the output queue size B (cells) will be properly engineered so as to provide a given cell loss performance target.

In the $N \times N$ tandem banyan switch, shown in Fig. 40.7, K banyan networks are serially arranged so that the total number of stages is now $nK(n = \log_2 N)$ each including 2×2 switching elements. Different topologies can be chosen for the basic banyan network of the tandem banyan switch, such as the omega [Lawrie, 1975], the baseline [Wu and Feng, 1980], and others. Each output queue is fed now by only K links, one from each banyan network. The first banyan network routes the received packets according to the very simple bit-by-bit self-routing (the specific bit to be used depends on the specific topology of the banyan network). In case of a conflict for the same SE outlet, the winner, which is chosen randomly, is routed correctly; the loser is deflected onto the other SE outlet, also by setting to 1 a proper field D of the routing tag, which is initially set to 0. Since each banyan network is a single I/O-path network, after the first deflection the packet cannot reach its addressed outlet in the same network. To prevent a deflected packet from causing the deflection of an undeflected packet at a later stage of the same network, the SE always correctly routes the undeflected packet if two packets with different values of the field D are received.

The output queue i is fed by the outlet i of each of the K banyan networks, through proper packet filters that select for acceptance only those packets with $D = 0$, whose routing tag matches the output queue address. The kth banyan network ($k = 2,\ldots,K$) behaves accordingly in handling the packets received from the upstream network $k - 1$: it filters out all of the of undeflected packets and accepts only the deflected packets ($D = 1$). Analogously to all of the previous architectures, packets that emerge from the network K with $D = 1$ are lost.

Unlike shuffleout, here a much smaller number of links enter each output queue (one per banyan network) so that the output queue, in general, does not need to be equipped with a concentrator. On the other hand, in general, cells cross a larger number of stages here since the routing of a deflected cell is not restarted just after a deflection, but rather when the cell enters the next banyan network.

Unlike ATM switches with input queueing where the head-of-the-line blocking limits the throughput performance, the maximum throughput $\rho_{max} = 1$ can be achieved in an ATM switch with deflection routing. Therefore, our attention will be focused on the cell loss probability. Cells can be lost in different parts of the switch, that is, in the interconnection network (a cell reaches the last stage without entering the addressed local outlet), the concentrator, and the output queue. Loss events in the concentrator and

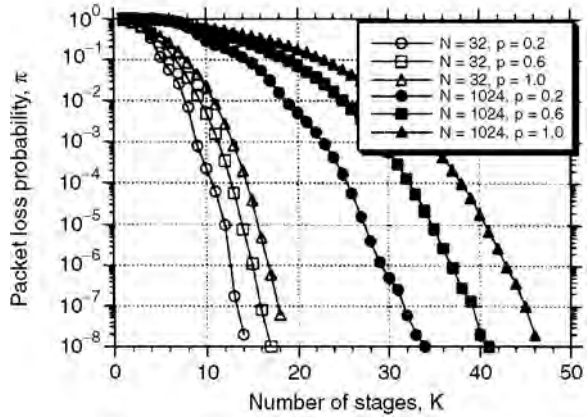


FIGURE 40.8 Loss performance in shuffleout.

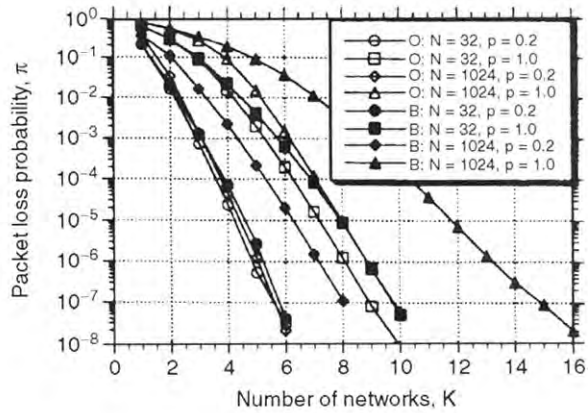


FIGURE 40.9 Loss performance in tandem banyan.

in the output queue can be easily constrained to arbitrarily low values by properly designing the number of concentrator outlets and the output queue size. A much more critical parameter to achieve a low loss is the selection of the number of stages in the interconnection network that gives a particular low cell loss probability. Figure 40.8 shows that less than 20 (50) stages are needed in shuffleout to provide a cell loss probability lower than 10^{-7} for a switch with size $N = 32$ (1024) under maximum input load. Less stages are required for lower input loads. Similar comments apply to the tandem banyan switch referred to the number of networks rather than to the number of stages. Figure 40.9 also shows that the banyan network topology significantly affects the overall performance. Adopting an omega topology (O) requires less networks than with a baseline topology (B).

40.6 ATM Switch with Nonblocking IN

The third class of ATM switching fabrics is characterized by a nonblocking IN with packet queueing capability placed in the IPCs (input queueing), the OPCs (output queueing), or placed within the IN itself (shared queueing). Mixed queueing strategies have been proposed as well. Requiring a nonblocking IN means referring to a network free from internal conflicts, that is, in principle, to a crossbar network (switching in an $N \times N$ crossbar network is accomplished by N^2 crosspoints, each dedicated to a specific I/O pair). Nevertheless, the centralized control characterizing a crossbar network makes this solution

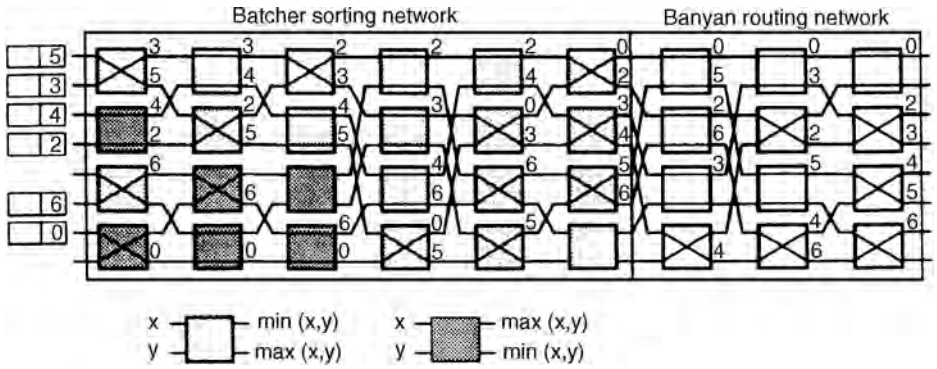


FIGURE 40.10 Example of cell switching in a Batcher-banyan network.

infeasible for a very high-speed ATM environment. As in the previous ATM switch classes, the only viable solutions are based on the adoption of multistage arrangements of very simple SEs, each capable of autonomously routing the cell to the required IN outlet.

The simplest ATM switch with nonblocking IN stores the cells in input queues located in the IPCs (IQ nonblocking IN). It is well known that a banyan network is nonblocking if the set of packets to be switched is *sorted* (by increasing or decreasing output addresses) and *compact* (packets received on adjacent network inlets). Therefore, the IN basically includes a sorting network and a routing banyan network, which are implemented as a multistage arrangement of very simple sorting and switching elements, respectively, whose size is typically 2×2. The number of stages of these networks is $n(n + 1)/2$ for a Batcher sorting network [Batcher, 1968] and n for a banyan routing network, where $n = \log_2 N$, each stage including $N/2$ elements. The simplicity of these SEs, each requiring a gate count on the order of a few tens of gates, is the key feature allowing the implementation of these INs on a small number of chips by relying on the current very large size integrated (VLSI) complementary metallic oxide (CMOS) technology. An example of sorting-routing IN is given in Fig. 40.10 for an 8×8 ATM switch, where the six cells received by the IN are sorted by the Batcher network and offered as a compact set with increasing addresses to the n-cube banyan network. Note that the nonblocking feature of the IN means that the network is free from internal conflicts. Therefore, additional means must be available to guarantee the absence of external conflicts occurring when more than one cell addresses the same switch outlet. Input queues hold the packets that cannot be transmitted immediately, and specific hardware devices arbitrate among the head-of-the-line (HOL) cells in the different input queues to guarantee the absence of external conflicts.

An OQ $N \times N$ nonblocking ATM switch provides cell storage capability only at OPCs. Now the IN is able to transfer up to K packets from K different inlets to each output queue without blocking due to internal conflicts (it accomplishes an internal speedup K). Nevertheless, now there is no way of guaranteeing the absence of external conflicts for the speedup K, as N packets per slot can enter the IN without any possibility of them being stored to avoid external conflicts. Thus, here the packets in excess of K addressing a specific switch outlet in a slot are lost.

According to the first original proposal of an ATM switch with pure output queueing, known as a *knockout* switch [Yeh et al., 1987], the IN includes a nonblocking $N \times N$ structure followed by as many output concentrators as the switch outlets. The $N \times N$ nonblocking structure is a set of N buses, each connecting one of the switch inlets to all of the N output concentrators. Each of these is an $N \times K$ network that is provided with N packet filters (one per attached bus) at its inputs, so as to drop all of the packets addressing different switch outlets.

A nonblocking $N \times N$ ATM switch with shared queueing behaves as a memory unit available in the IN and shared by all of the switch inlets to contain the packets destined for all of the switch outlets. The simplest implementation for this SQ switch model is given by a shared memory unit with N concurrent write accesses by the N inlets and up to N concurrent read accesses by the outlets. Clearly, such structure

shows implementation limits due to memory access speed, considering that the bit rate of the external ATM links is about 150 Mb/s. The current VLSI CMOS technology seems to allow the implementation of such $N \times N$ switches up to a moderate size, say, $N = 16$ or 32.

Multistage structures still provide a good answer for the implementation of larger size switches. The most important proposal is the *starlite* switch [Huang and Knauer, 1984], whose IN is a Batcher-banyan structure improved with additional hardware enabling feedback to the IN inlets those packets that cannot be switched to the addressed outlets because of external conflicts (the architecture does not have an internal speedup). Therefore, if P recirculation lines are available, then the size of the Batcher sorting network must be increased up to $(N+P) \times (N+P)$. These recirculation lines act as a shared buffer of size P.

The maximum throughput of an (unbuffered) crossbar network is $\rho_{max} = 0.632$ because of the unavoidable external conflicts, whereas it drops to $\rho_{max} = 0.586$ in an IQ nonblocking IN for an infinitely large switch ($N = \infty$). Thus, in spite of the queueing capability of this architecture, its capacity is even worse than in a crossbar switch. This phenomenon is explained by considering that holding in the HOL position those cells that cannot be transmitted because of external conflicts is a statistically worse condition than regenerating all of the cells slot by slot. Note that in a squared $N \times N$ switch with all active port controllers (PCs), a number h of blocked HOL cells imply h idle switch outputs, thus resulting in a carried load in the slot of only $(N - h)/N$. Nevertheless, behind a blocked HOL cell, the queue can hold other cells addressing one of the idle switch outlets in the slot. Such an *HOL blocking* phenomenon is said to be responsible for the poor throughput performance of pure IQ switches.

At first glance, a question could arise: what is the advantage of adding input queueing to a nonblocking unbuffered structure (the crossbar switch), since this results in a decrease of the switch capacity? The answer is provided in Fig. 40.11 showing the packet loss probability for a crossbar switch and an IQ switch with an input queue size B_i ranging from 1 to 32 cells. If the target loss probability is, say, 10^{-9}, we simply limit the offered load to $p = 0.3$ with $B_i = 8$, or to $p = 0.5$ for $B_i = 32$ with an IQ switch, whereas the packet loss probability of the crossbar switch is above 10^{-2} even for $p = 0.05$. Thus, input queueing does control the packet loss performance.

OQ nonblocking ATM switches show a much better traffic performance than IQ nonblocking switches as they do not have any HOL blocking phenomenon. In fact, the OQ switch capacity is $\rho_{max} = 1$ for $N = K$ and infinite output queue capacity. The loss performance is determined by the engineering of two parameters: the number of outlets per concentrator K, and the capacity of each output queue B_o. A speedup $K = 8$ is enough to guarantee a loss in the concentrator on the order of 10^{-8} with $p = 0.5$ and 10^{-6} with $p = 1.0$ for an arbitrary switch size. Moreover, the offered load p must be kept below a given threshold if we wish to keep the output queue capacity small by satisfying a given cell loss target (for example, $B_0 = 64$ guarantees loss figures below 10^{-7} for loads up to $p = 0.9$). The delay vs. throughput performance in OQ switches is optimal since the packet delay is only determined by the congestion for the access to the same output link.

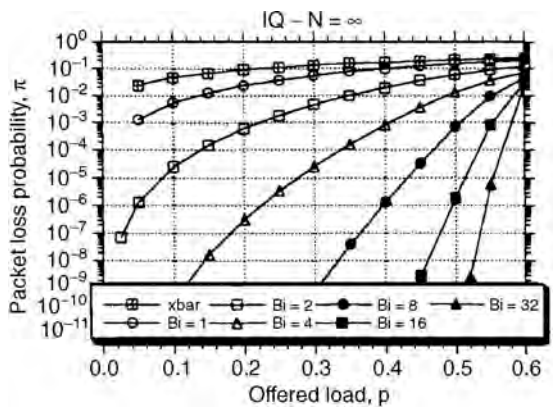


FIGURE 40.11 Loss performance with input queueing.

TABLE 40.1 Switch Capacity with Mixed Input/Output Queueing

x	$B_o = \infty$ $\rho_m(K)$	$K = 2$ $\rho_m(B_o)$	$K = 4$ $\rho_m(B_o)$
1	0.586	0.623	0.633
2	0.885	0.754	0.785
4	0.996	0.830	0.883
8	1.000	0.869	0.938
16		0.882	0.967
32		0.885	0.982
64		0.885	0.990
128			0.994
256			0.996
∞	1.000	0.885	0.996

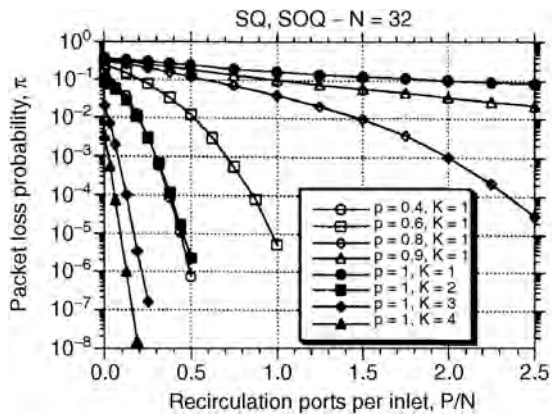


FIGURE 40.12 Loss performance with shared queueing.

Nonblocking ATM switches using multiple queueing strategies have also been proposed so as to overcome the performance limits of structures with input or shared queueing. Architectures with either mixed input/output queueing (IOQ) [Lee, 1990] or mixed shared/output queueing (SOQ) [Giacopelli et al., 1991] have been proposed. These architectures are basically an expansion of the respective basic Batcher-banyan structures in which K banyan networks are now available to switch up to K cells to the same OPC (the switch has an internal speedup K) and OPCs are equipped with output queues, each with a capacity B_o. The main performance results attained by these mixed queueing architectures are that the HOL blocking characterizing IQ switches is significantly reduced in an IOQ switch and the number of recirculators P that give a certain loss performance decrease considerably in an SOQ switch compared to the basic SQ switch.

Table 40.1 provides the switch capacity of an IOQ switch for different speedups and output queue sizes for $B_i = \infty$ and $N = \infty$. Note that the first column only contains the numerical values of the independent variable for which the switch capacity is evaluated, and its meaning is specified in each of the following columns. The switch capacity grows very fast with the speedup value K for an infinitely large output queue, $\rho_{max} = 0.996$ for $K = 4$, whereas it grows quite slowly with the output queue size given a switch speedup K. In general, it is cheaper to provide larger output queues rather than increase the switch speedup, unless the switching modules are provided in VLSI chips that would prevent memory expansions. Thus, these results suggest that a prescribed switch capacity can be obtained acting on the output buffer size, given a speedup compatible with the selected architecture and the current technology.

Analogously to OQ switches, pure shared queueing provides maximum switch capacity $\rho_{max} = 1$ for a shared queue with infinite capacity. The loss performance of SQ and SOQ architectures is given in Fig. 40.12 for $N = 32$ as a function of the normalized shared buffer size P/N for different offered loads.

It is seen that even a small shared queue of 32 cells ($P/N = 1$) requires the offered load to be limited to $p = 0.6$ if loss figures below 10^{-5} are required. It is interesting to note how the SOQ switch capacity approaches 1 even with a small shared buffer capacity by adopting a very little speedup ($\rho_{max} = 1 - 10^{-6}$ for $K = 3$ and $P = 7$).

40.7 Conclusions

The main architectural and performance features of switching fabrics for broadband networks based on asynchronous time-division switching have been presented. It has been shown that multistage arrangements of very simple switching elements is the typical solution adopted in ATM switches, which enables the processing required by ATM cells to be compatible with the bit rate on the incoming links. All of the different switch classes presented are able to provide the required performance target in an ATM network, typically the cell loss probability, by means of a suitable limitation of the offered load. Choosing one architecture rather than another is just a matter of finding an acceptable tradeoff between implementation costs and traffic performance.

Defining Terms

Arbitrary depth network: Interconnection network providing more than one internal path per inlet/outlet pair.

Asynchronous transfer mode (ATM): Transport technique defined for the lower layers of the broadband communication network.

ATM cell: Basic fixed-size unit defined for the transport of user information in a broadband ATM network.

Blocking network: Network in which internal conflicts occur for an arbitrary permutation that is free from external conflicts for a given network speedup.

Cell self-routing: Property of a multistage interconnection network that consists of the autonomous routing of the packets in each switching element only based on a routing tag carried by the cell without the use of many central control.

Deflection routing: Routing technique that deviates a packet from its intended route through the network because of an internal conflict.

External conflict: Conflict between packets for the access to the same network outlet.

Input queueing: Queueing strategy in which the information units are stored in buffers associated with the inlets of the interconnection network (or a switching element).

Internal conflict: Conflict between packets for the access to the same link internal to the interconnection network.

Minimum depth network: Interconnection network providing only one internal path per inlet/outlet pair.

Network speed-up: Number of packets that can be concurrently received at each output interface of an interconnection network.

Nonblocking network: Network free from internal conflicts for an arbitrary permutation that is free from external conflicts for a given network speedup.

Output queueing: Queueing strategy in which the information units are stored in buffers associated with the outlets of the interconnection network (or a switching element).

Shared queueing: Queueing strategy in which the information units are stored in a buffer shared by all of the inlets and outlets of the interconnection network (or a switching element).

References

Batcher, K.E. 1968. Sorting networks and their applications. In *AFIPS Proceedings of Spring Joint Computer Conference,* 307–314.

Decina, M., Giacomazzi, P., and Pattavina, A. 1991. Shuffle interconnection networks with deflection routing for ATM switching: the open-loop shuffleout. In *Proceedings of the 13th International Teletraffic Congress,* Copenhagen, Denmark, June, 27–34.

Giacopelli, J.N., Hickey, J.J., Marcus, W.S., Sincoskie, W.D., and Littlewood, M. 1991. Sunshine: a high performance self-routing broadband packet switch architecture. *IEEE J. Select. Areas Commun.,* 9(Oct.):1289–1298.

Huang, A. and Knauer, S. 1984. Starlite: A wideband digital switch. In *Proceedings of GLOBECOM 84,* Atlanta, GA, Nov., 121–125.

ITU. 1991. B-ISDN protocol reference model and its application. ITU-TI.321, International Telecommunications Union-Telecommunications Standardization Sector, Geneva, Switzerland.

ITU. 1993. B-ISDN asynchronous transfer mode functional characteristics. ITU-TI.150, International Telecommunications Union-Telecommunications Standardization Sector, Geneva, Switzerland.

Lawrie, D.H. 1975. Access and alignment of data in an array processor. *IEEE Trans. Comp.,* C-24(12):1145–1155.

Lee, T.T. 1990. A modular architecture for very large packet switches. *IEEE Trans. Commun.,* 38(7):1097–1106.

Tobagi, F.A., Kwok, T., and Chiussi, F.M. 1991. Architecture, performance and implementation of the tandem banyan fast packet switch. *IEEE J. Select. Areas Commun.,* 9(8):1173–1193.

Wu, C.L. and Feng, T.-Y. 1980. On a class of multistage interconnection networks. *IEEE Trans. Comp.,* C-29(Aug.) :694–702.

Yeh, Y.S., Hluchyj, M.G., and Acampora, A.S. 1987. The knockout switch: a simple, modular architecture for high-performance packet switching. *IEEE J. Select. Areas Commun.,* SAC-5(Oct.):1274–1283.

Further Information

A greater insight in the area of switching architectures for ATM networks can be obtained consulting the main journals and conference proceedings in the field of telecommunications in the last 15 years. A very comprehensive survey about architectures and traffic performance of ATM switches can be found in A. Pattavina, *Switching Theory, Architectures and Performance in Broadband ATM Networks,* John Wiley & Sons, 1998.

41

Internetworking

Harrell J. Van Norman
Unisys Corporation

41.1 Introduction

Engineering an internetwork is a lot like a farmer getting his apples delivered from an orchard to the supermarket. First, the farmer needs something to put the apples in: a bag, a box, a basket, or maybe a crate; second, the farmer needs to drive on some roads over a specific route. Those same two problems confront an engineer trying to get data delivered from one local area network (LAN) to another. Developing an internetwork involves finding something to put the apples in and providing a way to transport them from the orchard to the supermarket. The apples are the data and the container is the communications protocol that encapsulates or contains the data. Communications protocols envelope the application data and provide a container to ensure the goods are delivered safely and reliably. Internetworks require compatible protocols at both the transmitting and receiving sites. Various communications protocols are available such as transmission control protocol/Internet protocol (TCP/IP), system network architecture (SNA), DECnet, IPX, AppleTalk, and XNS, to name a few.

Roads and routes the farmer travels on from the orchard to the supermarkets represent telecommunications circuits the communications protocols use to route the data from source to destination. Engineering an internetwork involves sizing circuits and designing routes for the efficient and optimal transfer of information between LANs. Providing cost-effective and properly performing internetworks involves designing the best type of circuits, either dedicated or switched, and sizing bandwidth requirements.

Internetworking is defined in the IBM dictionary of computing as "communication between two or more networks." Miller [1991] defined the term as "communication between data processing devices on one network and other, possibly dissimilar devices on another network." For the purpose of this chapter, internetworking will take on the meaning of connecting two or more LANs to broaden the scope of communications.

41.2 Internetworking Protocols

Understanding, designing, and optimizing an internetwork would be too difficult unless the problem were broken down into smaller subtasks. Communications architectures partition functionality into several layers, each providing some service to the layer above and using the services of the layer below.

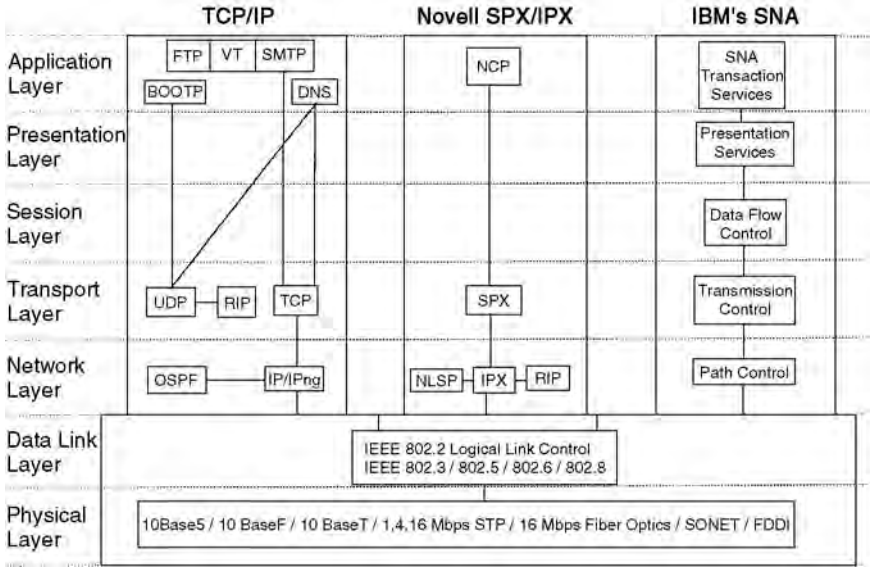


FIGURE 41.1 Internetworking communications protocol architectures.

Figure 41.1 illustrates the seven-layer International Organization for Standardization (ISO) open systems interconnection (OSI) reference model. TCP/IP, SNA, and sequence packet exchange/Internet packet exchange (SPX/IPX) are mapped into the seven-layer model. Each of the communications protocols converge at the data link and physical layers; however, the architectures all diverge at the network layer. Each communications protocol is represented uniquely at the upper five layers of the OSI communications model.

The data link layer is relevant to internetworking because this is where bridges operate. The service provided by the data link layer is important to routers—the key internetworking device, operating at the network layer. The transport layer is the user of network services and is influenced by the routing decisions made at the network layer. These three layers are the primary focus of internetworking.

Two models represent how an internetwork appears to endnodes. In the **connectionless** or datagram model, an endnode transmits packets containing data and the destination address to which the data is to be delivered. The network does the best to deliver the data, with some probability that the data will get lost, duplicated, or damaged. Packets are individually routed and the network does not guarantee the packets will be delivered in the same order they were sent.

The connectionless model is similar to the postal system in which each letter enters the postal system with the address where it should be delivered. The postal system does its best to deliver the data, with some probability that the letter will get lost or damaged. Letters are individually routed, and a letter posted on Monday might get delivered after a letter posted on Tuesday, even if both were posted from the same location and addressed to the same location.

The other model is the **connection-oriented** or virtual circuit. Here, an endnode first informs the network that it wishes to start a conversation with some other endnode. The network then notifies the destination that a conversation is requested, and the destination accepts or refuses. This is similar to the telephone system in which the caller informs the telephone system of the wish to start a conversation by dialing the telephone number of the destination. The telephone system establishes a path, reserves the necessary resources, contacts the destination by ringing its phone, and after the call is accepted, the conversation can take place. Often, a connection-oriented service has the following characteristics [Perlman, 1992]:

1. The network guarantees packet delivery.
2. A single path is established for the call and all data follows that path.
3. The network guarantees a certain minimal amount of bandwidth.
4. If the network becomes overly utilized, future call requests are refused.

TCP/IP Internetworking

TCP/IP has emerged as the primary communications protocol for developing internetworks. Why? The answer is simple: interoperability; almost every conceivable combination of computer hardware and operating system has a driver available for TCP/IP. International travelers know English is the speaking world's most common second language. A native language may be preferred when speaking with local family and friends, but English is the worldwide second language and is predominately used to communicate internationally. Similarly, TCP/IP has emerged as the primary protocol for communicating between networks.

The suite of TCP/IP protocols, adopted in the 1970s by the U.S. Department of Defense, is the Internet backbone. Now, over 3.1 million users are connected to the Internet, all speaking IP, up 81% from last year. The International Data Corporation (IDC), in Framingham, MA, estimates 4.5 million PCs are running TCP/IP and the cumulative installed base for TCP/IP is 8 million computers. TCP/IP is clearly the internetworking protocol of choice.

Endorsements of TCP/IP just keep coming. First, Microsoft tucked a TCP/IP stack into Windows NT Version 3.1 and declared TCP/IP its strategic wide-area network protocol family. Most recently, a federally commissioned panel recommended that the government sanction TCP/IP as a standard and surrender its open systems interconnection-only stance in the Government OSI Profile (GOSIP). TCP/IP, the preferred protocol for internetworking, is also becoming the primary backbone protocol for local communications.

One of the major forces driving TCP/IP to the PC desktop is UNIX, which usually includes TCP/IP. UNIX has become the de facto standard enterprise server OS. As users downsize from mainframe platforms and adopt UNIX-based servers, they must outfit their PCs with access to TCP/IP. Downsizing is a major trend as processing power becomes more cost effective in the minicomputer and microcomputer platforms. An Intel 286-based PC introduced in 1983 provided 1 MIPS performance, in 1986 the i386 offered 5 MIPS, in 1990 the i486 came with 20 MIPS, and more than 100 MIPS was available in 1993 with the Pentium. Intel plans to routinely increase desktop processing power with the i686 in 1995, offering 175 MIPS, the i786, slated to provide 250 MIPS in 1997, and the i886 in 1999, capable of approximately 2000 MIPS.

TCP/IP is a connectionless protocol. TCP operates the transport layer and IP is the network layer protocol based on 4 octet addresses. A portion of the address indicates a link number and the remaining portion a system or host on the network. A 32-b **subnet mask** has 1s in the bits corresponding to the link portion of the IP address and 0s in the bits belonging to the host identifier field. If an IP address is X and its mask is Y, to determine whether a node with IP address Z is on the same link, AND the mask Y with X and AND the mask Y with Z. If the result is the same, then Z is on the same link.

Originally, IP addresses were grouped according to three distinct classes: class A contained the network address field in the first octet and host identifier in the remaining 3 octets, class B contained the network address field in the first two octets and the host in the remaining two fields, and class C contained the network address field in the first three octets and the host identifier in the remaining octet. Today, IP allows a variable sized subnet mask and even permits noncontiguous 1s in the mask. This can result in incredibly confusing and computationally inefficient addressing structures. Regardless, each node on a link must know the subnet mask of its own IP address.

IP routers connect networks by routing packets based on destination link numbers, not destination host identifiers. If IP routers had to route packets to destination hosts, enormous memory would be necessary to contain information about every machine on the Internet. Consequently, IP routing does not route to the destination endnode, it routes only to the destination link.

TCP/IP does have a certain amount of overhead associated with address resolution protocol (ARP) queries and responses that can degrade network performance significantly. The ARP allows a host to find the physical address of a target host on the same link number, given only the target's IP address. For example, if host A wanted to resolve the IP address of host B, a broadcast query would be transmitted and received by all hosts, but only host B would respond with its physical address. When A receives the reply, it uses the physical address to send packets directly to B. Considerable bandwidth and memory in

routers are consumed by routers keeping track of endnodes rather than just keeping track of links [Comer, 1991].

Improvement of the TCP/IP protocol suite will come in new protocols and emerging classes of applications. The dynamic host configuration protocol (DHCP) is being standardized to automatically gather detail configuration data for a PC when it joins the network, creating a plug-and-play environment. Historically, network administrators manually set parameters on each workstation, such as IP addresses and domains, but BOOTP and DHCP have advanced TCP/IP to an automated installation process.

Two classes of emerging applications for TCP/IP include interenterprise and personal interenterprise. Interenterprise applications unite different organizations' networks; for example, a retailer could order automatically from his suppliers or check credit-card information, all by TCP/IP. This brings electronic data interchange transactions, typically separate dial-up links, into the TCP/IP network. Personal interenterprise applications customize a PC view of the Internet, so that when connected, specific information is automatically presented, such as stock portfolio data.

The Internet is the largest network in the world, a network of networks that unarguably demonstrates the usefulness of TCP/IP. But TCP/IP is suffering from its own success. By the year 2020, the Internet will likely have exhausted all available addressing space. The Internet's standards-setting body, the Internet Engineering Task Force, has adopted the next generation IP. Called **IPng**, this new version of the IP layer will solve the problem of dwindling address space by supporting 16-byte addresses that will allow billions more users to surf the Internet. The new protocol (to be assigned version number IP6) will not require any immediate changes to users' existing IP internetworks. However, users will need to upgrade their routers and eventually change applications if they want to take advantage of the advanced features of IPng, such as flow control, autoconfiguration, and encryption-based security.

SNA Internetworking

IBM's SNA traffic is difficult to integrate into a multiprotocol internetwork because it is a nonroutable protocol. NetBIOS and Digital's local area transport (LAT) are also nonroutable protocols. Although it is unfortunate, these nonroutable protocols are without network addresses. Without network layer functionality, only bridges can internetwork native SNA. There is no intrinsic reason why these protocols could not have been designed to run with network layer addresses, they just were not. Perhaps designers never believed LANs would be internetworked.

Routers, as opposed to bridges, are the primary internetworking devices used to direct traffic from its source to its destination. Bridging is useful in simple point-to-point networks where alternate paths are not available and where protocol overhead is not a significant factor. Disadvantages of bridges relative to routers include:

- Bridges can use only a subnet of the topology (a spanning tree). Routers can use the best path that physically exists between source and destination.
- Reconfiguration after a topological change is an order of magnitude slower in bridges than routers.
- The total number of stations that can be interconnected through bridges is limited to the tens of thousands. With routers, the total size of the network is, for all practical purposes, unlimited (at least with ISO).
- Bridges limit the intermediate hop count to seven, significantly restricting the size of an internetwork.
- Bridges offer no firewall protection against broadcast storms.
- Bridges drop packets too large to forward, since the data link layer cannot fragment and reassemble packets.
- Bridges cannot give congestion feedback to the endnodes.

For the past several years, most corporations relied on token ring bridges implementing SNA's data link protocol, Logical Link Control 2 (LLC2), to give PCs on remote LANs a way to access IBM mainframe data.

Frames are forwarded throughout the internetwork by token ring bridges using source-route bridging protocol to ensure data is delivered to the proper destination. The problem is that even a few seconds of congestion on the wide area network (WAN) link can drop the SNA/LLC2 sessions.

IBM's **data link switching (DLSw)** standard provides the reliability needed to integrate mission-critical SNA networks with IP backbones. DLSw products integrate SDLC with IP by stripping off the synchronous data link control (SDLC) header and replacing it with a TCP/IP header. Prior to DLSw routers, SNA data was encapsulated in TCP/IP, leaving the SDLC headers intact. Current testing of router vendors' DLSw implementations indicate they keep SNA sessions from being dropped over even the most congested internetwork links. Performance at full speed of the WAN link is another big improvement over token ring bridges. A successor to source-route bridging, DLSw routers deliver high-performance SNA internetworking even with significantly more complexity than source-route bridging.

The DLSw scheme allows routing IP and internetwork packet exchange (IPX) traffic alongside connection-oriented SNA. DLSw routers have the added complexity of blending connectionless LAN traffic with connection-oriented SNA traffic by sending messages to the stations on each end of the connections, acknowledging that a path is still open between them. Preventing the SNA connection from timing out and dropping sessions involves considerable overhead traffic, such as frequent receiver ready, keep alive broadcasts, and timing signals for constant responses.

DLSw routers are criticized by some SNA users because of the inconsistent performance associated with contention-based internetworks. IP response time may be nearly instantaneous one time and may take a second the next, frustrating SNA users accustomed to consistent performance. DLSw can benefit large SNA networks suffering from time outs, hop limits, alternate routing disruptions, and the lack of security mechanisms. DLSw routing gives non-routable SNA traffic the robustness of nondisruptive alternate routing when links or nodes fail, the benefits from open nonproprietary protocols, and the cost effectiveness of standard TCP/IP-based solutions.

For IBM users, there is the option of remaining all "blue" with IBM's advanced peer-to-peer networking (APPN) family of protocols, the next generation of SNA. IBM continues to develop, position, and promote its own APPN as the only viable next-generation networking technology for its mainframe customers. The need for guaranteed response time in their mission-critical mainframe applications makes APPN an attractive option. Other advantages to the APPN approach include full class of service application priorities, enhanced fault management through NetView monitoring, the added efficiency from not encapsulating SNA into a routable protocols like TCP/IP and thereby doubling the addressing overhead, and improved accounting from SNA usage statistics.

SPX/IPX Internetworking

The number one LAN protocol is undoubtedly Novell's sequence packet exchange/Internet packet exchange (SPX/IPX), but most network administrators use TCP/IP for internetworking. As a result, Novell has taken two significant steps to improve the performance of SPX/IPX for internetworking: **packet burst** and the network link services protocol (NLSP).

Before packet burst, the NetWare core protocol (NCP) was strictly a ping-pong discipline. The workstation and the server were limited to communicating by taking turns sending a single-packet request and receiving a single-packet response. Packet burst is an NCP call permitting a sliding window technique for multiple packet acknowledgment. By transmitting a number of packets in a burst, before waiting for an acknowledgment or request, packet burst eliminates the ping-pong effect and can provide up to a 60% reduction in the number of packets to read or write large files. Limited to only file reads and writes, packet burst cannot reduce the ping-pong method of all other NCP communications such as opening directory handles, locating and opening files, and querying the bindery.

Rather than reducing the number of packets when the receiving station or internetwork becomes congested, as do most transport layer protocols, packet burst actually varies the interpacket gap (IPG). The IPG is the time between the end of one packet and beginning of the next packet. Varying the IPG for flow control is called packet metering.

The amount of data sent during a burst is called the message. Message size is based on the number of packets in the burst, called the window size, and the media access method of the link. The default window size has been optimized for most communications to 16 packets for a read and 10 for a write. For example, 16 reads on a token ring link with a 4096 packet size gives a 53-KB message. On an ethernet link, a read message is about 22 KB (16×1500 bytes). The number of packets in the burst window can be increased on high-speed links and where the degree of transmission success is high. If a packet or packets are lost in transmission for any reason, the client will send a request for retransmission of the lost packets. The server will retransmit only those packets that were lost. Once the client has all of the packets, it requests another message; however, since packets were lost, the IPG is increased for the next transmission. The IPG will continue to be increased until packets are no longer being lost.

NLSP is another internetworking improvement to NetWare. Novell developed the **link-state router** protocol NLSP, to replace the older IPX distance-vector based router information protocol (RIP). IPX RIP is similar to Cisco's IGRP, IP RIP, and RTMP in the AppleTalk suite. **Distance-vector routing** requires each node to maintain the distance from itself to each possible destination. Distance vectors are computed using the information in neighbors' distance vectors and store this information in routing tables that show only the next hop in the routing chain rather than an entire map of the network. As a result, RIP frequently broadcasts these summaries to adjacent routers.

Distance-vector routing is efficient for routers with limited memory and processing power, but newer routers typically have reduced instruction set computer (RISC) processors with power to spare. Link-state routers can provide full knowledge of the network including all disabled links with a single query. With distance-vector routers, this would involve querying most, if not all, of the routers. In addition, link state routers, like NSLP, converge more quickly than distance-vector routers, which cannot pass routing information on until distance vector has been recomputed. NSLP supports large, complex IPX internetworks without the overhead present in NetWare distance-vector routers.

NSLP routers compress service advertizing protocol (SAP) information and only broadcast SAP packets every two hours unless a change is made to network services, such as print servers or file servers. SAP broadcast overhead can be significant, especially with many servers or services. Novell has also reduced SAP traffic on the network by using NetWare Directory Services (NDS) to keep all network services information on a central server. NDS is available in NetWare 4.x and is similar to TCP/IP's domain name server (DNS). Novell also provides the capability to communicate with different systems through NetWare for SAA, NetWare NFS, and NetWare TCP/IP Gateway.

41.3 The Total Network Engineering Process

This guide to internetworking provides a step-by-step engineering approach. Migrating from a LAN or several LANs to an internetwork is no small undertaking. Key to the process of analyzing and designing internetworking strategies is employing analytic heuristic algorithms and simulation techniques. Network design tools are essential to effectively evaluate alternative internetworking schemes while attempting to optimize communications resources. As new networking requirements surface, engineers design capacity and equipment to satisfy the demands. Facing requirements on an as-needed basis results in network evolution. Networks developed in an evolutionary manner provide incremental solutions to an array of individual requirements. With each new requirement, the most cost-effective configuration of equipment and communications lines is determined, focusing on the specific need. Unfortunately, design decisions are often made without rigorous analysis of the entire network, addressing only the incremental expansion. This approach seldom provides the best design for the total communications environment.

The preferred method of addressing network requirements is through a structured-design approach. Here, alternative design techniques for the entire network are evaluated in terms of response time, cost, availability, and reliability. While avoiding excessive redesign and allowing for orderly capacity expansion, the process of network engineering evaluates current networking requirements and plans for future network expansion.

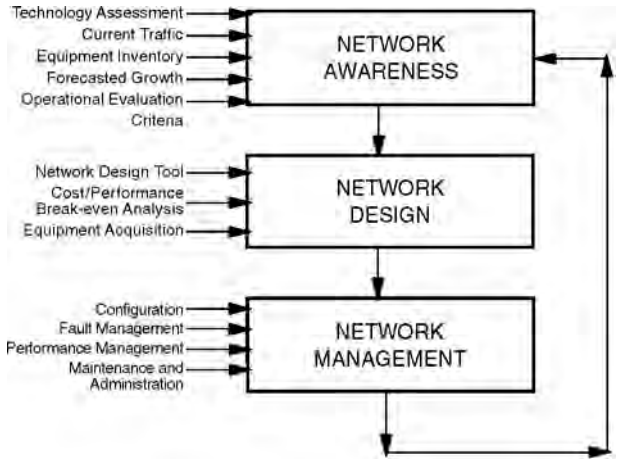


FIGURE 41.2 The total network engineering process.

Figure 41.2 shows the **total network engineering process** and its three phases: awareness, design, and management. **Network awareness** quantifies current traffic, equipment inventories, forecasted growth, and operational objectives. **Network design** uses design tools, performs a cost/performance breakeven analysis, and develops line configurations. **Network management** encompasses management of configuration, faults, performance, accounting, and security. Equipment and network software is acquired, verified, installed, maintained, and administered.

Network Awareness

The process of network awareness includes the following operations:

- conducting a traffic inventory by collecting measurements of current traffic
- mapping the physical inventory of networking equipment into the logical interconnections
- forecasting requirements for network growth through surveying line management
- qualifying through senior management the criteria used to evaluate operational tradeoffs

Many of these inputs are, of necessity, qualitative or conceptual in nature, whereas others can be quantified with varying degrees of accuracy. Using a network design process often requires a greater level of network awareness than most engineers generally possess. When an engineer has an accurate understanding of the network—including possible bottlenecks, response time spikes, and traffic patterns—the process of network design can begin. Gaining an accurate awareness of network traffic involves monitoring current loads via a network analysis tool, like the LANalyzer, Sniffer, or through the use of RMON probes. Traffic statistics for network planning should be collected, not only for all significant hours of each day, but also for a significant number of days in a year. CCITT recommends measuring carried traffic statistics for at least the 30 days (not necessarily consecutive) of the previous 12 months in which the mean busy hour traffic is the highest. A second-choice method recommended by CCITT, when the first method is not possible, consists of a measuring period of 10 consecutive normal working days during the busiest season of the year [CCITT, 1979]. Both average and peak loads are characterized hourly over the sampling period to understand current bandwidth requirements. Monitoring new networked applications predicts their impacts to the internetwork. Traffic must also be differentiated between what stays local, remaining only within a LAN, and traffic that is internetworked between remote locations via WAN links.

Interviews with line management help to forecast growth requirements. Many managers have hidden agendas for using the network for a variety of business purposes. Without surveys or interviews of key

management personnel, predicting future network requirements is impossible. Questions need to center around four key areas:

1. Future business communications requirements based on the use of advancing technologies
2. Planned facility changes and population shifts through developing new business partners
3. Forecasted needs for external connectivity with customers, vendors, and new markets
4. Changes in business practices based on new programs, applications, and remote systems access.

A little-publicized problem, but one that nevertheless causes big headaches, is developing an inventory of current network equipment. Obtaining an accurate image of the inventory database is helpful in the network engineering process to gain an improved network awareness. Based on network discovery techniques, like IP discovery and IPX SAP, routing nodes may be identified throughout the internetwork. These techniques only find network layer routing devices; however, internetworks are primarily based on routers rather than data link layer bridges. Once these devices are identified, MIB browsers are effective to understand what mix of protocol traffic is being routed throughout the internetwork. Finally, a topological map can be developed by translating the logical routing to physical locations.

Top level management must also be involved to quantify the operational evaluation criteria. Every business has established measures for effective operations. A network exists to meet the needs of those who use the network, not the preferences of those who maintain it. Therefore, objectives for the communications resources should be a reflection of those criteria held by the business as a whole. Five key operational evaluation criteria or network design goals characterize most networks: performance, availability, reliability, cost, and security. Manufacturing industries and start-up companies often choose cost minimization, financial markets mandate high reliability, security is of utmost importance throughout classified governmental networks like the Department of Energy (DOE) nuclear weapons complex, and high performance is necessary to mass merchandise retail companies. Key operational evaluation criteria are described below.

Cost minimization is a weighty design factor that includes both monthly recurring and nonrecurring fees. To compare competing alternatives with different cost outlays, the time value of money must be considered and the net present value method is usually the best method for financial analysis. In the name of cost minimization, depending on the specific operational objective, one may be led to different solutions given the same problem. Variations of cost minimization goals include: build the cheapest network, get the most network for the dollar, maximize profit by allowing the firm to reach new markets, minimize risk of loss, or maximize growth opportunity.

All network designs include an inventory of lines, equipment, and software needed to make them run. Including all of the cost elements (installation costs, shipping costs, purchase price, monthly lease costs, maintenance costs, and service costs) in the optimization process prevents unpleasant surprises. Total network costs C_{TOT} include the three major cost components in a network design: line costs, equipment costs, and software costs:

$$C_{TOT} = C_{AL} + C_{TL} + C_{IE} + C_{WAE} + C_S \tag{41.1}$$

where:

C_{AL} = access line costs
C_{TL} = interoffice channel (IXC) line costs
C_{IE} = internetworking equipment costs (remote bridge, router, gateway, etc.)
C_{WAE} = WAN access equipment costs [modems, computer software unit/data software unit (CSU/DSU), channel banks, etc.]
C_S = software costs

Too often, availability and reliability are mistakenly assumed to be synonymous. Reliability and serviceability are both components of availability. Network availability, a value between zero and one, reflects stability characteristics of communications systems connecting end users. Availability A is composed of two factors: reliability, the probability a system will not fail during a given time, is expressed as mean

time between failure (MTBF), and serviceability, including the time required to diagnosis and repair, is expressed as mean time to repair (MTTR),

$$A = \text{MTBF}/(\text{MTBF} + \text{MTTR}) \tag{41.2}$$

Performance is another key operational evaluation criteria. Represented by average transaction response times RT, it can be calculated by the following formula:

$$R_{RT} = I_T + H_T + O_T \tag{41.3}$$

where

I_T = input message transmission time: number of bits in input message/link speed bits per second + propagation delay
H_T = host processing time
O_T = output message transmission time: number of bits in output message/link speed bits per second + propagation delay

At the completion of the network awareness phase, the network planner must objectively evaluate whether or not this phase has been completed adequately for the purpose of the design process. Questions that might be addressed are:

- Do I have a reasonable assessment of current technology in this networking arena?
- Have I quantified the current network traffic?
- Is there a justifiable forecast of network growth?
- Are the operational objectives and goals quantified and agreed upon by upper level management?

Upon completion, quantifiable operational evaluation criteria are available as input into the network design phase.

Network Design

For too long, networks merely evolved to their present condition by adding users and applications without analyzing performance impacts. The better way is for engineers to describe the topology and media via a network modeling tool. In many network design tools, models are constructed by connecting icons representing specific network devices or segments. Behind the model, powerful software simulates and analyzes the behavior of the network. Also, network conditions such as protocols, traffic patterns, and number of nodes are specified in the model. Traffic can be characterized from statistical traffic models or actual traces of network traffic taken from an analyzer. Simulation and performance evaluation predict network performance and explore future design options. Graphical user interfaces (GUI) simplify changing assumptions on number of users, type of media protocol stacks, application mix, traffic patterns, and so on. When the network model is in place, proposals for adding or deleting resources (such as controllers, terminals, multiplexers, protocol converters, remote front-end processors), redesigning application software, or changing routing strategies may be evaluated.

Network design consists of applying network design tools, evaluating alternative network configurations through a cost/performance breakeven analysis, and optimizing communications lines. Because there are many ways to configure a network and interconnect transmission components, it is not realistic to evaluate exhaustively every possible option. However, by using certain shortcuts, called heuristics (procedures that approximate an optimal condition using algorithms), the number of feasible designs can be scaled down to a manageable few.

Network design tools are developed using heuristic algorithms that abbreviate the optimization process greatly by making a few well-chosen approximations. Queuing theory is incorporated in the algorithms

that calculate response times. A convenient notation used to label queues is:

$$A/B/n/K/p \tag{41.4}$$

where:

A = arrival process by which messages are generated
B = queue discipline, describing how messages are handled
n = number of servers available
K = number of buffering positions available
p = ultimate population size that may possibly request service

Some of the more typical queuing models pertaining to internetworking are as follows:

- Memoryless (exponential) interarrival distribution, memoryless service distribution, one server (M/M/1) is typical of situations involving a communications server or a file server of electronic mail.
- Memoryless (exponential) interarrival distribution, memoryless service distribution, *c* servers (M/M/*c*) are typical of situations involving dial-up ports for modem pools or private branch exchange (PBX) trunks.
- Memoryless (exponential) interarrival distribution, fixed (deterministic) service time, one server (M/D/1) is representative of a trunk servicing a packet switched network.

A number of internetworking situations require a more advanced model involving distributions other than exponential or deterministic. For the terminal nodes of a communications system, the queue type that is most appropriate is M/G/l distributions, referring to a generic service time represented by G. This means the process messages that arrive are modeled as Markovian. The M/G/1 solution of queuing delay is given by the Pollaczek–Khinchin (PK) formula:

$$d = m + l * m^2/[2(1-p)] \tag{41.5}$$

where

d = total delay
m = average time to transmit a message
l = number of arriving messages/second
p = utilization of the facility, expressed as a ratio

After a network model is produced, network engineers often use the tool to evaluate different scenarios of transmission components and traffic profiles. Through an iterative process of network refinement, "what if" questions may be addressed concerning various alternative configurations:

- What would be the impact on performance and utilization if more nodes are added to my network?
- How would response times change if my internetwork links were upgraded from voice grade private lines to 56-kb/s digital circuits; from fractional T-1 dedicated circuits to fast packet switched frame relay?
- How will downsizing to client/server technologies impact the network?
- Should I segment the network with high-speed switches or with traditional routers?
- How much additional capacity is available before having to reconfigure?
- What effect do priorities for applications and protocols have on user response times?

Each of the various alternative configurations will have associated costs and performance impacts. This produces a family of cost/performance curves that forms a solution set of alternative network designs. With the cost/performance curves complete, the final selection of an optimum configuration can be left up to the manager responsible for balancing budget constraints against user service levels.

At the completion of network design, the network planner must objectively evaluate whether or not this phase has been completed adequately for the purpose of the design process. Questions that might be asked are:

- Is the model verified by accurately representing the inputs and generating output that satisfy the operational evaluation criteria?
- Is the cost/performance breakeven analysis selected from the family of feasible solutions the one best configuration based on the operational evaluation criteria?
- Is the equipment rated to function well within the design objectives?

When this phase is complete, an optimized implementable network design is available as input into the network management phase.

Network Management

Network management includes equipment acquisition, verification, installation, maintenance, and administration. This phase of the network engineering process feeds information into the network awareness component, which completes the cycle and begins the iterations of design and refinement.

Network design tools use statistical techniques to make certain assumptions about the behavior of a network and the profiles of transactions in the network. These assumptions in terms of lengths, distributions, and volumes of traffic make the mathematics easier, at the expense of accuracy. Operational networks often violate, to a certain extent, many design assumptions. The optimized network designs confront the real world of actual network performance during the network management phase. Detected variances between optimized network designs and actual network performance provide an improved awareness concerning host processing requirements, traffic profiles, and protocol characteristics.

In a nutshell, network management improves awareness. Design iterations and refinement based on improved network awareness progress toward the optimum. Network management refines operational evaluation criteria. Cost/performance ratios are evaluated against the operational objectives and goals an organization is employing. Network management avoids crises by reacting proactively to changing trends in utilization and performance.

At the completion of network management, the network planner must objectively evaluate whether or not this phase has been completed adequately for the purpose of the design process. Questions that might be asked are:

- Is the network operating error free or are there periodic or isolated errors?
- Is the network performing properly by satisfying response time and availability goals?

When this phase is complete, an operational network is available.

Finally, the network engineer must verify and validate the model by asking: is the model validated by the performance of the operational network? If this objective evaluation results in discrepancies between the model and the real world operational network, then there must be a return to the network awareness phase for model refinement and, possibly, improved optimization. Whenever new needs are placed on the network, the process must be reinitiated, with a revised needs analysis reflecting these new requirements. When the model is fully validated and verified, the network planner can exit the total network engineering decision process with an optimized operational network.

41.4 Internetwork Simulation

Today, with an overwhelming variety of alternative internetwork implementation options, engineers are often uncertain which technology is best suited for their applications. A few years ago, there were relatively few systems from each of several vendors. Times have changed, and now a multitude of interconnection options exist. LAN media choices involve twisted pair, coaxial cable, fiber optics (single mode and multimode), and even wireless systems. LAN media access protocols include token ring (4 or 16 Mb/s),

token bus, ethernet's carrier sense multiple access/collision delay (CSMA/CD), 100 Mb/s ethernet, 100 base VGAnyLAN, full duplex ethernet, switched ethernet, full duplex switched ethernet, fiber distributed data interface (FDDI), and distributed queue dual bus (DQDB.) Cabling strategies include star-shaped intelligent hub wiring, distributed cabling interconnected through a high-speed backbone, and centralized tree-type cabling alternatives.

Discrete event **simulation** is the preferred design technique for internetworks involving varied LAN architectures and high-speed routers connected via WAN links. Here, performance prediction is the primary objective. Engineers specify new or proposed LAN configurations with building blocks of common LAN components. LAN simulation tools predict measures of utilization, queuing buffer capacity, delays, and response times. Proposed configurations of departmental LANs or enterprise internetworks can be modeled, evaluating the capabilities of competing designs involving hardware components, protocols, and architectures. This can be achieved without learning complex programming languages or possessing a depth of queuing theory background. Several excellent design tools have been developed and marketed in recent years for simulation-based modeling and analysis of internetworks.

Can your present configuration handle an increased workload or are hardware upgrades needed? What part of the internetwork hardware should be changed: The backbone? The server? The router or bridge? How reliable is the internetwork, and how can it recover from component failure? Predictive methods preclude the unnecessary expense of modifying a network only to discover it still does not satisfy user requirements. What-if analysis applied to network models helps evaluate topologies and plot performance against workloads. Performance and reliability tradeoffs can be effectively identified with simulation. This is essential in service level planning.

Simulation is beneficial throughout the life cycle of a communications system. It can occur before significant development, during the planning stage, as an ongoing part of development, through installation, or as a part of ongoing operations. Alternatives of different configurations are compared and accurate performance is predicted before implementing new network systems. Presizing networks based on current workloads and anticipated growth demands and pretesting application performance impact are just a few of the benefits of simulation design.

Simulation tools expedite investigating internetwork capacities and defining growth. Knowing whether circuits, servers, node processing, or networked applications are the bottlenecks in performance is useful to any engineer. Managing incremental growth and predicting the impact of additional networking demands are other advantages. Justifying configuration designs to upper level management is another big asset of internetwork simulation.

41.5 Internetwork Optimization

Optimization tools balance performance and cost tradeoffs. Acceptable network performance is essential, but due to the nonlinear nature of tariffs, cost calculations are critical. Tariff calculations are not trivial. Designs based on various routing, multiplexing, and bridging approaches can yield widely varying prices. Tariff data changes weekly, further complicating cost calculations. With new tariff filings, there may be new optimal designs. The best topology of circuit routes and circuit speeds can significantly differ from previously optimal designs when re-evaluating with new tariff data. All optimization tools evaluate cost calculations against performance requirements.

Figure 41.3 graphically illustrates this cost/performance relationship. The ideal cost/performance point is where traffic matches capacity, with acceptable zones around this ideal. Acceptable oversizing is due to economies of scale or forecasted growth, and undersizing is due to excessive rebuilding costs. Outside this acceptable zone, networks need redesign. Where capacity significantly exceeds traffic, the network feels the effects of overbuilding (performance goals satisfied at unnecessarily high costs). Conversely, where traffic is greater than capacity, the network has been underbuilt (cost goals satisfied yet at the expense of performance objectives) and rebuilding is necessary.

Network optimization techniques have become even more complex due to recent introductions of new service offerings, such as frame relay, integrated services digital network (ISDN), synchronous digital

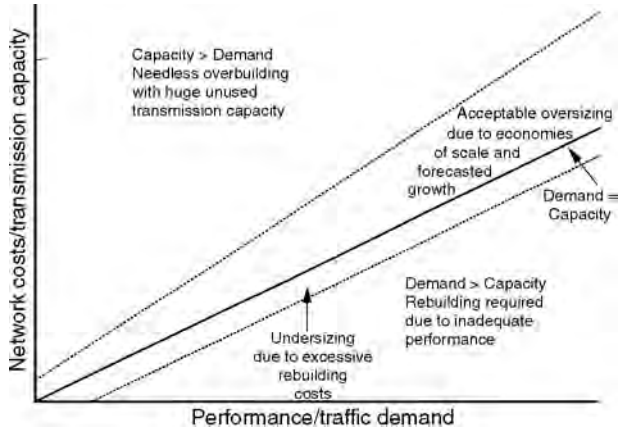


FIGURE 41.3 The cost/performance and capacity/demand relationship.

network (SONET), SMDS and asynchronous transfer mode (ATM), and hubless digital services. These new services are offered by the major interexchange carriers, specialized fiber carriers, and regional Bell operating companies (RBOCs). In addition to reducing costs, new services when designed properly can also increase network reliability.

The enormous variability in circuit costs is illustrated in pricing a single 56-kb/s circuit. An interexchange circuit between Atlanta and Kansas City varies based on service types (DDS and fractional T-1) and service providers (AT&T, MCI, and WTG) between a high of $2751/month and a low of $795/month. Similarly, an intraLATA circuit between San Jose and Campbell, California varies based on service types (DDS and new hubless digital) and service providers (AT&T, MCI, U.S. Sprint, and LEC-PacBell) between a high of $1339/month and a low of $136/month.

Furthermore, a variety of customer equipment-based multiplexing alternatives run the gamut from simple point-to-point multiplexing, to sophisticated dynamic bandwidth contention multiplexing, to multiplexer cascading, to colocating the CPE-based multiplexing equipment in the serving central office. Numerous networking options make the network engineer's job more complex, requiring optimization tools to evaluate these many new design techniques.

41.6 Summary

Information management today calls for effectively accommodating expansion of interconnected networks and accurately forecasting networking requirements. Only those who have an accurate awareness of their current communications requirements who can apply a structured design approach when evaluating competing networking techniques can confront this formidable challenge. James Jewett, cofounder of Telco Research, states, "Today, U.S. businesses frequently spend up to 40% more than is necessary on telecommunications services. Managers must be willing to scrutinize current communications networks and alternative networking techniques."

This is an age when communications budgets are being tightened, head counts are being trimmed, and major capital outlays are being delayed. Vendors selling LANs, internetworking gear, and long distance services feel cutbacks in flat revenues and intensely competitive bids. Network engineers feel the pinch as they try to meet increasingly demanding enduser performance objectives with dwindling budgets. Many network managers spend hundreds of thousands of dollars monthly on communications resources. These high costs reflect the importance of networks. Yet, with the high priority placed on LAN/WAN internetworking and the significant expenses devoted to maintain them, insufficient emphasis is placed on rigorous design and analysis. Network engineers tend to focus solely on operational capabilities. As long as the network is up and running, analysis or design is rarely conducted.

An engineering process of analysis and design for internetworks was presented. The process includes three interrelated phases: network awareness, network design, and network management. Each phase is part of a feedback control mechanism, progressing toward a network optimum. The industry emphasizes developing internetworking technologies; however, modeling techniques incorporating these technologies into operational networks have received little attention. Engineers applying this design and analysis process will be able to optimize their internetworks.

Defining Terms

Connectionless: A best-effort packet delivery model where the network does not guarantee delivery and data may be lost, duplicated, damaged, or delivered out of order; also called datagram services.

Connection-oriented: A guaranteed packet delivery model where a virtual circuit with minimal bandwidth is established before any data transfer occurs and packets are delivered sequentially over a single path.

Data link switching (DLSw): A routing standard for integrating SNA data into IP-based internetworks by stripping off the SDLC header and replacing it with a TCP/IP header.

Distance-vector routing: An early routing information protocol implementation common to IPX RIP, IGRP, IP RIP, and AppleTalk's RTMP.

Internetwork optimization: This class of network design tools balance performance, cost, and reliability tradeoffs. Optimization algorithms apply heuristic procedures to greatly abbreviate the optimization process by making a few well-chosen approximations. Queuing theory is also incorporated in the algorithms to calculate response times.

Internetwork simulation: Discrete event simulation tools predict network performance and identify bottlenecks. The nondeterministic nature of internetworking communications is reflected in models by pseudorandom numbers that generate interarrival times of messages, lengths of message, time between failures, and transmission errors.

IPng: The new version of the IP layer that supports 16-byte addresses and offers advanced features such as flow control, autoconfiguration, and encryption-based security.

Link state routing: A new breed of efficient routing information protocols designed for large, complex internetworking environments and implemented in OSPF, IS-to-IS, and Novell's NSLP.

Network awareness: The first component in the network design process, in which existing technology is reviewed, future technology trends are evaluated, current traffic is determined, equipment is inventoried, and future growth is forecasted. An organization's unique operational evaluation criteria are also chosen during this phase.

Network design: The second component of the network design process, which includes network design tools, performing a cost-performance breakeven analysis, and developing link configurations.

Network management: The third component of the network design process, which encompasses management of configuration, faults, performance, accounting, and security. Equipment and network software is acquired, verified, installed, maintained, and administered.

Packet burst: A sliding window implementation that allows transmitting multiple packets and receiving a multiple packet acknowledgment, greatly improving the efficiency of SPX/IPX internetworking.

Subnet mask: A four octet field used to identify the link portion of an IP address by setting bits to 1 where the network is to treat the corresponding bit in the IP address as part of the network address.

Total network engineering process: A process to engineer and optimize internetworks including the three phases of awareness, design, and management, with a feedback control mechanism, progressing toward a network optimum.

References

CCITT. 1979. Telephone operation: quality of service and tariffs. *Orange Book,* Vol. II.2, International Telegraph and Telephone Consultative Committee. Geneva, Switzerland.

Comer, D.E. 1991. *Internetworking with TCP/IP,* Vol. I, Prentice-Hall, Englewood Cliffs, NJ.

Comer, D.E. and Stevens, D.L. 1991. *Internetworking with TCP/IP,* Vol. II, Prentice-Hall, Englewood Cliffs, NJ.
Ellis, R.L. 1986. *Designing Data Networks,* Prentice-Hall, Englewood Cliffs, NJ.
Marney-Petix, V.C. 1992. *Mastering Internetworking,* Numedia, Freemont, CA.
Marney-Petix, V.C. 1993. *Mastering Advanced Internetworking,* Numedia, Freemont, CA.
Miller, M.A. 1991. *Internetworking: A Guide to Network Communications,* M&T Books, San Mateo, CA.
Perlman, R. 1992. *Interconnections. Bridges and Routers,* Addison–Wesley, Reading, MA.
Shilling, G.D. and Miller, P.C. 1991. Performance modeling. *DataPro Netwk. Mgt.,* NM20(400): 101–112.
Van Norman, H.J. 1992. *LAN/WAN Optimization Techniques,* Artech House, Boston, MA.

Further Information

The most complete reference is the three volume set, *Internetworking with TCP/IP*, by Douglas E. Comer and David L. Stevens. Volume one covers principles, protocols, architectures design, and implementation; volume two covers design, implementation, and interfaces; volume three covers client/server programming and applications.

An excellent self-paced learning series covering the basics in a highly readable style with insights addressing an often misunderstood arena of communications technology is *Mastering Internetworking and Mastering Advanced Internetworking*, by Victoria C. Marney-Petix.

The best technical treatment of the internal workings of bridges and routers is *Interconnections: Bridges and Routers*, by Randia Perlman. It not only presents the algorithms and protocols but points out deficiencies, compares competing approaches, and weighs engineering tradeoffs.

42

Architectural Framework for Asynchronous Transfer Mode Networks: Broadband Network Services

Gerald A. Marin
University of Central Florida

Raif O. Onvural
Orologic

42.1 Introduction

The main feature of the integrated services digital network (ISDN) concept is the support of a wide range of audio, video, and data applications in the same network. Broadband ISDN (B-ISDN) is based on ISDN concepts and has been evolving during the last decade by progressively incorporating directly into the network additional B-ISDN functions enabling new and advanced services. B-ISDN supports switched, semipermanent and permanent, and point-to-point and point-to-multipoint connections and provides on-demand reserved and permanent services. B-ISDN connections support both circuit-mode and packet-mode services of a mono- and multimedia type with connectionless or connection-oriented nature, in bidirectional and unidirectional configurations.

Since the completion of the initial standardization work on B-ISDN in the late 1980s, there has been a tremendous amount of standards work toward defining B-ISDN interfaces, physical layers, signalling, and transfer mode. Asynchronous transfer mode (ATM) is the transfer mode of choice for B-ISDN. ATM is a connection-oriented, packet switching technique. The ATM architecture and related standards define the basic framework on which different services with varying source characteristics and quality of service

(QoS) requirements can be supported in an integrated manner. However, enabling B-ISDNs requires the design and development of a complete architectural framework that complements the ATM standards and includes network services such as congestion-control and path-selection frameworks, directory services, and group management protocols.

In this chapter, we first review the B-ISDN architecture, the architectural drivers such as high-bandwidth application requirements, and the changes in the networking infrastructure that need to be taken into account to enable ATM networks. We then present a high-level overview of a comprehensive, high-speed multimedia architecture developed by IBM to address the challenges of enabling ATM networks.

42.2 Broadband Integrated Services Digital Network (B-ISDN) Framework

B-ISDN is envisioned as providing universal communications based on a set of standard interfaces and scaleability in both distance and speed. B-ISDN physical layer interfaces, including the ones currently being defined, scale from 1.5 to 622 Mb/s. We have already started to see ATM being deployed in the local area, while a number of network providers have started to experiment with ATM in the wide area, providing seamless integration from local to wide area networks. Furthermore, B-ISDN services conceptually include every application that we can think of as we go forward into the future:

- Applications with or without timing relationship among the end users
- Variable and continuous bit rate services
- Connectionless and connection-oriented services

Voice and video are two applications that require an end-to-end timing relationship whereas there is no such requirement for data services. Most applications generate traffic at varying rates. Voice traffic alternates between talk spurts and silent periods. The amount of data generated per frame in a video service varies, depending on the compression algorithm used and how much information from the previous frame(s) can be used to reconstruct a new frame at the destination station while providing the required QoS. Variable bit rate (VBR) service refers to a source traffic behavior in which the rate of traffic submitted to the network varies over time. Early video and audio services in early ATM networks are expected to be supported using continuous bit rate (CBR) traffic in which the traffic is submitted to the network at a constant rate. This is mainly because of the fact that the bandwidth scalability provided by ATM is a relatively new concept, and our previous experience in supporting such services is based on circuit switching, that is, PCM voice and H.222 video. As we understand traffic patterns and source characteristics better, VBR service is expected to replace CBR service for most real-time applications.

ATM is a connection-oriented service, that is, end-to-end connections must be established prior to data transmission. However, most network applications in current packet switching networks are developed on top of a connectionless service. In order to protect customer investment in current applications, ATM should support connectionless service. This is provided mainly by providing an overlay on top of ATM.

Enabling Asynchronous Transfer Mode Networks

The main question in enabling ATM networks is how to make it all work as an integrated network. The answer to this challenge is the ATM standards and the development of network services such as congestion control, bandwidth management, path selection, and multicast and group-management mechanisms that complement the ATM standards while achieving high-resource utilization in the network.

The ATM standards have been developed by the International Telecommunications Union–Telecommunications Standardization Sector (ITU-TS) and various national standardization organizations such as American National Standards Institute (ANSI) and ETSI. The ATM Forum, on the other hand, is a consortium of more than 500 companies established to speed up the development and deployment of ATM products through interoperability specifications based on standards whenever available. Hereafter, we do not differentiate between the standards and the ATM Forum's specifications.

User-network signalling Q.2931
ATM adaptation layer (AAL)
ATM layer
Physical layer

FIGURE 42.1 Current ATM framework.

Figure 42.1 illustrates the current ATM architecture. A variety of physical layer interfaces are defined, including different transmission mediums such as multimode and single-mode fiber, shielded twisted pair, and coax. Transmission speeds now vary from 45 Mb/s (DS-3) to 622 Mb/s synchronous optical network (SONET).

The ATM layer is based on the use of 53-byte fixed-size cells composed of a 5-byte header and a 48-byte payload. Of the 5-byte cell header, 28 b are used for routing. There is an 8-b header error check field to ensure the integrity of the cell header. This leaves only 4 b, of which 1 b is used for cell loss priority (high or low), and 3 b are used to identify the type of payload information and carry the explicit forward-congestion indicator. As discussed next, there is not really much time to do any processing in the network nodes as the transmission speeds increase. Some networking experts may have desired an extra byte or two to provide more features at the ATM layer. However, it would have been difficult and/or expensive to process this additional information as transmission speeds reach several gigabits per second.

Nevertheless, more features than the ATM cell header supports are required to meet application requirements. Considering a wide variety of services with different requirements, the ITU-TS grouped functions are common to most applications at the ATM adaptation layer (AAL). AAL type 1 is defined for CBR services that require an end-to-end timing relationship. The VBR counterpart of these services will be supported via AAL type 2, which is currently being defined. AAL type 3/4 is defined for the initial connection-oriented and connectionless services. AAL type 5 is a more efficient way of supporting data services. All AAL services are defined among end stations only; they are transparent to intermediate switches in the network.

ATM connections may be pre-established or may be established dynamically on request. Connections that are pre-established are defined by the network's management layer (not discussed here). Switched connections, on the other hand, require the definition of interfaces and signalling at these interfaces. Signalling at an ATM user–network interface (UNI) provides the capabilities needed to signal the network and to request a connection to a destination. Various capabilities included in UNI version 3.0 [ATM, 1994] can be summarized as follows:

- Establishment of point-to-point and point-to-multipoint virtual channel connections
- Three different private ATM address formats
- One public ATM address format
- Symmetric and asymmetric QoS connections with declared QoS class
- Symmetric and asymmetric bandwidth connections with declared bandwidth
- Transport of network transparent parameters among end nodes
- Support of error handling

An ATM user signals to request a connection from the network to a particular destination(s) with a specified QoS. Information elements such as peak cell rate, sustainable cell rate, cell delay tolerance, maximum burst size, AAL type, and address of destination end node(s) are included in the signalling messages exchanged across a UNI. Despite the tremendous effort that has gone into the defining these specifications, it will take a lot more work to enable high-bandwidth, multimedia applications in ATM networks. Various additional services are required:

- Locating destination end node(s): The addresses of end nodes are provided in a related signalling message. Determining the exact location of these end nodes, however, is not a part of the current standards.

- Determining how much network resources are needed to provide the required QoS: Guaranteeing the QoS needs of network applications requires, in general, that resources in the network be reserved. Although the traffic characteristics would be included in the connection request, a way of determining the commitment of resources needed to support different applications is not a part of today's ATM.
- Determining a path in the network that can support the connection: Path selection framework in ATM networks is quite different from path selection in current packet switching networks. While meeting the QoS requirements of applications, it is highly desirable to minimize resources used by the network in order to achieve economic resource utilization. This requires the solution of a constrained optimization problem unlike the shortest path problem posed by traditional packet networks.
- Establishing and managing connections in the network: Routing in ATM networks remains largely an unresolved issue. Routing metrics and the decision place (i.e., source routing) are the two areas that are being studied currently at the ATM Forum. However, the efficient distribution of control information in a subnetwork and the quick dissemination of connection establishment and termination flows are two areas that are not addressed in the ATM standards.

42.3 Architectural Drivers

Before discussing how some of these issues may be addressed, let us look at various application requirements and the changes in the networking infrastructure imposed by advances in the transmission technology.

Perhaps one of the main drivers for emerging ATM networks is the deployment of multimedia applications. Although initial ATM networks may be deployed as data-overlay networks that keep voice separate, it is unavoidable that voice, video, and data will come together as multimedia applications begin to emerge (i.e., videoconferencing, video distribution, distant learning). After all, multimedia networks supporting multimedia applications have been the main driving force behind the B-ISDN initiative.

Multimedia applications differ from monoapplications in various ways. They often take place among a group involving more than two users. This necessitates the existence of group-management and multicast services. Group-management functions would basically allow users to join and leave the group as they wish and would manage the connection among the group's members. Multicast trees provide a connection path among these users efficiently compared with point-to-point connections. Yet, the construction of a multicast tree that satisfies end-to-end delay constraints while achieving high-resource utilization is a hard problem with an exponentially growing time complexity. Addressing this problem in real time requires the development of efficient heuristics.

Multimedia applications also combine the different QoS requirements of individual applications into a single application. This requires the integration of the service demands of different applications and the definition of network services that can support such requirements. As an example, voice and lip motion need to be tightly synchronized in a video service, whereas the synchronization of subtitles (i.e., data) to video does not need to be as stringent.

ATM networks will support a wide variety of traffic with different traffic characteristics. Using the source parameters defined in the ATM Forum's UNI 3.0 specification, a user may pass its peak and sustainable cell rates to the network during the connection-establishment phase. The burstiness (defined in this context as the ratio of peak to average bit rates) of various applications envisioned in ATM networks varies from 1 (e.g., for CBR service) to 100 s (e.g., for distributed computing). As the burstiness of a source increases, the predictability of its traffic decreases. Accordingly, it becomes harder for the network to accommodate highly bursty sources while achieving high resource utilization. Another challenge is the integration of highly bursty sources in the network with less bursty sources, say, CBR services, while meeting the mean-delay and jitter requirements. Let us consider two applications at the opposite ends of the spectrum with respect to their burstiness, voice, and distributed computing. When a distributed computing source becomes active, it generates traffic at the rate of, say, 100Mb/s. An active period is

followed by a silent period during which the source processes data in its memory and does not generate any network traffic. The high burstiness of the source would mean that the periods of activity persist for considerably less time than the periods of inactivity. On the other hand, consider a PCM voice application generating frames at the constant rate of 64 Kb/s. The challenge is to provide network services such that when the two sources are active simultaneously on one or more physical links, the QoS provided to the voice source does not degrade due to a sudden spurt of cells belonging to the distributed computing application. The problem of dealing with applications with widely varying burstiness in an integrated manner cannot be avoided as we move toward VBR multimedia services; network services that meet this challenge are needed to deploy integrated services.

In addition to having to deal with varying source characteristics, ATM networks should support services with quite different QoS requirements. For example, most data applications require fully reliable service (i.e., a banking transaction). However, these services can use the functions provided by a higher layer protocol (i.e., a transport service) on top of an ATM connection that will recover lost data by retransmission. Voice applications can tolerate moderate cell losses without affecting the quality of the service. Interactive VBR video services have very stringent loss requirements since a lost cell could cause the loss of frame synchronization, which in turn would cause the service to degrade for seconds until the stream was refreshed. All of these applications with different requirements are integrated in the network. Network services designed to support these applications should take into consideration their QoS requirements explicitly if high resource utilization in the network is desired.

In addressing the challenges of supporting application requirements, it is necessary to consider the constraints inherent to high-speed networking. As transmission speed increases, two considerations start to play important roles in designing network services: in-flight data becomes huge, and available processing time at the intermediate nodes decreases.

With high-speed transmission links in the network, the amount of in-flight data increases dramatically as the link propagation delay increases. As an example, when the one-way propagation delay on a T1 link is 20 ms (i.e., across the U.S.), there can be up to 3750 bytes simultaneously on the link at any particular time. This number increases to 2.5 million bytes on a 1 Gb/s link. If hop-by-hop flow control were applied between the two ends of such a link, which works well in current packet networks with low-speed links, millions of bytes of data could be lost by the time a control packet traversed the network from one end to the other. This phenomenon dramatically changes the congestion-control framework in high-speed networks. Another aspect of this is the challenge of providing efficient and accurate distribution of control information in the network, that is, link failures, resource reservation levels, actual link utilization, etc.

Another key challenge high-speed links introduce is to keep the functions at intermediate network nodes simple while providing enough functionality for high-resource utilization. Consider a 9.6-kb link and a processor with one million instructions per second capability. Then, up to 44, 166 instructions can be executed for each cell while keeping the link full. When the transmission rate reaches gigabits per second, however, this number goes below one. Even with processors that run at 30 MIPS, the functionality at an intermediate node should be minimized to keep up with high-speed links. Most required functions need to be simple enough to be implemented in hardware. This, perhaps, is one of the main reasons for keeping the ATM cell header so simple.

Simple Solutions: At a Cost

Most of these challenges could be addressed rather easily if the network resource utilization were not taken into consideration. For example, all applications can be treated in the network with the most stringent loss and delay requirements. Though simple, this approach might prove to be quite costly, in that an artificially large amount of network resources might be reserved in the network to support applications that could tolerate moderate loss and/or delay.

As another example, a very small amount of buffering can be provided at each network node with deterministic multiplexing (i.e., peak rate bandwidth allocation). Although this scheme minimizes queueing

delays at each node and guarantees practically no loss, it causes resource utilization in the network to be unacceptably low. As an illustration, consider a source with peak and average bit rate of 10 Mb/s and 1 Mb/s, respectively. On a 100-Mb/s link, only 10 of these connections can be multiplexed, which results in an average link utilization of 10%.

Similarly, point-to-point connections can be used to provide communication among a group of users. This might, however, artificially increase the number of links used to provide a multipoint communication compared with a multicast tree and thereby reduce the availability of extra links for use in point-to-point connections to support traffic from other sources.

In summary, simple things can be done to enable ATM networks; the challenge, however, is to provide network services that will meet the application requirements economically.

42.4 Review of Broadband Network Services

This section presents a high-level review of IBM's broadband network services (BBNS) architecture. BBNS is a comprehensive approach to providing network services for high-speed, multimedia networking based on existing ATM standards and ATM Forum specifications. Conceptually, the BBNS architecture is broken down into the following three functional groups:

- *Access services* provide the framework for supporting standards-based and industry-pervasive interfaces and protocols.
- *Transport services* provide the pipes and the trees needed to switch voice, video, and data across the network in an integrated structure with appropriate QoS guarantees.
- *Control point services* manage network resources by providing the distributed brain power that enables all network services.

Access services can be thought of as locators and translators. They interface with the foreign protocols outside the BBNS network and translate them into what the network understands inside.

Transport services are the transfer modes defined in the architecture and related services developed to support different types of applications with varying service requirements. In addition to ATM, the BBNS architecture supports variable length packets to provide a migration path from the current state of the art in networking today to an all-ATM network. Other services developed include the datagram and multicast services and mechanisms to support different delay priorities.

Control point services mainly provide congestion control, path selection, group management for multimedia applications, directory services, and distribution of network control information.

42.5 How Does It All Fit Together?

Before proceeding with the details of the BBNS architecture, we discuss how the various functions required to enable high-speed networking relate to each other. This discussion uses can setup as a vehicle.

A new connection request is first received by an access agent at the edge of the network. The agent runs an address resolution mechanism to find the destination end node. Then, using the source traffic characteristics and application requirements, the amount of bandwidth required by the application is determined. Bandwidth calculation takes into consideration the cell loss requirement of the application, uses the source traffic characteristics, and determines a guaranteed upper bound on the bandwidth required to support the application when this connection is multiplexed with the connections already established in the network. Then, the path-selection algorithm is run to determine a path from the origin node to the destination node(s). In determining this path, the end-to-end delay requirement of the application is taken into consideration explicitly. In particular, the algorithm finds a minimum-hop path in the network that can support the bandwidth end-to-end requirements of the application.

Once a path is found, an end-to-end connection is established using internal signalling. The traffic can start flowing after every node along the path accepts the connection establishment request. In this

framework, source routing is used; paths are determined by the nodes the connections originate. This approach requires that each node that supports an access service knows the topology of the network (which is subject to change due to link and node failures) and the utilization and current reservation levels of network resources. Although the topology of the network does not change that often, utilization levels change frequently. This necessitates an efficient means of distributing network control information that maximizes the amount of information available to each node while minimizing the amount of network control traffic overhead. These two objectives are contradictory, and the design and development of this feature is a complex task.

At this stage, let us consider the network operating with traffic flowing through a large number of connections. At the edge of the network, it is necessary to monitor the traffic generated by each source to ensure that each source stays within the parameters negotiated at the call establishment phase. Otherwise, nonconforming sources would cause the QoS provided to conforming sources to degrade. A leaky bucket mechanism is used for this purpose. The main challenge in using a leaky bucket is to determine its parameters such that its operation would be transparent to conforming sources while ensuring that nonconforming traffic is detected.

Let us now focus on the nonconforming traffic. There are various reasons that a source might not stay within the parameters negotiated at the call setup time: the source might not be able to characterize its traffic behavior accurately, there might be an equipment malfunction, or the source might simply be cheating. There are two mechanisms developed in the BBNS architecture to provide some amount of forgiveness to nonconforming sources. The first one is the adaptation function which filters the traffic and estimates the amount of bandwidth required by connections based on the actual traffic generated by each source. The second mechanism is the use of so called red/green marking. In this mechanism, conforming traffic is marked green before it is transmitted. Red marking allows (some) nonconforming traffic to enter the network to achieve higher utilization of network resources. Doing so, however, requires the development of mechanisms so that the services provided to conforming traffic in the traffic are not affected by the non-conforming traffic when a node becomes congested.

Another feature required in the network is to minimize the negative effects of resource failures to the service provided to applications. A closely associated function is that of providing support for different priority connections. Nondisruptive path switching is used to reroute connections established on a link (node) around failures in a way that the impact of the failure to the service provided to these affected connections is minimal. Similarly, different connections may have different priorities. Depending on the availability of resources in the network, it may be necessary to take down some low-priority connections in order to accommodate a high-priority connection in the network. The main challenge here is to minimize the cascading effect that may occur when a connection that was taken down tries to re-establish a new connection.

A multimedia network should support connections with different delay requirements. The architecture classifies the multimedia traffic into three categories: real-time, nonreal-time, and nonreserved. Real-time traffic has the stringent delay and loss requirements and requires bandwidth reservation to guarantee its QoS metrics. Nonreal-time traffic also requires bandwidth reservation but only for the loss guarantee, that is, this type of traffic can tolerate moderate delay. Nonreserved traffic, on the other hand, can tolerate both loss and delay. Accordingly, there is no bandwidth reserved for this type of traffic.

42.6 Broadband Network Services

Although the availability of high-speed links in the network infrastructure is important, what really matters is what customers are able to do with them. Two important steps need to be taken to realize fully the promise of ATM. They are:

- The development of high-bandwidth applications
- The development of networking services based on the evolving standards

So far, we have discussed various application requirements and types of services that are required to enable high-bandwidth applications in ATM networks. In this section, we will present a more detailed view of IBM's broadband network services.

Current predominant networking technologies are designed to emphasize service for one or two of these QoS classes, for example, telex (delay and loss tolerant); telephony (delay intolerant); TV and community antenna television (CATV) (delay intolerant); and packet-switched networks like X.25, frame relay, APPN, transport control protocol/Internet protocol (TCP/IP), and SNA (delay tolerant). The promise of ATM is that networks can be built on a single technology to accommodate all QoS classes. BBNS is a comprehensive approach to providing network services for high-speed, multimedia networks based on the existing ATM standards and ATM Forum implementation agreements. Various services developed in this architecture are presented next.

Access Services

While providing users with a high-speed, multimedia networking infrastructure for emerging high-bandwidth applications, it is necessary to support existing user applications that have been developed based on multiple protocols. Broadband network services achieve this through access services that support multiple protocols including frame relay, ATM, voice, IP, and HDLC data, to name a few. External nodes connected to BBNS do not require any special knowledge of any BBNS features, that is, these stations communicate with BBNS access services using their native protocol.

Functions performed by access services include the following:

- Understanding and interpreting the external service or protocol
- Performing an address resolution function to locate the target resource
- Maintaining and taking down connections across the network in response (for example) to connection requests received at the ATM UNI
- Selecting the switching mode and network connection parameters that are optimal for the external protocol and the service requested
- Mapping the QoS parameters of the external protocol the parameters used by the architecture to ensure the availability of sufficient network resources to guarantee that the QoS requirements are met
- Ensuring fairness among users by calling on BBNS bandwidth management services.

Transport Services

BBNS transport services provide a transport facility across the network for user traffic generated at the edges of the network and accommodated by access services. These include transmission scheduling, hardware-based switching with multicast capability, and provision of a switching mode that meets the requirements of external protocols efficiently. In particular, the architecture supports the following switching modes:

- ATM switching
- Automatic network routing, which provides source-routing capability
- Label swap (with tree support)
- Remote access to label-swap tree

The architecture supports ATM natively. Depending on product implementation and customer choice, a network built based on this architecture can support network trunks that carry either ATM cells only or a mix of ATM cells and variable length packets, with the former being transported in standard ATM cell format. Just as ATM cells provide native transport for the ATM UNI, BBNS offers support for other interfaces such as HDLC, frame relay, and IP. This feature allows customers to introduce early ATM applications using ATM transport without any changes or performance penalty in supporting their

existing applications. However, any customer committed to ATM-only networking is fully supported through the BBNS architecture.

To provide service appropriate to the QoS classes discussed previously, BBNS supports two types of traffic priorities: delay and loss. Three different delay priorities are defined: real-time reserved bandwidth, nonreal-time reserved bandwidth, and best-effort service. Logically, each delay class is routed through a corresponding queue in each node. Similarly, four different loss priorities are defined for real-time traffic, and two different loss priorities are defined for nonreal-time traffic. Combined with the power of BBNS bandwidth management, these delay and loss priorities offer the flexibility to support a broad range of QoS requirements and, thus, to handle efficiently the emerging multimedia applications.

Control Point Services

BBNS control point services control, allocate, and manage network resources. Among other things, these services include bandwidth reservation, route computation, topology updates, address resolution, and the group management support needed for multipoint connections. Some of these services are briefly described next.

Set Management

Set management provides the mechanism to group user resources for the purpose of defining multiple logical (i.e., virtual private) networks. This is a critical requirement for service-provider networks, where resources belonging to one customer may have to be isolated from those of other customers. The same mechanism can also be used to define multiple virtual local area networks (LANs) on a network.

Set management also gives a network the ability to learn about new resources and users automatically, thereby reducing system definition and increasing productivity. In essence, set management allows a network function or a user port to be defined only once to make it available to the entire network. Furthermore, it provides full flexibility and mobility in naming/addressing. A user can be moved from one point to another without being forced to change his external address (phone number, for example).

Bandwidth Management

BBNS bandwidth management, congestion control, and path selection enable a network to provide quality of service guarantees on measures such as end-to-end delay, delay jitter, and packet (or cell) loss ratio, while significantly reducing recurring bandwidth cost. This is achieved through efficient bandwidth reservation, load balancing, and dynamic bandwidth allocation.

For each reserved bandwidth connection requiring QoS guarantees, the architecture reserves a sufficient amount of bandwidth on the links traversed by the connection. The BBNS approach, which depends on the connection's traffic characteristics and QoS requirements, computes sufficient bandwidth to satisfy the QoS requirements, taking into account the statistical multiplexing at the links. BBNS path selection finds a path from source to destination that satisfies the connection's QoS requirements using algorithms that minimize the number of hops in the path and balance the network traffic load.

The architectural approach supports both reserved and nonreserved bandwidth transport of user frames or cells. Note that no QoS support is provided for nonreserved traffic. When reserved bandwidth connections are used to support bursty connections (like traditional data traffic), the challenge of setting network connection parameters can be significant. When a user has a good understanding of his traffic characteristics, the user may be unable to signal those characteristics to the network until the workstation implements a new interface, such as the ATM UNI. Even when it is possible to set traffic parameters correctly initially, the bandwidth requirements may change dramatically on the order of seconds to minutes. To mitigate these problems, the architecture offers bandwidth adaptation (dynamic bandwidth allocation) as an optional added value. First, BBNS bandwidth management will enforce an existing traffic contract on a cell time scale of milliseconds or lower. Then, at user option, bandwidth management will learn the actual bandwidth requirements of a connection and adapt by rereserving (and possibly rerouting) the connection using a longer time scale on the order of the roundtrip delay (seconds).

Efficient Distribution of Control Information

BBNS hardware-based multicast switching capability allows fast and efficient distribution of network control messages. Messages such as topology and link-state updates are multicast on the control-point (CP) spanning tree (a multipoint connection, or tree, that extends to every node in the network) at close to propagation speed. Furthermore, since these update messages are carried only by links on the CP spanning tree (n - 1 links in an n-node network), the bandwidth needed by the control traffic is greatly reduced. The BBNS CP spanning tree algorithm maintains the integrity of the spanning tree. The algorithm builds the tree and automatically repairs it in case of partitioning caused by link or node failures. The topology algorithm provides reliable delivery of multicast messages. Control messages such as connection setup requests and connection maintenance messages also benefit significantly from the hardware multicast capability. These messages are multicast linearly to every intermediate node along a connection's path, thus allowing connection setups and takedowns much faster than through hop-by-hop signalling methods.

Nondisruptive Path Switching (NDPS)

Nondisruptive path switching (NDPS) makes it possible for a network to reroute network connections automatically in the face of node and/or link failures while minimizing user disruption. Although such a feature is common to other networking architectures, the BBNS version is extremely efficient. First, changes to the network planned or unplanned, are nearly instantly known throughout the network because of the CP spanning tree. Second, all rerouting decisions are made in parallel in a distributed fashion, rather than sequentially at a central control point. This allows for rapid network stabilization.

Call Pre-emption

Call pre-emption enables a network to provide different levels of network availability to users with different relative priorities. When required, call pre-emption will reroute existing connections to accommodate higher priority connections. BBNS call pre-emption strategy intelligently selects connections to be pre-empted, minimizing the possible rippling effects. NDPS is used to reroute the pre-empted connections; thus, disruption is kept to a minimum. The architecture gives the network operator the flexibility of assigning a wide range of priorities to network users.

42.7 Conclusions

There are various issues that need to be satisfactorily addressed in order for ATM networks to become a reality. In this chapter, we reviewed some of these issues based on application requirements and changing networking infrastructure. Then, we presented a high-level review of IBM's architecture for high-speed, multimedia networking that addresses some of these challenges while achieving high resource utilization in the network. ATM, related standards, and IBM's broadband network services architecture are evolving over time. Despite the tremendous amount of work completed on the standards and various networking issues, and the fact that vendors have been working on the design and development of an architectural framework to enable ATM networks, there still remain several issues that need to be satisfactorily resolved. This includes the design and development of applications that can take advantage of high bandwidth in customer premises and public networks as well as network services that can meet the application requirements while maximizing the usage of network resources. ATM is positioned as the transport method with the greatest potential to provide unified voice, video, and data service for high-speed networks supporting multimedia applications. Much work has been done in defining the B-ISDN standards in which ATM transport is featured. Both international standards bodies and the ATM Forum continue the prolific work to enhance the standards and implementers' agreements that are fundamental to realizing customer expectations. Yet, various vendors are expected to provide value-added services that complement the ATM standards, thereby distinguishing their products from others. Broadband network services

are IBM's architecture for value addition to various ATM networking that includes:

- A set of transport services that allow efficient integration of traffic generated by different sources with varying delay and loss requirements
- Bandwidth management framework that provides quality of service guarantees while achieving high-resource utilization in the network
- Efficient distribution of control information in the network minimizing the resources used for network control overhead while updating control information at the network nodes (at propagation delay speeds)
- Set management techniques that define and maintain logical groups and allow the network to learn about new resources and users automatically
- Nondisruptive path switching that minimizes disruption to users in case of network resource failures
- Call pre-emption that allows priority handling of connections and minimizes possible cascading of connection takedowns
- A path selection framework that addresses the requirements of different types of applications with varying delay and loss demands while achieving high utilization of network resources (this framework includes both point-to-point and point-to-multipoint path construction)
- The flexibility to help users evolve their existing applications and networks to ATM transport by matching the switching mode provided by the network to the native transport mode of existing protocols and applications.

References

ATM. 1994. User-network specification version 3.0, ATM Forum.

ITU-TS Recommendations, International Telecommunications Union-Telecommunications Standards Sector.

Onvural, R.O. 1994. *ATM: Performance Issues,* Artech House, Norwood, MA.

Prycker, Q. *Asynchronous Transfer Mode.*

IBM. *Broadband Network Services,* IBM Blue Book, Research Triangle Park, NC.

43

Control and Management in Next Generation Networks: Challenges and Opportunities

Abdi R. Modarressi
BellSouth Science and Technology

Seshadri Mohan
Comverse Network Systems

This chapter discusses the challenges and opportunities associated with a fundamental transformation of current networks towards a multiservice ubiquitous infrastructure with a unified control and management architecture. After articulating the major driving forces behind network evolution, we outline the fundamental reasons why neither the service control infrastructure of the public switched telephone network (PSTN), nor that of the present-day Internet, are adequate to support the myriad of new application services in next generation networks. Although NGN will inherit heavily from both the Internet and the PSTN, its fundamental service control and management architecture is likely to be radically different from both and will be anchored on a clean separation between a QoS-enabled connectivity network domain and a component-oriented application service infrastructure domain, with a distributed processing environment that glues things together by providing universal support for distribution, redundancy, and communication between distributed components. Finally, we allude to the transition issues and show how voice-over-packet services are emerging as the bootstrap application for marshaling in the NGN architecture.

43.1 Introduction

The future vision of "information and communication anytime, anywhere, and in any form" is starting to come into focus as major players begin to position themselves for a fundamental transformation of their connectivity network and application service infrastructures (ASI). It has become increasingly

clear that a prerequisite for realization of such a vision is the convergence of the current multiple networks—each employing different transport and control technologies—into a unified, multiservice, data-centric network offering services at different qualities and costs on open service platforms. The evolution towards such a vision is undeniably influenced by the growing trends of deregulation in the telecommunication industry on the one hand, and the rapidly evolving technological convergence between distributed computing and communication (singularly embodied in the Internet and its associated technologies) on the other hand. In a real sense, the necessary technological and environmental underpinnings are coming together for next generation service providers to begin the process of transforming their infrastructures in ways that will enable provision of many new innovative services rapidly, economically, on a mass scale, and with the required quality of service. System, hardware, and especially software vendors will also need to revamp and retune their production processes to meet the challenges of convergence in next generation networks.

At a high level, the fundamental driving forces for next generation networks can be categorized as follows.

Environmental drivers reflect changes that have been happening in the telecommunication business environment in the course of the past decade and a half. In this period, the global telecommunication industry as a whole has been gradually moving away from the model of state-owned and/or regulated monopolies into the model of a competitive industry operating in an open market. The process started with the breakup of AT&T in 1984, picked up speed with privatization of state-owned enterprises in the late 1980s and early 1990s, was greatly accelerated by the US Telecom act of 1996, and has picked up momentum by the progressive opening up of global telecommunication markets in the late 1990s. The rising tide of deregulation and privatization and the emerging competitive landscape, hand-in-hand with market and technology drivers (to be mentioned later), have caused unprecedented regrouping and consolidation of service providers and, in parallel, of system and equipment vendors that serve them.

Market/business drivers on the one hand reflect the continuously expanding set of capabilities and features that various markets demand to satisfy an expanding set of personal and professional needs, either as end users of services (consumers), or as intermediaries (wholesalers) who enhance the acquired services and offer them to their customers. For example, mobility of various kinds (user, access device, or application) has become a paramount requirement of services in different markets. On the other hand, the never-ending market pressure to minimize costs leads to transformational changes in the core of the network that may not even be visible to customers. Transformation of the TDM core of the circuit switched network into packet is an example of a business-driven initiative to minimize cost. Other market-driven needs include ready access to information, easy-to-use communication capabilities involving more than one medium (of both real-time and non-real-time vintage), greater and more granular end-user control over services, and progressively higher-quality content delivered for purposes of entertainment and/or education. Market drivers, even as other factors modulate them, are unquestionably the decisive factors for architecture evolution because **services are what customers use and pay for** and service providers constantly strive to reduce the cost of providing such services.

Technology drivers include all the technological enablers that a service provider, in partnership with its vendors, can take advantage of in the course of architecting and composing its repertoire of services. In today's information society, technology drivers have a lot to do with shaping customer expectations, thereby modulating market and environmental drivers. It is interesting to recall how common it was not too long ago to relegate technology drivers to the backseat and position business considerations and short-term customer needs as the only driver for network evolution. However, the spectacular rise of the Internet on the heels of the convergence of distributed computing and communication technologies has underlined the critical importance of technology drivers in elevating and reshaping customer expectations and in pushing out the restrictive envelope of regulatory constraints. The Internet, with the technologies it has spawned, is perhaps the single most prominent technology driver of our time.

These drivers, operating against the backdrop of some mega trends—including the growing diversity and complexity of new application services, the increasing variety and power of end user devices, and

the competitive push to minimize time-to-market—unmistakably underline the urgency of fundamental transformation in communication network and application service infrastructures. This chapter attempts to draw the broad outline and requirements of a target next generation architecture, which can serve the role of the distant "lighthouse" that guides the overall process of architecture evolution over the long haul. Any target NGN architecture would necessarily have to abstract away the current product limitations to provide a somewhat idealized, top-down perspective of next-generation networks. Throughout the discussion, it will be assumed without further elaboration that the architectural principles of scalability, reliability, evolvability, controllability, and efficiency will be applied to any concrete intermediate instantiation of the architecture.

Before outlining the fundamental requirements and salient features of service control and management in next generation networks, it may be appropriate to sketch an outline of signaling and control in modern PSTN and elaborate on its strengths and weaknesses in the context of evolution towards next generation networks.

43.2 Signaling and Control in PSTN

Since the advent of digital networks a quarter of a century ago, control of the TDM-based public switched telephone network (PSTN) has by and large been delegated to a physically separate signaling network with its own distinct transport infrastructure and protocol suite. The initial signaling network was the Common Channel Interoffice Signaling 6 (CCIS6) network deployed by AT&T in the mid-1970s. This network eventually evolved to the Signaling System No. 7 (SS7) network about a decade later, which is the current most widely deployed signaling network in the world [1]. The SS7 network provides a real-time highly reliable means for switches to communicate and exchange call/connection/service control information. PSTN switch software and SS7 call control protocols (e.g., ISUP) also typically provide support for a number of supplementary services beyond simple call setup. Call forwarding, call waiting, and calling number identification are some of the better-known switch-based services of this type.

To analyze a simplified call setup (involving only two local switches), the central office (CO) switch can be decomposed into four functional modules (Fig. 43.1). When a telephone connected to the originating CO switch goes off-hook, the line module is responsible for detecting the off-hook event, providing a dial tone to the telephone over the analog local loop, and collecting the dialed digits such as the called party's telephone number. The line module then hands over the dialed digits for processing to the call processing module. After analyzing the collected digits and determining the appropriate route to complete the call, the call processing module asks the signaling module to initiate appropriate

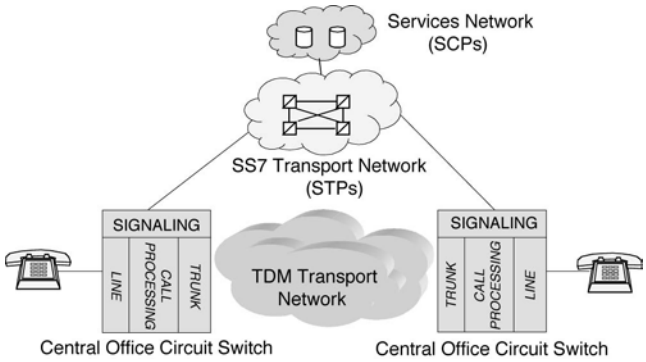


FIGURE 43.1 The fundamental structure of signaling and control in modern PSTN.

signaling over the SS7 network. The SS7 network carries the call setup message to the CO switch serving the called party and initiates establishment of a bearer channel between the originating and terminating switches over the TDM-based transport network. The line module of the terminating switch applies a ringing tone to the called party's telephone. When the called party answers, the call is established between the two parties. The trunk module converts analog speech to digital and appropriately formats it for transmission over the TDM trunks. At the terminating CO switch, the reverse process takes place. When one of the telephones goes on-hook, the line module serving that telephone detects the on-hook event and sets off appropriate signaling through the SS7 network to release the resources allocated for the call.

The complexity of the monolithic switch software, coupled with the inherent lack of flexibility and control over its capabilities by the service providers, led to the inception and deployment of intelligent networks (IN) in the early 1980s and advanced intelligent networks (AIN) [2] about a decade later. Here, some of the call processing logic and data were moved to separate centralized network nodes called service control points (SCPs). SCPs were also connected to the highly reliable SS7 transport network of signal transfer points (STP) just like switches (Fig. 43.1). The signaling network then evolved to support a client–server protocol for communication between the switches and SCPs in support of IN/AIN services. When the call processing module of the originating switch detects a "trigger" (e.g., a free-phone dialed number), it passes control of the call to the SCP through a series of signaling messages sent over the SS7 network. The SCP continues processing the call and instructs the switch, for example, by providing it a routable number, on how to proceed completing the service. The connection setup phase of the call would proceed in the same manner described above for a simple call [2]. Note that Fig. 43.1 does not show other resources that may be attached to the switch, such as intelligent peripherals to provide additional services like announcements and tone detection for enhanced services. Nor does it show any tandem switches that may be involved in the call.

IN/AIN effected a partial decoupling between call/connection control and enhanced service control through the use of SCPs, greatly reducing the time-to-market for new SCP-hosted services and features. This enriched the service repertoire of the telephone companies, a process that is still going strong albeit at a slower pace. As some services are switch-based while others are provided by IN/AIN, and because even IN/AIN services are dependent on triggers embedded in switch software, the decoupling of service control and call/connection control in PSTN/IN is only partial, with substantial dependence still remaining on switch vendors for service provisioning.

Some of the attributes and characteristics of the signaling and control infrastructure in modern PSTN that are particularly relevant to architecture evolution towards NGN are as follows:

- The fundamental assumption about the underlying transport network is that it is a TDM-based network supported by (local and tandem) circuit switches and used almost exclusively for voice and/or voice-band communication.
- The signaling and control infrastructure of PSTN and its associated protocols are optimized to provide voice-centric services.
- The control infrastructure of PSTN is proprietary and by and large closed to third-party service providers. This, among other things, implies a fairly high degree of operational complexity when one network or service provider has to interface to another provider. It also means that the knowledge required for service creation in this environment is narrowly specialized and relatively hard to find or nurture, further contributing to delays in service rollout.
- The customer premise equipment (CPE) used for PSTN access has very limited capabilities and is predominantly analog in nature (the "dumb phone"). Even when this access device is digital, as in the case of an ISDN terminal, its capabilities are still quite restrictive and limited.
- Control and management in PSTN are totally separate domains with different requirements. PSTN signaling and control requirements are much more real-time sensitive, and call for far greater reliability and robustness, than PSTN management (operation, administration, and maintenance, or OA&M) requirements.

43.3 General Attributes and Requirements of NGN

Consider some of the attributes and characteristics of NGN and the services that it needs to support. Following the same thread as the one at the end of the previous section, one can note the following:

- The underlying transport network in NGN is unquestionably packet based. This packet infrastructure is expected to support a large variety of new QoS-enabled (broadband and narrowband) application services involving an arbitrary combination of voice, video, and data. This is in sharp contrast to the predominantly narrowband, single QoS, voice-centric services of PSTN and the best effort, non-QoS-enabled, data services of the Internet.
- The control infrastructure of NGN has to be optimized to support a large variety of broadband multimedia applications on a QoS-enabled packet services. Nonetheless, it will still have to exhibit a high degree of reliability and robustness, just as signaling does in PSTN.
- The application service infrastructure of NGN has to be open in order to allow third-party application service providers to easily integrate their services into it. This is simply recognition of the fact that, as the Internet has proven, there is no way a single service provider can create all or even the majority of the application services that its customers need. This openness also makes the whole architecture more easily evolvable, with a much larger pool of technical staff who can quickly become productive in creating new application services and features.
- Given the nature of services supported by NGN, the CPE involved in service delivery can become far more sophisticated than the low functionality phone in PSTN. PCs of various kinds will probably constitute the dominant kind of powerful CPE in next generation networks. NGN, however, will need to support a wide range of other wired and wireless devices ranging from screen phones to sophisticated multimedia workstations and media centers. Sophisticated CPE gear will not only facilitate end-to-end multimedia communication and information retrieval, but will provide a hitherto unprecedented computing platform for local execution of logic and storage of data in concert with network-based servers. This, in turn, will have enormous implications for the service control infrastructure of NGN.
- Finally, the demarcation between control and management is already beginning to fade and will completely disappear in NGN. Application service, and to a large extent network, management functions are beginning to require real-time performance, thanks to the rising expectations brought about by the Internet. Also, because of an increasing level of reliance by customers on their application (and network) services, the reliability of the associated management functions needs to converge to the level of reliability of real-time control functions. In short, management and control functions in NGN will not be distinguishable any longer and need to be supported by the same architecture.

As we highlight some of the fundamental differences between PSTN and NGN, it may be instructive to also highlight the major differences between the Internet and NGN. While the Internet and NGN are both packet-based and support sophisticated CPE for service delivery, two fundamental differences exist between them. One is that the current Internet by and large provides services only on a best-effort basis. This means that neither reliability nor quality of the delivered service can be guaranteed. Such will clearly not be the case for NGN, where QoS and reliability of services (both network services and application services) will have to be engineered and guaranteed (and duly paid for). The other major difference is that the service architecture in the Internet is ad-hoc: application servers are at the edges of the network and each application service is architected in its own way without fitting into a unified service infrastructure. NGN, on the other hand, needs to have a fairly sophisticated connectivity network architecture supporting a unified application service infrastructure that optimizes the distribution of intelligence between network servers and end systems and ties into a powerful service creation environment. In the absence of such a unified application service architecture, it is highly doubtful that the Internet, even with enhanced capabilities to address QoS, can become the single ubiquitous infrastructure for delivery of sophisticated multimedia multi-party services of the future. Nonetheless, there is no doubt that NGN

TABLE 43.1 Some of the Attributes of PSTN/IN, Current Internet, and NGN

Parameter	PSTN/IN	Internet	NGN
Multimedia services	No	Yes	Yes
QoS-enabled	Yes (voice)	No	Yes
Network intelligence	Yes	No	Yes
Intelligent CPE	No	Yes	Yes
Underlying transport network	TDM	Packet	Packet
Service architecture	Semi-distinct	Ad-hoc	Distinct
Integrated control and management	No	Yes	Yes
Service reliability	High	Low	High
Service creation	Complex	Ad-hoc	Systematic
Ease of use of services	Medium	High	High
Evolvability/modularity	Low	Medium	High
Time to market of services	Long	Short	Short
Architecture openness	Low	High	High

will inherit far more from the Internet paradigm than from the PSTN paradigm. Some of the similarities and differences between NGN, PSTN/IN, and the current Internet are highlighted in Table 43.1.

43.4 A Broad Outline of the NGN Architecture

As implied by Table 43.1, next generation networks will have both similarities and differences with PSTN/IN, as well as with the Internet. The fact that the connectivity network in NGN is going to be packet-based opens up vast possibilities in terms of new application services, services that could not even be imagined in a TDM environment. Application services, and service interactions, will be much more diverse and complex, requiring a far more sophisticated system of control and management. In the interest of reducing time to market, as well as management of complexity through separation of concerns, NGN will require a clean separation of functions and domains with maximum degree of reuse built into the architecture and its components. Three distinct, separable, and interacting domains will exist in NGN with clearly defined boundaries. These are the **connectivity network** domain, the **application services** domain, and the **application services infrastructure (ASI)** domain. Although these domains functionally exist in PSTN/AIN as well, their capabilities are quite limited and boundaries between them are hard to distinguish and maintain.

These three domains are pictorially shown in Fig. 43.2. The two types of customers (residences and businesses) are each depicted to have their own internal networks and devices with gateways connecting them to the service provider's infrastructure. The mobile users shown in the figure by cars can be either business or residential customers. Business networks could also contain servers, which can be thought of as third-party servers. Note that all four types of relationships (residence–residence, residence–business, business–residence, and business–business) have to be supported, with the service provider directly or more likely indirectly through third parties, enabling rollout of all kinds of broadband application services to its customers. This enablement on the part of the next generation service provider consists primarily of providing and operating a highly reliable and large-scale ASI, and to a lesser extent and more selectively a connectivity network infrastructure. Let us now consider in more detail what constitutes the three domains.

Connectivity Network Domain

This domain provides connectivity requested by other domains using an underlying packet-based transport infrastructure. Such connectivity is characterized by communication associations that meet the requirements of QoS, scalability, traffic engineering, and reliability. These associations can be based on either connection-oriented or connectionless network services, although connectionless network services

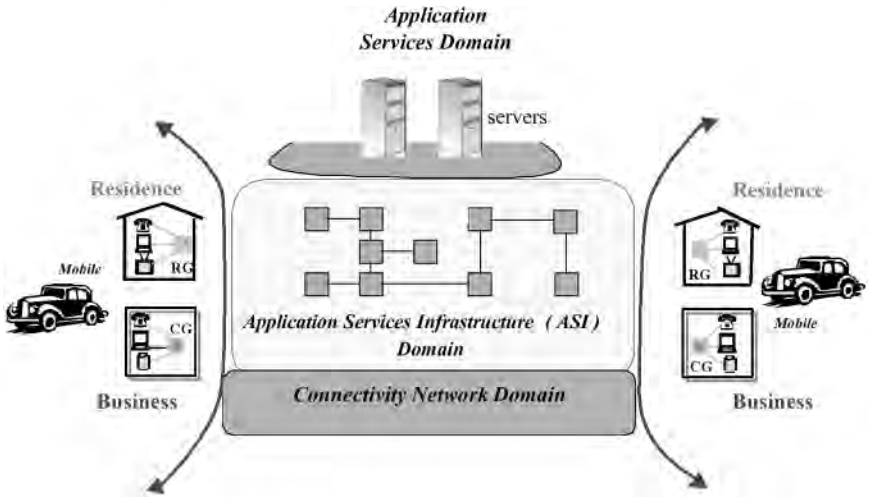


FIGURE 43.2 A high-level view of various domains in next-generation networks.

using IP will become dominant. The connectivity network domain would provide all functions generally attributed to the lower four layers of the OSI model.

A communications session in the ASI domain maps to a **connectivity session** in the connectivity network domain. ASI can initiate two types of connectivity sessions. A **default connectivity session** gives a customer access to the network. Activities such as web browsing, notification, and sending and receiving email can occur in such a session, and do not require ASI to establish new sessions specific to such activities. Only limited, provisioned QoS is available for such activities (which could in most cases use a best-effort service). An **enhanced connectivity session** would go beyond the default connectivity session to support some combination of QoS guarantees, policy, and billing. It is within an enhanced connectivity session, mapped to a communication session, that ASI can exert control over establishing and managing connections and/or associations.

Application Services Domain

In broad terms and for the purposes of this chapter, an **application service** is defined as a software application that provides a useful and well-defined functionality to the user. An application service is realized by a set of interacting software components distributed across multiple computing elements.

The user of the application service can be any human or machine entity. It may be useful to classify application services into a few major categories to gain insight into architectural issues that impact ASI requirements in next generation networks.

The first, and perhaps most prominent, class of application services offered by the NGN will be interactive *communication services.* These include real-time communication services involving multiple parties interacting in real-time and using multiple media (voice, video, and data). Multiple qualities or grades of these services will need to be supported. Services of this kind will constitute the evolution of today's voice telephony into next-generation, multimedia, multiparty, communication. Because of their real-time performance and dynamic control requirements, these services are likely to become the most complex set of services that NGN will offer to its customers. Communication services also include non-real-time services involving multiple parties and multiple media. These will constitute the evolution of today's messaging (predominantly e-mail and voice mail) into a unified set of store-and-forward multimedia services that have elaborate terminal adaptability, control, and media conversion capabilities.

The second major class of application services can be broadly labeled *information/data services*. These services may be thought of as the evolution of today's Internet services: browsing, information retrieval,

on-line directories, scheduled information delivery, e-commerce, advertising, productivity and other network computing services that the Internet can ubiquitously provide today (excluding e-mail, which we classified as a communication service). They also include a rich set of remote access and control services such as telemetry, monitoring, security, network management, and other data-oriented services. The evolution of this class of services will not only include their rudimentary versions that exist today, but major improvements to enhance response time and reliability and provide sophisticated billing, quality of service, and policy enforcement options. What will need to be preserved, however, and indeed extended to all other NGN services, is the ease of use inherent in today's Internet paradigm.

The third class of services that a next generation service provider will inevitably need to offer and/or enable is delivery of content. Typically, such content delivery is for purposes of entertainment, education, training, etc. We label these *content delivery services.* These services can be offered on demand, nearly on demand, on a broadcast or multicast basis, or on a deferred-delivery basis for use at a later time. The various flavors of on-demand and/or multicast services (video-on-demand, high quality music-on-demand, etc.) can pose interesting technical challenges from the point of view of scalability. The NGN's application services infrastructure must address these challenges and provide an efficient and economical way to support them.

Finally, there is another class of services that has to do with management of other services. These services may or may not have revenue-generating potentials by themselves, but they are otherwise every bit as useful and necessary as any other service. These are categorized as *ancillary/management services*. This class includes services such as subscription, customer provisioning, customer network management, customer service management, etc. The dominant mode of accessing these services is likely to be through a Web-based interface. Many services in this category typically are needed in support of other end-user services and, hence, are ancillary in nature. Other services in this class include configuration management, performance monitoring and management, billing and accounting management, service security management, policy enforcement, and similar services. The users of some of these services may be internal customers (typically IT personnel). Offering efficient and user-friendly versions of these application services, in conjunction with primary services, will become a strong service differentiator in next-generation networks. Although these services have much in common with information and data services, they are separately classified to draw attention to their ancillary and/or management nature. As alluded to earlier, and unlike in PSTN/IN, the same application service infrastructure domain that supports primary services can support ancillary/management services, thereby virtually eliminating some of the biggest historical obstacles in the path of drastic reduction in time-to-market.

The different application service categories and representative examples are shown in Table 43.2.

As shown in Fig. 43.2, the application-specific parts of these services will be hosted in application servers, content servers, and media servers that may belong to and be operated by the next generation service provider or, what is more likely, by different kinds of third party application service providers.

Application Services Infrastructure (ASI) Domain

As mentioned earlier, the PSTN/IN infrastructure is optimized for narrowband, circuit-switched, voice-centric services. The application services infrastructure of next generation networks seeks to provide a universal, large-scale, open component-oriented software platform on which a large variety of broadband

TABLE 43.2 NGN Application Service Categories and Examples

Application Service Category	Representative Examples
Communication services	Multimedia, multiparty conferencing; unified multimedia messaging
Information/data services	Internet browsing; e-commerce; directory services; push services; office productivity services
Content delivery services	Video-on-demand; music-on-demand; distance learning; multiplayer network games
Ancillary/management services	Subscription services; Web-based provisioning; service management

(as well as narrowband) services can be architected. Software components distributed on various computing platforms and communicating with one another constitute the "service plumbing" that the application services need for their realization. This service plumbing is the essence of the ASI. Fundamental to ASI and its support of the application service catagories mentioned above is the notion of a **session**. A session represents the evolution and outgrowth of the notion of a "call" in telecommunication networks. A session can be defined as the **context** that brings together related activities, resources, and interactions in an application service. A service session can be further specialized into three distinguishable types. An **access session** is the starting point of interaction between the user of a service and the service provider. Authentication, authorization, and similar access-control functions would take place within the access session. A **usage session** is the actual context in which application service execution takes place and constitutes the utilization phase of the service. Clearly, multiple usage sessions can be associated with the same access session over time. A **communication session** provides an abstract view by ASI of connection-related resources necessary for the delivery of a particular service. A communication session is typically associated with a unique usage session while a usage session can involve multiple communication sessions. The actual connection-related resources are provided by the connectivity network through a connectivity session.

In order to support the development and deployment of distributed applications in the next generation networks, mechanisms are needed for load-balanced and reliable distribution of software components among computing platforms, and for communication among those components. These mechanisms are embodied in a distributed processing environment (DPE). DPE is the invisible "ether" that enables service components (and potentially transport components) to interact with one another. It relieves application services of having to deal with and solve the difficult problems of distribution, currency control, and communication on a per service basis. Components communicate with other components as though all components were resident on the same local host. Thus, application service developers are by and large relieved of explicit concerns about details and complexities associated with distributed processing, communication protocol design, concurrency control, fail-over strategy, platform redundancy and reliability, etc. These are all addressed by DPE experts, resulting in a separation of concerns that is crucial for evolvability as well as for minimizing time to market and managing complexity. Signaling System No. 7 is the closest approximation to a DPE in the voice-centric narrowband network of today. Clearly, DPE in next generation networks will inherit the high-level functions of today's SS7 and add much more functionality to satisfy the requirements of next-generation broadband data-centric application services. Common object request broker architecture (CORBA) may be a good approximation of what NGN's DPE may look like, provided real-time and reliability issues of CORBA are successfully addressed. Extended markup language (XML), simple object access protocol (SOAP), and related technologies may provide another way of realizing DPE functionality in NGN.

43.5 Evolution Towards NGN: Trials and Tribulations

We have briefly sketched a broad outline of the target NGN architecture and discussed how service control in NGN differs from signaling and control in legacy PSTN networks. The fundamental question that needs to be addressed now is how the transition from PSTN/IN (and the current Internet) to NGN is likely to take place, given the fact that such a transition has to preserve the huge investment that the incumbent carriers have in their legacy networks. New entrants clearly have less of a problem in that respect but still need to proceed intelligently as all the technologies to support NGN are not yet ready for prime time. This is a monumental challenge no matter how it is viewed, making the transition to NGN decidedly prolonged and painful.

Curiously, the bootstrap engine for evolution towards NGN seems to have become voice-over-packet (VoP) applications. This may at first seem strange since voice services are currently provided ubiquitously and in a more or less optimal fashion by PSTN/IN. Enter the Internet again! The advent and rapid proliferation of a universal packet network infrastructure, in conjunction with the changing environmental

constraints, prompted initially new and subsequently incumbent service providers to turn what, at one point, was a decidedly low quality (international) long-distance bypass technology into a system that will parallel the service capabilities of PSTN/IN in an economical manner. But unlike PSTN/IN, this new infrastructure will have almost unlimited potential in providing innovative new broadband application services, including multimedia, thereby becoming the beginning of, and the nexus to, the next generation network. Loss of focus on this enormous potential, and looking at packet voice application as a be-all-end-all application, can cost prospective service providers their viability in NGN space.

NGN as a Voice-over-Packet Network

Imagine separating all the functional components of a CO switch, as shown in Fig. 43.1, into the following components of an NGN network:

- A residential gateway (RG) or, equivalently, a customer gateway (CG) that includes the functions of the line module
- A trunk gateway that replaces the trunk module
- A signaling gateway that replaces the signaling module
- A softswitch or media gateway controller (also known as a call agent) that replaces the call processing module

Now, if a packet switching network such as an IP/ATM network underlies platforms that host these four components, and with an appropriate DPE in place, the result is a network similar to the one shown in Fig. 43.3. Interconnection to the PSTN/IN is required in order to allow arbitrary point-to-point communication. When the CO switch functions are distributed as shown, one can view the combined set of functions as a distributed switch. But now protocols are needed for communication between the distributed entities. And, as can be seen from the figure, there is no shortage of possibilities here! What has been achieved in this transition architecture, however, is the historically unprecedented complete separation of bearer/connection control (connectivity network domain) from call/session control (application service domain). This is no small step in the overall evolutionary path of network transformation along the lines that we have sketched towards the NGN target architecture. Note that initially, a ubiquitous DPE capability may not exist due to the immaturity of the middleware products. This, unfortunately, forces adoption of a growing number of protocols to address specific inter-component communication and application needs.

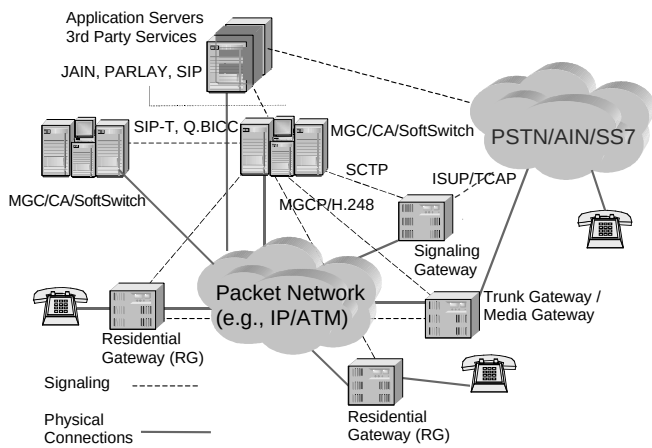


FIGURE 43.3 Towards NGN with packet voice as the bootstrap engine.

When a DPE eventually gets deployed, only the application-related parts of these protocols will be needed as the DPE will take care of lower levels of communication in a distributed environment.

To provide voice-over-packet services, the protocol between the residential gateway (or trunk gateway) and the softswitch becomes critical. This is identified as the IETF Media Gateway Control Protocol (MGCP) evolving to Megaco/ITU-T's H.248 protocol [3]. A simple call setup scenario proceeds as follows:

- The telephone goes off-hook.
- The residential gateway (RG) serving the telephone detects the off-hook event, applies a dial tone, collects the dialed digits, and notifies the softswitch using MGCP messages. The RG informs the softswitch that it is able to receive real-time transport protocol (RTP) streams at a certain port address and the type of audio coding it is able to support.
- The softswitch processes the digits, identifies an egress signaling gateway (if the called party is connected to the PSTN) or RG (if the called party is connected to the IP network) that can send signaling messages to the called party. It also provides the egress gateway with the port and coding information provided by the calling RG.
- Assuming the simple case of the called party being served directly by another RG, the egress RG informs the softswitch that it can receive audio streams from the called party on a certain port. It also applies a ringing tone to the called party's telephone.
- The softswitch informs the calling RG of the port and audio coding information provided by the egress RG.
- Since the RGs at both ends of the call now know which port to send audio RTP packets to and which one to receive on, once the called party answers, the bearer channels are set up and the parties begin to communicate.

Note that the circuit-switched connection in PSTN is now replaced by routing RTP packets over the user datagram protocol (UDP). RTP refers to the IETF real-time transport protocol standard for the transport of real-time data such as audio and video. It provides time stamping, sequence numbering, and delivery monitoring. RTP does not guarantee real-time delivery but must be used in conjunction with other protocols to guarantee quality of service (QoS). RTP control protocol (RTCP) supports feedback from receivers that can be used for monitoring the quality of service. It also supports media stream synchronization for support of multimedia applications. Another protocol that appears in Fig. 43.3 is the streaming control transmission protocol (SCTP), which is designed for the purpose of reliably transporting SS7 application-part messages (ISUP and TCAP) over an IP network. It may find broader applications beyond signaling transport [4].

We have described only a basic two-party call setup over NGN. Intelligent network (IN) services such as free-phone service or local number portability (LNP) and other supplementary services such as call forwarding and three-way calling may be implemented in a separate server. Alternatively, services in the PSTN/AIN that are implemented in an SCP may be directly accessed by NGN entities, while others that are currently switch-based may be integrated directly into the softswitch. The figure shows that the softswitch can provide an open application programming interface (API) to facilitate provision of enhanced services by the service provider or by third parties.

Transition to NGN Beyond Voice

Service providers may begin transitioning to NGN by introducing softswitching technologies in an overlay fashion. Initially, the core of the TDM switching network may be progressively transformed into packet with softswitch control to carry tandem and interoffice voice traffic. Such transformations are generally triggered by tandem switch exhaust. In the second step, packet voice services, enhanced by innovative data features, are extended to the end users in the form of additional lines consuming part of the broadband access bandwidth. Going beyond this stage into offers of integrated voice, data, and multimedia application

services to intelligent and terminals requires introduction of session control functions and an extensive infrastructure of application servers supported by ubiquitous DPE (the true beginnings of ASI). The current state of technology in this space includes activities to validate key architectural assumptions through modeling, simulation, and prototyping. Session control capabilities in support of true NGN application services are only beginning to be understood and prototyped. The pull of the market for new advanced services, coupled with strategic visions of NGN, will determine how fast providers of large scale application service infrastructures and their technology vendors will move in this direction.

43.6 Conclusions

We have discussed the major driving forces for evolution towards the next-generation networks and the challenges and opportunities that characterize this evolution. Limitations of the current control and management infrastructures in pre-NGN networks have been discussed, and various classes of application services that NGN will have to support have been elaborated on. True to the principle of separation of concerns to maximize evolvability and minimize time to market, the basic building blocks of NGN have been identified in three domains: the connectivity network domain, the application services infrastructure (ASI) domain, and the application services domain. We have also shown how the softswitching technology can be used as a stepping stone towards NGN by providing packet voice services initially in the core and subsequently to end devices. Evolution beyond voice-centric softswitching to support NGN application services will require development and deployment of session control functions and their integration into the current softswitching paradigm.

References

1. A.R. Modarressi and R.A. Skoog, Overview of Signaling System No. 7 and its role in the evolving information age network, *Proc. IEEE,* April 1992, 590.
2. Bellcore AIN Generic Requirements.
3. Internet Draft, Megaco Protocol, draft-ietf-megaco-protocol-08, April 2000.
4. Internet Draft, Stream Control Transmission Protocol, <draft-ietf-sigtran-sctp-13.txt>, July 11, 2000.
5. Multiservices Switching Forum, System Architecture Implementation Agreement, MSF-ARCH-001.00-FINAL.
6. International Softswitch Consortium, http://www.softswitch.org/.

IV

Optical

44

Fiber Optic Communications Systems

Joseph C. Palais
Arizona State University

44.1 Introduction

Transmission via beams of light traveling over thin glass fibers is a relative newcomer to communications technology, beginning in the 1970s, reaching full acceptance in the early 1980s, and continuing to evolve since then [Chaffee, 1988]. Fibers now form a major part of the infrastructure for national telecommunications information highways in the U.S. and elsewhere and serve as the transmission media of choice for numerous local area networks. In addition, short lengths of fiber serve as transmission paths for the control of manufacturing processes and for sensor applications.

This section presents the fundamentals of fiber communications and lays foundations for the more detailed descriptions to follow in this chapter. Many full-length books (e.g., Hoss [1990], Jeunhomme [1990], Palais [1998], and Keiser [1991]) exist for additional reference.

Earlier chapters in this handbook covered data rates and bandwidth requirements for telephone systems and local area networks. The steadily increasing demand for information capacity has driven the search for transmission media capable of delivering the required bandwidths. Optical carrier transmission has been able to meet the demand and should continue to do so for many years. Reasons for this are given in the following paragraphs.

Optical communications refers to the transmission of information signals over carrier waves that oscillate at optical frequencies. Optical fields oscillate at frequencies much higher than radio waves or microwaves, as indicated on the abbreviated chart of the electromagnetic spectrum in Fig. 44.1. Frequencies and wavelengths are indicated on the figure.

For historical reasons, optical oscillations are usually described by their wavelengths rather than their frequencies. The two are related by

$$l = c/f \tag{44.1}$$

where f is the frequency in hertz, l is the wavelength, and c is the velocity of light in empty space (3×10^8 m/s). A frequency of 3×10^{14} Hz corresponds to a wavelength of 10^{-6} m (a millionth of a meter is often called a micrometer). Wavelengths of interest for optical communications are on the order of a

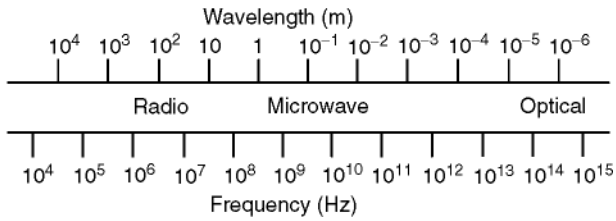


FIGURE 44.1 Portion of the electromagnetic spectrum.

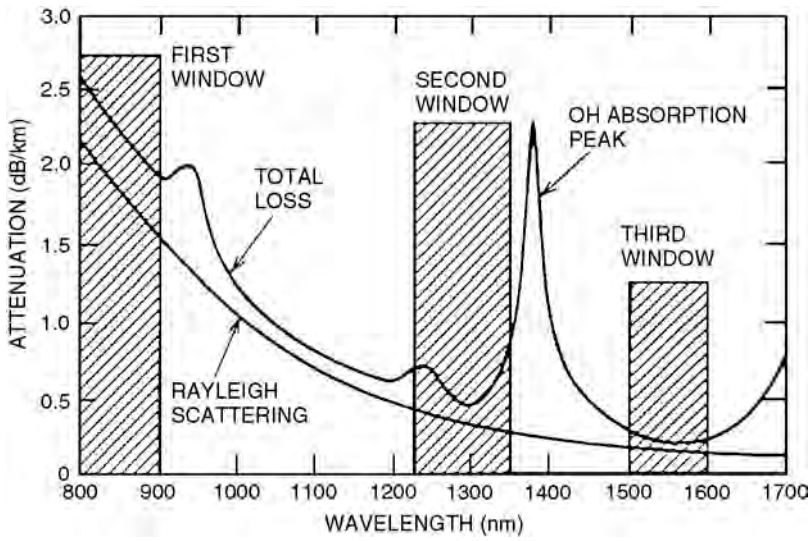


FIGURE 44.2 Attenuation of a silica glass fiber showing the three wavelength regions where most fiber systems operate. (*Source:* Palais, J.C. 1998. *Fiber Optic Communications,* 4th ed., Prentice-Hall, Englewood Cliffs, NJ, p. 112. With permission.)

micrometer. Glass fibers have low loss in the three regions illustrated in Fig. 44.2, covering a range from 0.8 to 1.6 μm (800 to 1600 nm). This corresponds to a total bandwidth of almost 2×10^{14} Hz. The loss is specified in decibels, defined by

$$\text{dB} = 10 \log P_2/P_1 \tag{44.2}$$

where P_1 and P_2 are the input and output powers. Typically, fiber transmission components are characterized by their loss or gain in decibels. The beauty of the decibel scale is that the total decibel value for a series of components is simply the sum of their individual decibel gains and losses.

Although the dB scale is relative, it can be made absolute by specifying a reference level. Two popular references are the milliwatt and microwatt. The respective scales are the dBm (decibels relative to a milliwatt) and the dBm (decibels relative to a microwatt). That is,

$$P(\text{dBm}) = 10 \log P(\text{mW}) \tag{44.3a}$$

$$P(\text{dBm}) = 10 \log P(\text{mW}) \tag{44.3b}$$

Optical power measuring meters are often calibrated in dBm and dBm. Light emitter power and receiver sensitivity (the amount of power required at the receiver) are often given in these units.

Losses in the fiber and in other components limit the length over which transmission can occur. Optical amplification and regeneration are needed to boost the power levels of weak signals for very long paths.

The characteristically high frequencies of optical waves (on the order of 2×10^{14} Hz) allow vast amounts of information to be carried. A single optical channel utilizing a bandwidth of just 1% of this center frequency would have an enormous bandwidth of 2×10^{12} Hz. As an example of this capacity, consider frequency-division multiplexing of commercial television programs. Since each TV channel occupies 6 MHz, over 300,000 television programs could be transmitted over a single optical channel. Although it would be difficult to actually modulate light beams at rates as high as those suggested in this example, tens of gigahertz rates have been achieved.

In addition to electronic multiplexing schemes, such as frequency-division multiplexing of analog signals and time-division multiplexing of digital signals, numerous optical multiplexing techniques exist for taking advantage of the large bandwidths available in the optical spectrum. They include **wavelength-division multiplexing (WDM)** and **optical frequency-division multiplexing (OFDM)**. These technologies allow the use of large portions of the optical spectrum. The total available bandwidth for fibers approaches 2×10^{14} Hz (corresponding to the 0.8-1.6 μm range). Such a huge resource is hard to ignore.

Although atmospheric propagation is possible, the vast majority of optical communications utilizes the waveguiding glass fiber.

A key element for optical communications, a coherent source of light, became available in 1960 with the demonstration of the first **laser**. This discovery was quickly followed by plans for numerous laser applications, including atmospheric optical communications. Developments on empty space optical systems in the 1960s laid the groundwork for fiber communications in the 1970s. The first low-loss optical waveguide, the glass fiber, was fabricated in 1970. Soon after, fiber transmission systems were being designed, tested, and installed. Fibers have proven to be practical for path lengths of under a meter to distances as long as needed on the Earth's surface and under its oceans (for example, almost 10,000 km for transpacific links).

44.2 Optical Communications Systems Topologies

Fiber communications are now common for telephone, local area, and cable television networks. Fibers are also found in short data links (such as required in manufacturing plants), closed-circuit video links, and sensor information generation and transmission.

A block diagram of a point-to-point fiber optical communications system appears in Fig. 44.3. This is the structure typical of the telephone network. The various components will be described later in this section and in succeeding sections of this chapter.

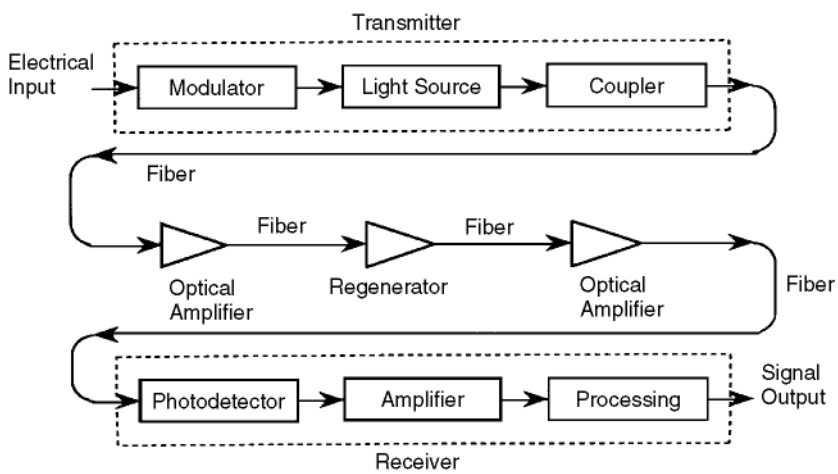


FIGURE 44.3 Major components in a point-to-point fiber transmission system.

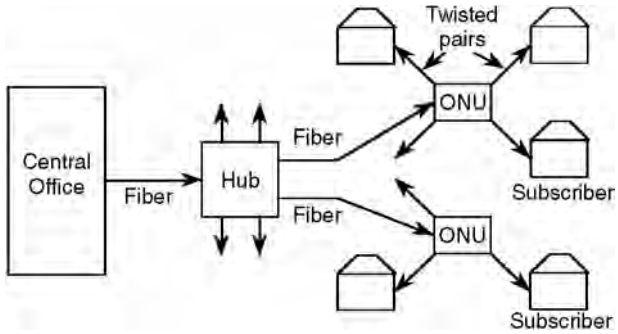


FIGURE 44.4 Fiber-to-the-curb network.

The fiber telephone network is digital, operating at data rates from a few megabits per second up to 2.5 Gb/s and beyond. At the 2.5-Gb/s rate, several thousand digitized voice channels (each operating at 64 kb/s) can be transmitted along a single fiber using time-division multiplexing (TDM).

Because cables may contain more than one fiber (in fact, some cables contain hundreds of fibers), a single cable may be carrying hundreds of thousands of voice channels. Rates in the tens of gigabit per second are attainable, further increasing the potential capacity of a single fiber.

Telephone applications may be broken down into several distinctly different areas: transmission between telephone exchanges, long-distance links, undersea links, and distribution in the local loop (that is, to subscribers). Although similarities exist among these systems, the requirements are somewhat different. Between telephone exchanges, large numbers of calls must be transferred over moderate distances. Because of the moderate path lengths, optical amplifiers or regenerators are not required. On the other hand, long-distance links (such as between major cities) require signal boosting of some sort (either regenerators or optical amplifiers). Undersea links (such as transatlantic or transpacific) require multiple boosts in the signal because of the long path lengths involved [Thiennot, Pirio, and Thomine, 1993].

The local loop does not involve long path lengths but does include division of the optical power in order to share fiber transmission paths over all but the last few tens of meters or so into the subscriber's premises. One architechure for the subscriber distribution network, called fiber-to-the-curb (FTTC), is depicted in Fig. 44.4. Signals are transmitted over fibers through distribution hubs into the neighborhoods. The fibers terminate at optical network units (ONUs) located close to the subscriber. The ONU converts the optical signal into an electrical one for transmission over copper cables for the remaining short distance to the subscriber. Because of the power division at the hubs, optical amplifiers are needed to keep the signal levels high enough for proper signal reception.

Cable television distribution remained totally conducting for many years. This was due to the distortion produced by optical analog transmitters. Production of highly linear laser diodes [such as the distributed feedback (DFB) laser diode] in the late 1980s allowed the design of practical analog television fiber distribution links.

Conversion from analog to digital cable television transmission will be facilitated by the vast bandwidths that fibers make available and by signal compression techniques that reduce the required bandwidths for digital video signals.

Applications such as local area networks (LANs) require distribution of the signals over shared transmission fiber. Possible topologies include the passive star, the active star, and the ring network [Hoss, 1990]. These are illustrated in Figs. 44.5 through 44.7.

The major components found in optical communications systems (called out in Figs. 44.3 through 44.7) are modulators, light sources, fibers, photodetectors, connectors, splices, directional couplers, star couplers, regenerators, and optical amplifiers. They are briefly described in the remainder of this section. More complete descriptions appear in the succeeding sections of this chapter.

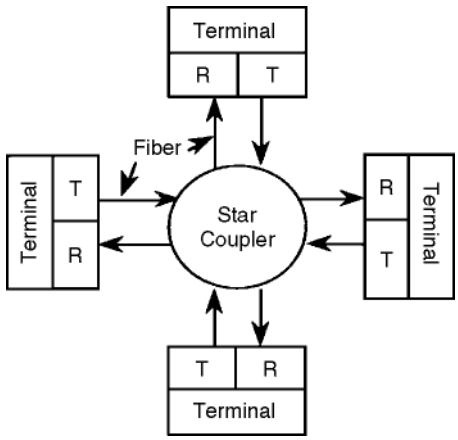


FIGURE 44.5 Passive star network: *T* represents an optical transmitter and *R* represents an optical receiver.

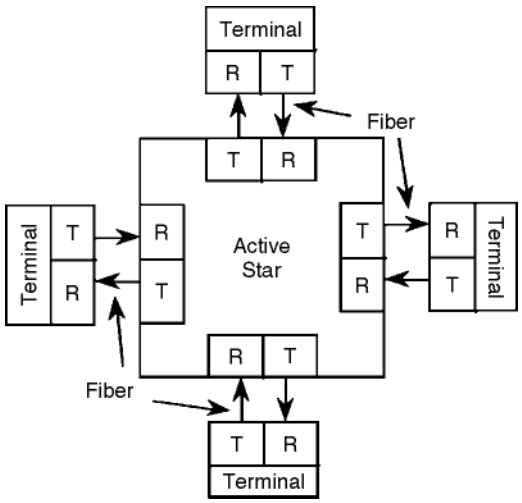


FIGURE 44.6 Active star network: *T* represents an optical transmitter and *R* represents an optical receiver. The active star consists of electronic circuits, which direct the messages to their intended destination terminals.

Fibers

Fiber links spanning more than a kilometer typically use silica glass fibers, as they have lower losses than either plastic or plastic cladded silica fibers. The loss properties of silica fibers were indicated in Fig. 44.2.

Material and waveguide **dispersion** cause pulse spreading, leading to intersymbol interference. This limits the fiber's bandwidth and, subsequently, its data-carrying capability. The amount of pulse spreading is given by

$$\mathrm{D}t = (M + M_g)L\mathrm{D}l \tag{44.4}$$

where M is the material dispersion factor and M_g is the waveguide dispersion factor, L is the fiber length, and $\mathrm{D}\lambda$ is the spectral width of the emitting light source. Because dispersion is wavelength dependent, the spreading depends on the chosen wavelength and on the spectral width of the light source. The total dispersion $(M + M_g)$ has values near 120, 0, and 15 ps/(nm × km) at wavelengths 850, 1300, and 1550 nm, respectively.

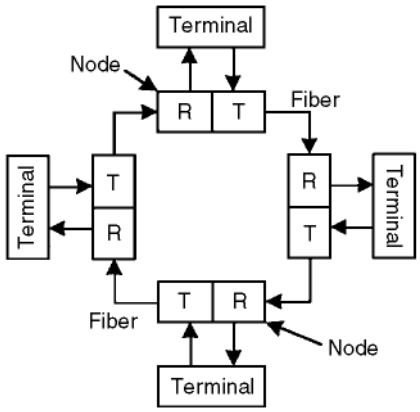


FIGURE 44.7 Ring network: *T* represents an optical transmitter and *R* represents an optical receiver. The nodes act as optical regenerators. Fibers connect the nodes together, while the terminals and nodes are connected electronically.

The fiber's bandwidth and subsequent data rate is inversely proportional to the pulse spreading. As noted, dispersion is a minimum in the second window, actually passing through zero close to 1300 nm for conventional fibers. Special fibers are designed to compensate for the material dispersion by adjusting the waveguide dispersion to cancel it at the low-loss 1550-nm third window. Such fibers (called *dispersion shifted*) are ideal for long fiber links because of their low loss and low dispersion.

Multimode fibers allow more than one mode to simultaneously traverse the fiber. This produces distortion in the form of widened pulses because the energy in different modes travels at different velocities. Again, intersymbol interference occurs. For this reason, multimode fibers are only used for applications where the bandwidth (or data rate) and path length are not large.

Single-mode fibers limit the propagation to a single mode, thus eliminating multimode spreading. Since they suffer only material and waveguide dispersive pulse spreading, these fibers (when operating close to the zero dispersion wavelength) have greater bandwidths than multimode fibers and are used for the longest and highest data rate systems.

Table 44.1 lists bandwidth limits for several types of fibers and Table 44.2 illustrates typical fiber sizes. Step index fibers (SI) have a core having one value of refractive index and a cladding of another value. Graded-index (GRIN) fibers have a core index whose refractive index decreases with distance from the axis and is constant in the cladding. As noted, single-mode fibers have the greatest bandwidths. To limit the number of modes to just one, the cores of single-mode fibers must be much smaller than those of multimode fibers.

Because of the relatively high loss and large dispersion in the 800-nm first window, applications there are restricted to moderately short path lengths (typically less than a kilometer). Because of the limited length, multimode fiber is practical in the first window. Light sources and photodetectors operating in this window tend to be cheaper than those operating at the longer wavelength second and third windows.

The 1300-nm second window, having moderately low losses and nearly zero dispersion, is utilized for moderate to long path lengths. Nonrepeatered paths up to 70 km or so are attainable in this window. In this window, both single-mode and multimode applications exist. Multimode is feasible for short lengths required by LANs (up to a few kilometer) and single-mode for longer point-to-point links.

Fiber systems operating in the 1550-nm third window cover the highest rates and longest unamplified, unrepeated distances. Lengths on the order of 200 km are possible. Single-mode fibers are typically used in this window. Erbium-doped optical amplifiers operate in the third window, boosting the signal levels for very long systems (such as those traversing the oceans).

TABLE 44.1 Typical Fiber Bandwidths

Fiber	Wavelength, nm	Source	Bandwidth, MHz ¥ km
Multimode SI	850	LED	30
Multimode GRIN	850	LD	500
Multimode GRIN	1300	LD or LED	1000
Single mode	1300	LD	>10,000
Single mode	1550	LD	>10,000

TABLE 44.2 Glass Fiber Sizes

Core, μm	Cladding, μm	Numerical Aperture	Operation
8	125	0.11	Single mode
50	125	0.2	Multimode
62.5	125	0.27	Multimode
85	125	0.26	Multimode
100	140	—	Multimode
200	280	—	Multimode

Other Components

Semiconductor **laser diodes (LD)** or **light-emitting diodes (LED)** serve as the light sources for most fiber systems. These sources are typically modulated by electronic driving circuits. The conversion from signal current i to optical power P is given by

$$P = a_0 + a_1 i \tag{44.5}$$

where a_0 and a_1 are constants. Thus, the optical power waveform is a replica of the modulation current.

For very high-rate modulation, external **integrated optic devices** are available to modulate the light beam after its generation by the source. Laser diodes are more coherent (they have smaller spectral widths) than LEDs and thus produce less dispersive pulse spreading, according to Eq. (44.4). In addition, laser diodes can be modulated at higher rates (tens of gigabit per second) than LEDs (which are limited to rates of just a few hundred megabit per second). LEDs have the advantage of lower cost and simpler driving electronics.

Photodetectors convert the optical beam back into an electrical current. Semiconductor **PIN photodiodes** and **avalanche photodiodes (APD)** are normally used. The conversion for the PIN diode is given by the linear equation

$$i = \rho P \tag{44.6}$$

where i is the detected current, P is the incident optical power, and ρ is the photodetector's **responsivity**. Typical values of the responsivity are on the order of 0.5 mA/mW.

Avalanche photodiodes follow the same equation but include an amplification factor that can be as high as several hundred. They improve the **sensitivity** of the receiver.

According to Eq. (44.6), the receiver current is a replica of the optical power waveform (which is itself a replica of the modulating current). Thus, the receiver current is a replica of the original modulating signal current, as desired.

An optical regenerator (or repeater) consists of an optical receiver, electronic processor, and an optical transmitter. Regenerators detect (that is, convert to electrical signals) pulse streams that have weakened because of travel over long fiber paths, electronically determine the value of each binary pulse, and transmit a new optical pulse stream replicating the one originally transmitted. Using a series of regenerators spaced

at distances of tens to hundreds of kilometers, total link lengths of thousands of kilometers are produced. Regenerators can only be used in digital systems.

Optical amplifiers simply boost the optical signal level without conversion to the electrical domain. This simplifies the system compared to the use of regenerators. In addition, optical amplifiers work with both analog and digital signals.

Splices and connectors are required in all fiber systems. Many types are available. Losses tend to be less than 0.1 dB for good splices and just a few tenths of a decibel for good connectors. Fibers are spliced either mechanically or by actually fusing the fibers together.

Directional couplers split an optical beam traveling along a single fiber into two parts, each traveling along a separate fiber. The splitting ratio is determined by the coupler design. In a star coupler (see Fig. 44.5), the beam entering the star is evenly divided among all of the output ports of the star. Typical stars operate as 8×8, 16×16, or 32×32 couplers. As an example, a 32×32 port star can accommodate 32 terminals on a LAN.

44.3 Signal Quality

Signal quality is measured by the **signal-to-noise ratio** (***S/N***) in analog systems and by the **bit error rate** (**BER**) in digital links.

The signal-to-noise ratio in a digital network determines the error rate and is given by

$$\frac{S}{N} = \frac{(M\rho P)^2 R_L}{M^n 2eR_L B(I_D + \rho P) + 4kTB} \tag{44.7}$$

where P is the received optical power, ρ is the detector's unamplified responsivity, M is the detector gain if an APD is used, n (usually between 2 and 3) accounts for the excess noise of the APD, B is the receiver's bandwidth, k is the Boltzmann constant ($k = 1.38 \times 10^{-23}$ J/K), e is the magnitude of the charge on an electron (1.6×10^{-19} C), T is the receiver's temperature in kelvin, I_D is the detector's dark current, and R_L is the resistance of the load resistor that follows the photodetector.

The first term in the denominator of Eq. (44.7) is caused by shot noise, and the second term is attributed to thermal noise in the receiver. If the shot noise term dominates (and the APD excess loss and dark current are negligible), the system is *shot-noise limited*. If the second term dominates, the system is *thermal-noise limited*.

In a thermal-noise limited system, the probability of error P_e (which is the same as the bit error rate) is

$$P_e = 0.5 - 0.5 \text{ erf}\left(0.354\sqrt{S/N}\right) \tag{44.8}$$

where erf is the error function, tabulated in many references [Palais, 1998]. An error rate of 10^{-9} requires a signal-to-noise ratio of nearly 22 dB ($S/N = 158.5$).

44.4 System Design

System design involves ensuring that the signal level at the receiver is sufficient to produce the desired signal quality. The difference between the power available from the transmitting light source (e.g., P_t in dBm) and the receiver's sensitivity (e.g., P_r in dBm) defines the system power budget L. Thus, the power budget is the allowed accumulated loss for all system components and is given (in decibels) by

$$L(\text{dB}) = P_t(\text{dBm}) - P_r(\text{dBm}) \tag{44.9}$$

In addition to ensuring sufficient available power, the system must meet the bandwidth requirements for the given information rate. This requires that the bandwidths of the transmitter, the fiber, and the receiver are sufficient for transmission of the message.

Defining Terms

Avalanche photodiode: Semiconductor photodetector that has internal gain caused by avalanche breakdown.

Bit error rate: Fractional rate at which errors occur in the detection of a digital pulse stream. It is equal to the probability of error.

Dispersion: Wavelength-dependent phase velocity commonly caused by the glass material and the structure of the fiber. It leads to pulse spreading because all available sources emit light covering a (small) range of wavelengths. That is, the emissions have a nonzero spectral width.

Integrated optics: Technology for constructing one or more optical devices on a common waveguiding substrate.

Laser: A source of coherent light, that is, a source of light having a small spectral width.

Laser diode: A semiconductor laser. Typical spectral widths are on the order of 1–5 nm.

Light-emitting diode: A semiconductor emitter whose radiation typically is not as coherent as that of a laser. Typical spectral widths are on the order of 20–100 nm.

Multimode fiber: A fiber that allows the propagation of many modes.

Optical frequency-division multiplexing: Multiplexing many closely spaced optical carriers onto a single fiber. Theoretically, hundreds (and even thousands) of channels can be simultaneously transmitted using this technology.

PIN photodiode: Semiconductor photodetector converting the optical radiation into an electrical current.

Receiver sensitivity: The optical power required at the receiver to obtain the desired performance (either the desired signal-to-noise ratio or bit error rate).

Responsivity: The current produced per unit of incident optical power by a photodetector.

Signal-to-noise ratio: Ratio of signal power to noise power.

Single-mode fiber: Fiber that restricts propagation to a single mode. This eliminates modal pulse spreading, increasing the fiber's bandwidth.

Wavelength-division multiplexing: Multiplexing several optical channels onto a single fiber. The channels tend to be widely spaced (e.g., a two-channel system operating at 1300 nm and 1550 nm).

References

Chaffee, C.D. 1988. *The Rewiring of America,* Academic, Orlando, FL.

Hoss, R.J. 1990. *Fiber Optic Communications,* Prentice-Hall, Englewood Cliffs, NJ.

Jeunhomme, L.B. 1990. *Single-Mode Fiber Optics,* 2nd ed., Marcel Dekker, New York.

Keiser, G. 1991. *Optical Fiber Communications,* McGraw-Hill, New York.

Palais, J.C. 1998. *Fiber Optic Communications,* 4th ed., Prentice-Hall, Englewood Cliffs, NJ.

Thiennot, J., Pirio, F., and Thomine, J.B. 1993. Optical undersea cable systems trends. *Proc. IEEE,* 81(11):1610–11.

Further Information

Information on optical communications is included in several professional society journals. These include the *IEEE Journal of Lightwave Technology* and the *IEEE Photonics Technology Letters*. Valuable information is also contained in several trade magazines such as *Lightwave* and *Laser Focus World*.

45

Optical Fibers and Lightwave Propagation

Paul Diament
Columbia University

45.1 Transmission Along Fibers

Light that finds itself at one end of an **optical fiber** may or may not find its way to the other end. For a randomly chosen combination of the properties and configuration of the light, the optical power has only a small chance of propagating significant distances along the fiber. With proper design of the fiber and of the mechanism for launching the light, however, the signal carried by the light can be transmitted over many kilometers without severe leakage or loss [Diament, 1990; Palais, 1988; Jones, 1988; Cheo, 1985; Basch, 1987; Green, 1993; Hoss and Lacy, 1993; Buckman, 1992; Sterling, 1987].

In its simplest form, an optical fiber consists of a long cylindrical **core** of glass, surrounded by a **cladding** of slightly less **optically dense** glass. See Fig. 45.1. Since both the core and cladding are transparent, propagation depends on the ability of the combination to confine the light and thwart its escape across the side of the fiber. It is the interface between the core and cladding that acts as a reflector to confine any light that would otherwise escape the core and leak away from the cladding into the outside world. The mechanism for confinement to the core is that of **total internal reflection**, a process that is easy to understand for a planar interface between two dissimilar dielectric media [Diament, 1990; Born and Wolf, 1965]. See Fig. 45.2. That process applies as well to the cylindrical geometry of the interface between the core and cladding of the optical fiber.

45.2 Total Internal Reflection

When light strikes the planar interface between two dissimilar transparent media, part of the light is transmitted and the rest is reflected. The proportions of each depend on the impedance mismatch between the two media; this, in turn, depends on the dissimilarity of the indices of refraction of the two materials. The **index of refraction** n measures how much more slowly light travels in the medium than it does in a vacuum.

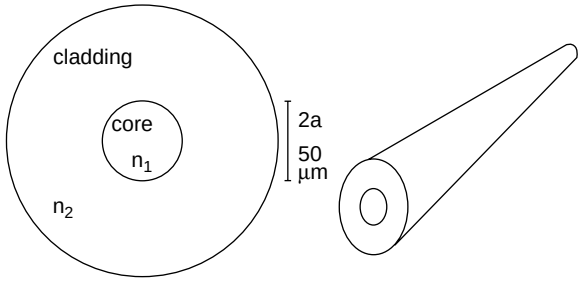


FIGURE 45.1 Structure of an optical fiber: core of radius a, refractive index n_1, surrounded by cladding of index n_2. Typical core diameter is 50 μm.

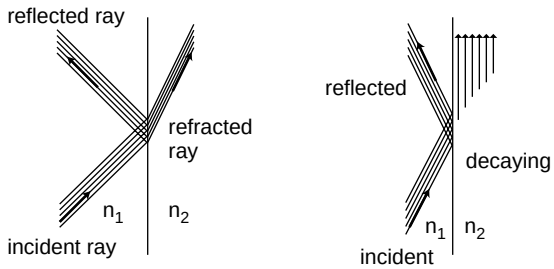


FIGURE 45.2 Ordinary refraction at a planar interface between two media (left), with angle of incidence less than the critical angle, compared with total internal reflection (right), with angle of incidence exceeding the critical angle. Light is incident from the denser medium ($n_1 > n_2$).

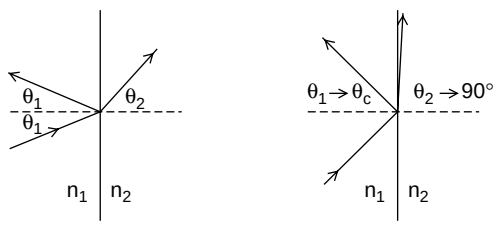


FIGURE 45.3 Snell's law determines the angle of refraction θ_2 at an interface; the angle of reflection θ_1 equals the angle of incidence. The case $n_1 > n_2$ is illustrated. When the angle of incidence θ_1 reaches the critical value $\theta_c = \sin^{-1}(n_2/n_1)$, refraction is at $\theta_2 = 90\infty$. Beyond the critical angle, there is total internal reflection.

When light strikes the interface at an angle to the normal, rather than head-on, the light that is transmitted is **refracted**, or bent away from its original direction, in accordance with **Snell's law**,

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \tag{45.1}$$

where n_1 and n_2 are the refractive indices of the two media and θ_1 and θ_2 are the angles of the ray to the normal to the interface in the two regions. See Fig. 45.3. If the light is incident onto the interface from the denser medium and if the ray's direction is close to grazing, so that the angle of incidence θ_1 is sufficiently large, then it may happen that $n_1 \sin \theta_1$ exceeds n_2 and there is then no solution for the angle of refraction θ_2 from Snell's law, since the quantity $\sin \theta_2$ can never exceed unity. Instead of partial reflection into the denser medium and partial refraction into the other region, there is then no transmission into the less dense medium. Rather, all of the light is reflected back into the denser medium. Total internal

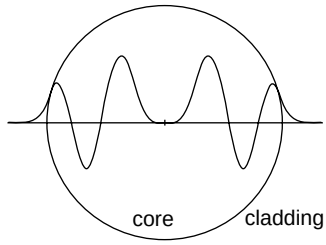


FIGURE 45.4 Standing wave across the core, formed by a repeatedly reflecting ray, with decaying field strength in the cladding, formed by the process of total internal reflection at the cylindrical boundary. The entire field pattern propagates along the axis of the fiber.

reflection occurs (internal to the denser medium in which the light originated), and the escape of the light into the less dense medium has been thwarted.

There is more to the process of total internal reflection at a planar boundary than merely reflection from the interface. Although the light does not leak into the less dense medium and thereby get transmitted outward, it does seep across the interface and propagate, but only parallel to that interface surface [Diament, 1990]. In the direction normal to the interface, there is no propagation on the less dense side of the surface but the strength of the light that appears there decays exponentially. The spatial rate of decay increases with the angle of incidence, beyond the critical angle at which total internal reflection begins. If the decay is sufficiently strong, the light that seeps across is effectively constrained to a thin layer beyond the interface and propagates along it, while the light left behind is reflected back into the denser region with its original strength. The **reflection coefficient** is 100%. There is undiminished optical power flow along the reflected ray in the denser medium, but only storage of optical energy in the light that seeps across the interface into the less dense region.

Much the same process occurs along the curved, cylindrical interface between the core and cladding of the optical fiber. When the light strikes this interface at a sufficiently high grazing angle, total internal reflection occurs, keeping the optical power propagating at an angle within the core and along the axial direction in the cladding. The reflected light strikes the core again on the other side of the axis and is again totally internally reflected there. This confines the light to the core and to a thin layer beyond the interface in the cladding and keeps it propagating indefinitely along the axial direction. The overall result is that the light is guided by the fiber and not permitted to leak away across the side of the fiber, provided the light rays within the core encounter the interface at a sufficiently high grazing angle. See Fig. 45.4.

45.3 Modes of Propagation

For the cylindrical geometry of core and cladding in a fiber, the equivalent of a ray of light in an unbounded medium is any one of a set of **modes** of optical propagation. These refer to configurations of electromagnetic fields that satisfy **Maxwell's equations** and the boundary conditions imposed by the core–cladding interface and also conform to the total internal reflection requirement of grazing incidence to allow guided-wave propagation of the mode along the fiber [Diament, 1990; Born and Wolf, 1965].

These modes are nearly all **hybrid**, meaning that neither the electric nor the magnetic field of the mode is fully transverse to the axial direction of propagation, as can be the case for **waveguides** formed by hollow conducting pipes. Whereas the field configuration in a rectangular metallic waveguide varies as a sine function across the guide, that of an optical fiber follows a **Bessel function** variation. Bessel functions oscillate, much like trigonometric ones, but not with a constant amplitude. In the cladding, the fields decay in the radial direction away from the interface, but not exponentially as at a planar interface under total internal reflection. Rather, the fields in the cladding decay as **modified Bessel functions**, but the effect is the same: confinement to a thin layer just beyond the interface.

Just as the modes of hollow metal pipes are classified as **transverse electric** (TE) and **transverse magnetic** (TM), the hybrid modes of optical fibers also come in two distinguishable varieties [Schlesinger, Diament, and Vigants, 1961], labeled HE and EH, with pairs of subscripts appended to indicate the number of oscillations of the fields azimuthally (around the axis) and radially (across the core). For practical designs of optical fibers, dictated by the smallness of optical wavelengths, the indices of refraction of the core and cladding are very close to each other, and there exist combinations of the hybrid modes that are very nearly of the more familiar transverse, **linearly polarized** types and of greatly simplified description; these are designated LP modes [Gloge, 1971]. The **dominant mode**, which has no constraint of a minimum frequency, is the HE_{11} mode, also designated the LP_{01} mode in the simplified version. All other modes have a **cutoff** frequency, a minimum frequency for axial propagation of a confined mode. Even the dominant mode, in practical cases, has fields that are virtually unconfined when the frequency is too low.

In frequency, then, optical fibers behave as **high-pass filters**. In angular distribution, fibers allow propagation of only those rays that are directed in a sufficiently narrow cone about the axis to satisfy the conditions of total internal reflection.

45.4 Parameters of Fibers

A few parameters are in common use to describe the ability of the fiber to transmit light. One is the **numerical aperture**, *NA*, defined for the fiber as

$$NA = \sqrt{(n_1^2 - n_2^2)} \tag{45.2}$$

and measuring the dissimilarity of the core and cladding indices, n_1 and n_2. This number corresponds more generally for any optical interface to

$$NA = n \sin\theta_{\max} \tag{45.3}$$

a quantity that is conserved across any such interface, by Snell's law, and that measures how great the deviation of an incoming ray from the axial direction can get and still allow capture by, and propagation along, the optical system.

The numerical aperture enters into the formula for the key fiber parameter *V*, often referred to as the **normalized frequency** or sometimes as the **mode volume parameter**. This is defined for a **step-index fiber** of core radius *a* and for light of vacuum wavelength λ as

$$V = (2\pi a/\lambda)\, NA \tag{45.4}$$

The *V* parameter combines many physical attributes of the fiber and the light source; it is proportional to the frequency of the light and to the size of the core, as well as to the numerical aperture. It determines whether a particular mode of the fiber can propagate; in effect, it tells whether the light in that mode strikes the core–cladding interface at a sufficiently high grazing angle to be captured by the fiber. Consequently, it also furnishes an estimate of how many modes can propagate along the fiber. For large values of *V*, the approximate number of distinguishable modes that are above cutoff and can therefore propagate is given by $V^2/2$.

For many purposes, including efficiency of launching and detection of the light and maintenance of the integrity of the signal carried by the light, it is undesirable to have a multiplicity of modes able to propagate along the fiber. For single-mode operation, the parameter *V* must be small, less than 2.4048 [which is the argument at which the Bessel function $J_0(x)$ first goes through zero]. This is not easy to achieve, because wavelengths of light are so tiny: for a core diameter of $2a = 50\ \mu$m and an operating wavelength of $\lambda = 0.8\ \mu$m, we have $V \approx 200(NA)$ and the core and cladding indices have to be extremely

close, yet dissimilar, to attain **single-mode operation**. A much thinner core is needed for a single-mode fiber than for **multimode operation**. For the more readily achieved refractive index differences and core diameters, *NA* may be about 0.15 and several hundred modes can propagate along the fiber. A consequence of this multiplicity of modes is that each travels along the fiber at its own speed and the resulting confusion of arrival times for the various fractions of the total light energy carried by the many modes contributes to signal distortion ascribed to **dispersion**.

45.5 Attenuation

Dispersion limits the operation of the fiber at high data rates; losses limit it at low data rates. Light attenuation due to losses is ascribable to both **absorption** and **scattering**. Absorption occurs when impurities in the glass are encountered; the hydroxyl ion of water present in the glass as an impurity contributes heavily to absorption, along with certain metal ions and oxides. Careful control of impurity levels during fabrication of fibers has allowed light attenuation due to absorption to be kept below 1 dB/km.

Scattering is caused by imperfections in the structure of the fiber. A lower limit on losses is due to **Rayleigh scattering**, which arises from thermal density fluctuations in the glass, frozen into it during manufacture. The typical attenuation constant a from Rayleigh scattering varies as the inverse fourth power of the optical wavelength λ,

$$\alpha = \alpha_0(\lambda_0/\lambda)^4; \quad \alpha_0 = 1.7 \text{ dB/km}, \ \lambda_0 = 0.85\ \mu\text{m}. \tag{45.5}$$

This makes operation at longer wavelengths less subject to losses, but molecular absorption bands at wavelengths beyond about 1.8 μm must be avoided. Rayleigh scattering is excessive below about 0.8 μm and, between these limits, the hydroxyl ion absorption peak near 1.39 μm must also be avoided. This has left three useful transmission windows for fiber optic systems, in the ranges 0.8–0.9 μm, 1.2–1.3 μm, and 1.55–1.6 μm.

45.6 Dispersion

Dispersion is a major concern in a communication system. It refers to processes that cause different components of the light signal to be affected differently. In particular, the speeds at which the different components are propagated along the fiber can vary enough to result in significantly different transmission delays or arrival times for the various portions of the signal. The outcome is distortion of the signal when it is reconstructed from all its components at the receiver or **repeater**. The signal components referred to may be the various frequencies in its spectrum, or they can comprise the multiplicity of modes that each carry their respective fraction of the total optical power to be transmitted.

For optical fibers, there are two aspects to dispersion effects. One is **material dispersion**, also called **intramodal** or **chromatic dispersion**; the other is termed **multimode dispersion**, or also **intermodal dispersion**. Material dispersion arises from the properties of the fiber material that cause the index of refraction to depend on frequency. Since the light source is never truly **monochromatic** but, rather, has a **linewidth**, the various frequencies in the light signal are subject to slightly different refractive index values, which entails different speeds of transmission along the fiber even if there is only one mode of propagation. It is the spread in transmission delays that distorts the signal, and this spread depends on the second derivative of the refractive index with respect to wavelength. One way to minimize this source of dispersion is to operate at about 1.3 μm because silica glass has a second derivative that goes through zero at that wavelength.

Multimode dispersion arises from the fact that the optical energy is distributed among the many modes of propagation of the fiber and each mode has its own **group velocity**, the rate at which a signal carried by that mode is transmitted. The relevant measure of dispersion is then the difference between

the propagation delays of the slowest and the fastest of the participating modes. For a step-index fiber, multimode dispersion is much more severe than material dispersion, although it is somewhat mitigated by the fact that different modes do not propagate independently in a real, imperfect fiber. The phenomena of **mode mixing** and of greater attenuation of high-order modes (those with many field oscillations within the core) tend to reduce the differences in arrival times of the various modes that carry the signal. Nevertheless, intermodal dispersion imposes a severe limitation on data transmission along a step-index fiber.

45.7 Graded-Index Fibers

The way to reduce the deleterious effects of multimode dispersion on the propagation of a signal along a fiber becomes clear when the mechanism of this sort of dispersion is understood. That different modes should exhibit significantly different propagation delays follows from consideration of where the bulk of the field energy resides in each mode. As mentioned, high-order modes have field profiles that oscillate more than those of low-order modes. This includes oscillations in both the radial and the azimuthal directions. Examination of the behavior of Bessel functions indicates that high-order modes have most of their energy nearer to the core–cladding boundary than do the low-order ones. The azimuthal variation of the fields corresponds to propagation at an angle to the fiber axis so that there is progression azimuthally, around the axis, even as the mode propagates along the axis.

For the high-order modes, the azimuthal component of the propagation is more significant than it is for low-order modes, with the result that the overall progression of the high-order mode is in a spiral around the axis, whereas the low-order modes tend to progress more nearly axially than helically. The spiral trajectories of the energy also tend to concentrate far from the axis, nearer to the cladding, as compared to the straighter ones of the low-order modes, whose energy is massed near the axis and travels more nearly straight along it, as suggested in Fig. 45.5. The difference in propagation delay is then readily understood as arising from the longer, spiral path of the modal energy for high-order modes, as compared with the shorter, more direct path taken by the energy in the low-order modes.

This picture of the underlying mechanism for multimode dispersion suggests an equalization scheme aimed at slowing down the low-order modes that are concentrated near the axis, while speeding up the high-order ones that spiral around as they progress and are to be found primarily near the core–cladding interface, farther from the axis. This is achieved by varying the index of refraction of the core gradually, from a high value (for slower speed) near the axis to a lower one (for higher speed) near the cladding. Fibers so fabricated are termed **graded-index** (or **GRIN**) fibers. It is found that a refractive index profile that varies approximately (but not precisely) quadratically with radius can be effective in reducing multimode dispersion dramatically by equalizing the speeds of propagation of various modes of that fiber. Figure 45.6 illustrates the radial variation of the refractive index $n(r)$ for both a step-index fiber and a GRIN fiber.

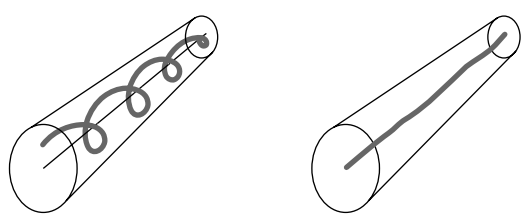


FIGURE 45.5 Spiral path of light energy in higher order modes, compared to more direct trajectory for lower order ones. High-order modes have their energy concentrated near the core–cladding interface; lower order ones are concentrated closer to the axis. The spiral path takes longer to traverse than the direct one.

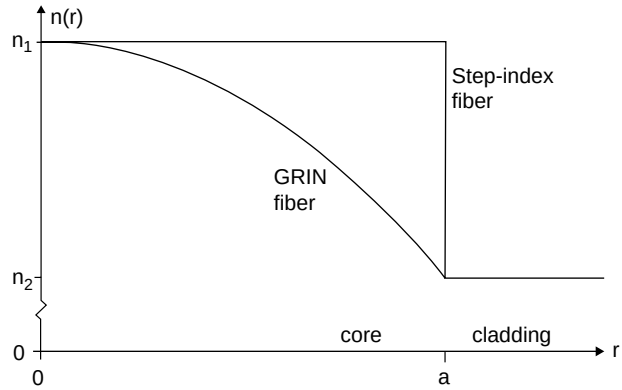


FIGURE 45.6 Radial index profiles of step-index and graded-index fibers. The broken vertical axis is a reminder that the maximum core index n_1 is quite close to the cladding index n_2 for practical fibers.

45.8 Mode Coupling

One more consideration affecting propagation along an optical fiber is the possibility of **mode coupling.** Although the theoretical modes of propagation travel independently along fibers of perfect constitution and perfect geometry, real fibers inevitably have imperfections that can cause different modes to couple and exchange energy between them as they propagate.

A fiber may be imperfect in geometry in that it may not be perfectly straight axially, or its core may not be perfectly round, or it may have been bent. The core and cladding may deviate from their nominal refractive indices. More importantly, there may be fluctuations in their material properties, or the graded-index profile may not be the ideal one. All such imperfections give rise to coupling between modes of propagation of the ideal version of the fiber, such that some of the power in one mode is drained and feeds another mode. When the other mode becomes the stronger one, energy is fed back to the first mode and thereafter sloshes back and forth repeatedly between the modes as the light travels along the fiber. How many modes participate in this periodic exchange of energy depends on the nature of the fluctuations of the imperfections.

A **perturbation analysis** of the effects of fluctuations in some parameter that would be constant in an ideal fiber reveals that the relevant property is the **spatial Fourier spectrum** of the fluctuations. The strength of the coupling between any two modes is proportional to the statistical average of the **power spectrum** of the fluctuations, evaluated at the difference between the **propagation constants** of the two modes in the ideal structure.

This implies that deviations on a large scale, such as gradual bending of the fiber, will affect only a certain group of modes. In this case, the spatial power spectrum exhibits a relatively narrow bandwidth, near zero **spatial frequency**. This allows only modes whose ideal propagation constants are very close to each other to be coupled by the large-scale fluctuations. Such deviations, therefore, couple only modes that are nearly **degenerate** (have nearly equal propagation constants).

On the other hand, periodic fluctuations, such as striations in the material properties, exhibit a spatial spectrum that is peaked in a spectral region corresponding to the period of the striations. In that case, coupling will be significant for modes whose propagation constants differ by the **spatial wavenumber** associated with the periodicity of the fluctuations.

One effect of coupling among modes is to increase losses along the fiber, because energy can get transferred to modes that are not trapped within the core. Some of that energy can then leak out into the cladding and be lost before it can return to the propagating mode. Another effect, however, is to average the dispersion associated with many modes and thereby reduce the overall **pulse spreading**. For example, for closely coupled modes, propagation delays may vary as only the square root of the distance traveled, rather than with the more severe variation, proportional to distance, for uncoupled modes.

45.9 Summary

The step-index optical fiber comprises a dielectric cylindrical core surrounded by a cladding of slightly lower refractive index. This constitutes a dielectric cylinder, which is capable of acting as a waveguide. The mechanism involved in confining and guiding the optical power is that of total internal reflection at the cylindrical boundary between the core and the cladding. This phenomenon results in a **standing wave** pattern within the core and a decaying field in the cladding, with the entire field pattern propagating along the cylinder. The general modes of propagation are hybrid modes, having axial components of both the electric and the magnetic fields, and they advance along the fiber by spiraling around the axis. The exceptional transverse electric or transverse magnetic types are the azimuthally symmetric ones, which do not twist about the axis as they progress along the fiber. The dominant mode is hybrid, however.

Because the optical wavelengths are so short, the core is normally many wavelengths across, with the result that a large number of modes can propagate along the fiber, even when the refractive indices of the core and cladding are kept quite close to each other. Unless single-mode fiber operation is achieved, this entails multimode dispersion effects. Efforts to overcome the limitations this imposes on data transmission have led to the development of graded-index fibers, with an advantageous refractive index profile of the core that tends to equalize the propagation delays of different modes. With careful fabrication of the fibers and with selections of appropriate source wavelengths and bandwidths, attenuation along the fiber can be kept quite low and dispersive effects can be controlled, making optical fibers the medium of choice for propagation of optical signals.

Defining Terms

Absorption: Process by which light energy is drained from a wave and transferred to the material in which it propagates.

Bessel function: A type of mathematical function that oscillates and also decays in amplitude; descriptive of the radial dependence of the fields in a uniform-index core.

Chromatic dispersion: The variation of a material's index of refraction with frequency or wavelength.

Cladding: The outer, optically less dense portion of an optical fiber, surrounding the core.

Core: The inner, optically dense portion of an optical fiber, surrounded by the cladding.

Cutoff: The condition (usually too low an operating frequency) under which a mode is not capable of propagating; the cutoff frequency is the minimum frequency for propagation of a mode.

Degenerate modes: Modes whose propagation constants are equal or nearly equal.

Dispersion: The property of a fiber (or any wave-propagating system) that causes different frequencies (or operating wavelengths) to propagate at different speeds; it is a cause of distortion of signals transmitted along the structure.

Dominant mode: The mode that can propagate under conditions (usually a specific frequency range) that do not allow any other modes to propagate.

Graded-index fiber: The type of optical fiber whose core has a nonuniform index of refraction.

GRIN fiber: Graded-index fiber.

Group velocity: The velocity of propagation of a signal whose spectrum occupies a narrow range of frequencies about some carrier frequency; the rate of change of frequency with respect to propagation constant.

High-pass filter: A circuit or structure that passes high-frequency signals but blocks low-frequency ones.

Hybrid: Describing a mode for which neither the electric nor the magnetic field vector is fully transverse to the direction of propagation.

Index of refraction: Ratio of the vacuum speed of light to its speed in the medium.

Intermodal dispersion: Dispersion effect that arises from a signal's being carried by many modes that travel at different speeds.

Intramodal dispersion: The variation of a material's index of refraction with frequency or wavelength, resulting in dispersive effects for a single mode.

Linearly polarized: Descriptive of a mode of propagation whose field vectors each maintain a fixed orientation in space as they oscillate in time.

Linewidth: The range of frequency or wavelength contained within a nominally monochromatic signal.

Material dispersion: The variation of a material's index of refraction with frequency or wavelength.

Maxwell's equations: The mathematical laws governing the propagation of light and other electromagnetic waves.

Mode: Field configuration of light with well-defined propagation characteristics.

Mode coupling: Process by which different modes of propagation can interchange their energy instead of propagating independently of each other.

Mode mixing: Process of transfer or interchange of light energy among different modes of propagation.

Mode volume parameter: The V parameter, which determines how many modes can propagate along an optical fiber; it is proportional to the frequency of the light, to the size of the core, and to the numerical aperture.

Modified Bessel function: A type of mathematical function that either grows or decays in amplitude; descriptive of the radial dependence of the fields in a uniform-index cladding.

Monochromatic: Of a single frequency or wavelength.

Multimode dispersion: Dispersion effect that arises from a signal's being carried by many modes that travel at different speeds.

Multimode operation: Operation of an optical fiber under conditions that allow many modes to propagate along it.

Normalized frequency: The V parameter, which determines whether a mode can propagate along an optical fiber; it is proportional to the frequency of the light, to the size of the core, and to the numerical aperture.

Numerical aperture (*NA*): A measure of the ability of an optical structure or instrument to capture incident light; $NA = n \sin \theta$, where n is the refractive index and θ is the maximum angle of deviation from the optical axis of a ray that can still be captured by the aperture.

Optical fiber: A long, thin glass structure that can guide light.

Optically dense: Transmitting light at a relatively slow speed; having a high refractive index.

Perturbation analysis: Mathematical process of determining approximately the effects of a relatively weak disturbance of some system that is otherwise in equilibrium.

Power spectrum: Description of some quantity that varies in space or time in terms of the decomposition of its power into constituent sinusoidal components of different spatial or temporal frequencies or periodicities.

Propagation constant: The spatial rate of phase progression of a wave.

Pulse spreading: Process whereby the width of a pulse increases as it propagates.

Rayleigh scattering: A scattering process associated with thermal density fluctuations in the constitution of the glass.

Reflection coefficient: Ratio of strength of reflected light to that of the incident light.

Refraction: Process of bending a ray of light away from its original direction.

Repeater: Electronic circuitry that takes in a degraded signal and restores or reconstitutes it for further transmission.

Scattering: Process by which portions of light energy are redirected (by encounters with microscopic obstacles or inhomogeneities) to undesired directions and thereby lost from the propagating wave.

Single-mode operation: Operation of an optical fiber under conditions that allow only one mode to propagate along it.

Snell's law: Law governing reflection (reflection angle equals incidence angle) and refraction (component of propagation vector tangential to the interface is preserved) of a ray of light encountering the interface between media of different optical density.

Spatial Fourier spectrum: Description of some quantity that varies in space in terms of its decomposition into constituent sinusoidal components of different spatial frequencies or periodicities.

Spatial frequency: The number of cycles per unit length of a quantity that varies periodically in space.

Spatial wavenumber: Reciprocal of a quantity's interval of periodicity in space.

Standing wave: Resultant of counterpropagating waves, with a spatial distribution that does not move.

Step-index fiber: The type of optical fiber whose core and cladding each have a uniform index of refraction.

Total internal reflection: Process of reflecting light incident from a denser medium toward a less dense one, at an angle sufficiently grazing to avert transmission into the less dense medium. All of the incident light power is reflected, although optical fields do appear beyond the interface.

Transverse electric: Descriptive of a mode of propagation whose electric field vector is entirely transverse to the direction of propagation.

Transverse magnetic: Descriptive of a mode of propagation whose magnetic field vector is entirely transverse to the direction of propagation.

Waveguide: A structure that can guide the propagation of light or electromagnetic waves along itself.

References

Basch, E.E.B., Ed. 1987. *Optical-Fiber Transmission,* Howard W. Sams, Indianapolis, IN.

Born, M. and Wolf, E. 1965. *Principles of Optics,* Pergamon Press, Oxford, England.

Buckman, A.B. 1992. *Guided-Wave Photonics,* Saunders College Publishing, Fort Worth, TX.

Cheo, P.K. 1985. *Fiber Optics Devices and Systems,* Prentice-Hall, Englewood Cliffs, NJ.

Diament, P. 1990. *Wave Transmission and Fiber Optics,* Macmillan, New York.

Gloge, D. 1971. Weakly guiding fibers. *Applied Optics,* 10:2252–2258.

Green, P.E., Jr. 1993. *Fiber Optic Networks,* Prentice-Hall, Englewood Cliffs, NJ.

Hoss, R.J. and Lacy, E.A. 1993. *Fiber Optics,* Prentice-Hall, Englewood Cliffs, NJ.

Jones, W.B., Jr. 1988. *Introduction to Optical Fiber Communication Systems,* Holt, Rinehart and Winston, New York.

Palais, J.C. 1988. *Fiber Optic Communications,* Prentice-Hall, Englewood Cliffs, NJ.

Schlesinger, S.P., Diament, P., and Vigants, A. 1961. On higher-order hybrid modes of dielectric cylinders. *IRE Trans. Micro. Theory Tech.,* MTT-8(March):252–253.

Sterling, D.J., Jr. 1987. *Technician's Guide to Fiber Optics,* Delmar, Albany, NY.

Further Information

Periodicals that can help one keep up to date with developments in fiber optics and related systems include the following.

Journal of Lightwave Technology, The Institute of Electrical and Electronics Engineers, Inc., New York.

IEEE Journal of Quantum Electronics, The Institute of Electrical and Electronics Engineers, Inc., New York.

Journal of the Optical Society of America, Optical Society of America, New York.

Optics Letters, Optical Society of America, New York.

Applied Optics, Optical Society of America, New York.

Bell System Technical Journal, AT&T, New York.

Optical Engineering, Society of Photo-Optical Instrumentation Engineers, Bellingham, WA.

Optics Communications, North-Holland, Amsterdam.

Photonics Spectra, Optical Publishers, Pittsfield, MA.

Lightwave, the Journal of Fiber Optics, Howard Rausch Associates, Waltham, MA.

46

Optical Sources for Telecommunication

Niloy K. Dutta
University of Connecticut

Niloy Choudhury
Network Elements, Inc.

46.1 Introduction

Phenomenal advances in research results, development, and application of optical sources have occurred over the last decade. The two primary optical sources used in telecommunications are the semiconductor laser and the light emitting diode (LED). The LEDs are used as sources for low data rate (<200 Mb/s) and short distance applications, and lasers are used for high data rate and long distance applications. The fiber optic revolution in telecommunication which provided several orders of magnitude improvement in transmission capacity at low cost would not have been possible without the development of reliable semiconductor lasers. Today, semiconductor lasers are used not only for fiber optic transmission but also in optical reading and recording (e.g., CD players), printers, fax machines, and in numerous applications as a high power laser source. Semiconductor injection lasers continue to be the laser of choice for various system applications primarily because of their small size, simplicity of operation, and reliable performance. For most transmission system applications, the laser output is encoded with data by modulating the current. However, for some high data rate applications which require long distance transmission, external modulators are used to encode the data. This chapter describes the fabrication, performance characteristics, current state of the art, and research directions for semiconductor lasers and integrated lasers with modulators. The focus of this chapter is laser sources needed for fiber optic transmission systems. These devices are fabricated using the InP material system. For early work and thorough discussion of semiconductor lasers, see Refs. [1–4].

The semiconductor injection laser was invented in 1962 [5–7]. With the development of epitaxial growth techniques and the subsequent fabrication of double heterojunction, the laser technology advanced rapidly in the 1970s and 1980s [1–4]. The demonstration of CW operation of the semiconductor laser in the early 1970s [8] was followed by an increase in development activity in several industrial laboratories.

This intense development activity in the 1970s was aimed at improving the performance characteristics and reliability of lasers fabricated using the AlGaAs material system [1]. These lasers emit near 0.8 μm and were deployed in early optical fiber transmission systems (in the late 1970s and early 1980s).

The optical fiber has zero dispersion near 1.3 μm wavelength and has lowest loss near 1.55 μm wavelength. Thus, semiconductor lasers emitting near 1.3 μm and 1.55 μm are of interest for fiber optic transmission application. Lasers emitting at these wavelengths are fabricated using the InGaAsP/InP materials system and were first fabricated in 1976 [9]. Much of the fiber optic transmission systems around the world that are in use or are currently being deployed utilize lasers emitting near 1.3 μm or 1.55 μm.

Initially, these lasers were fabricated using liquid phase epitaxy (LPE) growth technique. With the development of metal-organic chemical vapor deposition (MOCVD) and gas source molecular beam epitaxy (GSMBE) growth techniques in the 1980s, not only the reproducibility of the fabrication process has improved but it has also led to advances in laser designs such as quantum well lasers and very high speed lasers using semi-insulating Fe doped InP current blocking layers [10].

46.2 Laser Designs

A schematic of a typical double heterostructure used for laser fabrication is shown in Fig. 46.1. It consists of n-InP, undoped $In_{1-x}Ga_xP_yAs_{1-y}$, p-InP and p-InGaAsP grown over (100) oriented n-InP substrate. The undoped $In_{1-x}Ga_xP_yAs_{1-y}$ layer is the light emitting layer (active layer). It is lattice matched to InP for $x \sim 0.45\,y$. The band gap of the $In_{1-x}Ga_xP_yAs_{1-y}$ material (lattice matched to InP) which determines the laser wavelength is given by [11]

$$\mathrm{Eg(eV)} = 1.35 - 0.72y + 0.12y^2$$

For lasers emitting near 1.3 μm, $y \sim 0.6$. The double heterostructure material can be grown by the LPE, GSMBE, or MOCVD growth technique. The double heterostructure material can be processed to produce lasers in several ways. Perhaps the simplest is the broad area laser (Fig. 46.1) which involves putting contacts on the p- and n-side and then cleaving. Such lasers do not have transverse mode confinement or current confinement, which leads to high threshold and nonlinearities in light vs. current characteristics. Several laser designs have been developed to address these problems. Among them are the gain guided laser, weakly index guided laser, and buried heterostructure (strongly index guided) laser. A typical version of these laser structures is shown in Fig. 46.2. The gain guided structure uses a dielectric layer for current confinement. The current is injected in the opening in the dielectric (typically 6 to 12 μm wide), which produces gain in that region and, hence, the lasing mode is confined to that region. The weakly index guided structure has a ridge etched on the wafer, and a dielectric layer surrounds the ridge. The current is injected in the region of the ridge, and the optical mode overlaps the dielectric (which has a low index) in the ridge. This results in weak index guiding.

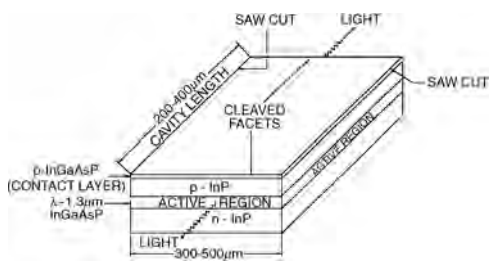


FIGURE 46.1 Schematic of a double heterostructure laser.

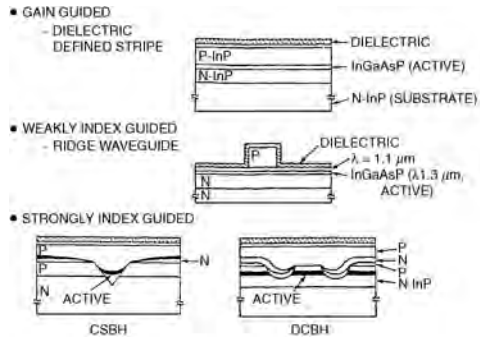


FIGURE 46.2 Schematic of a gain guided, weakly index guided, and strongly index guided buried heterostructure laser.

The buried heterostructure design shown in Fig. 46.2 has the active region surrounded (buried) by lower index layers. The fabrication process of the DCPBH (double channel planar buried heterostructure) laser, involves growing a double heterostructure, etching a mesa using a dielectric mask, and then re-growing the layer surrounding the active region using a second epitaxial growth step. The second growth can be a single Fe doped InP (Fe:InP) semi-insulating layer or a combination of p-InP, n-InP, and Fe:InP layer. Generally, the MOCVD growth process is used for the growth of the regrown layer. Researchers have often given different names to the particular buried heterostructure laser design that they discovered. These are described in detail in Ref. [12]. For the structure of Fig. 46.2, the Fe doped InP layer provides both optical confinement to the lasing mode and current confinement to the active region. Buried heterostructure lasers are generally used in communication system applications because a properly designed strongly index guided buried heterostructure design has superior mode stability, higher bandwidth, and superior linearity in light vs. current (L vs. I) characteristics compared to the gain guided and weakly index guided designs. Early recognition of these important requirements of communication grade lasers led to intensive research on InGaAsP BH laser designs all over the world in the 1980s. It is worth mentioning that BH lasers are more complex and difficult to fabricate compared to the gain guided and weakly index guided lasers. A scanning electron micrograph of a capped mesa (CMBH) buried heterostructure laser along with the laser structure is shown in Fig. 46.3. The current blocking layers in this structure consist of i-InP (Fe doped InP), n-InP, i-InP, and n-InP layers. These sets of blocking layers have the lowest capacitance and are therefore needed for high-speed operation. An optimization of the thickness of these layers is needed for highest speed performance. The laser fabrication involves the following steps. One-micron wide mesas are etched on the wafer and the current blocking layers consisting of i-InP, n-InP, i-InP, and n-InP layers are grown on the wafer with an oxide layer on top of the mesa in place. The oxide layer is then removed and a third growth of p-InP cladding layer and p-InGaAs contact layer is carried out. The wafer is then processed using standard lithography, metallization, and cleaving techniques to produce the lasers.

The light vs. current characteristics at different temperatures of an InGaAsP BH laser emitting at 1.3 μm is shown in Fig. 46.4. Typical threshold current of a BH laser at room temperature is in the 5 to 10 mA range. For gain guided and weakly index guided lasers, typical room temperature threshold currents are in the 25–50 mA and 50–100 mA range, respectively. The external differential quantum efficiency defined as the derivative of the L vs. I characteristics above threshold is ~0.25 mW/mA/facet for a cleaved uncoated laser emitting near 1.3 μm.

An important characteristic of the semiconductor laser is that its output can be modulated simply and easily by modulating the injection current. The relative magnitude of the modulated light output is plotted as a function of the modulation frequency of the current in Fig. 46.5 at different optical output powers. The laser is of the BH type (shown in Fig. 46.3), has a cavity length of 250 μm, and the modulation current amplitude was 5 mA. Note that the 3 dB-frequency to which the laser can be modulated increases

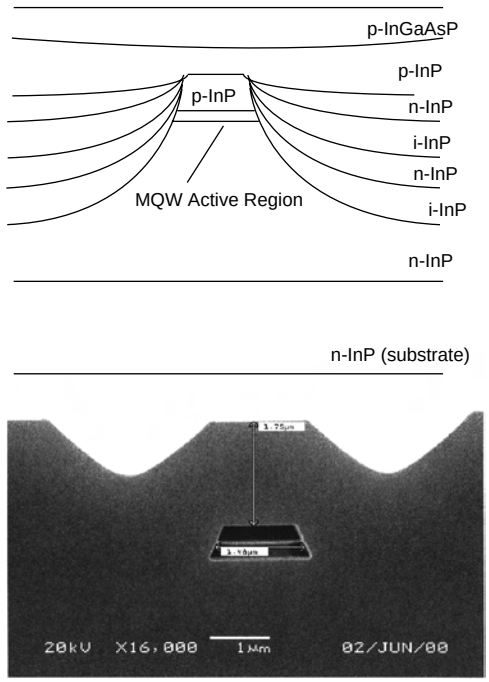


FIGURE 46.3 Schematic of a BH laser and scanning electron photomicrograph of the same laser.

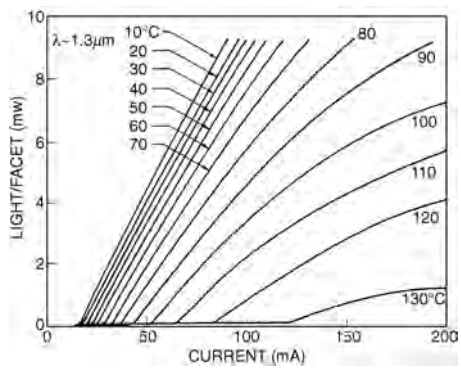


FIGURE 46.4 Light vs. current characteristics of an InGaAsP buried heterostructure laser emitting at 1.3 μm.

with increasing output power and the modulation response is maximum at a certain frequency (ω_r). The resonance frequency ω_r is proportional to the square root of the optical power. The modulation response determines the data transmission rate capability of the laser; for example, for 10 Gb/s data transmission, the 3-dB bandwidth of the laser must exceed 10 GHz. However, other system level considerations, such as allowable error-rate penalty, often introduce much more stringent requirements on the exact modulation response of the laser.

A semiconductor laser with cleaved facets generally emits in a few longitudinal modes of the cavity. A typical spectrum of a laser with cleaved facets is shown in Fig. 46.6. The discrete emission wavelengths are separated by the longitudinal cavity mode spacing which is ~10 A for a laser (λ ~ 1.3 μm) with 250 μm cavity length. Lasers can be made to emit in a single frequency using frequency selective feedback, for example, using a grating internal to the laser cavity as described above.

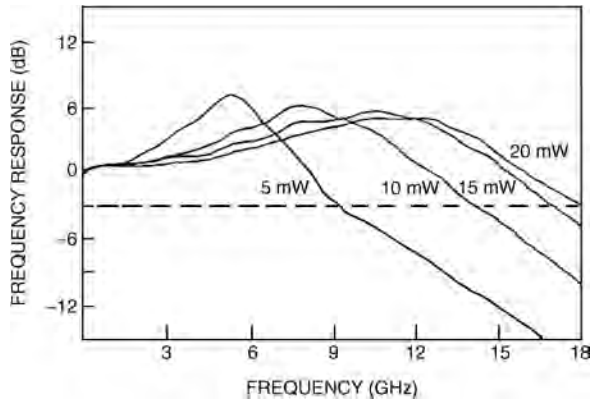


FIGURE 46.5 Modulation response of a laser at different optical output powers.

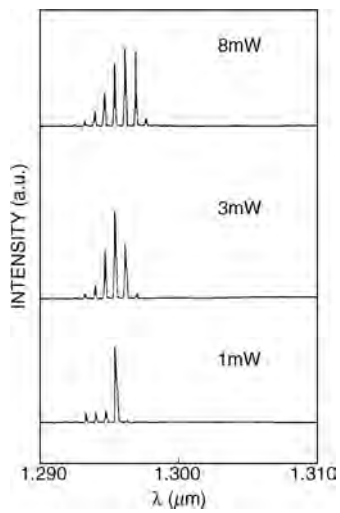


FIGURE 46.6 Emission spectrum of a laser with cleaved facets.

46.3 Quantum Well Lasers

So far we have described the fabrication and performance characteristics of regular double heterostructure (DH) laser which has an active region ~0.1 to 0.2 μm thick. Beginning with the 1980s, lasers with very thin active regions, quantum well lasers, were being developed in many research laboratories [13–22]. Quantum well (QW) lasers have active regions of ~100 A thick which restricts the motion of the carriers (electrons and holes) in a direction normal to the well. This results in a set of discrete energy levels and the density of states is modified to a "two dimensional-like" density of states. This modification of the density of states results in several improvements in laser characteristics such as lower threshold current, higher efficiency, higher modulation bandwidth, and lower CW and dynamic spectral width. All of these improvements were first predicted theoretically and then demonstrated experimentally [23–32].

The development of InGaAsP QW lasers were made possible by the development of MOCVD and GSMBE growth techniques. The transmission electron micrograph (TEM) of a multiple QW laser structure is shown in Fig. 46.7. Shown are four InGaAs quantum wells grown over n-InP substrate. The well thickness is 70 A, and the wells are separated by barrier layers of InGaAsP ($\lambda \sim 1.1$ μm). Multi-quantum well (MQW) lasers with threshold current densities of 600 A/cm^2 have been fabricated [33].

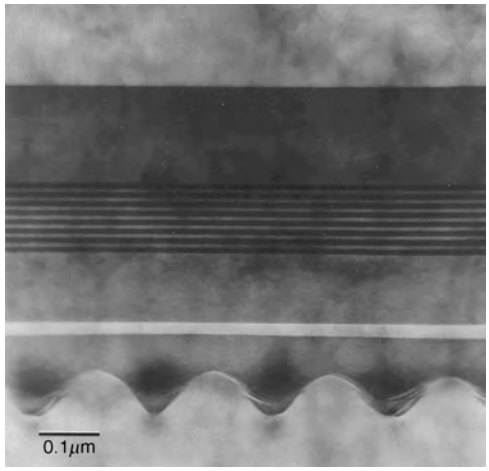


FIGURE 46.7 The transmission electron micrograph of a multiquantum well laser structure.

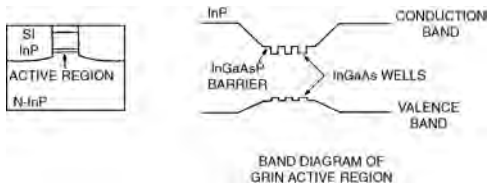


FIGURE 46.8 Schematic of multiquantum well buried heterostructure laser.

The schematic of a MQW BH laser is shown in Fig. 46.8. The composition of the InGaAsP material from the barrier layers to the cladding layer (InP) is gradually varied in this structure over a thickness of ~0.1 μm. This produces a graded variation in index (GRIN structure) which results in a higher optical confinement of the fundamental mode than that for an abrupt interface design. A larger mode confinement factor results in lower threshold current. The laser has an MQW active region and it utilizes Fe doped semi-insulating (SI) InP layers for current confinement and optical confinement. The light vs. current characteristics of an MQW BH laser is shown in Fig. 46.9. The laser emits near 1.5 μm. The MQW lasers have lower threshold currents than regular DH lasers. Also, the two-dimensional like density of states of the QW lasers makes the transparency current density of these lasers significantly lower than that for regular DH lasers [30]. This allows the fabrication of very low threshold lasers using high reflectivity coatings.

The optical gain (g) of a laser at a current density J is given by

$$g = a(J - J_0) \tag{46.1}$$

where a is the gain constant and J_0 is the transparency current density. Although a logarithmic dependence of gain on current density [23] is often used in order to account for gain saturation, a linear dependence is used here for simplicity. The cavity loss α is given by

$$\alpha = \alpha_c + (1/L)\ln(1/R_1R_2) \tag{46.2}$$

where α_c is the free carrier loss, L is the length of the optical cavity, and R_1, R_2 are the reflectivity of the two facets. At threshold, gain equals loss, hence it follows from Eqs. (46.1) and (46.2) that the threshold

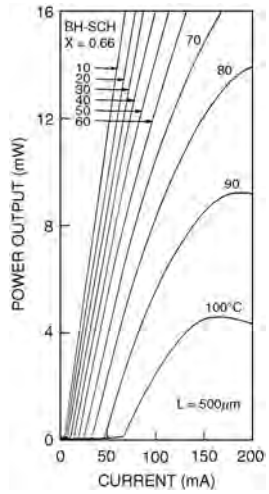


FIGURE 46.9 Light vs. current characteristics of a multiquantum well buried heterostructure laser at different temperatures.

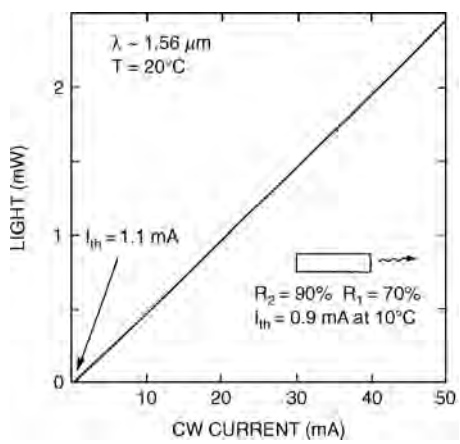


FIGURE 46.10 Light vs. current of a quantum well laser with high reflectivity coatings on both facets.

current density (J_{th}) is given by

$$J_{th} = \alpha_c/a + (1/La)\ln(1/R_1R_2) + J_0 \tag{46.3}$$

Thus, for a laser with high reflectivity facet coatings (R_1, $R_2 \sim 1$) and with low loss ($\alpha_c \sim 0$), $J_{th} \sim J_0$. For a QW laser, $J_0 \sim 50$ A/cm^2 and for a DH laser, $J_0 \sim 700$ A/cm^2, hence it is possible to get much lower threshold current using QW as the active region.

The light vs. current characteristics of a QW laser with high reflectivity coatings on both facets is shown in Fig. 46.10 [33]. The threshold current at room temperature is ~1.1 mA. The laser is 170 μm long and has 90% and 70% reflective coating at the facets. This laser has a compressively strained MQW active region. For a lattice matched MQW active region, a threshold current of 2 mA has been reported [34]. Such low threshold lasers are important for array applications. Recently, QW lasers were fabricated which have higher modulation bandwidth than regular DH lasers. The current confinement and optical confinement in this laser is carried out using MOCVD grown Fe-doped InP lasers similar to that shown

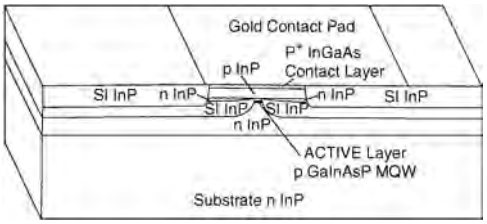


FIGURE 46.11 Schematic of a laser designed for high speed (Morton et al. [35]).

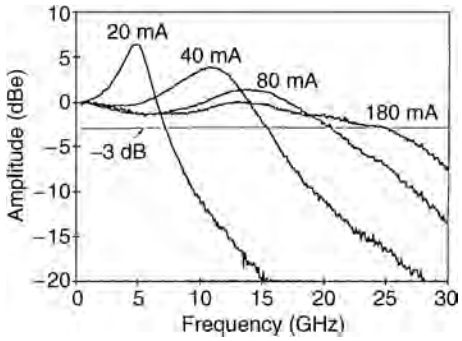


FIGURE 46.12 Modulation response of multiquantum well high speed lasers (Morton et al. [35]).

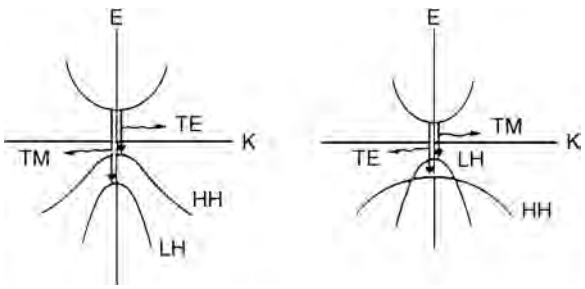


FIGURE 46.13 Band structures under stress. The figures on the left and right represent situations under compressive and tensile strain, respectively.

in Fig. 46.2. The laser structure is then further modified by using a small contact pad and etching channels around the active region mesa (Fig. 46.11). These modifications are designed to reduce the capacitance of the laser structure. The modulation response of the laser is shown in Fig. 46.12. A 3-dB bandwidth of 25 GHz is obtained [35].

Strained Quantum Well Lasers

Quantum well lasers have also been fabricated using an active layer whose lattice constant differs slightly from that of the substrate and cladding layers. Such lasers are known as strained quantum well lasers. Over the last few years, strained quantum well lasers have been extensively investigated all over the world [36–43]. They show many desirable properties such as (1) a very low threshold current density and (2) a lower linewidth than regular MQW lasers both under CW operation and under modulation. The origin of the improved device performance lies in the band-structure changes brought about by the mismatch-induced strain [44,45]. Figure 46.13 shows the band structure of a semiconductor under tensile and

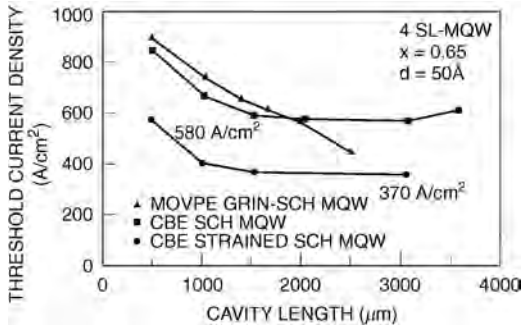


FIGURE 46.14 Broad-area threshold current density as a function of cavity length for strained and lattice-matched InGaAs/InP MQW lasers (Tsang et al. [39]).

compressive strains. Strain splits the heavy-hole and the light-hole valence bands at the Γ point of the Brillouin zone where the bandgap is minimum in direct bandgap semiconductors.

Two material systems have been widely used for strained quantum-well lasers: (1) InGaAs grown over InP by the MOCVD or the CBE growth technique [36–40] and (2) InGaAs grown over GaAs by the MOCVD or the MBE growth technique [41–43]. The former material system is of importance for low-chirp semiconductor lasers for lightwave system applications, while the latter material system has been used to fabricate high-power lasers emitting near 0.98 μm, a wavelength of interest for pumping erbium-doped fiber amplifiers.

The alloy $In_{0.53}Ga_{0.47}As$ has the same lattice constant as InP. Semiconductor lasers with an $In_{0.53}Ga_{0.47}As$ active region have been grown on InP by the MOCVD growth technique. Excellent material quality is also obtained for $In_{1-x}Ga_xAs$ alloys grown over InP by MOCVD for nonlattice-matched compositions. In this case, the laser structure generally consists of one or many $In_{1-x}Ga_xAs$ quantum well layers with InGaAsP barrier layers whose composition is lattice matched to that of InP. For $x < 0.53$, the active layer in these lasers is under tensile stress, while for $x > 0.53$, the active layer is under compressive stress.

Superlattice structures of InGaAs/InGaAsP with tensile and compressive stress have been grown by both MOCVD and CBE growth techniques over an n-type InP substrate. Figure 46.14 shows the broad-area threshold current density as a function of cavity length for strained MQW lasers with four $In_{0.65}Ga_{0.35}As$ [39] quantum wells with InGaAsP ($\lambda \sim 1.25$ μm) barrier layers. The active region in this laser is under 0.8% compressive strain. Also shown for comparison is the threshold current density as a function of cavity length of MQW lattice-matched lasers with $In_{0.53}Ga_{0.47}As$ wells. The entire laser structure, apart from the quantum well composition, is identical for the two cases. The threshold current density is lower for the compressively strained MQW structure than that for the lattice-matched MQW structure.

Buried-heterostructure (BH) lasers have been fabricated using compressive and tensile strained MQW lasers. The threshold current of these lasers as a function of the In concentration is shown in Fig. 46.15 [46]. Lasers with compressive strain have a lower threshold current than that for lasers with tensile strain. This can be explained by splitting of the light hole and heavy hole bands under stress [47,48]. However, more recent studies have shown that it is possible to design tensile strained lasers with lower thresholds [37,42].

Strained quantum well lasers fabricated using $In_{1-x}Ga_xAs$ layers grown over a GaAs substrate have been extensively studied [41,43,49–54]. The lattice constant of InAs is 6.06 A and that of GaAs is 5.654 A. The $In_{1-x}Ga_xAs$ alloy has a lattice constant between these two values, and to a first approximation it can be assumed to vary linearly with x. Thus, an increase in the In mole fraction x increases the lattice mismatch relative to the GaAs substrate and therefore produces larger compressive strain on the active region.

A typical laser structure grown over the n-type GaAs substrate is shown in Fig. 46.16 [41] for this material system. It consists of an MQW active region with one to four $In_{1-x}Ga_xAs$ wells separated by

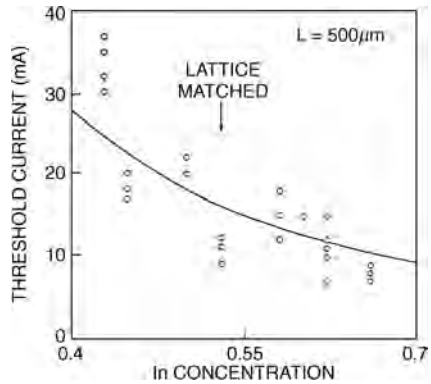


FIGURE 46.15 Threshold current of buried heterostructure $In_xGa_{1-x}As/InP$ MQW lasers plotted as a function of In concentration x (Temkin et al. [46]).

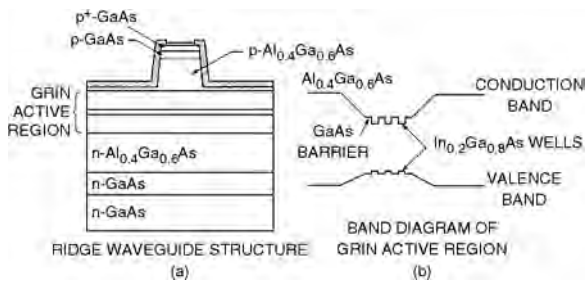


FIGURE 46.16 Typical $In_{1-x}Ga_xAs/GaAs$ MQW laser structure.

GaAs barrier layers. The entire MQW structure is sandwiched between n- and p-type $Al_{0.3}Ga_{0.7}As$ cladding layers, and the P-cladding layer is followed by a p-type GaAs contact layer. Variations of the above structure with different cladding layers or large optical cavity designs have been reported. Emission wavelength depends on the In composition, x. As x increases, the emission wavelength increases, and for x larger than a certain value (typically ~0.25), the strain is too large to yield high-quality material. For $x \sim 0.2$, the emission wavelength is near 0.98 μm, a wavelength region of interest for pumping fiber amplifiers [49]. Threshold current density as low as 47 A/cm^2 has been reported for $In_{0.2}Ga_{0.8}As/GaAs$ strained MQW lasers [52]. High-power lasers have been fabricated using an $In_{0.2}Ga_{0.8}As/GaAs$ MQW active region. Single-mode output powers of greater than 200 mW have been demonstrated using a ridge-waveguide-type laser structure.

Frequency chirp of strained and unstrained QW lasers has been investigated. Strained QW lasers (InGaAs/GaAs) exhibit the lowest chirp (or dynamic linewidth) under modulation. The lower chirp of strained QW lasers is consistent with a small linewidth enhancement factor (α-factor) measured in such devices. The α-factor is the ratio of the real and imaginary part of the refractive index. A correlation between the measured chirp and linewidth enhancement factor for regular double-heterostructure, strained and unstrained QW lasers is shown in Table 46.1. The high efficiency, high power, and low chirp of strained and unstrained QW lasers make these devices attractive candidates for lightwave transmission applications.

Other Material Systems

A few other material systems have been reported for lasers in the 1.3 μm wavelength range. These are the AlGaInAs/InP and InAsP/InP materials grown over InP substrates and, more recently, the InGaAsN material grown over GaAs substrates. The AlGaInAsP/InP system has been investigated with the aim of producing lasers with better high temperature performance for uncooled transmitters [55]. This material

TABLE 46.1 Linewidth Enhancement Factor and Chirp of Lasers

Laser Type	Linewidth Enhancement Factor	FWHM Chirp at 50 mA and 1 Gb/s (A)
DH Laser	5.5	1.2
MQW Laser	3.5	0.6
Strained MQW Laser InGaAs/GaAs, $\lambda \sim 1\ \mu$m	1.0	0.2
Strained MQW Laser InGaAsP/InP, $\lambda \sim 1.55\ \mu$m	2.0	0.4

Note: FWHM = full width at half maximum.

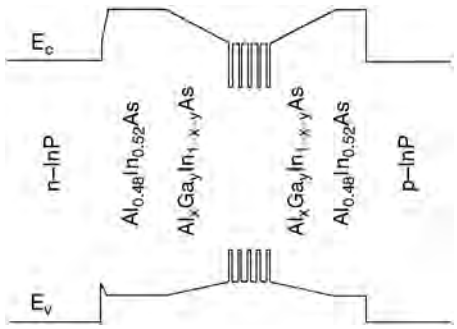


FIGURE 46.17 Band diagram of an AlGaInAs GRINSCH with five quantum wells (Zah et al. [55]).

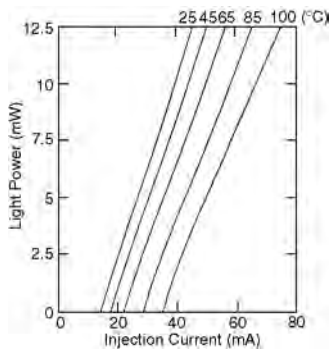


FIGURE 46.18 Light vs. current characteristics of a AlGaInAs quantum well laser with five wells (Zah et al. [55]).

system has a larger conduction band offset than the InGaAsP/InP material system which may result in lower electron leakage over the heterobarrier and, thus, better high temperature performance. The energy band diagram of a GRINSCH (graded index separate confinement heterostructure) laser design is shown in Fig. 46.17. The laser has five compressively strained quantum wells in the active region. The 300 μm long ridge-waveguide lasers typically have a threshold current of 20 mA. The measured light vs. current characteristics of a laser with 70% high reflectivity coating at the rear facet is shown in Fig. 46.18. These lasers have somewhat better high temperature performance than InGaAsP/InP lasers.

The InAsP/InP material system has also been investigated for 1.3 μm lasers [56]. InAsP with an arsenic composition of 0.55 is under 1.7% compressive strain when grown over InP. Using the MOCVD growth technique, buried heterostructure lasers with InAsP quantum well, InGaAsP ($\lambda \sim 1.1\ \mu$m) barrier layers, and InP cladding layers have been reported. The schematic of the laser structure is shown in Fig. 46.19. Typical threshold currents of the BH laser diodes are ~20 mA for 300 μm cavity length.

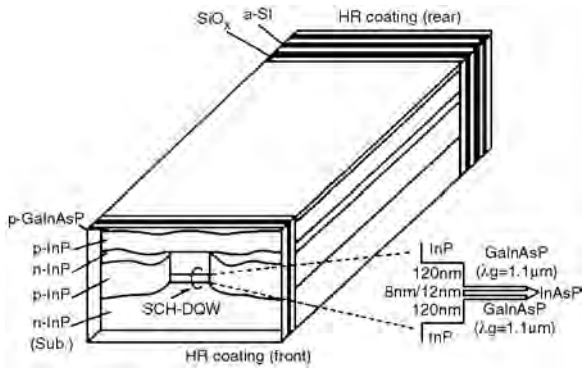


FIGURE 46.19 Schematic of a buried heterostructure InAsP/InGaAsP quantum well laser (Kusukawa et al. [56]).

The material InGaNAs, when grown over GaAs, can have very large (~300 meV) conduction band offset which can lead to much better high temperature performance than the InGaAsP/InP material system [57]. The temperature dependence of threshold is characterized by $I_{th}(T) = I_o \exp(T/T_o)$ where T_o is generally called the characteristic temperature. Typical T_o values for an InGaAsP/InP laser is ~60–70 K in the temperature range of 300–350 K. The predicted T_o value for the InGaNAs/GaAs system is ~150 K, and, recently, T_o = 126 K has been reported for an InGaNAs laser emitting near 1.2 μm [57].

46.4 Distributed Feedback Lasers

Semiconductor lasers fabricated using the InGaAsP material system are widely used as sources in many lightwave transmission systems. One measure of the transmission capacity of a system is the data rate. Thus, the drive towards higher capacity pushes the systems to higher data rates where the chromatic dispersion of the fiber plays an important role in limiting the distance between regenerators. Sources emitting in a single wavelength help reduce the effects of chromatic dispersion and are, therefore, used in most systems operating at high data rates (>1.5 Gb/s).

The single wavelength laser source used in most commercial transmission systems is the distributed feedback (DFB) laser, where a diffraction grating etched on the substrate close to the active region provides frequency selective feedback which makes the laser emit in a single wavelength. This section reports on the fabrication, performance characteristics, and reliability of DFB lasers [58].

The schematic of our DFB laser structure is shown in Fig. 46.20. The fabrication of the device involves the following steps. First, a grating with a periodicity of 2400 A is fabricated on a (100) oriented n-InP substrate using optical holography and wet chemical etching. Four layers are then grown over the substrate. These layers are (1) n-InGaAsP ($\lambda \sim 1.3$ μm) waveguide layer, (2) undoped InGaAsP ($\lambda \sim 1.55$ μm) active layer, (3) p-InP cladding layer, and (4) p-InGaAsP ($\lambda \sim 1.3$ μm) contact layer. Mesas are then etched on the wafer using an SiO_2 mask and wet chemical etching. Fe doped InP semi-insulating layers are grown around the mesas using the MOCVD growth technique. The semi-insulating layers help confine the current to the active region and also provide index guiding to the optical mode. The SiO_2 stripe is then removed and the p-InP cladding layer and a p-InGaAsP contact layer is grown on the wafer using the vapor phase epitaxy growth technique. The wafer is then processed to produce 250 μm long laser chips using standard metallization and cleaving procedures. The final laser chips have antireflection coating (<1%) at one facet and high reflection coating (~65%) at the back facet. The asymmetric facet coatings help remove the degeneracy between the two modes in the stop band.

The CW light vs. current characteristics of a laser are shown in Fig. 46.21. Also shown is the measured spectrum at different output powers. The threshold current of these lasers are in the 15 to 20 mA range. For high fiber coupling efficiency, it is important that the laser emit in the fundamental transverse mode. The measured far field pattern parallel and perpendicular to the junction plane at different output powers

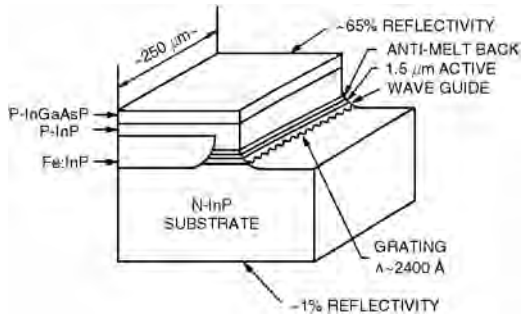


FIGURE 46.20 Schematic of a capped mesa buried heterostructure (CMBH) distributed feedback laser.

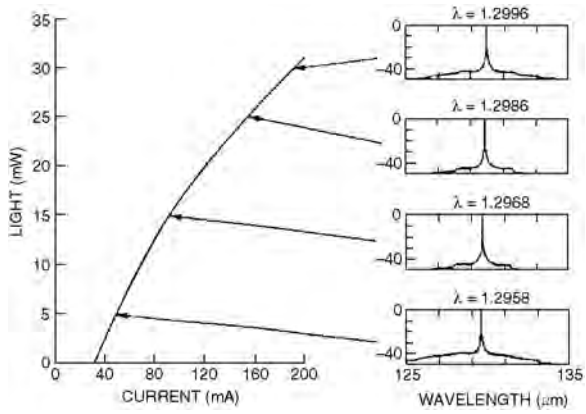


FIGURE 46.21 CW light vs. current characteristics and measured spectrum at different output powers. Temperature = 30°C.

of a device is shown in Fig. 46.22. The figure shows that the laser operates in the fundamental transverse mode in the entire operating power range from threshold to 60 mW. The full width at half maximum of the beam divergences parallel and normal to the junction plane are 40° and 30°, respectively.

The dynamic spectrum of the laser under modulation is an important parameter when the laser is used as a source for transmission. The measured 20 dB full width is shown in Fig. 46.23 at two different data rates as a function of bias level. Note that for a laser biased above threshold, the chirp width is nearly independent of the modulation rate.

Tunable Lasers

Tunable semiconductor lasers are needed for many applications. Examples of application in lightwave transmission systems are (1) wavelength-division multiplexing where signals at many distinct wavelengths are simultaneously modulated and transmitted through a fiber and (2) coherent transmission systems where the wavelength of the transmitted signal must match that of the local oscillator. Several types of tunable laser structures have been reported in the literature [59–64]. Two principle schemes are (1) multisection DFB laser and (2) multisection distributed Bragg reflector (DBR) laser. The multisection DBR lasers generally exhibit higher tunability than the multisection DFB lasers. The design of a multisection DBR laser is shown schematically in Fig. 46.24 [59]. The three sections of this device are (1) the active region section that provides the gain, (2) the grating section that provides the tunability, and (3) the phase-tuning section that is needed to access all wavelengths continuously. The current through each of these sections can be varied independently. The tuning mechanism can be understood by noting that the emission wavelength λ of a DBR laser is given by $\lambda = 2n\Lambda$, where Λ is the grating period and n is

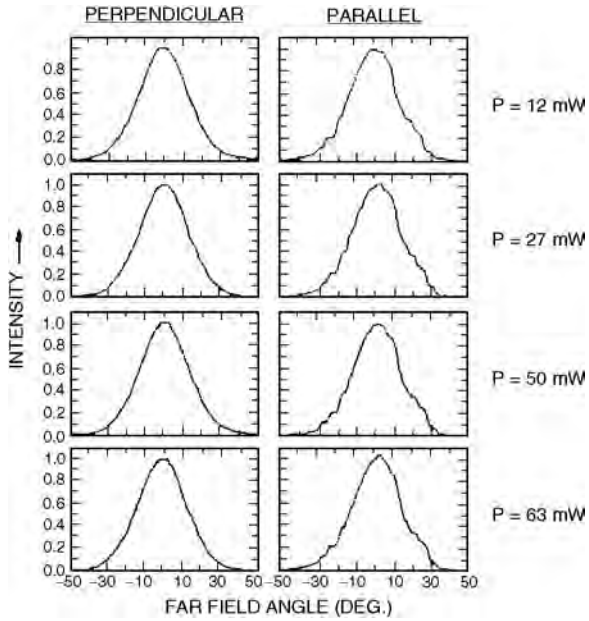


FIGURE 46.22 Measured far field pattern parallel and perpendicular to the junction plane.

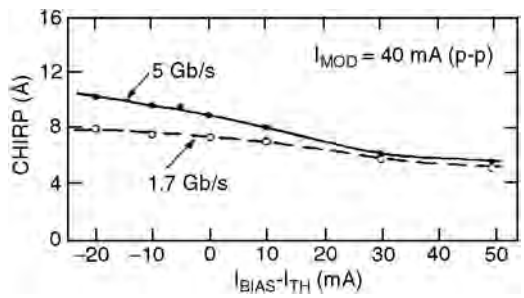


FIGURE 46.23 Measured chirp as a function of bias.

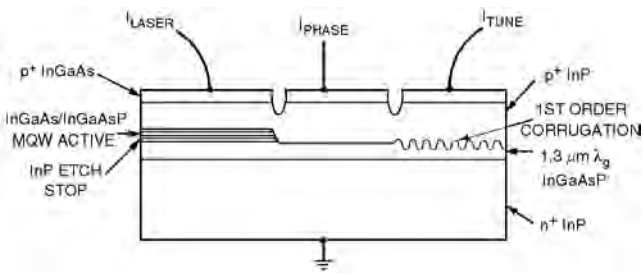


FIGURE 46.24 Schematic of a multisection DBR laser. The laser has an MQW active region. The three sections are optically coupled by the thick waveguide layer below the MQW active region (Koch et al. [59]).

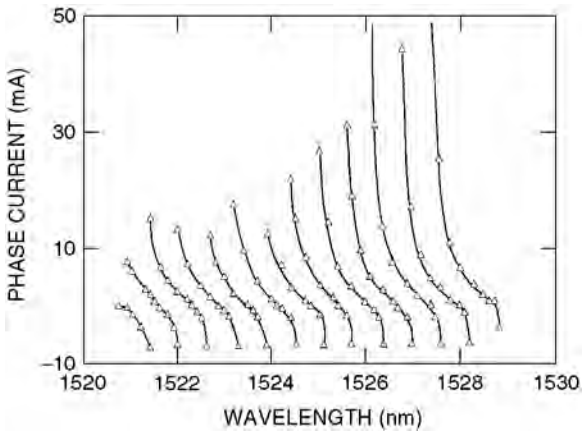


FIGURE 46.25 Frequency tuning characteristics of a three-section MQW DBR laser (Koch et al. [59]).

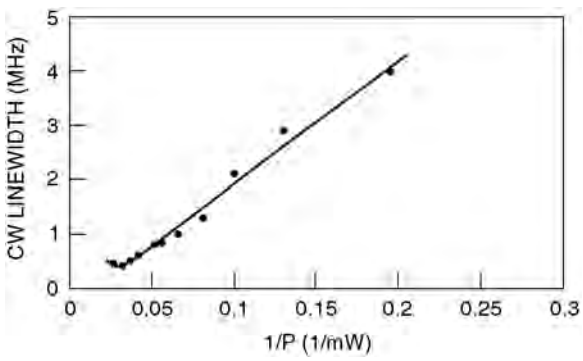


FIGURE 46.26 Measured CW linewidth plotted as a function of the inverse of the output power for an MQW DFB laser with a cavity length of 850 μm.

the effective refractive index of the optical mode in the grating section. The latter can be changed simply by varying the current in the grating section.

The extent of wavelength tunability of a three-section DBR laser is shown in Fig. 46.25 [61]. Measured wavelengths are plotted as a function of phase-section current for different currents in the tuning section. A tuning range in excess of 6 nm can be obtained by controlling currents in the grating and phase-tuning sections.

An important characteristic of lasers for applications requiring a high degree of coherence is the spectral width (linewidth) under CW operation. The CW linewidth depends on the rate of spontaneous emission in the laser cavity. For coherent transmission applications, the CW linewidth must be quite small. The minimum linewidth allowed depends on the modulation format used. For differential phase-shift keying (DPSK) transmission, the minimum linewidth is approximately given by $B/300$, where B is the bit-rate. Thus, for 1-Gb/s transmission rate, the minimum linewidth is 3 MHz. The CW linewidth of a laser decreases with increasing length and varies as α^2, where α is the linewidth enhancement factor. Since α is smaller for a multiquantum well (MQW) laser, the linewidth of DFB or DBR lasers utilizing an MQW active region is smaller than that for lasers with a regular DH active region. The linewidth varies inversely with the output power at low powers (<10 mW) and shows saturation at high powers. The measured linewidth as a function of output power of a 850 μm long DFB laser with an MQW active region is shown in Fig. 46.26. The minimum linewidth of 350 kHz was observed for this device at an operating power of 25 mW. For multisection DBR lasers of the type shown in Fig. 46.24, the linewidth varies with

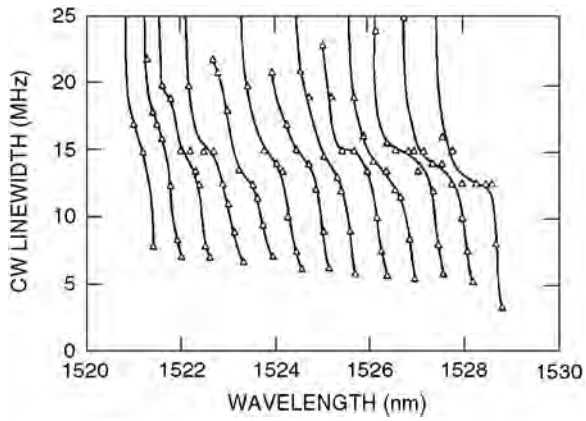


FIGURE 46.27 Measured CW linewidth as a function of wavelength for a 3-section MQW DBR laser (Koch et al. [61]).

changes in currents in the phase-tuning and grating sections. The measured data for an MQW three-section DBR laser is shown in Fig. 46.27. Measured linewidths are plotted as a function of phase-section current for different currents in the tuning section.

46.5 Surface Emitting Lasers

Semiconductor lasers have cleaved facets that form the optical cavity. The facets are perpendicular to the surface of the wafer and light is emitted parallel to the surface of the wafer. For many applications requiring a two-dimensional laser array or monolithic integration of lasers with electronic components (e.g., optical interconnects), it is desirable to have the laser output normal to the surface of the wafer. Such lasers are known as surface-emitting lasers (SEL). A class of surface emitting lasers also have optical cavity normal to the surface of the wafer [65–72]. These devices are known as vertical-cavity surface-emitting lasers (VCSEL) in order to distinguish them from other surface emitters.

A generic SEL structure utilizing multiple semiconductor layers to form a Bragg reflector is shown in Fig. 46.28. The active region is sandwiched between n- and p-type cladding layers which are themselves sandwiched between the two n- and p-type Bragg mirrors. This structure is shown using the AlGaAs/GaAs material system which has been very successful in the fabrication of SELs. The Bragg mirrors consist of alternating layers of low index and high index materials. The thicknesses of each layer is one-quarter of the wavelength of light in the medium. Such periodic quarter wave thick layers can have very high reflectivity. For normal incidence, the reflectivity is given by [73]

$$R = \frac{\left(1 - n_4/n_1(n_2/n_3)^{2N}\right)^2}{\left(1 + n_4/n_1(n_2/n_3)^{2N}\right)^2} \tag{46.4}$$

where n_2 and n_3 are the refractive indices of the alternating layer pairs, n_4 and n_1 are the refractive indices of the medium on the transmitted and incident sides of the DBR mirror, and N is the number of pairs. As N increases, R increases. Also, for a given N, R is larger if the ratio of n_2/n_3 is smaller. For an $AlAs/Al_{0.1}Ga_{0.9}As$ set of quarter wave layers, typically 20 pairs are needed for a reflectivity of ~99.5%. Various types of AlGaAs/GaAs SELs have been reported [74–82].

For an SEL to have a threshold current density comparable to that of an edge-emitting laser, the threshold gains must be comparable for the two devices. The threshold gain of an edge emitting laser is ~100 cm^{-1}. For an SEL with an active-layer thickness of 0.1 μm, this value corresponds to a single-pass gain of ~1%. Thus, for the SEL device to lase with a threshold current density comparable to that of an edge emitter, the mirror reflectivities must be >99%.

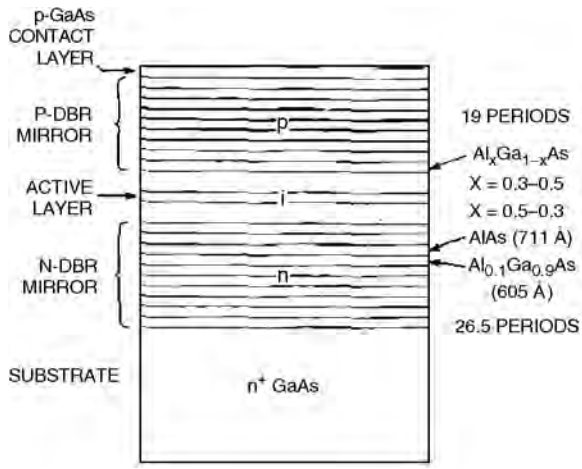


FIGURE 46.28 Schematic illustration of a generic SEL structure utilizing distributed Bragg mirrors formed by using multiple semiconductor layers. DBR pairs consist of AlAs (711 A thick) and $Al_{0.1}Ga_{0.9}As$ (605 A thick) alternate layers. Active layer could be either a quantum well or similar to a regular double-heterostructure laser.

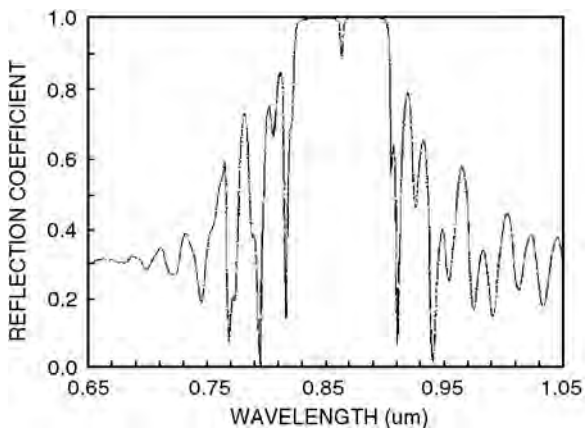


FIGURE 46.29 Typical reflectivity spectrum of an SEL stack.

The reflectivity spectrum of an SEL structure is shown in Fig. 46.29. The reflectivity is >99% over a 10 nm band. The drop in reflectivity in the middle of the band is due to the Fabry–Perot mode.

The number of pairs needed to fabricate a high reflectivity mirror depends on the refractive index of layers in the pair. For large index differences, fewer pairs are needed. For example, in the case of CaF_2 and ZnS for which the index difference is 0.9, only six pairs are needed for a reflectivity of 99%. By contrast, for an InP/InGaAsP ($\lambda \sim 1.3\ \mu$m) layer pair for which the index difference is 0.3, more than 40 pairs are needed to achieve a reflectivity of 99%.

Five principal structures (Fig. 46.30) used in SEL fabrication are (1) etched mesa structure, (2) ion-implanted structure, (3) dielectric isolated structure, (4) buried heterostructure, and (5) metallic reflector structure. Threshold currents of ~0.3 mA have been reported for InGaAs/GaAs SEL devices. An SEL design has been demonstrated whose output can be focussed to a single spot [80]. The laser has a large area (~100 μm dia) and it has a Fresnel zone-like structure etched on the top mirror (Fig. 46.31). The lasing mode with the lowest loss has π phase shift in the near field as it traverses each zone. The laser emits 500 mW in a single mode.

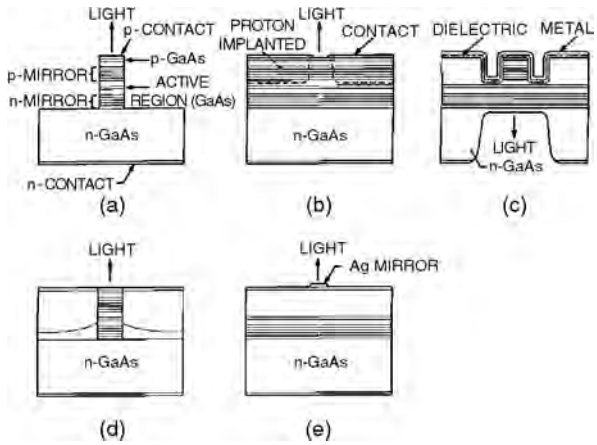


FIGURE 46.30 Schematic of several SEL designs.

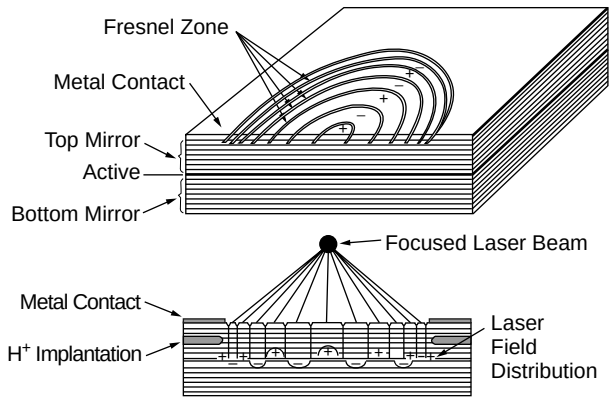


FIGURE 46.31 Schematic of a high power SEL with a Fresnel zone.

An important SEL structure for the AlGaAs/GaAs material system is the oxide aperture device [81] (Fig. 46.32). AlAs has the property that it oxidizes rapidly in the presence of oxygen to aluminum oxide which forms an insulating layer. Thus, by introducing a thin AlAs layer in the device structure, it is possible to confine the current to a very small area. This allows the fabrication of very low threshold (<0.2 mA) and high bandwidth (~14 GHz) lasers.

Central to the fabrication of low-threshold SELs is the ability to fabricate high-reflectivity mirrors. In the late 1970s, Soda et al. [83] reported on an SEL fabricated using the InP material system. The surfaces of the wafer form the Fabry–Perot cavity of the laser. Fabrication of the device involves the growth of a double heterostructure on an n-InP substrate. A circular contact is made on the p-side using an SiO_2 mask. The substrate side is polished, making sure that it is parallel to the epitaxial layer, and ring electrodes (using an alloy of Au–Sn) are deposited on the n-side. The laser had a threshold current density of ~11 kA/cm^2 at 77 K and operated at output powers of several milliwatts.

InGaAsP/InP SELs have been investigated by many researchers over the last few years [83–89]. Many of the schemes utilize alternating layers of InP and InGaAsP to produce Bragg mirrors (Fig. 46.33)[86]. The refractive index difference between InGaAsP ($\lambda \sim 1.3\ \mu$m) and InP layers is smaller than that in GaAs SELs, hence InGaAsP/InP SELs utilize more pairs (typically 40 to 50) to produce a high reflectivity (>99%) mirror. Such mirror stacks have been grown by both chemical beam epitaxy (CBE) and MOCVD growth techniques and have been used to fabricate InGaAsP SELs. Room temperature pulsed operation of InGaAsP/InP SELs using these mirror stacks and emitting near 1.5 μm have been reported [88].

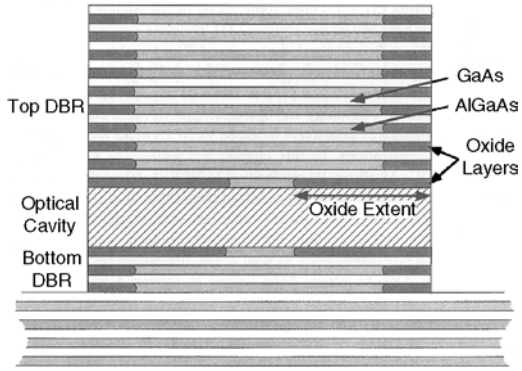


FIGURE 46.32 Schematic of a selectivaly oxidized SEL consisting of AlGaAs/GaAs multilayers and buried aluminum oxide layers. AlGaAs layers with higher Al content are oxidized more.

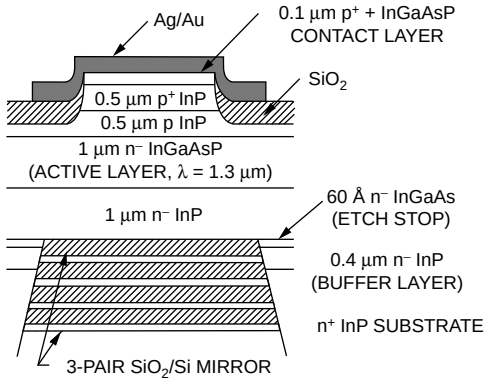


FIGURE 46.33 Schematic of an InGaAsP SEL fabricated using multilayer mirrors (Yang et al. [86]).

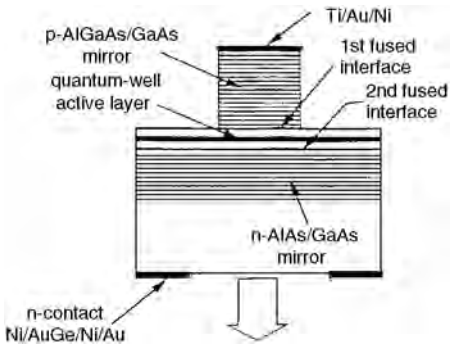


FIGURE 46.34 Schematic of an InGaAsP SEL fabricated using wafer fusion (Babic et al. [89]).

An alternative approach is using the technique of wafer fusion [89,90]. In this technique, the Bragg mirrors are formed using a GaAs/AlGaAs system grown by MBE, and the active region of InGaAsP bounded by thin InP layers is formed by MOCVD. The post type structure of a 1.5 μm wavelength SEL formed using this technique is shown in Fig. 46.34 [89]. The optical cavity is formed by wafer fusion of the InGaAsP quantum well between the Bragg mirrors. Room temperature CW threshold currents of 2.3 mA have been reported for the 8 μm dia post device [89].

46.6 Laser Reliability

The performance characteristics of injection lasers used in lightwave systems can degrade during their operation. The degradation is generally characterized by a change in the operational characteristics of the lasers and, in some cases, is associated with the formation and/or multiplication of defects in the active region. The degraded lasers usually exhibit an increase in the threshold current which is often accompanied by a decrease in the external differential quantum efficiency. For single wavelength lasers, the degradation may be a change in the spectral characteristics, e.g., the degraded device may no longer emit in a single wavelength although the threshold or the light output at a given current has changed very little.

The dominant mechanism responsible for the degradation is determined by any or all of the several fabrication processes including epitaxial growth, wafer quality, device processing, and bonding [91–101]. In addition, the degradation rate of devices processed from a given wafer depends on the operating conditions, i.e., the operating temperature and the injection current. Although many of the degradation mechanisms are not fully understood, extensive amounts of empirical observations exist in the literature which have allowed the fabrication of InGaAsP laser diodes with extrapolated median lifetimes in excess of 25 years at an operating temperature of 20°C [93].

The detailed studies of degradation mechanisms of optical components used in lightwave systems have been motivated by the desire to have a reasonably accurate estimate of the operating lifetime before they are used in practical systems. Since for many applications, the components are expected to operate reliably over a period in excess of 10 years, an appropriate reliability assurance procedure becomes necessary, especially for applications such as an undersea lightwave transmission system where the replacement cost is very high. The reliability assurance is usually carried out by operating the devices under a high stress (e.g., high temperature), which enhances the degradation rate so that a measurable value can be obtained in an operating time of a few hundred hours. The degradation rate under normal operating conditions can then be obtained from the measured high temperature degradation rate using the concept of an activation energy [93].

The light output vs. current characteristics of a laser change after stress aging. There is generally a small increase in threshold current and a decrease in external differential quantum efficiency following the stress aging. Aging data for 1.3 μm InGaAsP lasers used in the first submarine fiber optic cable is shown in Fig. 46.35 [91].

Some lasers exhibit an initial rapid degradation after which the operating characteristics of the lasers are very stable. Given a population of lasers, it is possible to quickly identify the "stable" lasers by a high

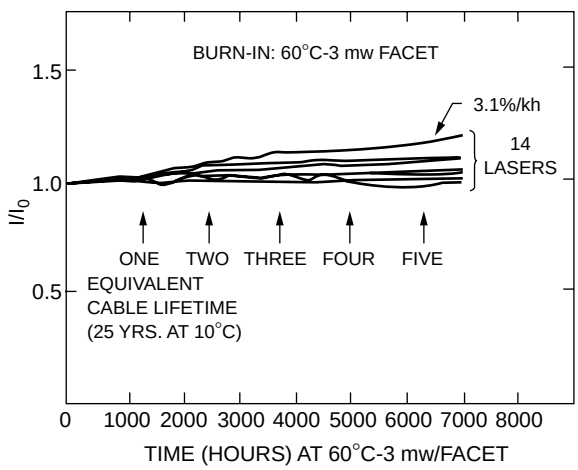


FIGURE 46.35 Operating current for 3 mW output at 60°C as a function of operating time. These data were generated for 1.3 μm InGaAsP lasers used in the first submarine fiber optic cable (Nash et al. [91]).

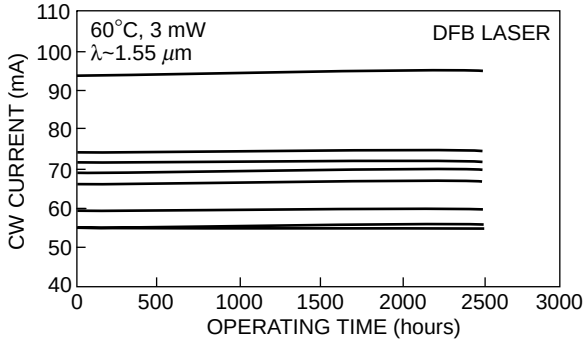


FIGURE 46.36 Operating current as a function of operating time for 3 mW output power at 60°C for lasers emitting near 1.55 μm (courtesy of F. R. Nash).

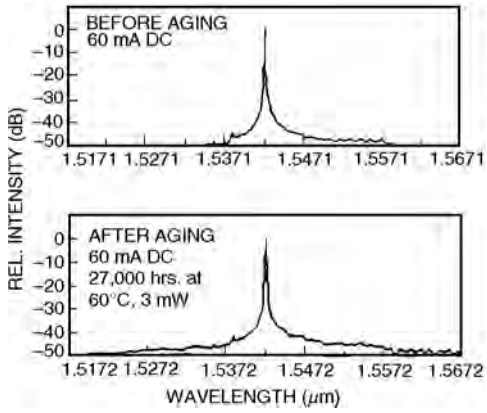


FIGURE 46.37 Spectrum of a DFB laser before and after aging.

stress test (also known as the purge test) [91,92,102,103]. The stress test implies that operating the laser under a set of high stress conditions (e.g., high current, high temperature, high power) would cause the weak lasers to fail and stabilize the possible winners. Observations on the operating current after stress aging have been reported by Nash et al. [91]. It is important to point out that the determination of the duration and the specific conditions for stress aging are critical to the success of this screening procedure.

The expected operating lifetime of a semiconductor laser is generally determined by accelerated aging at high temperatures and using an activation energy. The lifetime (t) at a temperature T is experimentally found to vary as $\exp(-E/kT)$, where E is the activation energy and k is the Boltzmann constant [104,105]. The measured injection current for 3 mW output power at 60 C for buried heterostructures DFB lasers as a function of operating (or aging) time is shown in Fig. 46.36. The operating current increases at a rate of less than 1%/khr of aging time. Assuming a 50% change in operating current as the useful lifetime of the device and an activation energy of 0.7 eV, this aging rate corresponds to a light emitting lifetime of greater than 100 years at 20 C.

A parameter that determines the performance of the DFB laser is the side mode suppression ratio (SMSR), i.e., the ratio of the intensity of the dominant lasing mode to that of the next most intense mode. An example of the spectrum before and after aging is shown in Fig. 46.37 [105]. The SMSR for a laser before and after aging as a function of current is plotted in Fig. 46.38. Note that the SMSR does not change significantly after aging, which confirms the spectral stability of the emission.

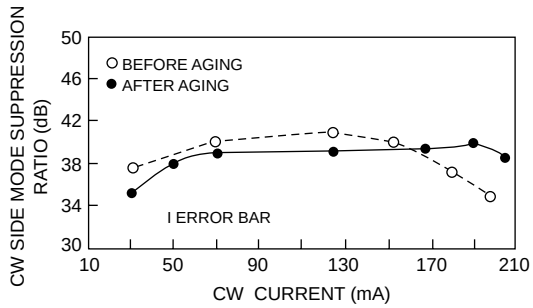


FIGURE 46.38 Side mode suppression ratio (SMSR) as a function of current before and after aging.

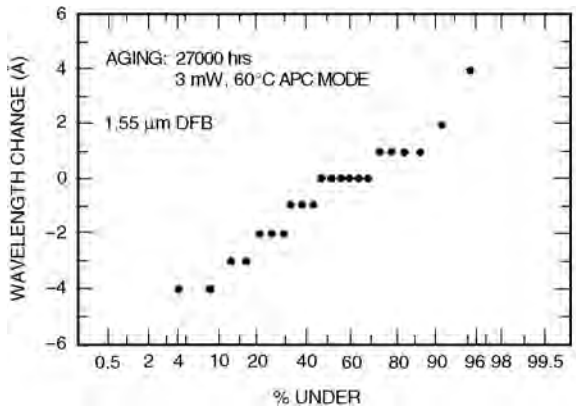


FIGURE 46.39 Change in emission wavelength after aging is plotted in the form of a normal probability distribution.

For some applications such as coherent transmission systems, the absolute wavelength stability of the laser is important. The measured change in emission wavelength at 100 mA before and after aging of several devices is shown in Fig. 46.39. Note that most of the devices do not exhibit any change in wavelength, and the standard deviation of the change is less than 2 A. This suggests that the absolute wavelength stability of the devices is adequate for coherent transmission applications.

Another parameter of interest in certifying the spectral stability of a DFB laser is the change in dynamic linewidth or chirp with aging. Since the chirp depends on data rate and the bias level, a certification of the stability of the value of the chirp is, in general, tied to the requirements of a transmission system application. A measure of the effect of chirp on the performance of a transmission system is the dispersion penalty. We have measured the dispersion penalty of several lasers at 600 Mb/s for a dispersion of 1700 ps/nm before and after aging. The median change in dispersion penalty for 42 devices was less than 0.1 dB and no single device showed a change larger than 0.3 dB. This suggests that the dynamic linewidth or the chirp under modulation is stable.

46.7 Integrated Laser Devices

There have been a significant number of developments in the technology of optical integration of semiconductor lasers and other related devices on the same chip. These chips allow higher levels of functionality than that achieved using single devices. For example, laser and optical modulators have been integrated, serving as simple monolithic transmitters.

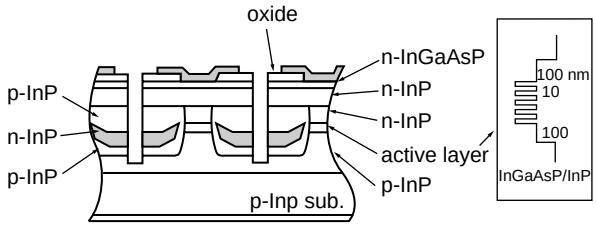


FIGURE 46.40 Schematic of two adjacent devices in a laser array.

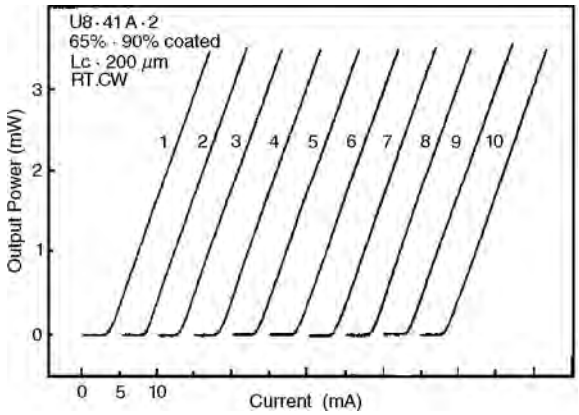


FIGURE 46.41 Light vs. current characteristics of all the lasers in a ten element array.

Laser Arrays

The simplest of all integrated laser devices are one-dimensional arrays of lasers, LEDs, or photodetectors. These devices are fabricated exactly the same way as individual devices except the wafers are not scribed to make single-device chips but left in the form of a bar. The main required characteristics of a laser array are low threshold current and good electrical isolation between the individual elements of the array. The schematic of two adjacent devices in a 10-channel low threshold laser array is shown in Fig. 46.40 [106]. These lasers emit near 1.3 μm and are grown by MOCVD on a p-InP substrate. The lasers have multi-quantum well active region with five 7 nm thick wells, as shown in the insert of Fig. 46.40. The light vs. current characteristics of all the lasers in a 10 element array are shown in Fig. 46.41. The average threshold current and quantum efficiency are 3.2 mA and 0.27 W/A, respectively. The cavity length was 200 μm and the facets of the lasers were coated with dielectric to produce 65% and 90% reflectivity, respectively [106].

The vertical cavity surface emitting laser (SEL) design is more suitable for the fabrication of two dimensional arrays than the edge emitting laser design. Several researchers have reported two-dimensional arrays of SELs. Among the individual laser designs used are the proton implanted design and the oxide confined design. These SELs and SEL arrays have been fabricated so far using the GaAs/AlGaAs material system for emissions near 0.85 μm.

Integrated Laser Modulator

Externally modulated lasers are important for applications where low spectral width under modulation is needed [107–110]. The two types of integrated laser modulator structures that have been investigated are the integrated electroabsorption modulated laser (EML) and the integrated electrorefraction modulated laser. The electrorefraction property is used in a Mach–Zehnder configuration to fabricate a low chirp modulated light source.

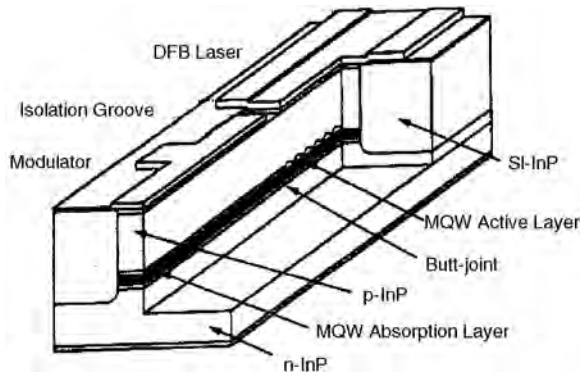


FIGURE 46.42 Schematic of an electroabsorption modulated laser structure (Takeuchi et al. [111]).

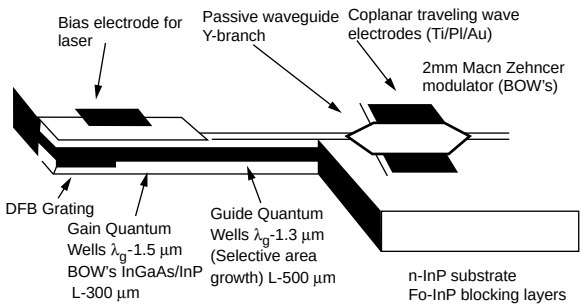


FIGURE 46.43 Schematic of an integrated laser and Mach–Zehnder modulator.

For some applications it is desirable to have the laser and the modulator integrated on the same chip. Such devices, known as electroabsorption modulated lasers (EMLs), are used for high data rate transmission systems with large regenerator spacing. The schematic of an EML is shown in Fig. 46.42 [111]. In this device, the light from the DFB laser is coupled directly to the modulator. The modulator region has a slightly higher band gap than that of the laser region, which results in very low absorption of the laser light in the absence of bias. However, with reverse bias, the effective band gap decreases, which results in reduced transmission through the modulator. For very high-speed operation, the modulator region capacitance must be sufficiently small, which makes the modulator length small resulting in low on/off ratio. Recently very high-speed EMLs have been reported using a growth technique where the laser and modulator active regions are fabricated using two separate growths, thus allowing independent optimization of the modulator band gap and length for high on/off ratio and speed. 40 Gb/s operation has been demonstrated using this device [111].

EML devices have also been fabricated using the selective area epitaxy growth process [112]. In this process, the laser and the modulator active region are grown simultaneously over a patterned substrate. The patterning allows the materials grown to have slightly different band gaps resulting in separate laser and modulator regions. EMLs have been fabricated with bandwidths of 15 GHz and have operated error free over 600 km at 2.5 Gb/s data rate.

An integrated laser Mach–Zehnder device is shown in Fig. 46.43. This device has a ridge-waveguide type DFB laser integrated with a Mach–Zehnder modulator, which also has a lateral guiding provided by a ridge structure. The Mach–Zehnder traveling wave phase modulator is designed so that the microwave and optical velocities in the structure are identical. This allows good coupling of the electrical and optical signal.

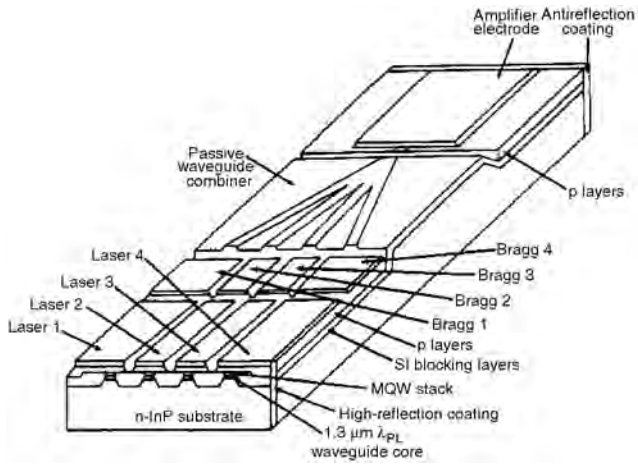


FIGURE 46.44 Schematic of a photonic integrated circuit with multiple lasers for a WDM source.

Multichannel WDM Sources

An alternative to single channel very high speed (>20 Gb/s) data transmission for increasing transmission capacity is multichannel transmission using wavelength division multiplexing (WDM) technology. In WDM systems, many (4, 8, 16, or 32) wavelengths carrying data are optically multiplexed and simultaneously transmitted through a single fiber. The received signal with many wavelengths is optically demultiplexed into separate channels which are then processed electronically in a conventional form. Such a WDM system needs transmitters with many lasers at specific wavelengths. It is desirable to have all of these laser sources on a single chip for compactness and ease of fabrication like electronic integrated circuits.

Figure 46.44 shows the schematic of a photonic integrated circuit with multiple lasers for a WDM source [113]. This chip has four individually addressable DFB lasers, the output of which are combined using a waveguide based multiplexer. Since the waveguide multiplexer has an optical loss of ~8 dB, the output of the chip is further amplified using a semiconductor amplifier. The laser output in the waveguide is TE polarized and, hence, an amplifier with a multiquantum well absorption region which has a high saturation power is integrated in this chip.

Spot Size Converter (SSC) Integrated Laser

A typical laser diode has too wide (~30° × 40°) an output beam pattern for good mode matching to a single mode fiber. This results in a loss of power coupled to the fiber. Thus, a laser whose output spot size is expanded to match an optical fiber is an attractive device for low loss coupling to the fiber without a lens and for wide alignment tolerances. Several researchers have reported such devices [114,115]. Generally, they involve producing a vertically and laterally tapered waveguide near the output facet of the laser. The tapering needs to be done in an adiabatic fashion so as to reduce the scattering losses. The schematic of an SSC laser is shown in Fig. 46.45 [115]. The laser is fabricated using two MOCVD growth steps. The SSC section is about 200 μm long. The waveguide thickness is narrowed along the cavity from 300 nm in the active region to ~100 nm in the region over the length of the SSC section. The laser emits near 1.3 μm, has a multiquantum well active region, and a laser section length of 300 μm. The light vs. current characteristics of a SSC laser [115] at various temperatures are shown in Fig. 46.46. A beam divergence of 13° was obtained for this device. Beam divergences of 9° and 10° in the lateral and vertical direction have been reported for similar SSC devices [115].

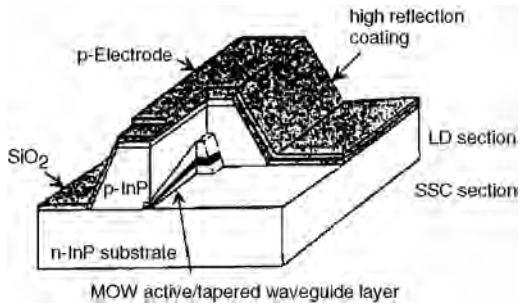


FIGURE 46.45 Schematic of a spot size converter (SSC) laser (Yamazaki et al. [115]).

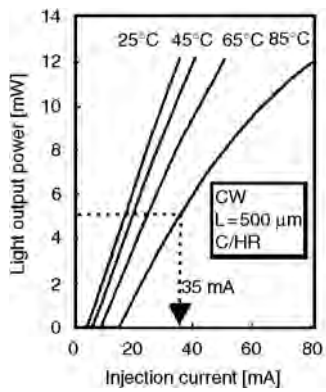


FIGURE 46.46 Light vs. current characteristics of an SSC laser at different temperatures (Yamazaki et al. [115]).

46.8 Summary and Future Challenges

Tremendous advances in semiconductor lasers have occurred over the last decade. The advances in research and many technological innovations have led to the worldwide deployment of fiber optic communication systems that operate near 1.3 μm and 1.55 μm wavelengths and compact storage disks which utilize lasers for read/write purposes. Although most of these systems are based on digital transmission, lasers have also been deployed for carrying high quality analog cable TV transmission systems. However, many challenges remain.

The need for higher capacity is pushing the deployment of WDM-based transmission which needs tunable or frequency settable lasers. An important research area continues to be the development of lasers with very stable and settable frequency. Integration of many such lasers on a single substrate would provide the ideal source for WDM systems.

The laser-to-fiber coupling is also an important area of research. Recent development in spot size converter integrated lasers is quite impressive, but some more work, perhaps, needs to be done to make them easy to manufacture. This may require more process developments.

Lasers with better high temperature performance constitute an important area of investigation. The InGaAsN based system is a promising candidate for making lasers that emit near 1.3 μm with good high temperature performance. New materials investigation is important.

Although WDM technology is currently being considered for increasing the transmission capacity, the need for sources with very high modulation capability remains. Hence, research on new mechanisms for very high-speed modulation is important.

The surface emitting laser is very attractive for two-dimensional arrays and for single wavelength operation. Several important advances in this technology have occurred over the last few years. An important challenge is the fabrication of a device with characteristics superior to that of an edge emitter.

Finally, many of the advances in the laser development would not have been possible without the advances in materials and processing technology. The challenges of much of the current laser research are intimately linked with the challenges in materials growth, which include not only the investigation of new material systems but also improvements in existing technologies to make them more reproducible and predictable.

References

1. H. Kressel and J. K. Butler, *Semiconductor Lasers and Heterojunction LEDs,* Academic Press, NY, 1977.
2. H. C. Casey, Jr. and M. B. Panish, *Heterostructure Lasers,* Academic Press, NY, 1978.
3. G. H. B. Thompson, *Physics of Semiconductor Lasers,* John Wiley and Sons, NY, 1980.
4. G. P. Agrawal and N. K. Dutta, *Long Wavelength Semiconductor Lasers,* Van Nostrand Reinhold Co., New York, 1986.
5. N. Holonyuk, Jr. and S. F. Bevacqua, *Appl. Phys. Lett.,* 1, 82, 1962.
6. M. I. Nathan, W. P. Dumke, G. Burns, F. H. Dill, Jr., and G. Lasher, *Appl. Phys. Lett.,* 1, 63, 1962.
7. T. M. Quist, R. H. Retiker, R. J. Keyes, W. E. Krag, B. Lax, A. L. McWhorter, and H. J. Ziegler, *Appl. Phys. Lett.,* 1, 91, 1962; R. N. Hall, G. E. Fenner, J. D. Kingsley, T. J. Soltys, and R. O. Carlson, *Phys. Rev. Lett.,* 9, 1962.
8. I. Hayashi, M. B. Panish, P. W. Foy, and S. Sumsky, *Appl. Phys. Lett.,* 47, 109, 1970.
9. J. J. Hsieh, *Appl. Phys. Lett.,* 28, 283, 1976.
10. See Chapter 4, Ref. 4.
11. R. E. Nahory, M. A. Pollack, W. D. Johnston, Jr., and R. L. Barns, *Appl. Phys. Lett.,* 33, 659, 1978.
12. See Chapter 5, Ref. 4.
13. N. Holonyak, Jr., R. M. Kolbas, R. D. Dupuis, and P. D. Dapkus, *IEEE J. Quantum Electron.,* QE-16, 170, 1980.
14. N. Holonyak, Jr., R. M. Kolbas, W. D. Laidig, B. A. Vojak, K. Hess, R. D. Dupuis, and P. D. Dapkus, *J. Appl. Phys.,* 51, 1328, 1980.
15. W. T. Tsang, *Appl. Phys. Lett.,* 39, 786, 1981.
16. W. T. Tsang, *IEEE J. Quantum Electron.,* QE-20, 1119, 1986.
17. S. D. Hersee, B. DeCremoux, and J. P. Duchemin, *Appl. Phys. Lett.,* 44, 1984.
18. F. Capasso and G. Margaritondo, Eds., *Heterojunction Band Discontinuation: Physics and Applications,* North Holland, Amsterdam, 1987.
19. N. K. Dutta, S. G. Napholtz, R. Yen, R. L. Brown, T. M. Shen, N. A. Olsson, and D. C. Craft, *Appl. Phys. Lett.,* 46, 19, 1985.
20. N. K. Dutta, S. G. Napholtz, R. Yen, T. Wessel, and N. A. Olsson, *Appl. Phys. Lett.,* 46, 1036, 1985.
21. N. K. Dutta, T. Wessel, N. A. Olsson, R. A. Logan, R. Yen, and P. J. Anthony, *Electron. Lett.,* 21, 571, 1985.
22. N. K. Dutta, T. Wessel, N. A. Olsson, R. A. Logan, and R. Yen, *Appl. Phys. Lett.,* 46, 525, 1985.
23. Y. Arakawa and A. Yariv, *IEEE J. Quantum Electron.,* QE-21, 1666, 1985.
24. Y. Arakawa and A. Yariv, *IEEE J. Quantum Electron.,* QE-22, 1887, 1986.
25. A. Yariv, C. Lindsey, and V. Sivan, *J. Appl. Phys.,* 58, 3669, 1985.
26. A. Sugimura, *IEEE J. Quantum Electron.,* QE-20, 336, 1984.
27. A. Sugimura, *Appl. Phys. Lett.,* 43, 728, 1983.
28. N. K. Dutta and R. J. Nelson, *J. Appl. Phys.,* 53, 74, 1982.
29. L. C. Chiu and A. Yariv, *IEEE J. Quantum Electron.,* QE-18, 1406, 1982.
30. N. K. Dutta, *J. Appl. Phys.,* 54, 1236, 1983.
31. A. Sugimura, *IEEE J. Quantum Electron.,* QE-19, 923, 1983.
32. C. Smith, R. A. Abram, and M. G. Burt, *J. Phys.,* C 16, L171, 1983.

33. H. Temkin, N. K. Dutta, T. Tanbun-Ek, R. A. Logan, and A. M. Sergent, *Appl. Phys. Lett.,* 57, 1610, 1990.
34. C. Kazmierski, A. Ougazzaden, M. Blez, D. Robien, J. Landreau, B. Sermage, J. C. Bouley, and A. Mirca, *IEEE JQE,* 27, 1794, 1991.
35. P. Morton, R. A. Logan, T. Tanbun-Ek, P. F. Sciortino, Jr., A. M. Sergent, R. K. Montgomery, and B. T. Lee, *Electronic Lett*.
36. P. J. A. Thijs, L. F. Tiemeijer, P. I. Kuindersma, J. J. M. Binsma, and T. van Dongen, *IEEE J. Quantum Electron.,* 27, 1426, 1991.
37. P. J. A. Thijs, L. F. Tiemeijer, J. J. M. Binsma, and T. van Dongen, *IEEE J. Quantum Electron.,* QE-30, 477, 1994.
38. H. Temkin, T. Tanbun-Ek, and R. A. Logan, *Appl. Phys Lett.,* 56, 1210, 1990.
39. W. T. Tsang, L. Yang, M. C. Wu, Y. K. Chen, and A. M. Sergent, *Electron Lett.,* 2033, 1990.
40. W. D. Laidig, Y. F. Lin, and P. J. Caldwell, *J. Appl. Phys.,* 57, 33, 1985.
41. S. E. Fischer, D. Fekete, G. B. Feak, and J. M. Ballantyne, *Appl. Phys. Lett.,* 50, 714, 1987.
42. N. Yokouchi, N. Yamanaka, N. Iwai, Y. Nakahira, and A. Kasukawa, *IEEE J. Quantum Electron.,* QE-32, 2148, 1996.
43. K. J. Beernik, P. K. York, and J. J. Coleman, *Appl. Phys. Lett.,* 25, 2582, 1989.
44. J. P. Loehr and J. Singh, *IEEE J. Quantum Electron.,* 27, 708, 1991.
45. S. W. Corzine, R. Yan, and L. A. Coldren, Optical gain in III-V bulk and quantum well semiconductors, in *Quantum Well Lasers,* P. Zory, Ed., Academic Press, NY, to be published.
46. H. Temkin, T. Tanbun-Ek, R. A. Logan, D. A. Coblentz, and A. M. Sergent, *IEEE Photonic Technol. Lett.,* 3, 100, 1991.
47. A. R. Adams, *Electron. Lett.,* 22, 249, 1986.
48. E. Yablonovitch and E.O. Kane, *J. Lightwave Technol.,* LT-4, 50, 1986.
49. M. C. Wu, Y. K. Chen, M. Hong, J. P. Mannaerts, M. A. Chin, and A. M. Sergent, *Appl. Phys. Lett.,* 59, 1046, 1991.
50. N. K. Dutta, J. Lopata, P. R. Berger, D. L. Sivco, and A. Y. Cho, *Electron. Lett.,* 27, 680, 1991.
51. J. M. Kuo, M. C. Wu, Y. K. Chen, and M. A. Chin, *Appl. Phys. Lett.,* 59, 2781, 1991.
52. N. Chand, E. E. Becker, J. P. van der Ziel, S. N. G. Chu, and N. K. Dutta, *Appl. Phys. Lett.,* 58, 1704, 1991.
53. H. K. Choi and C. A. Wang, *Appl. Phys. Lett.,* 57, 321, 1990.
54. N. K. Dutta, J. D. Wynn, J. Lopata, D. L. Sivco, and A. Y. Cho, *Electron. Lett.,* 26, 1816, 1990.
55. C. E. Zah, R. Bhat, B. N. Pathak, F. Favire, W. Lin, N. C. Andreadakis, D. M. Hwang, T. P. Lee, Z. Wang, D. Darby, D. Flanders, and J. J. Hsieh, *IEEE J. Quantum Electron.,* QE-30, 511, 1994.
56. A. Kusukawa, T. Namegaya, T. Fukushima, N. Iwai, and T. Kikuta, *IEEE J. Quantum Electron.,* QE-27, 1528, 1993.
57. J. M. Kondow, T. Kitatani, S. Nakatsuka, Y. Yazawa, and M. Okai, *Proc. OECC*'97, July 1997, Seoul, Korea, 168.
58. See Chapter 7, Ref. 4.
59. T. L. Koch, U. Koren, R. P. Gnall, C. A. Burrus, and B. I. Miller, *Electron. Lett.,* 24, 1431, 1988.
60. Y. Suematsu, S. Arai, and K. Kishino, *J. Lightwave Tech.,* LT-1, 161, 1983.
61. T. L. Koch, and U. Koren, *IEEE J. Quantum Electron.,* QE-27, 641, 1991.
62. N. K. Dutta, A. B. Piccirilli, T. Cella, and R. L. Brown, *Appl. Phys. Lett.,* 48, 1501, 1986.
63. K. Y. Liou, N. K. Dutta, and C. A. Burrus, *Appl. Phys. Lett.,* 50, 489, 1987.
64. T. Tanbun-Ek, R. A. Logan, S. N. G. Chu, and A. M. Sergent, *Appl. Phys. Lett.,* 57, 2184, 1990.
65. H. Soda, K. Iga, C. Kitahara, and Y. Suematsu, *Japan J. Appl. Phys.,* 18, 2329, 1979.
66. K. Iga, F. Koyama, and S. Kinoshita, *IEEE J. Quantum Electron.,* 24, 1845, 1988.
67. J. L. Jewell, J. P. Harbison, A. Scherer, Y. H. Lee, and L. T. Florez, *IEEE J. Quantum Electron.,* 27, 1332, 1991.
68. C. J. Chang-Hasnain, M. W. Maeda, N. G. Stoffel, J. P. Harbison, and L. T. Florez, *Electron. Lett.,* 26, 940, 1990.
69. R. S. Geels, S. W. Corzine, and L. A. Coldren, *IEEE J. Quantum Electron.,* 27, 1359, 1991.

70. R. S. Geels and L. A. Coldren, *Appl. Phys. Lett.,* 5, 1605, 1990.
71. K. Tai, G. Hasnain, J. D. Wynn, R. J. Fischer, Y. H. Wang, B. Weir, J. Gamelin, and A. Y. Cho, *Electron. Lett.,* 26, 1628, 1990.
72. K. Tai, L. Yang, Y. H. Wang, J. D. Wynn, and A. Y. Cho, *Appl. Phy. Lett.,* 56, 2496, 1990.
73. M. Born and E. Wolf, *Principles of Optics,* Pergamon Press, NY, 1977.
74. J. L. Jewell, A. Scherer, S. L. McCall, Y. H. Lee, S. J. Walker, J. P. Harbison, and L. T. Florez, *Electron Lett.,* 25, 1123, 1989.
75. Y. H. Lee, B. Tell, K. F. Brown-Goebeler, J. L. Jewell, R. E. Leibenguth, M. T. Asom, G. Livescu, L. Luther, and V. D. Mattera, *Electron. Lett.,* 26, 1308, 1990.
76. K. Tai, R. J. Fischer, C. W. Seabury, N. A. Olsson, D. T. C. Huo, Y. Ota, and A. Y. Cho, *Appl. Phys. Lett.,* 55, 2473, 1989.
77. A. Ibaraki, K. Kawashima, K. Furusawa, T. Ishikawa, T. Yamayachi, and T. Niina, *Japan J. Appl. Phys.,* 28, L667, 1989.
78. E. F. Schubert, L. W. Tu, R. F. Kopf, G. J. Zydzik, and D. G. Deppe, *Appl. Phys. Lett.,* 57, 117, 1990.
79. R. S. Geels, S. W. Corzine, J. W. Scott, D. B. Young and L. A. Coldren, *IEEE Photonic Tech. Lett.,* 2, 234, 1990.
80. D. Vakhshoori, J. D. Wynn, and R. E. Liebenguth, *Appl. Phys. Lett.,* 65, 144, 1994.
81. D. Deppe and K. Choquette, in *Vertical Cavity Surface Emitting Lasers,* J. Cheng and N. K. Dutta, Eds., Gordon Breach, NY, 2000.
82. N. K. Dutta, L. W. Tu, G. J. Zydzik, G. Hasnain, Y. H. Wang, and A. Y. Cho, *Electron. Lett.,* 27, 208, 1991.
83. H. Soda, K. Iga, C. Kitahara, and Y. Suematsu, *Japan J. Appl. Phys.,* 18, 2329, 1979.
84. K. Iga, F. Koyama, and S. Kinoshita, *IEEE J. Quantum Electron.,* QE-24, 1845, 1988.
85. K. Tai, F. S. Choa, W. T. Tsang, S. N. G. Chu, J. D. Wynn, and A. M. Sergent, *Electron. Lett.,* 27, 1514, 1991.
86. L. Yang, M. C. Wu, K. Tai, T. Tanbun-Ek, and R. A. Logan, *Appl. Phys. Lett.,* 56, 889, 1990.
87. T. Baba, Y. Yogo, K. Suzuki, F. Koyama, and K. Iga, *Electron. Lett.,* 29, 913, 1993.
88. Y. Imajo, A. Kasukawa, S. Kashiwa, and H. Okamoto, *Japan J. Appl. Phys. Lett.,* 29, L1130, 1990.
89. D. I. Babic, K. Streubel, R. Mirin, N. M. Margalit, J. E. Bowers, E. L. Hu, D. E. Mars, L. Yang, and K. Carey, *IEEE Photonics Tech. Letts.,* 7, 1225, 1995.
90. Z. L. Liau and D. E. Mull, *Appl. Phys. Letts.,* 56, 737, 1990.
91. F. R. Nash, W. J. Sundburg, R. L. Hartman, J. R. Pawlik, D. A. Ackerman, N. K. Dutta, and R. W. Dixon, *AT&T Tech. J.,* 64, 809, 1985.
92. The reliability requirements of a submarine lightwave transmission system are discussion in a special issue of *AT&T Tech. J.,* 64, 3, 1985.
93. B. C. DeLoach, Jr., B. W. Hakki, R. L. Hartman, and L. A. D'Asaro, *Proc. IEEE,* 61, 1042, 1973.
94. P. M. Petroff and R. L. Hartman, *Appl. Phys. Lett.,* 2, 469, 1973.
95. W. D. Johnston and B. I. Miller, *Appl. Phys. Lett.,* 23, p. 1972, 1973.
96. P. M. Petroff, W. D. Johnston, Jr., and R. L. Hartman, *Appl. Phys. Lett.,* 25, 226, 1974.
97. J. Matsui, R. Ishida, and Y. Nannichi, *Japan J. Appl. Phys.,* 14, 1555, 1975.
98. P. M. Petroff and D. V. Lang, *Appl. Phys. Lett.,* 31, 60, 1977.
99. O. Ueda, I. Umebu, S. Yamakoshi, and T. Kotani, *J. Appl. Phys.,* 53, 2991, 1982.
100. O. Ueda, S. Yamakoshi, S. Komiya, K. Akita, and T. Yamaoka, *Appl. Phys. Lett.,* 36, 300, 1980.
101. S. Yamakoshi, M. Abe, O. Wada, S. Komiya, and T. Sakurai, *IEEE J. Quantum Electron.,* QE-17, 167, 1981.
102. K. Mizuishi, M. Sawai, S. Todoroki, S. Tsuji, M. Hirao, and M. Nakamura, *IEEE J. Quantum Electron.,* QE-19, 1294, 1983.
103. E. I. Gordon, F. R. Nash, and R. L. Hartman, *IEEE Electron. Device Lett.,* ELD-4, 465, 1983.
104. R. L. Hartman and R. W. Dixon, *Appl. Phys. Lett.,* 26, 239, 1975.
105. W. B. Joyce, K. Y. Liou, F. R. Nash, P. R. Bossard, and R. L. Hartman, *AT&T Tech. J.,* 64, 717, 1985.
106. S. Yamashita, A. Oka, T. Kawano, T. Tsuchiya, K. Saitoh, K. Uomi, and Y. Ono, *IEEE Photonic Tech. Lett.,* 4, 954, 1992.

107. I. Kotaka, K. Wakita, K. Kawano, H. Asai, and M. Naganuma, *Electron. Lett.,* 27, 2162, 1991.
108. K. Wakita, I. Kotaka, K. Yoshino, S. Kondo, and Y. Noguchi, *IEEE Photonic Tech. Letts.,* 7, 1418, 1995.
109. F. Koyama and K. Iga, *J. Lightwave Tech.,* 6, 87, 1988.
110. J. C. Cartledge, H. Debregeas, and C. Rolland, *IEEE Photonic Tech. Lett.,* 7, 224, 1995.
111. H. Takeuchi, K. Tsuzuki, K. Sato, M. Yamamoto, Y. Itaya, A. Sano, M. Yoneyama, and T. Otsuji, *IEEE Photonic Tech. Lett.,* 9, 572, 1997.
112. M. Aoki, M. Takashi, M. Suzuki, H. Sano, K. Uomi, T. Kawano, and A. Takai, *IEEE Photonic Tech. Letts.,* 4, 580, 1992.
113. T. L. Koch and U. Koren, *AT&T Tech. J.,* 70, 63, 1992.
114. R. Y. Fang, D. Bertone, M. Meliga, I. Montrosset, G. Oliveti, and R. Paoletti, *IEEE Photonic Tech. Lett.,* 9, 1084, 1997.
115. H. Yamazaki, K. Kudo, T. Sasaki, and M. Yamaguchi, *Proc. OECC'97,* Seoul, Korea, paper 10C1-3, 440–441.

47

Optical Transmitters

Alistair J. Price
Corvis Corporation

Ken D. Pedrotti
University of California

47.1 Introduction

In the last 15 years, optical fiber transmission systems have progressed enormously in terms of information handling capacity and link distance. Some major advances in transmitter technology, both electronic and optoelectronic, have made this possible. The single longitudinal mode infrared emission from semiconductor distributed feedback (DFB) laser diodes can be coupled efficiently into today's low-loss silica single-mode fibers. In addition, these lasers can be switched on and off with transition times on the order of 100 ps to provide data rates up to approximately 10 Gb/s. The outstanding performance of these laser diodes has, in fact, provoked considerable effort to improve the switching speed of digital integrated circuits in order to fully utilize the capability of the optoelectronic components.

The performance of directly modulated laser transmitters greatly surpasses that of light emitting diode (LED) transmitters both in terms of optical power coupled into the fiber and switching speed. Short-distance (<10 km), low-capacity (<1 Gb/s) links use LED transmitters. For these applications, such as for local area networks, for example, LED transmitters can provide a cost-effective alternative to the laser.

Where high bit rates and a large regenerator spacing are required, optical pulse distortion occurring in the fiber between the transmitter and receiver is very important. Because of the advent of the optical amplifier, these requirements are now commonplace in terrestrial as well as submarine communication systems because amplifiers can cost-effectively eliminate the loss of very long fiber links. With loss out of the way, distortion becomes a major issue. Phase/frequency shifts accompany the switching of semiconductor lasers. Commonly referred to as **chirp**, these shifts cause rapid deterioration of the pulse shape in dispersive optical fibers, thereby limiting the reach of directly modulated transmitters. For example, at a bit rate of 2.5 Gb/s, a distance of 200 km represents the limit of what is currently achievable using conventional single-mode fiber (assuming an operating wavelength in the 1550 nm low-loss window). Higher performance transmitters use external modulators to exceed these limits. External modulators exist which produce little or no switching phase shift.

The following sections include some information about the characteristics of laser diodes used in directly modulated optical transmitters and the electronic circuits to drive them. A brief description of packaging technology is also provided. A section is included on the performance characterization of high-speed digital lightwave transmitters. The use of external modulators introduces a new set of design requirements. These are also discussed.

Whereas digital transmission dominates the field of optical fiber transmission systems, analog systems are also important for the transmission of cable TV signals and microwave signals for remote antenna location and phase array radar systems, for example. The requirements for analog transmitters depend on the type of signal to be transmitted and are generally quite different from those of digital systems. Properties which are not very important in digital systems, such as linearity and signal-to-noise ratio, are very important in analog systems because, once degraded, the signal quality cannot be restored. More information about analog transmission is contained in Kolner and Dalfi [1987], Atlas [1996], and Linder et al. [1995].

47.2 Directly Modulated Laser Transmitters

The design of a state-of-the-art laser transmitter requires the combining of several different technologies. The laser diode is, of course, the critical component whose properties govern the transmitter performance. Its electrical and optical properties are both very important. The electrical driver circuit and the link between it and the laser are also critical design areas. Finally, the package design is a specialized combination of a microwave and optical enclosure.

Laser Diode Characteristics

The laser diode must launch optical pulses into the optical fiber at the required wavelength, with high spectral purity, with a high **extinction ratio**, and with a low-intensity noise. These properties have been discussed in a previous chapter. The optical format is typically **non-return to zero (NRZ)**, but **return-to-zero (RZ)** format is also used. The ideal optical signal will reproduce the data stream faithfully without **jitter**, ringing, or chirp. These cannot be eliminated entirely, and each contributes to the overall system performance.

A look at the **light, current (L/I) characteristic** of a DFB laser reveals a lot about its basic character (see Fig. 47.1). The light output is relatively small until the current reaches a certain value, called the **threshold current**. Thereafter, the optical intensity rises approximately linearly with increasing current. For digital modulation, the current to the laser switches between two levels, the 0 level current being near the threshold current and the 1 level current being higher. The problems associated with typical

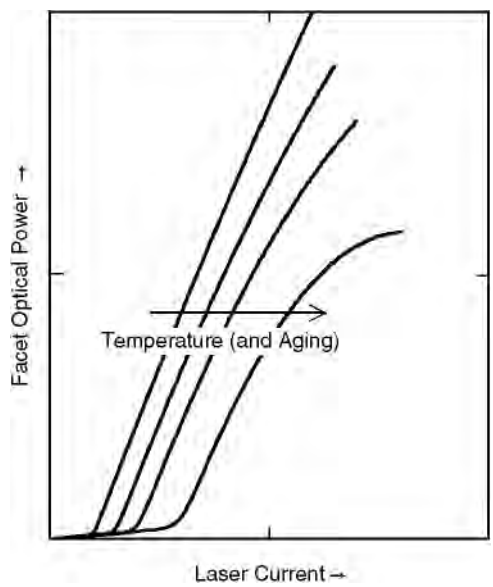


FIGURE 47.1 L/I (light vs. current) characteristic of a DFB laser.

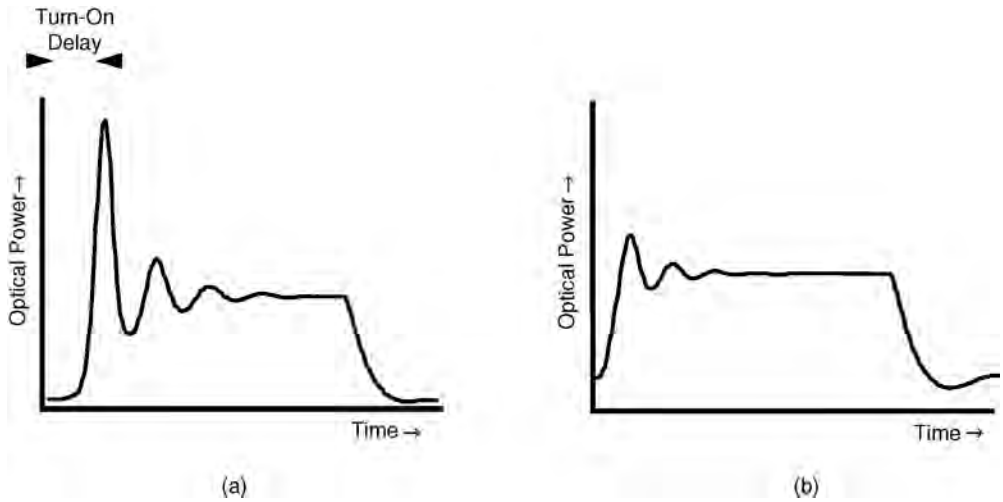


FIGURE 47.2 Switching waveforms of DFB lasers: (a) logic 0 bias below the threshold current and (b) logic 0 bias above the threshold current.

laser diodes are that the curve bends over at high current and tends to shift and bend to the right with increasing temperature and as the laser ages. In addition, the L/I characteristic often exhibits small kinks, which are usually related to the level of optical power reflected back into the device [Twu et al., 1992]. Wavelength fluctuations accompany these kinks. In high bit rate systems, an **optical isolator** prevents reflections from disturbing the laser. The changes due to temperature and time can, to a large extent, be countered using electronic feedback circuitry, as described in the next section. Often, a temperature-controlled laser mount is also used. Keeping the temperature constant also reduces wavelength drift, and a low temperature extends the laser's lifetime.

Measurement of laser output while switching the input current between two levels indicates that the dynamic response of these diodes is not straightforward. When switching from low optical power to a higher power, there is first a turn-on delay, which has a random component. A rapid increase in power follows, ringing at the **relaxation resonance** frequency. An unwanted frequency modulation accompanies the ringing. The turn-off is a smooth exponential decay, which is somewhat slower than the turn-on transition time. Typical waveforms for various current levels are shown in Fig. 47.2. The turn-on delay decreases as the off-state (bias) current is increased. For many digital applications, the bias is carefully set close to the laser threshold current to minimize the turn-on delay while at the same time providing a high-extinction ratio.

The dynamic behavior just described can be modeled using the **laser rate equations**. The single-mode rate equations are frequently presented as:

$$\frac{d}{dt}N = \frac{I_A}{q \cdot V_{\text{act}}} - \frac{N}{\tau_n} - g_0 \cdot (N - N_{0g}) \cdot (1 - \varepsilon \cdot S) \cdot S \tag{47.1}$$

$$\frac{d}{dt}S = \Gamma \cdot g_0 \cdot (N - N_{0g}) \cdot (1 - \varepsilon \cdot S)\frac{1}{\tau_p} \cdot S + \Gamma \cdot \beta \cdot \frac{N}{\tau_n} \tag{47.2}$$

where N = electron density; q = electron charge; τ_n = electron lifetime; N_{0g} = electron density for zero gain; S = photon density; τ_p = photon lifetime; I_A = current injected into active layer; V_{act} = volume of active layer; g_0 = differential gain constant; ε = gain compression factor; Γ = optical confinement factor; β = spontaneous emission factor.

Computer programs intended for nonlinear circuit analysis, such as the Berkeley SPICE simulation program, can be used to solve these equations by defining a circuit to represent them [Tucker, 1985; Wedding, 1987]. The electron and photon densities are represented by circuit node voltages. The technique is powerful because it enables additional electronic circuitry to be included in the analysis. For example, the driver circuit, bond wires, and the laser electrical parasitics, as well as the rate equations, can all be modeled in the same analysis. This provides a versatile tool for analyzing the dynamic properties of laser transmitters.

Electronic Driver Circuit Design

A simple approach to driving a laser diode is to connect the output of a current driver circuit directly to the laser diode, as shown in Fig. 47.3. This approach is frequently adopted when the operating frequency is below 1 GHz. However, the inductance of the bond wires combined with the capacitance associated with the driver circuit and the laser can significantly degrade the risetime and cause ringing in the current pulses. This approach is difficult at higher frequencies because it might be impossible to reduce the inductance sufficiently. For narrow-bandwidth designs, sometimes a matching circuit can be used to cancel the effect of the inductance, but for very high-speed digital designs, a different approach is required. Flip-chip laser mounting on top of the driver circuit [Pedrotti et al., 1992] and fully integrated lasers and drivers [Eriksson et al., 1996] are the subjects of ongoing research, but typically some form of transmission line is needed between driver and laser. This leads to new requirements for the driver circuit. It is generally difficult to implement a characteristic transmission line impedance as low as the typical dynamic resistance of a laser (3–8 Ω), and so a technique frequently used is to increase its load resistance to a more manageable value using a closely mounted series resistor. The series resistor increases the power dissipation of the resultant load for a given modulation current and also the required voltage swing. The former has adverse consequences related to laser thermal effects, and the latter affects the driver circuit requirements. Low impedances are normally used (in the range 12–25 Ω), although, when the driver circuit is completely separated from the laser package, the standard 50-Ω transmission line impedance is frequently used for convenience, allowing the use of standard coaxial connectors (such as sub-miniature A (SMA) type) and commercial amplifiers to drive the laser. Testing is also simplified because most high-frequency test equipment uses 50-Ω interconnects as standard. An example of a current switch driver circuit with transmission line link to the laser is shown in Fig. 47.4. The modulation current is equal to the current in the tail of the differential output pair, which is controlled by the voltage at the *Vgmod* input. The laser bias current (the logic 0 current) is determined by the current in the separate bias source. One example of a high-speed GaAs hetero-junction bipolar transistor (HBT) laser driver circuit is that reported in Runge et al. [1992], which is a monolithic design with several stages of current amplification

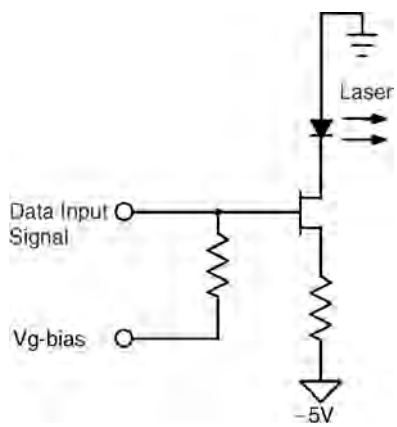


FIGURE 47.3 FET laser driver directly connected to laser diode.

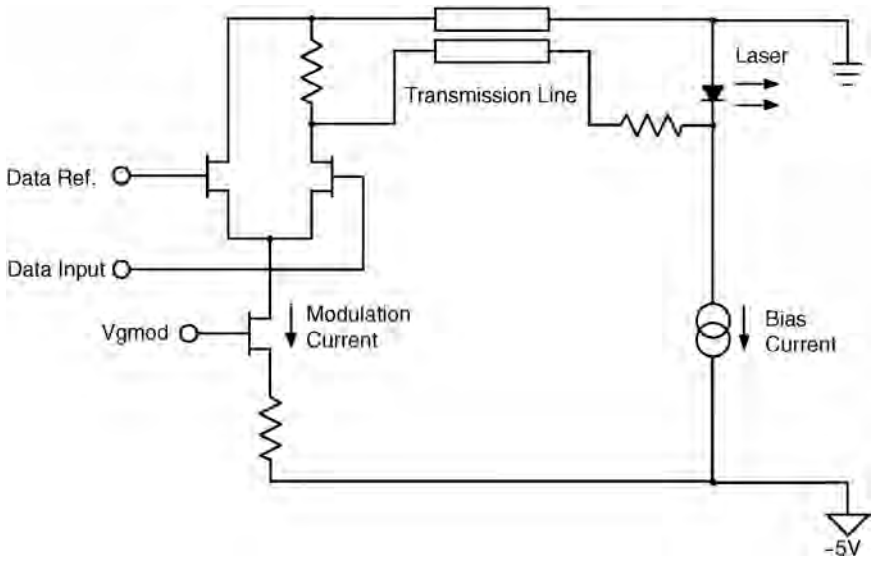


FIGURE 47.4 Current switch laser driver with transmission line coupling to laser diode.

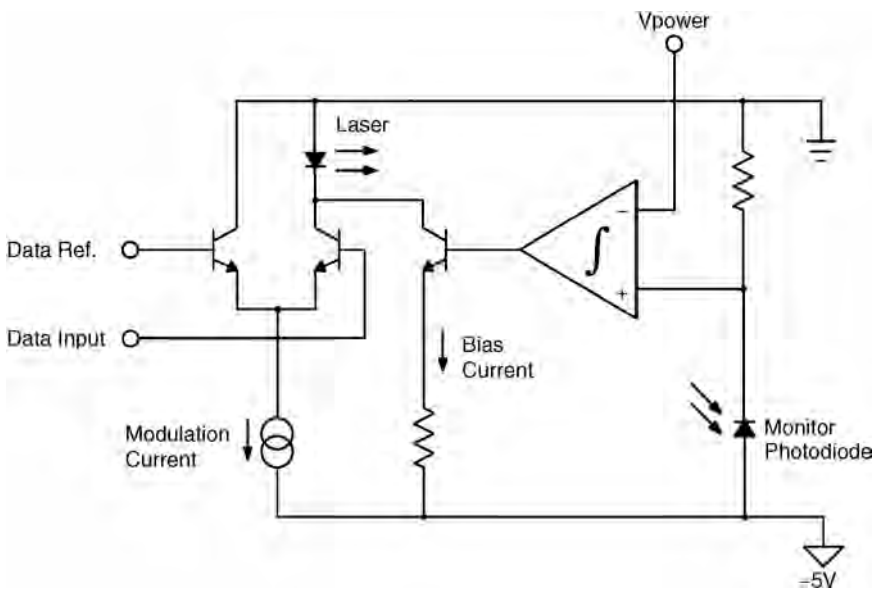


FIGURE 47.5 Average power control circuit.

for the laser drive. This circuit includes a Cherry–Hooper amplifier stage driving a cascaded differential pair output driver.

Typically, various feedback loops are used in practical laser transmitter designs to compensate for the behavior of laser diodes when influenced by variations in data rate, temperature, and aging effects. The simplest average power control circuit is shown in Fig. 47.5. The laser is connected in the collector lead of a typical differential pair driver stage. The data signal is applied to one side of the differential pair and a reference voltage is applied to the other. In addition, the drive transistor is connected in parallel with another transistor, which serves as a bias current source. This current source is controlled by the output of a monitor photodiode. A portion of the output power from the laser, typically taken from the laser's

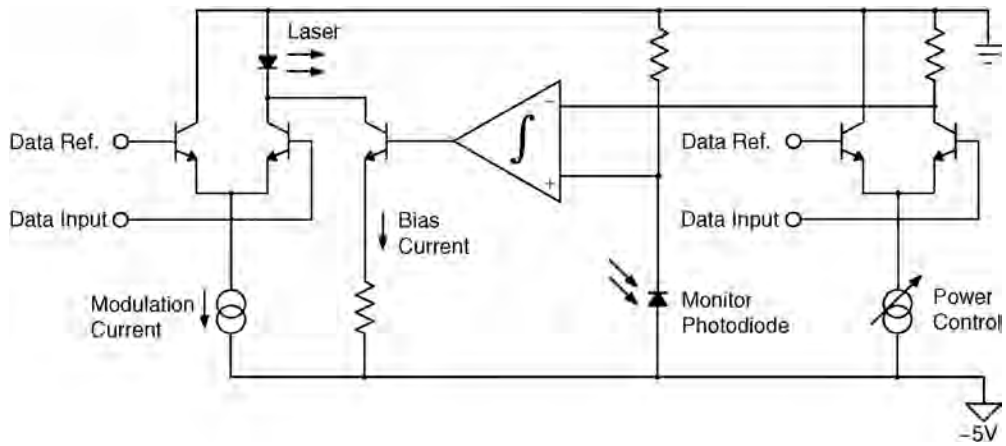


FIGURE 47.6 Average power control with mark density compensation.

rear facet, is incident on this detector. The circuit keeps the 0 and 1 levels constant with variations in threshold current as long as the slope of laser's L/I characteristic does not change.

The average number of 1s relative to the number of 0s, the mark density, is linearly related to the average power. If the mark density of the data changes, then the laser bias point will shift using the circuit of Fig. 47.5. In the extreme case of a long string of zeros, the average power feedback circuit will continue to increase this zero level up to the setpoint for the average power. To alleviate this problem, a circuit similar to that shown in Fig. 47.6 can be used. The V_{power} reference input of the average power feedback circuit is replaced with a signal proportional to the mark density. Now, in the case of a long string of zeros, the reference voltage increases, causing the laser bias current to decrease, maintaining the correct bias current to the laser. This approach now compensates for threshold current shifts and mark density variation but not for changes in the slope of the L/I curve.

An approach that allows slope correction is given in Smith [1978]. To detect whether the zero level is near threshold, the amplitude of the current is modulated at a low frequency when the data input is zero. By detecting the slow ripple from the photodiode output at this modulation frequency, a signal whose amplitude is sensitive to the L/I curve slope at the logical zero level is obtained. Since the slope changes abruptly around the threshold current, the logical zero can be maintained close to the threshold current using a feedback circuit. When used in conjunction with an average power controller, a constant amplitude high-extinction ratio optical signal can be maintained for changes in both threshold current and slope.

For high bit rate transmitters, sometimes the 0 level current must be set significantly above the threshold current in order to reduce turn-on delay, ringing, and chirp. The resulting extinction ratio penalty is often a small price to pay for the improved system performance when chromatic dispersion is present. In such circumstances, the circuit shown in Fig. 47.7 may be used [Chen, 1980]. Here, in addition to average power control, the AC component of the detected current is fed back to the modulation control of the laser driver. By keeping this constant, slope variations are compensated, and the extinction ratio is held approximately constant. Mark density compensation can be incorporated, as shown in Fig. 47.6. A disadvantage of the circuit shown is that a broadband monitor photodiode might be needed. This can be avoided by adding to the bias current a small low-frequency tone with an amplitude proportional to the modulation current. This can be detected accurately with low-frequency circuitry to obtain a measure of the optical modulation amplitude.

Various other methods have been investigated for use in the stabilization of laser output. Among these are encoding to ensure a suitable balance of 1s and 0s in the circuit, encoding both clock and data, and using the clock component to adjust the laser drive condition. Albanese [1978] developed a method to monitor the laser threshold without the use of a monitor photodiode by monitoring the change in laser electrical characteristics near the threshold.

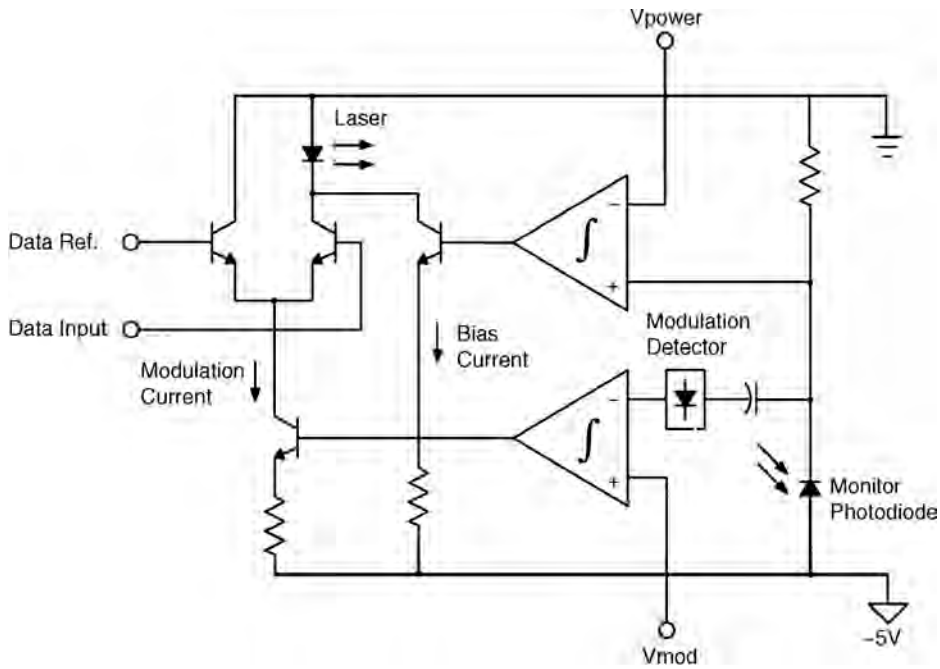


FIGURE 47.7 Average power and modulation amplitude control circuit.

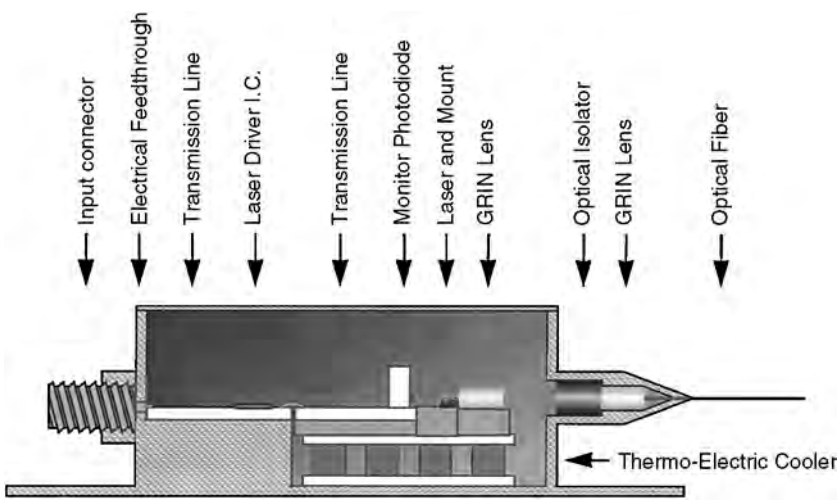


FIGURE 47.8 Laser package.

Optoelectronic Package Design

An illustrative example of a laser diode package is shown in Fig. 47.8. In this particular example, the laser diode is mounted on a diamond heat spreader, which is mounted on a ceramic carrier. Lasers such as these, which may be purchased from many vendors, allow ease of handling, testing, and replacement without handling the device itself. Laser and carrier are mounted on a **thermoelectric cooler** with a thermistor to monitor the temperature and a large-area optical detector, which is used to monitor the optical power by intercepting the backward wave emission of the laser. It is generally preferred to limit heat generating components mounted on the cooler to a minimum to keep down the cooling requirements.

Frequently, the temperature controller will be set to keep the temperature near 25∞C or a little higher in order to maximize the range over which stabilization can be maintained (the cooler can be used to heat as well as cool). The electrical link from the driver circuit is provided by a transmission line, as shown in the diagram. This critical link is designed carefully, taking into account dimensional tolerances in order to repeatedly obtain good performance. To couple power into the optical fiber, a graded index (GRIN) or spherical lens is mounted close to the front facet of the laser to produce a collimated beam. The beam passes through a small optical isolator and is focused into the fiber using a second lens. An angled optical arrangement is used to reduce reflections. The alignment and stability of this arrangement is critical to achieve a low coupling loss of about 3 dB. For applications that do not require the use of an isolator to reduce reflections, adequate coupling efficiency may be obtained by directly coupling from the laser into the fiber. The end of the fiber is normally tapered, and the tip is formed into a convex lens to match the laser light into the fiber.

Transmitter Performance Evaluation

The methods of evaluating digital optical transmitters and systems have, to a certain extent, been standardized in recent years due to the introduction of synchronous optical network (SONET) and synchronous digital hierarchy (SDH) optical systems standards [Bell, 1994]. One of the purposes of these standards is to specify a set of requirements to allow interoperability of equipment from different manufacturers. An additional objective is that transmitters from one manufacturer operate with receivers made by another manufacturer. Clearly, to achieve this, their characteristics must be sufficiently well specified.

Evaluating the performance of an optical transmitter requires certain measurements of the optical signal produced. Using a broadband optical detector, an electrical signal proportional to the optical intensity is obtained, allowing the intensity modulation to be examined using standard electrical measurement techniques. The signal is filtered using a low-pass filter with well-defined characteristics to provide a standard receiver response. For example, a fourth-order Bessel–Thompson filter with a 3-dB bandwidth equal to 0.75 of the bit rate is specified for testing SONET transmitters. Using an oscilloscope, the signal can then be displayed in the form of an **eye pattern**. The signal shape must meet certain criteria defined by an eye pattern mask (see Fig. 47.9). Measurement software overlays these masks on the oscilloscope screen, measures violations of the limits specified by the masks, and also measures the other characteristics such as extinction ratio and jitter.

Eye pattern measurements of the transmitter output signal do not give any indication of phase or frequency modulation of the optical signal. These, however, are critically important for long-distance

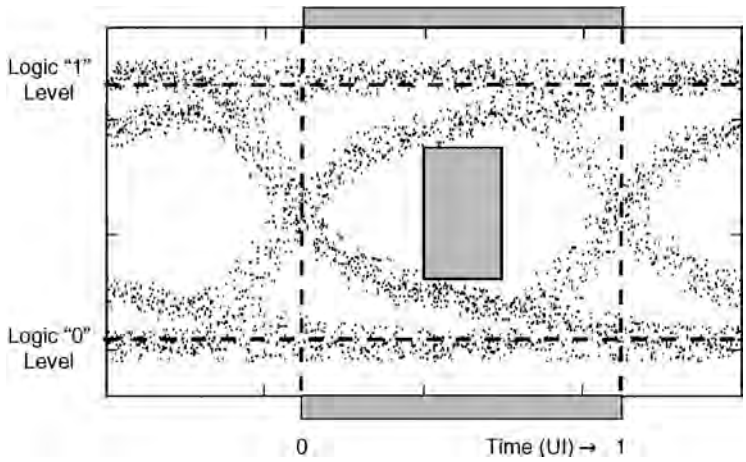


FIGURE 47.9 Eye pattern mask test.

transmission through dispersive fiber. The spectral width is one indicator which is helpful here. A wide spectrum is very susceptible to the effects of chromatic dispersion in the transmission medium. For this reason, a specification is imposed on the spectral width of the transmitted signal. In addition, degradation of the signal must be tested by passing it through a length of optical fiber. Typically, the signal power required at a reference optical receiver input, for example, defined by a 10^{-10} error rate, when the signal is passed through a given length of optical fiber, is compared with the power required when the transmitted signal is passed through a very short length of fiber. The optical power going into the receiver is adjusted using an optical attenuator. The measured power ratio is usually expressed in decibels of **dispersion power penalty** for a given chromatic dispersion. A small value is indicative of a high-performance, low-chirp transmitter.

As transmission rates increase, the impairments due to laser chirp combined with chromatic dispersion as well as other signal degradation mechanisms become more severe. One solution to this problem is to use lower chirp modulator sources, as investigated in the following section. Another that can be used in conjunction with either lasers or external modulators is to use forward error correction (FEC) encoding. The fundamentals of FEC are dealt with in detail elsewhere in this volume. Encoding the signal by the addition of extra bits at the transmitter, their values depending on the bit pattern to be sent, allows the receiver to not only detect the existence of errors but also to correct them. For example, using the ITU-G.975 optical coding standard which is a Reed–Solomon (255b/239b) code, a received signal with a bit error rate of $\sim 8 \times 10^{-5}$ with no coding can be received with $\sim 1 \times 10^{-15}$ BER using the FEC code. This improvement is sometimes expressed in terms of coding gain, which is the difference in received power with and without FEC necessary to achieve a certain BER. This gain depends on the BER used for comparison, but improvements of ~5 dB are common. This transmission improvement can be used to increase the transmission distance or to allow the use of less expensive components or fiber with wider specification variations. This improvement in transmission comes at the cost of the addition of extra bits and requires that the transmitter operate at a higher bandwidth, 7% higher in the case of the ITU standard. At 40 Gbit/s, the problems associated with limited electrical bandwidth, device imperfections, and fiber nonlinearities require the use of FEC in the transmitter to allow reasonable transmission distances with practical components. The application of other coding schemes with greater coding gain but more complicated coding and decoding algorithms are the topics of current investigation.

47.3 Externally Modulated Optical Transmitters

Absorptive modulators and interferometric types are considered here. Because the former type can be made of the same materials used in the fabrication of lasers, the popularity of these lies in the possibility of integrating them with the laser diode to save cost [Watanabe, Sato, and Soda, 1992]. The most commonly used interferometric type, the **Mach–Zehnder**, normally uses the electro-optic properties of lithium niobate to provide very precise control over the amplitude and phase of the signal, for high-performance applications.

Characteristics of External Modulators

Electroabsorption (EA) modulators are monolithic semiconductor devices, usually indium phosphide (InP) based. They rely on either the Franz–Keldysh effect (FKE) or the quantum-confined Stark effect (QCSE) and have optical waveguide p-i-n structures grown epitaxially, typically using the InGaAsP-InP or InGaAs-InAlAs material systems [Chin, 1995; Kawano et al., 1992]. Compatibility with laser structures makes the integration of laser and modulator possible and very attractive. The advantages of using these modulators compared with direct laser modulation are that the chirp associated with the laser relaxation resonance is avoided (although there is some inherent chirp in the modulators (Koyama and Iga, 1988]), the input impedance is high, and these modulators are capable of very high-speed modulation.

A comprehensive comparison of the relative merits of different types of EA modulators is beyond the scope of this text. Some of the factors that affect their suitability in different applications are drive level

requirements, extinction ratio, bandwidth, operating wavelength range, chirp, temperature stability, and, for analog systems, linearity.

Electrically, EA modulators are diodes that operate in reverse bias. The absorption is dependent on the applied voltage. For high-frequency applications, the capacitance of the device must be considered. A sufficiently low-driving impedance is used so that the low-pass pole caused by the capacitor does not limit the bandwidth too much. A transmission line link to the drive circuit will frequently be required. In that case, the line could be terminated using a resistor connected in parallel with the modulator at the end of the transmission line to prevent reflections. For bandpass applications, microwave matching techniques are used to improve the modulation efficiency. The bandwidth can also be improved by incorporating the device capacitance into a lumped transmission line structure. The electrical techniques appropriate to the use of these devices are those that also apply to many other high-impedance devices such as FETs and varactor diodes.

Mach–Zehnder modulators operate using interference by splitting the input into two paths, which are later recombined (see Fig. 47.10). Using lithium niobate or another electro-optic material for the optical waveguides, the phase delay is dependent on the electric field applied to the waveguides. Phase modulators

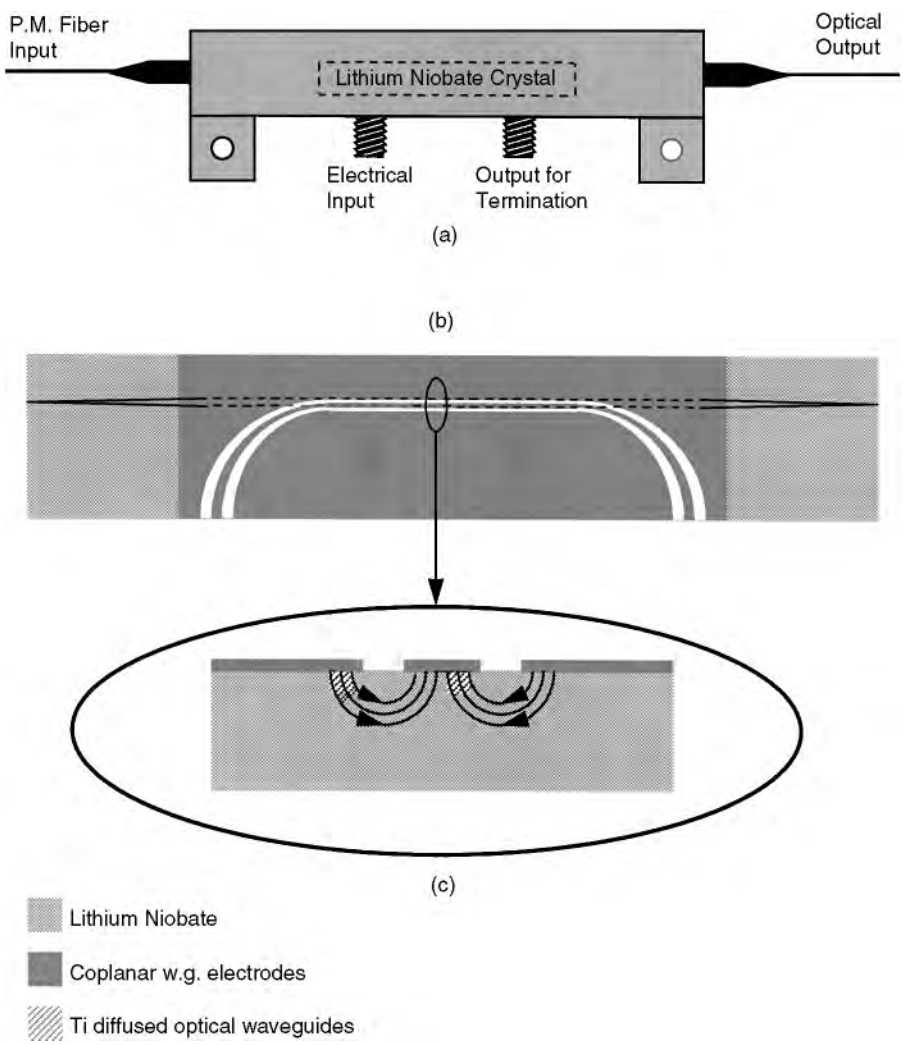


FIGURE 47.10 Mach–Zehnder modulators: (a) package outline, (b) lithium niobate crystal, and (c) detail.

can, therefore, be incorporated in one or both of the split paths allowing electrical control of the phase of the two signals. Depending on their relative phases when they are recombined, constructive or destructive interference occurs, resulting in amplitude modulation. The phase modulator electrodes are usually formed into transmission lines following the path of the optical waveguides. In this way, over a certain path length, an applied electrical pulse will continue to modulate an optical signal as both waves progress along their respective waveguides. If the velocities are not well matched, which is usually the case, the optical and electrical waves will not track each other over a long distance and the high-frequency response is affected. For a given required bandwidth, the length of interaction must be limited, which in turn limits the phase shift induced by a given applied voltage. Typically, therefore, high-bandwidth designs require relatively high-drive voltages compared with lower bandwidth designs. The drive voltage requirements are usually denoted by a value for $V\pi$, the change in voltage required to produce π-rad phase shift between the two paths. This corresponds to the voltage needed to go from maximum to minimum amplitude. Typical values are in the order of 10 V for a device with a 10-GHz modulation bandwidth, though this is continually being improved. The extinction ratio for these devices, the ratio of maximum to minimum value expressed in decibels, is dependent on the amplitude balance in the two paths, as well as the level of other unwanted transmission modes through the device. Typical values are >20 dB.

In modulators for optical transmitters, the path lengths may be several centimeters, equivalent to many wavelengths of the light passing through them. Although great care is taken to maintain device symmetry, the path lengths are not generally identical, and so the required voltage for maximum amplitude transmission is usually some random value. In addition, temperature gradients, stresses, and the presence of impurities in the optical waveguides all have an effect on the relative phase. For these reasons, care is taken in the packaging and package mounting to avoid mechanical stresses. The terminating resistors for the electrical transmission lines are usually isolated from the lithium niobate crystal to prevent heating. Measures are taken to minimize the absorption and subsequent diffusion of impurities in order to reduce long term drifts in the phase difference.

Various transmission line electrode arrangements are possible. Commonly, a single 50-Ω coplanar waveguide is designed to apply electric fields to both optical waveguides simultaneously such that one optical path is retarded while the other is advanced. If the phase shifts are equal but opposite in the two paths, the resultant amplitude modulation is free of phase modulation (chirp); whereas if only one path is phase modulated, the resultant modulation contains both amplitude and phase modulation components. Researchers have discovered [Gnauch et al., 1991] that under certain conditions of chromatic dispersion, transmission distances are actually enhanced if a small amount of phase modulation accompanies the amplitude modulation by comparison with a completely chirp-free modulation, so now some commercial modulators provide this optimum amount of chirp. Another type of modulator simply provides two separate electrical inputs, each one controlling one arm of the Mach–Zehnder. By controlling the relative amplitude and sign of the two inputs, the amount and direction of the accompanying chirp can be precisely controlled. In addition, a differential output driver circuit need supply only half the voltage swing to each input compared with single-ended drive.

Chirp control is a very important issue when high bit rates are combined with long-distance transmission through dispersive fiber. Another approach to phase control is to add a phase modulator in cascade with the Mach–Zehnder. It can, of course, be integrated with the Mach–Zehnder by extending it and incorporating another electrode on the input or output waveguide. The drive waveforms must then be coordinated between the Mach–Zehnder and the phase modulator to obtain the required simultaneous phase and amplitude modulation. Allowance must be made for the optical delay between the two modulators, and delay must be provided in the drive signals to compensate.

Whichever electrode configuration is used, the output intensity vs. input voltage is usually the same shape for all Mach–Zehnder modulators. It is shown in Fig. 47.11, and is sinusoidal in character. For digital modulation, the 1 level is aligned with a peak and the 0 level is aligned with an adjacent trough. Restated, the modulator is biased to a half-way point and a modulation swing equal to $V\pi$ (peak-to-peak) is applied. If the modulator has chirp, the sign of the chirp will depend on the sign of the slope, and so the correct side of the peak must be used. Some advantages of the sinusoidal characteristic are

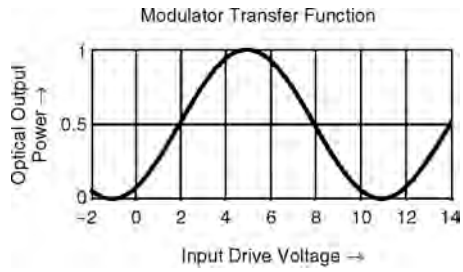


FIGURE 47.11 Electrical/optical (E/O) transfer function of Mach–Zehnder modulator.

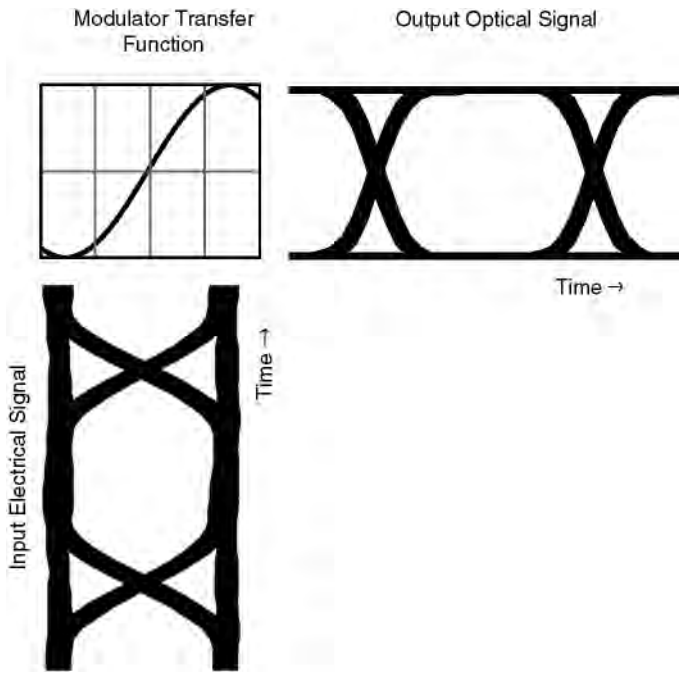


FIGURE 47.12 Effect of E/O transfer function on optical waveform.

readily apparent. First, because the transfer function is flat at both the peaks and the troughs, noise or spreading of the electrical logic levels will have little effect on the optical intensity levels. Second, the gradient of the transfer function is steepest at the bias point. Provided the bandwidth of the modulator is sufficiently high, this can result in a reduction in the output optical transition times compared with the electrical transition times of the drive waveform. The two factors together give rise to an optical output waveform, which is superior in many ways to the electrical waveform used to generate it. This is in marked contrast with the case of direct laser modulation, where the output optical waveforms not only contain the imperfections of the drive waveform, but also include some added aberrations caused by the laser itself. The waveform improvement is demonstrated in Fig. 47.12.

For most intensity modulation (IM) transmitters, the bias point is at the half-maximum power point, corresponding to phase quadrature between the two paths. The intensity vs. drive voltage characteristic is symmetrical, and the curve is most linear at this point. For most direct detection systems (those in which the received optical power is directly converted to a proportional electrical signal), this is the preferred operating point. However, for some applications, the modulator can be biased at the minimum

amplitude point. The modulator then behaves very much like the balanced mixers used, for example, as upconverters in microwave radio systems. An electrical signal applied to the modulator will be reproduced at optical frequencies in two sidebands on either side of the (suppressed) optical carrier. This means that many of the techniques used to generate radio signals can also be applied to the generation of optical signals. The upconversion is quite linear (sinusoidal in fact) so that the electrical signal properties, both time domain and frequency domain, are preserved in the conversion. If the electrical signal is filtered to reduce its electrical bandwidth, the optical sidebands will have the same bandwidth. Two references provide examples of such a use, Izutsu et al. [1981] and Price and Le Mercier [1995].

Modulator Driver Circuit Design

A common feature of EA modulators and Mach–Zehnder modulators is that electrically, by contrast with laser diodes, they present a high impedance to the driving circuit. A low-output impedance circuit, such as an emitter follower, is therefore appropriate for situations where the driver is connected directly to the modulator. However, as is true for laser driver circuits, the requirements change somewhat when the driver must be separated from the modulator, or when the bond wire inductance becomes a significant factor. Then resonances must be dampened using an appropriate driving resistance or load resistor. In the case of a transmission line link between driver and modulator, or when the modulator uses transmission line electrodes, it is essential to either resistively terminate the line, drive the line using a resistive source, or do both. Again, a 50-Ω transmission line is frequently chosen for convenience and standardization.

When transmission lines are used, the main driver amplifier requirements for both laser drivers and modulator drivers become very similar: they must generate electrical pulses of the required power in the transmission line. The current switch topology of Fig. 47.4 is appropriate for modulators as well as lasers. The main difference is that the required voltage swings are significantly greater for modulators, especially existing Mach–Zehnder designs, compared with laser drivers. For this reason, relatively highpower broadband 50-Ω amplifiers are frequently used to drive modulators. Power output requirements between 20 and 30 dB (1 mW) are typical. Hybrid GaAs metal-semiconductor field-effect transistor (MESFET) designs are used. Integrated circuit solutions include distributed amplifiers.

Automatic control algorithms for Mach–Zehnder modulators contrast those of other optical transmitters because the transfer characteristic (see Fig. 47.11) is quite different and there exists the potential for multiple operating points and an unlimited lateral drift due to temperature and aging. Typically, some optical power is tapped from the output of the modulator and detected in a low-bandwidth optical detector. One method to automatically control the bias point of the modulator is to apply a low-frequency modulation to the gain of the driver amplifier. A signal will be generated in the detector if the bias point is not at the quadrature point as intended. By comparing the phase of the detected signal with the original low-frequency modulation, using a synchronous detector, an error signal can be obtained, which is used to force the bias point to the quadrature point. Inverting the sense of the feedback loop forces the bias point to move to the adjacent quadrature point of opposite slope, causing an inversion of the optical data and the chirp. Methods also exist to optimize the modulation amplitude in order to maintain a good optical output waveform. A problem with some Mach–Zehnder modulators is that, over a period of time, the phase shift might continue to drift over many cycles of phase. An automatic bias circuit is not able to continue to apply higher and higher voltages to track this drift, so a warning circuit is sometimes included to indicate that the voltage might soon reach a limit, after which it would fail to track. At a convenient time, the circuit is reset by remote control before failure occurs.

Most fielded optical systems today rely on NRZ encoding. However, for long distance transmission such as for undersea links, soliton transmission has for many years been a topic of serious research interest [Agrawal, 1997; Mollenauer et. al., 1997]. A soliton is a particular type of pulse that maintains its shape during propagation in a dispersive medium. This is the result of a balancing between the effects of nonlinear self phase modulation due to the Kerr nonlinearity in the fiber with fiber dispersion. To launch a soliton, a rather high power (sufficient to bring the self phase modulation to an appropriate level) and

narrow pulse of a particular shape is launched into the fiber. Transmitters to create sequences of suitable pulses can be realized using mode-locked fiber ring lasers, gain switched DFB lasers, mode-locked semiconductor lasers in hybrid or monolithic extended cavities, and modulators injected with CW beams. Information is typically encoded onto these pulse trains using another Mach–Zehnder interferometer to either pass or block the individual pulses, resulting in an RZ encoded pulse train. The spacing between these pulses is much greater than their width so that many such pulse trains can be interleaved in time to create very high bit rate time-division multiplexed signals. New fibers, transmitters, and receivers are resulting in commercialization of the first such soliton systems.

Another related application of RZ modulation is emerging for high bit rate systems operating at rates of 10 Gb/s and higher. Here, the use of a similar approach to externally modulate a train of short pulses provides greater link lengths at high bandwidths although true solitons are not used. The short pulses are generated by the use of CW light injected into an external or monolithically integrated modulator. With appropriate modulator bias and modulation [Veselka, et. al., 1996], a series of short pulses with accurately controlled repetition rate can be generated. These can then be modulated using another amplitude modulator to create the RZ signal. The advantage for high bit rate systems is that the first modulator can operate narrowband at the pulse repetition frequency using a sinusoidal electrical signal. The resulting narrow pulses define the edge rates of the transmitted pulses; the second modulator, which must operate with broadband data, can now have relaxed bandwidth and edge rate response without incurring excessive inter-symbol interference (ISI), which can also ease the design of the modulator driver electronics. Further, the launch power can be increased, increasing regenerator spacing, and some advantage can be obtained from the interplay of the self phase modulation and fiber dispersion. Additionally, 2-5 dB of receiver sensitivity improvement can be realized by the use of this impulsive encoding, depending on the sources of noise in the transmission channel and receiver [Winzer, Kalmar, 1999; Boivin et. al. 1997].

Defining Terms

Chirp: A frequency or phase shift associated with the switching on or off of a laser diode or modulator.

Dispersion power penalty: The power penalty in decibels associated with the increase in receiver optical power level required to achieve a specified error rate, for example, 10^{-10}, when the signal is subjected to a specified amount of chromatic dispersion.

Electroabsorptive (EA) modulator: Semiconductor modulators using the Franz–Keldysh effect (FKE) or the quantum-confined Stark effect (QCSE).

Extinction ratio: The ratio of maximum to minimum insertion loss for an amplitude modulator. The term also applies to the ratio (in decibels) of the average optical energy in a logic 1 level to the average optical energy in a logic 0 level of a digitally modulated optical signal.

Eye pattern: Picture built up by overlaying traces corresponding to all possible sequences of a digital waveform (see Fig. 47.9).

Jitter: Random and systematic deviations of the arrival times of pulses or transitions in digitally transmitted data. Observed as a horizontal spreading of an eye pattern displayed on an oscilloscope (see Fig. 47.12).

Laser rate equations: Set of equations representing the dynamic behavior of electron and photon population densities in a laser.

Laser threshold current: Current threshold at which the gain inside the laser cavity exceeds the losses. Below the threshold current, light emission is spontaneous. Above the threshold current, stimulated emission also occurs.

L/I characteristic: Graph of light output vs. injected current, which exhibits a characteristic knee at the laser threshold current (see Fig. 47.1).

Mach–Zehnder modulator: A two-path optical interferometer that uses phase shifters in one or both paths to provide amplitude modulation.

Non-return to zero (NRZ): Digital format in which the logic 1 pulses occupy the whole bit period (cf. return to zero).

Optical isolator: A device that utilizes the Faraday effect to allow transmission through the device in one direction only. These devices are used to isolate lasers from the effects of reflections.

Relaxation resonance: Internal resonance between electrons and photons in a laser diode.

Return to zero (RZ): Digital format in which the logic 1 pulses occupy a fraction (usually 50%) of the bit period.

Thermoelectric (TE) cooler: Device utilizing the Peltier effect to pump heat from one surface to the other. Reversing the direction of current flow reverses the direction of heat flow, so that the same device may be used to heat or cool as necessary.

Vπ: When referred to a phase modulator, it is the voltage change required to give a phase shift of π rad (180∞). For a Mach–Zehnder modulator, this corresponds to the voltage required to change from maximum to minimum extinction.

References

Albanese, A. 1978. An automatic bias control (ABC) circuit for injection lasers. *Bell Syst. Tech. J.,* 57:1533-1544.

Agrawal, G. P., Chang, K. (editors). 1997. Soliton communication systems, in *Fiber-Optic Communication Systems,* chap. 10, John Wiley & Sons, NY.

Atlas, D.A. 1996. On the overmodulation limit in externally modulated lightwave AMVSB CATV systems. *IEEE phot. Tech. Lett.,* 8(5):697-699.

Bell. 1994. Physical layer, in *Synchronous Optical Network (SONET) Transport Systems: Common Generic Criteria,* chap. 4, Issue 1, GR-253-CORE, Bell Communications Research.

Boivin, L., Nuss, M.C., Shah, J. Miller, D.A.B., and Haus H.A. 1997. Receiver sensitivity improvement by impulsive coding. *IEEE Photon. Tech. Lett.,* 9(5):684-686.

Chen, F.S. 1980. Simultaneous feedback control of bias and modulation currents for injection lasers. *Elect. Lett.,* 16(l):7-8.

Chin, M.K. 1995. Comparative analysis of the performance limits of Franz–Keldysh effect and quantum-confined Stark effect electroabsorption waveguide modulators. *IEEE Proc. Optroelectron.,* 142(2): 109-114.

Eriksson, U., Evaldsson, P., Donegan, J.F., Jordan, C., Hegarty, J., Hiei, F., and Ishibashi, A. 1996. Vertical integration of an InGaAs/InP HBT and a 1.55 um strained MQW p-substrate laser. *IEEE Proc. Optoelectron.,* 143(1):107-109.

Gnauck, A.H., Korotky, S.K., Veselka, J.J., Nagel, J., Kemmerer, C.T., Minford, W.J., and Moser, D.T. 1991. Dispersion penalty reduction using an optical modulator with adjustable chirp. *Elec. Lett.,* 28:954-955.

Izutsu, M. et al. 1981. Integrated optical SSB modulator/frequency shifter. *IEEE J. Quant. Electron.,* QE-17(11):2225.

Kawano, K., Wakita, K., Mitomi, O., Kotaka, I., and Naganuma, M. 1992. Design of InGaAs-InAlAs multiple-quantum-well (MQW) optical modulators. *IEEE J. Quant. Electron.,* 28(1):224-230.

Koyama, F. and Iga, K. 1988. Frequency chirping in external modulators. *J. Lightwave Tech.,* 6:87-93.

Kolner, B.H. and Dolfi, D.W. 1987. Intermodulation distortion and compression in an integrated electro-optic modulator. *Appl. Optics,* 26(17):3676-3680.

Linder, N., Kiesel, P., Kneissl, M., Knuepfer, B., Quassowski, S., Doehler, G.H., and Traenkle, G. 1995. Linearity of double heterostructure electroabsorptive waveguide modulators. *IEEE J. Quant. Electron.,* 31(9):1674-1681.

Mollenauer, L.F., Gordon, J. P., Mamyshev, P. V. 1997. Solutions in high bit rate, long distance transmission, In *Optical Fiber Telecommunications IIIA,* Eds. I.P. Kaminow and T. L. Koch, Academic Press, NY.

Pedrotti, K.D., Seabury, C.W., Sheng, N.H., Lee, C.P., Agarwal, R., Chen, A.D.M., and Renner, D. 1992. 6-GHz operation of a flip-chip mounted 1.3-um laser diode on an AlGaAs/GaAs HBT laser driver circuit. *OFC' 92,* ThJ7, 241.

Price, A.J. and Le Mercier, N. 1995. Reduced bandwidth optical digital intensity modulation with improved chromatic dispersion tolerance. *Elec. Lett.,* 31(1):58-59.

Runge, K., Detlef, D., Standley, R.D., Gimlett, J.L., Nubling, R.B., Peterson, R.L., Beccue, S.M., Wang, K.C., Sheng, N.H., Chang, M.F., Chen, D.M., and Asbeck, P.M. 1992. AlGaAs/GaAs HBT IC's for high-speed lightwave transmission systems. *IEEE JSSC,* 27(10):1332-1341.

Smith, D.W. 1978. Laser level-control circuit for high-bit-rate systems using a slope detector. *Elec. Lett.,* 14(24):775-776.

Tucker, R.S. 1985. High-speed modulation of semiconductor lasers, *J. Lightwave Tech.,* LT-3(6):1180-1192.

Twu, Y., Parayanthal, P., Dean, B.A., and Hartman, R.L. 1992. Studies of reflection effects on device characteristics and system performances of 1.5 um semiconductor DFB lasers. *J. Lightwave Tech.,* 10(9):1267-1271.

Veselka, J.J., Korotky S.K. 1996. Pulse generation for soliton systems using lithium niobate modulators. *IEEE J. Sel. Topics. Quant. Electron.,* 2(2):300-310.

Watanabe, T., Sato, K., and Soda, H. 1992. Low drive voltage and low chirp modulator integrated DFB laser light source for multiple-gigabit systems. *Fuj. Sci. Tech. J.,* 28(1):115-121.

Wedding, B. 1987. SPICE simulation of laser diode modules. *Elect. Lett.,* 23(8):383-384.

Winzer, P.J. and Kalmar, A. 1999. Sensitivity enhancement of optical receivers by impulsive coding. *J. Lightwave Tech.,* 17(2):171-177.

Further Information

Alferness, R.C. 1982. Waveguide electrooptic modulators. *IEEE Trans. Microwave Th. Tech.,* MTT-30:1121-1137.

Gu, X. and Blank, L.C. 1993. 10 Gbit/s unrepeatered three-level optical transmission over 100 km of standard fiber. *Elec. Lett.,* 29(25):2209-2211.

Henry, P.S. 1985. Lightwave primer. *IEEE J. Quant. Electron.,* QE-21:1862-1879.

Ortel. 1992. *RF/Microwave Fiber Optic Link Design Guide,* Ortel Corp.

Patel, B.L. et al. 1992. Transmission at 10 Gb/s over 100 km using a high performance electroabsorption modulator and the direct prechirping technique. In *ECOC'92,* postdeadline paper Th PD 1.3, Berlin, Germany.

Pedrotti, K.D., Zucca, F., Zampardi, P.J., Nary, S.M., Beccue, S.M. Runge, K., Meeker, D., Penny, J., and Wang, K.C. 1995. HBT transmitter and data regenerator arrays for WDM optical communications application. *IEEE J. Solid-State Circ.,* 30(10):1141-1144.

Shumate, P.W. 1988. Lightwave transmitters, in *Optical Fiber Telecommunications II,* Eds. S.E. Miller and I.P. Kaminow, Academic Press, NY, chap. 19.

Wedding, B., Franz, B., and Junginger, B. 1994. 10-Gb/s optical transmission up to 253 km via standard single-mode fiber using the method of dispersion-supported transmission. *J. Lightwave Tech.,* 12(10):1720-1727.

48

Optical Receivers

Richard G. Smith
Cenix, Inc.

Byron L. Kasper
Agere Systems

48.1 Introduction

The optical receiver is an integral part of an optical link. Its function is to detect the received optical power, amplify the signal, retime, regenerate, and reshape the signal where it may remain as an electrical signal (terminal), or be used to drive an optical transmitter which then feeds another fiber section.

Goal

The goal of the optical receiver is to achieve the desired system performance while requiring the minimum amount of received optical power. System performance criteria include bit error ratio (BER) and eye margin for digital systems, signal-to-noise ratio (SNR) for analog systems, and dynamic range (ratio of largest to smallest signal levels for which the BER or SNR can be achieved).

Over the past 20 years there has been a large amount of work on optical receivers. This work is well covered by review papers and book chapters [Kasper, 1988; Muoi, 1984; Smith and Personick, 1982]. Although the field is fairly mature, there are still advances being made, such as the use of *optical preamplifiers* to improve receiver sensitivity [Olsson, 1989].

48.2 The Receiver

A block diagram of the optical receiver is shown in Fig. 48.1. The receiver consists of a photodetector, a preamplifier, postamplifier, filter, timing recovery, and decision circuit.

The Photodetector

The purpose of the photodetector is to convert the received optical signal into a photocurrent. It is essentially a photon-to-electron converter. The most widely used forms of photodetectors used in optical communications are the p-i-n photodiode and the avalanche photodiode (APD). In the case of the p-i-n, a single electron-hole pair is generated for each photon absorbed by the detector. In the case of the APD, electronic gain within the detector itself generates secondary electron-hole pairs, multiplying the photocurrent and generally improving the receiver sensitivity. (See chapter on optical detectors for a discussion of photodetectors.)

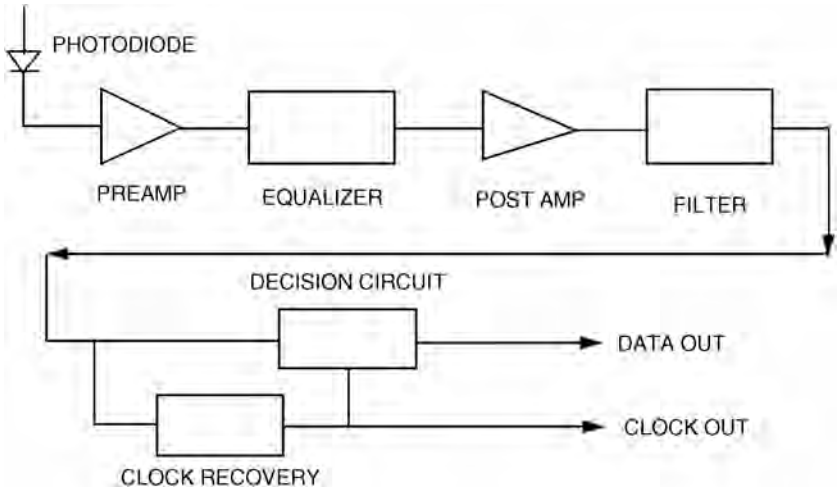


FIGURE 48.1 Schematic diagram of an optical receiver.

Preamplifier

The preamplifier or front end is the most important element of the receiver in determining the optical sensitivity. Its principal function is to amplify the generally small photocurrent with the least amount of added noise. Under most circumstances (especially with a p-i-n detector), the preamplifier noise is dominant in determining the sensitivity. The preamplifier must also have sufficient frequency response so as not to severely distort the received signal shape (*eye margin*). In addition to noise level and frequency response, the front end must be able to respond to a wide range of received optical power levels without severely distorting the signal (dynamic range). Often, these objectives are in conflict and compromises must be made.

Optimized optical receivers differ in several ways from receivers used at RF or microwave frequencies. This difference arises from the fact that the photodetectors generally used (p-i-n, APD) are current sources, and therefore have inherently high impedance. Optimal signal generation therefore requires the front end to have a high impedance, which differs from other types of receivers which are typically driven by low-impedance detectors.

A second major difference between optical systems derives from the relatively large energy of a photon. For a given received power, there are orders of magnitude fewer photons at optical frequencies than at microwave or lower frequencies. In the limit of low received optical power, and, hence, low number of received photons and photogenerated hole–electron pairs, the random arrival rate results in signal related shot noise, which is often important in determining receiver sensitivity.

Optical receiver front ends can be broadly classed into one of three types: high-impedance or *integrating* front end; *transimpedance* front end; and low-impedance front end, typically a 50-Ω impedance level. Although the distinction is not hard and fast, these categories characterize the various approaches to receiver design.

The high-impedance front end achieves the best sensitivity for a given detector. This design approach minimizes the thermal noise generated by resistors, given by $4kT/R$, by increasing their values such that the noise is reduced to a level where it is small compared to other noise sources. As a result of the increased impedance level, the frequency response, determined by the combination of the resistor and the capacitance at the input node of the receiver, is generally very much limited, typically being a factor of 10 or more smaller than that required to reproduce the signal shape. Such a receiver integrates the signal, hence the sometimes used description as an integrating front end. The accepted approach to the limited

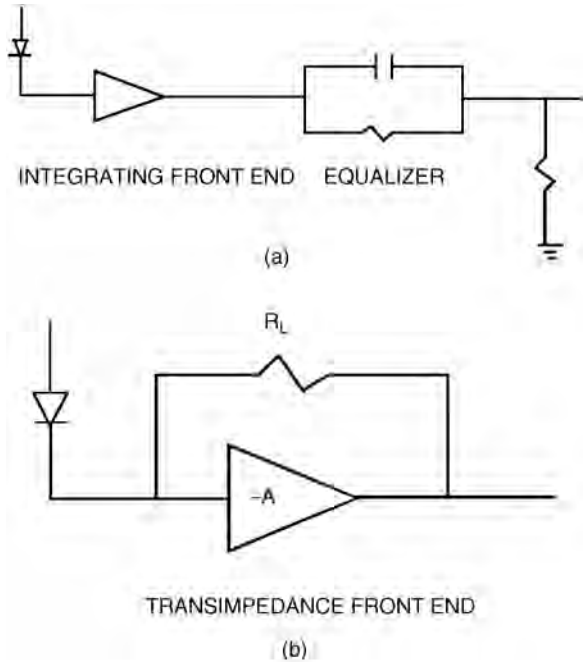


FIGURE 48.2 Schematic diagrams: (a) an integrating front end with equalization and (b) a transimpedance front end.

bandwidth is to equalize the signal after sufficient amplification. A simplified schematic of an equalized, high-impedance front end is shown in Fig. 48.2(a).

The high-impedance front end reduces the thermal noise associated with the input resistor, which has two major consequences. The first is related to the dynamic range of the receiver. Although the signal fidelity is restored by the equalizer, the low-frequency components of the received signal generate large voltages at the input node of the receiver. At received optical signals well in excess of the minimum, this signal buildup will eventually cause the active device (input transistor) to overload, distorting the signal, or perhaps cutting off the receiver. For many applications, dynamic ranges of 20 to 30 dB of optical power (40 to 60 dB electrical) are required. This requirement and the limited dynamic range of the high-impedance type of front end limit the usefulness of this approach.

The second consequence of the equalization is to accentuate the noise generated at the output of the amplifying stage. Because of the effect of the equalization, the noise spectral density associated with the output stage is proportional to the square of the input capacitance. Optimization (reduction) of the equivalent noise of the output stage, and, hence, the receiver front end, requires the minimization of the input capacitance, a goal of all optical receiver designs. The details of such optimization and the relevant formulas giving the analytical expressions for the resulting noise can be found in Kasper [1988], Muoi [1984], and Smith and Personick [1982], and are summarized subsequently.

At the other extreme of receiver design is the low-impedance front end typified by a 50-Ω input impedance amplifier. The low-impedance front end is the least sensitive of the receivers as its noise level is much larger than that of both the high-impedance and transimpedance approaches. Low-impedance receivers find use in laboratory experiments and are used in practical optical communication systems only in conjunction with optical preamplifiers.

The transimpedance receiver is schematically shown in Fig. 48.2(b). It is essentially a current-to-voltage converter. The bandwidth of the receiver is often dominated by the parasitic capacitance of the feedback resistor,

which can be very small. Hence, for a given resistor value, the bandwidth of the transimpedance receiver will be greater than the high-impedance receiver. Under practical conditions, the transimpedance receiver is usually designed with a bandwidth sufficient to reproduce the signal without recourse to equalization. Typically, this results in a higher noise level and hence a poorer sensitivity compared to a high-impedance receiver. It does, however, have an improved dynamic range as the input node is effectively a low impedance due to the electrical feedback. The transimpedance receiver is generally used in optical systems as it offers a good compromise between the various receiver requirements.

Early receivers were realized with discrete transistors. At the present time, many designs are realized using integrated circuits, especially Si bipolar and GaAs field-effect transistor (FET) designs. The GaAs receiver designs generally provide superior receiver sensitivity but are more expensive than their Si bipolar counterparts. Metal-oxide-semiconductor (MOS) integrated circuit (IC) technology has not found much application in optical receiver front ends, but is used in associated circuitry where the speed is sufficient. Generally speaking, the best sensitivities are still achieved with discrete FET front ends, but for economic considerations, integrated circuits are widely used.

Postamplifier

Front ends, as usually employed, typically generate relatively small output signal levels. The output from the front end is usually amplified by a postamplifier. Because the range of optical input signals can be as much as 30 dB (60-dB electrical), the postamplifier usually employs some form of output level control. This is typically achieved by an automatic gain control (AGC) function or by using a limiting amplifier. The output signal level is typically of the order of 1 V, which is sufficient to toggle the decision circuit.

Filtering

The main purpose of the filter shown in Fig. 48.1 is to improve the SNR by reducing the noise. Under ideal conditions, where the effective bandwidths of the components in the system prior to the filter are large, the filter function is such that the signal in a given time slot has zero signal level in all adjacent time slots. The effective bandwidth of such a filter is approximately 0.6 to 0.7 times the signalling rate, B. In a practical system, the pulse shape incident upon the filter is not ideal as a result of limited bandwidth in the transmitter, front end, postamplifier, and dispersion in the transmission fiber. The actual filter function required is one which transforms the waveform prior to the filter into one that produces an acceptable waveform at the decision circuit while minimizing the noise. In many practical lightwave systems, there is little or no actual filtering done due to the limited bandwidth of components preceeding the filter. Often, the filtering is accomplished by the preamplifier or front end.

Timing Recovery

In nearly all lightwave systems it is necessary to recover the clock in the process of regenerating the signal. Several commonly used methods of clock recovery make use of phase locked loops or surface acoustic wave filters, with the latter being the most widely used for high-frequency systems which have strict requirements on the phase transfer function. Because the commonly used nonreturn to zero (NRZ) format does not contain energy at the clock frequency, it is necessary to use some form of nonlinear device, such as a rectifier or exclusive OR gate, to produce frequency components at or near the clock frequency. This signal is then passed through a narrowband filter such as a surface acoustic wave (SAW) filter or dielectric resonator to extract the clock. The extracted clock, suitably phased, is used to clock the decision circuit and may be used elsewhere in the terminal or regenerator as needed.

Decision Circuit

The decision circuit receives the signal (plus noise) from the postamplifier along with the recovered clock. The decision circuit produces an output which is reshaped and retimed.

48.3 Receiver Sensitivity: General

The sensitivity of a receiver is usually measured in terms of the average optical power required to produce the desired level of system performance. For digital systems, the measure of performance is the BER or bit error ratio (often referred to as the bit error rate, although incorrectly). The performance measure commonly used is for a $BER = 10^{-9}$, although many systems require significantly better performance. For analog applications, the measure of performance is usually the carrier-to-noise ratio (CNR). Power is measured in terms of decibels referred to 1 mW of received power. Further, for digital systems, sensitivity is measured in terms of the *average* power received, taking into account the population of marks and spaces (i.e., 1s and 0s). The average power is thus approximately one-half of the peak power or 3 dB below the peak received power. In order to evaluate the receiver sensitivity, it is necessary to evaluate both the photogenerated signal current and the sources of noise which corrupt the signal.

Signal Current

The photocurrent generated by a received optical power P is given by

$$I_s = \left[\frac{q\eta}{h\upsilon}\right]P \quad \text{(p-i-n)} \tag{48.1}$$

where h is Planck's constant (6.623E-34 J/s), υ is the optical frequency, q is the electronic charge, and η is the quantum efficiency of the detector (defined in terms of the number of photoelectrons generated per photon received). In well-designed detectors, the quantum efficiency is usually between 0.8 and 1.0. When a p-i-n detector is used, Eq. (48.1) becomes the signal current; for an avalanche photodiode, the signal current is given by

$$I = MI_s = M\left[\frac{q\eta}{h\upsilon}\right]P \quad \text{(APD)} \tag{48.2}$$

where M is the expectation value of the avalanche gain.

Sources of Noise

Amplifier or Circuit Noise

As the received signal level (photocurrent generated in the photodetector) is small it must be amplified by an electronic amplifier, which contributes noise to the system. The principal sources of noise in the amplifier are thermal noise from resistors and noise generated in the amplifying transistors. It is common practice in the analysis of optical receivers to refer all noise sources to the input node where they can be compared to the signal current, Eq. (48.1) or (48.2).

In evaluating the contributions of the various sources of noise, account is taken of the frequency response of the various components of the receiver. The details of this computation are too numerous to be included in this summary and may be found in Personick [1973, 1979] and in Smith and Personick [1982]. We will summarize here the essential results described in the aforementioned references.

The variance of the circuit noise for an FET front end is given by

$$\langle i^2 \rangle_{\text{ckt}} = \frac{4kT}{R_L}\Im_2 B + 2qI_g\Im_2 B + \frac{4kT\Gamma}{g_m}(2\pi C_T)^2\Im_3 B^3 \quad \text{(FET)} \tag{48.3}$$

where k is Boltzmann's constant, T is the absolute temperature, R_L is the load resistance at the input node, q is the electronic charge, I_g is the gate leakage current, g_m is the transconductance of the input transistor,

C_T is the total capacitance at the input node, B is the design bit rate, and Γ is a parameter characterizing the channel noise, which has a value ª1. The parameters $\Im_2$ and $\Im_3$, referred to as the *Personick integrals*, are dimensionless numbers, which derive from the details of the transfer function of the system. Typical values of ¡$_2$ and ¡$_3$ are 0.6 and 0.1, respectively. (A detailed explanation of the Personick integrals may be found in Personick [1973] and Smith and Personick [1982].) The B^3 dependence of the last term arises from the referral of the output noise source to the input.

The corresponding equation for the variance of the noise for a bipolar front end, neglecting the effect of noise due to base resistance, is given by

$$\langle i^2\rangle_{\chi\kappa\tau} = \frac{4kT}{R_L}\Im_2 B + 2qI_b\Im_2 B + \frac{2qI_c}{g_m^2}(2\pi C_T)^2\Im_3 B^3 \quad \text{(bipolar)} \tag{48.4}$$

For a bipolar transistor, $I_c = \beta I_b$ and the transconductance g_m is proportional to the collector current I_c; hence, the noise can be minimized by operating the input transistor at the optimum collector current. The minimum bipolar circuit noise is given by

$$\langle i^2\rangle_{\beta\iota\pi\chi\kappa\tau/\mu\iota\nu} = \frac{4kT}{R_L}\Im_2 B + 8\pi kTC_T\sqrt{\frac{\Im_2\Im_3}{\beta}} \tag{48.5}$$

Signal Shot Noise

The random arrival rates of photons at the detector result in an additional source of noise, which is proportional to the received photocurrent, referred to as *signal-dependent shot noise*. The spectral density of this noise source is given by

$$\frac{\mathrm{d}}{\mathrm{df}}\langle i^2\rangle_{\sigma\iota\gamma} = 2qI \tag{48.6}$$

where I is the average current. This signal-dependent noise constitutes a major difference in optical systems compared to radio frequency and microwave systems. The variance of this contribution to the noise is found by integrating the noise spectral density over the transfer function of the system giving

$$\langle i^2\rangle_{\sigma\iota\gamma} = 2qI\Im_1 B \tag{48.7}$$

where $\Im_1$ is another Personick integral, which has a value ª0.5, and B is the bit rate.

Shot Noise with an Avalanche Photodiode

In many receivers, an avalanche photodiode is used in place of a p-i-n photodiode. As discussed in the chapter on optical detectors in this volume, the APD multiplies the received primary photocurrent given by Eq. (48.1) by an average multiplication factor M,

$$I_{\mathrm{APD}} = MI = M\left[\frac{q\eta}{h\upsilon}\right]P$$

For receivers employing an APD, this becomes the signal current [see Eq. (48.2)].

If the avalanche multiplication process were ideal (i.e., deterministic), the shot noise spectral density associated with the signal would be given by

$$\frac{\mathrm{d}}{\mathrm{df}}\langle i^2\rangle = 2qIM^2 \tag{48.8}$$

The avalanche process is, however, a random process and not deterministic. This randomness adds additional noise, which is accounted for by a multiplicative noise factor, referred to as the excess noise

factor, and denoted $F(M)$. In this case, the signal-dependent shot noise spectral density becomes

$$\frac{\mathrm{d}}{\mathrm{df}}\langle i^2\rangle_{\sigma\gamma} = 2qIM^2F(M) \tag{48.9a}$$

and the mean square noise after integration with respect to frequency is

$$\langle i^2\rangle_{\sigma\gamma} = 2qIM^2F(M)\Im_1 B \tag{48.9b}$$

Many approximations for $F(M)$ have been given. Among the most commonly used are

$$F(M) = M^x \tag{48.10a}$$

and

$$F(M) = M\left[1-(1-k)\left(\frac{M-1}{M}\right)^2\right] \tag{48.10b}$$

In Eq. (48.10a), x is an empirical constant, usually between 0 and 1, and in Eq. (48.10b), k is the ratio of the ionization coefficients of holes and electrons (not to be confused with Boltzmann's constant). The value of k is between 0 and 1 if the avalanche is initiated by the most ionizing carrier, as it should be in a well-designed device. Equation (48.10b) provides a better approximation to the excess noise factor and is generally used.

Sensitivity of a Digital Receiver

In a digital system, the objective is to correctly identify transmitted marks (1s) and spaces (0s). Incorrectly identifying the transmitted signal produces an error. The probability of making an error in turn depends on the signal level and the probability distribution of the system noise. Under many circumstances, but not always, the probability distribution of the noise is Gaussian, with an associated standard deviation σ.

Gaussian Approximation

The assumption of a Gaussian distribution permits simplified analytical approximations to sensitivity calculations, which are generally close to the more rigorous computational approaches. In optical systems, the noise levels associated with the different signal states are not necessarily identical. The standard deviation of the noise associated with a transmitted 1 is denoted by σ_1 and for a transmitted 0 by σ_0. The ratios of the signal levels relative to the decision level D normalized to the associated root mean square (rms) noise levels are given by

$$Q_1 = \frac{S_1 - D}{\sigma_1} \tag{48.11a}$$

and

$$Q_0 = \frac{D - S_0}{\sigma_0} \tag{48.11b}$$

With these definitions, and using the Gaussian approximation, the probability of misinterpreting a transmitted 0 as a 1 is given by

$$P(0,1) = \frac{1}{2\pi\sigma_0}\int_D^\infty \exp\frac{-(S-D)^2}{2\sigma_0^2}\,\mathrm{d}S = \frac{1}{2\pi}\int_{Q_0}^\infty \exp -\frac{Q^2}{2}\,\mathrm{d}Q \approx \frac{1}{2\pi Q_0}\exp\left(-\frac{Q_0^2}{2}\right) \tag{48.12a}$$

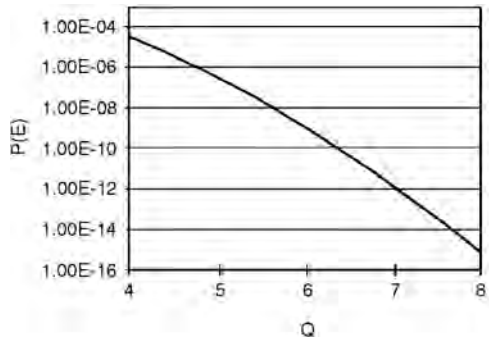


FIGURE 48.3 Probability of error vs. Q value.

and the probability of misinterpreting a transmitted 1 is given by

$$P(1,0) = \frac{1}{\sqrt{2\pi}\sigma_1}\int_{-\infty}^{D} \exp\left[\frac{-(S-D)^2}{2\sigma_1^2}\right] \mathrm{d}S$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{Q_1} \exp\left[-\frac{Q^2}{2}\right] \mathrm{d}Q \approx \frac{1}{\sqrt{2\pi}Q_1}\exp\left[-\frac{Q_1^2}{2}\right] \tag{48.12b}$$

To a good approximation, the minimum probability of error occurs when each probability of error is the same, that is, $P(0, 1) = P(1, 0)$, resulting in $Q_1 = Q_0 = Q$ with

$$Q = \frac{S_1 - S_0}{\sigma_1 + \sigma_0} \tag{48.13}$$

The probability of error as a function of the Q factor is shown in Fig. 48.3. The desired error performance determines the value of Q. The average required signal level is given by

$$\bar{S} = \frac{S_1 + S_0}{2} \tag{48.14}$$

with

$$S_1 - S_0 = Q(\sigma_1 + \sigma_0) \tag{48.15}$$

When the decision level is not optimally placed (Q_1 π Q_2), the probability of error becomes

$$P(E) = 0.5[P(1, 0) + P(0, 1)] \tag{48.16}$$

Receiver Sensitivity: p-i-n Detector

We are now in a position to calculate the sensitivity of a receiver using a p-i-n detector. Combining Eqs. (48.1) and (48.15), we find that the required *peak* signal current is

$$I_s = Q(\sigma_1 + \sigma_0) \tag{48.17}$$

where we assume the signal in the 0 state is identically zero and $S_1 = I_s$. The variance of the noise σ^2 is given by the sum of the circuit noise and the signal shot noise, Eq. (48.7).

$$\sigma^2 = \langle i^2\rangle_{\mathrm{ckt}} + \langle i^2\rangle_{\mathrm{sig}} = \langle i^2\rangle_{\mathrm{ckt}} + 2qI_s\Im_1 B \tag{48.18}$$

The circuit noise for an FET front end is calculated from Eq. (48.3) and for an optimally biased bipolar front end by Eq. (48.5). Since the signal current equals I_s in the 1 state and 0 in the 0 state, the variances become

$$\sigma_0^2 = \langle i^2 \rangle_{ckt} \tag{48.19a}$$

and

$$\sigma_1^2 = \langle i^2 \rangle_{ckt} + 2qI_s \mathfrak{I}_1 B \tag{48.19b}$$

Substituting Eqs. (48.19a) and (48.19b) into Eq. (48.17) gives

$$I_s = 2\sigma_0 Q + 2qQ^2 \mathfrak{I}_1 B \tag{48.20}$$

The peak optical power is found from Eq. (48.1), and the sensitivity, measured in terms of the average power, is given by

$$\eta \bar{P}_s = \left(\frac{h\upsilon}{q}\right) [\sigma_0 Q + qQ^2 \mathfrak{I}_1 B] \quad \text{(p-i-n)} \tag{48.21}$$

Under most circumstances, with a p-i-n detector, the second term in Eq. (48.21) is small compared to the first, leading to the commonly used expression for the p-i-n sensitivity

$$\bar{P}_s = \left(\frac{h\upsilon}{q}\right) Q \langle i^2 \rangle_{ckt}^{1/2} = \frac{1.24Q}{\lambda} \langle i^2 \rangle_{ckt}^{1/2} \tag{48.22}$$

where λ is the wavelength expressed in micrometers.

Receiver Sensitivity: Avalanche Photodiode Detector

For an APD receiver, the signal current is given by Eq. (48.2) and the variances of the noise in the two signal states become

$$\sigma_0^2 = \langle i^2 \rangle_{\text{ckt}} \tag{48.23a}$$

and

$$\sigma_1^2 = \langle i^2 \rangle_{\text{ckt}} + 2qI_s M^2 F(M) \mathfrak{I}_1 B \tag{48.23b}$$

These relations neglect the effect of photodetector dark current and also assume the received signal is identically zero when a 0 is received. A more general account maybe found in Smith and Personick [1982]. Substituting these equations into Eq. (48.17) gives the result

$$\eta \bar{P}_s = \left(\frac{h\upsilon}{q}\right) \left[\frac{\sigma_0 Q}{M} + qQ^2 \mathfrak{I}_1 B F(M)\right] \tag{48.24}$$

A plot of this expression vs. the gain of the APD, M, is shown in Fig. 48.4 for a receiver operating at 622 Mb/s [Tzeng, 1994]. For low gain, the sensitivity is seen to improve, reaching an optimum, and then declines for further increases in the gain. The initial improvement is associated with the amplification of the signal current by the APD, whereas the decrease at high gain results from the excess signal shot noise associated with the random nature of the avalanche process.

Receiver Sensitivity: Optical Preamplifier

A recent advance in optical fiber technology is the development of optical amplifiers. The most widely used is the Er-doped fiber amplifier EDFA (see Chapter 52 in this volume for a detailed discussion of

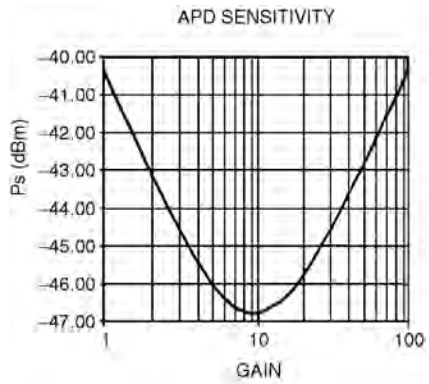


FIGURE 48.4 Sensitivity of an APD receiver at 622 Mb/s as a function of the APD gain.

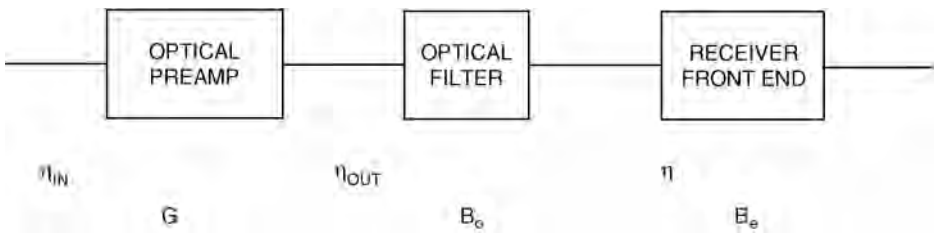


FIGURE 48.5 Schematic diagram of an optical preamplifier.

optical amplifiers). One application of optical amplifiers is their use as an optical preamplifier preceding the optical receiver previously described. The optical preamplifier amplifies the received optical signal prior to detection by the photodetector. As with other types of amplifiers, optical amplifiers introduce noise resulting from spontaneous emission in the amplifying region. This type of noise is referred to as *amplified spontaneous emission* (ASE).

Figure 48.5 is a schematic of an optical preamplifier showing the important parameters defined as follows:

η_{in} = coupling efficiency of transmission fiber to the optical amplifier
η_{out} = coupling of optical amplifier to photodetector input including transmission loss of the optical filter
η = detector efficiency
G = small signal gain of the optical amplifier
F_m = noise figure of the optical amplifier
B_o = bandwidth of the optical filter, Hz
B_e = electrical bandwidth of the receiver, Hz

The noise figure of the amplifier is defined as $2n_{sp}$ where n_{sp} is the spontaneous emission factor. See Chapter 52 on optical amplifiers in this volume for more detail.

Photocurrents

In receivers incorporating optical preamplifiers, p-i-n detectors are usually employed. For this case, the signal photocurrent is given by

$$I_s = \left(\frac{\eta q}{h\upsilon}\right) \eta_{in} \eta_{out} G P_s \tag{48.25}$$

and the photocurrent generated by the ASE is given by

$$I_{\text{ASE}} = F_m q \eta_{\text{out}} \eta (G-1) B_o \tag{48.26}$$

Noise Sources

The photodetector behaves as a square law detector. The generated photocurrent, which is proportional to the square of the incident optical field, contains components proportional to signal*signal, ASE*ASE, and signal*ASE. Following the analysis of Olsson [1989], the components of the noise are given by

$$\sigma^2_{\text{shot}} = 2q(I_s + I_{\text{ASE}})B_e \tag{48.27a}$$

$$\sigma^2_{\text{spon-spon}} = \frac{1}{2} F_m^2 q^2 \eta^2_{\text{out}} \eta^2 (G-1)^2 (2B_o - B_e) B_e \tag{48.27b}$$

$$\begin{aligned} \sigma^2_{\text{sig-spon}} &= [2qF_m \eta_{\text{out}} \eta (G-1)] \left[\left(\frac{\eta q}{h\upsilon} \right) P_s \eta_{\text{in}} \eta_{\text{out}} G \right] B_e \\ &= 2q^2 F_m \eta_{\text{in}} \eta^2_{\text{out}} \eta^2 G (G-1) \left(\frac{1}{h\upsilon} \right) P_s B_e \end{aligned} \tag{48.27c}$$

The first term, s^2_{shot}, is the shot noise generated by the amplified signal and ASE and contains a term proportional to the received signal. The second term, $s^2_{\text{spon-spon}}$, results from the ASE beating with itself in the detector. Since the bandwidth of the optical gain is large (a10^{12} Hz) compared to the signal bandwidth in common usage, the spectral width of the ASE is normally reduced by the use of an optical filter of bandwidth B_o. B_o can be reduced in principle until it approaches B_e. The problems of achieving a narrow optical filter bandwidth and of aligning the optical filter with the signal wavelength set the practical limit on B_o. The noise term $s^2_{\text{spon-spon}}$ is independent of the signal level and, hence, is present in both the 0 and 1 states of the signal.

The term $s^2_{\text{sig-spon}}$ results from the mixing of the signal and ASE fields. This noise term is proportional to the received optical signal power, and in an ideal system, when the received optical power is zero in the 0 state, $s^2_{\text{sig-spon}} = 0$. The total noise associated with the 0 state, assuming zero power received, is given by

$$s^2_0 = s^2_{\text{ckt}} + s^2_{\text{shot}}(ASE) + s^2_{\text{spon-spon}} \tag{48.28}$$

and for the 1 state the noise is given by

$$\sigma^2_1 = \sigma^2_{\text{ckt}} + \sigma^2_{\text{shot}}(ASE + \text{signal}) + \sigma^2_{\text{spon-spon}} + \sigma^2_{\text{sig-spon}} \tag{48.29}$$

The receiver sensitivity can be calculated by combining Eqs. (48.13) and (48.25–48.29), yielding

$$I_s = 2Q^2 \left[qB_e + qF_m \eta_{\text{out}} \eta (G-1) B_e + \frac{\sigma_0}{Q} \right] \tag{48.30}$$

Typically, G is greater than or equal to 30 dB (1000), in which case the first term in brackets is negligible. Further, since $s^2_{\text{spon-spon}}$ varies as the square of the gain while s^2_{shot} varies only as the first power of the gain and s^2_{ckt} is independent of the gain,

$$\sigma^2_0 \text{ a } \sigma^2_{\text{spon-spon}} \quad (G \gg 1) \tag{48.31}$$

Substituting this limit gives

$$I_s = 2Q^2\left\{qF_m\eta_{\text{out}}\eta(G-1)B_e\left[1+\frac{1}{Q}\left(\frac{B_o}{B_e}-\frac{1}{2}\right)^{\frac{1}{2}}\right]\right\} \tag{48.32}$$

and the sensitivity measured in terms of the average received power incident on the optical amplifier in the limit of large gain is

$$\eta_{\text{in}}\bar{P}_s = Q^2(h\upsilon)F_mB_e\left[1+\frac{1}{Q}\left(\frac{B_o}{B_e}-\frac{1}{2}\right)^{\frac{1}{2}}\right] \tag{48.33}$$

When the optical filter bandwidth is comparable to the electrical bandwidth (practical only at very high bit rates) and for very low BER ($Q \gg 1$), the limiting sensitivity becomes

$$\eta_{\text{in}}\bar{P}_s \approx Q^2(h\upsilon)F_mB_e \tag{48.34}$$

When the bandwidth of the optical filter is large compared to the electrical bandwidth, Eq. (48.33) applies.

The value of the noise figure F_m depends on the detailed characteristics of the optical amplifier. Under ideal conditions of perfect inversion of the gain medium, $F_m = 2$. Typical values are somewhat higher. With an electrical bandwidth $B_e = B/2$, the limiting sensitivity for a receiver using an optical preamplifier is

$$\eta_{\text{in}}\bar{P}_s = Q^2(h\upsilon)B \tag{48.35}$$

Sensitivity Limits

Quantum Limit. In the limit that the sensitivity is given by the statistics of the received photons which are characterized by a Poisson process, the ultimate sensitivity for a *BER* = 1 *E*-9 is 10 detected photons per bit received [Henry, 1985], which corresponds to

$$\eta\bar{P}_{\text{limit}} = 10(h\upsilon)B \tag{48.36}$$

Optical Amplifier Limit. With an optimized optical amplifier, the limiting sensitivity for BER = 1E-9 ($Q = 6$) is

$$\eta_{\text{in}}\bar{P}_{s-\text{opt.ampl}} = 36(h\upsilon)B \tag{48.37}$$

which is approximately four times the quantum limit. Demonstrated sensitivities are within a factor of three to four of this value at bit rates greater than 1 GB/s. See Park and Grandlund [1994] for a discussion of experimental results.

Practical Avalanche Photodiode Performance. The ultimate receiver sensitivity achievable with an APD is a function of the circuit noise and the excess noise factor associated with the avalanche process. Because the excess noise factor increases as the avalanche gain is increased, the optimum gain is in the range of 10–30 for existing APDs and is significantly below the gains used in optically preamplified receivers. The achievable sensitivity is accordingly poorer than the limit for an optical preamp, being in the range of a few hundred to as much as several thousand photons per bit and, hence, between a factor of 10–100 worse than the quantum limit.

Practical p-i-n Performance. In the absence of a gain mechanism, either avalanche or optical amplification, the receiver sensitivity is determined by the circuit noise, hence, depending on the design of the receiver front end. Typically, the sensitivity of a p-i-n-based receiver will be ≈10 dB poorer than an APD-based receiver and, hence, typically a few thousand photons per bit.

Defining Terms

Amplified spontaneous emission (ASE): The fundamental source of noise in optical amplifiers.

Avalanche photodiode (APD): A photodiode with internal current gain resulting from impact ionization.

Excess noise factor (ENF): A measure of the amount by which the short noise of an APD exceeds that of an ideal current amplifier.

Front end (FN): The input stage of the electronic amplifier of an optical receiver.

Integrating front end (IFE): A type of front end with limited bandwidth which integrates the received signal; requires subsequent equalization to avoid intersymbol interference.

Transimpedance front end (TFE): A type of front end which acts as a current to voltage converter; usually does not require equalization.

References

Henry, P.S. 1985. Lightwave primer. *IEEE J. Quantum Electr.,* QE-21:1879–1879.

Kasper, B.L. 1988. Receiver design. In *Optical Fiber Telecommunications II,* Eds. S.E. Miller and I.P. Kaminov, Eds., Academic Press, San Diego, CA, 689–722.

Muoi, T.V. 1984. Receiver design for high-speed optical-fiber systems. *IEEE J. Lightwave Tech.,* LT-2:243–267.

Olsson, N.A. 1989, Lightwave systems with optical amplifiers. *IEEE J. Lightwave Tech.,* LT-7:1071–1082.

Park, Y.K. and Grandlund 1994. Optical preamplifier receivers: application to long-haul digital transmission, *Optical Fiber Tech.,* 1:59–71.

Personick, S.D. 1973. Receiver design for digital fiber optic communication systems I. *Bell Syst. Tech. J.,* 52:843–874.

Personick, S.D. 1979. Receiver design. In *Optical Fiber Telecommunications,* Eds. S.E. Miller and A.G. Chenoweth, Academic Press, New York, 627–651.

Smith, R.G. and Personick, S.D. 1982. Receiver design for optical fiber communication systems. In *Semiconductor Devices for Optical Communication, Topics in Applied Physics,* Vol. 39, H. Kressel, Ed., Springer–Verlag, Berlin, 89–160.

Tzeng, L.D. 1994. Design and analysis of a high-sensitivity optical receiver for SONET OC-12 systems, *J. Lightwave Tech.,* LT-12:1462–1470.

Further Information

This section has focused on direct detection of digital signals, which constitute the most widely used format in optical communication systems. Coherent optical systems have been studied for some time but have not yet found practical application. A discussion of this technology and the associated receiver sensitivities may be found in Chapter 53 on coherent systems in this volume. Another application of optical fiber technology is in the distribution of analog video, which is being employed by community antenna television (CATV) distribution companies. A discussion of the sensitivity of analog receivers may be found in Smith and Personick [1982].

49

Fiber Optic Connectors and Splices

William C. Young
Bell Communications Research

49.1 Introduction

In recent years, the state of the art of optical fiber technology has progressed to where the achievable attenuation levels for the fibers are very near the limitations due to Rayleigh scattering. As a result, optical fibers, and particularly the single-mode fibers that are used in today's communications systems, can be routinely fabricated with attenuation levels below 0.5 dB/km at a wavelength of 1300 um and 0.25 dB/km at 1550 nm. Employing these fibers in optical communications systems requires precise jointing devices such as connectors and splices. Considering the small size of the fiber cores, less than 10 μm in diameter for single mode and 100 μm for multimode fibers, it is not surprising that these interconnecting devices can easily introduce significant optical losses. Furthermore, since single-mode fibers have practically unlimited bandwidth, it is also not surprising that they have become the choice for optical communications applications. To provide low-loss connectors and splices for these single-mode fibers, reliable and cost-effective alignment accuracies in the submicrometer range are required.

This chapter will review the fundamental technology that is presently used for optical connectors and splices. In particular, since single-mode fibers dominate optical communications systems and also require the greatest precision and performance, we will focus mainly on the jointing of these fibers. However, for completeness, we will also briefly review multimode fiber jointing technology as well.

Before reviewing the technology, we should define the implied differences between an optical connector and a splice. The term *connector* is commonly used when referring to the jointing of two optical fibers in a manner that not only permits but also anticipates the unjointing, or unconnecting, through the design intent. Optical connectors are commonly used for terminating components, system configuration and reconfiguration, testing, and maintenance. In contrast, the term *splice* is commonly used when referring to the jointing of two optical fibers in a manner that does not lend itself to unjointing.

There are two types of splices: mechanical splices that join two butt-ended fibers and use index matching material between their endfaces, and fusion splices where the two fibers are heated to just below their melting point and are pressed together to form a permanent joint. Splices are commonly used when the total optical fiber span length can be realized only by the concatenation of shorter sections of fiber, and also for repairing severed fiber/cables.

49.2 Optical Fiber Coupling Theory

The optical forward coupling efficiency of a connector or splice, that is, the insertion loss, and the optical backward coupling efficiency (reflectance) are the main optical performance criteria for optical fiber joints, and, therefore, these two parameters are used to evaluate the joint's optical performance over various climatic and environmental conditions [Miller, Mettler, and White, 1986].

The factors affecting the optical coupling efficiency of optical fiber joints can be divided into two groups, extrinsic and intrinsic factors (Table 49.1).

Factors extrinsic to the optical fiber, such as lateral or transverse offset between the fiber cores, longitudinal offset (endface separation), and angular misalignment (tilt), are directly influenced by the techniques used to join the fibers. In contrast, intrinsic factors are directly related to the particular properties of the two optical fibers that are joined. In the case of single-mode fibers, the important intrinsic factor is the fiber's mode field diameter. With multimode fibers, the intrinsic factors are the core diameter and numerical aperture (NA). The fiber's core concentricity and outside diameter are also commonly referred to as intrinsic factors, but practical connectors and splices must be designed to accommodate variations in these two parameters.

The effect that these factors have on coupling efficiency may also depend on the characteristics of the optical source, and in the case of multimode fiber joints, on the relative location of the joint in the optical path. For example, in the case of an incoherent optical source such as a light emitting diode (LED), the mode power distribution (MPD) along a continuous fiber's path largely depends on the length and curvature of the fiber. With these sources and fibers, the effect a particular offset will have on the coupling efficiency will decrease until a steady-state mode distribution is realized. Because of these MPD effects, a much practiced method of evaluating multimode fiber joints is to use an over-moded launch condition (uniform) followed by a mode filter that selectively filters the higher-ordered modes, thereby establishing a steady-state MPD condition. Characterizing multimode fiber joints in this manner eliminates the uncertainty of different MPD effects. Of course, in use, the coupling efficiency of the connector or splice is dependent on the particular MPD (a function of the optical source and location along the optical path). In the case of single-fiber joints, it is only necessary that the fiber is operated in the single-mode wavelength regime.

Multimode Fiber Joints

Various experiments [Chu and McCormick, 1978] and analytical models have been used to quantify the effects that extrinsic factors have on coupling efficiencies for butt-jointed multimode fibers. Based on these studies, we have plotted (Fig. 49.1) the coupling loss as a function of normalized lateral, longitudinal, and angular offset for graded-index multimode fibers.

TABLE 49.1 Factors Affecting Coupling Efficiency

Extrinsic factors:
Lateral fiber cone offset
Longitudinal offset
Angular misalignment
Reflections
Intrinsic factors:
Mismatch in fiber core diameters (mode field diameters)
Mismatch in index profiles

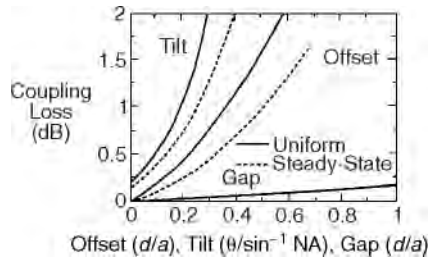


FIGURE 49.1 Coupling loss vs. offsets for butt-jointed graded-index multimode fibers (where a is the fiber's core radius and d is the offset in micrometers).

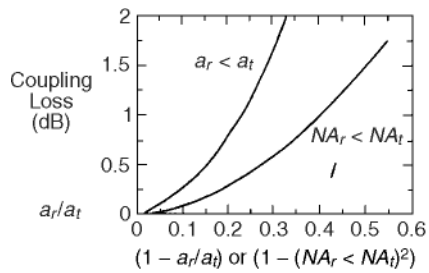


FIGURE 49.2 Coupling loss vs. mismatches in numerical aperture NA and fiber core radius a for graded-index multimode fiber having a steady-state MPD. The subscripts r and t represent the receiving and transmitting fibers.

In Fig. 49.1, the coupling losses for various offsets are plotted for both a uniform MPD as well as a steady-state MPD. From these curves, it can be seen that as the quality of the alignments improve, the effect of the presence of higher order modes diminishes. For example, today's connectors and splices have high control on angular (tilt) and longitudinal (gap) offsets, and, as a result, the coupling loss is mainly due to lateral offset. Therefore, as can be seen in Fig. 49.1, a connection of a typical graded-index multimode fiber having a 50-mm core diameter and having a lateral offset of about 10 μm has a coupling loss of 0.25 dB with a steady-state MPD, whereas with a uniform MPD the same joint has a loss of 0.5 dB. It is, therefore, very important, in the case of multimode fibers, that the particular MPD be known at the site of each joint for correct estimation of the loss incurred in the system. It should also be pointed out that due to cross-coupling effects, the offsets at a joint can also change the MPD directly after the joint. Thus, in summary, because of the many factors that can change the MPD in a multimode fiber, such as optical launch conditions, offsets at joints, fiber length and deployment conditions, mode mixing, and differential mode attenuation of the particular fibers, there is no unique coupling or insertion loss for multimode fiber joints that is based only on their physical fiber alignment. It therefore is very important for the system designer to analyze the system for these factors and then make the appropriate assignment for expected loss of the connector and splices.

Intrinsic factors can also have a large effect on multimode fiber joints. Therefore, when evaluating the joint loss of multimode fiber joints one also must consider the characteristics of the fibers on either side of the joint, as well as the direction of propagation of the optical power through the joint. Again, various experiments and analytical models have been used to quantify these effects. Based on these studies, the dependence of the coupling loss on mismatches in numerical apertures and core radii is summarized in Fig. 49.2.

As can be seen in Fig. 49.2, we have plotted the coupling loss for the cases when the transmitting fiber has the larger core radius and larger NA. For the opposite cases, there is no loss incurred at the joints. Referring to Fig. 49.2, for a mismatch in core diameters of 48 and 52 μm, the expected loss is about

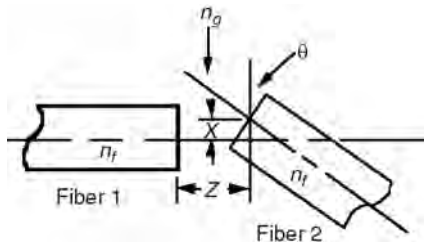


FIGURE 49.3 Schematic representation of a fiber butt joint having lateral and longitudinal offset and angular misalignment.

0.25 dB, and for a mismatch in numerical aperture of 0.20–0.23 it is about 0.35 dB. These values are significant when compared to expected loss due solely to extrinsic alignment inaccuracies and are also propagation direction dependent. Therefore, the performance of multimode fiber connectors and splices are very deployment dependent, requiring a complete understanding of the characteristics of the optical path before expected loss values can be calculated with confidence.

Single-Mode Fiber Joints

In the case of single-mode fiber joints, it has been shown that the fields of single-mode fibers being used today are nearly Gaussian, particularly those designed for use at a wavelength of 1310 nm. Therefore, the coupling losses for the joints can be calculated by evaluating the coupling between two misaligned Gaussian beams [Marcuse, 1977]. Based on this model, the following general formula has been developed [Nemota and Makimoto, 1979] for calculating the insertion loss (IL) between two single-mode fibers that have equal or unequal mode field diameters (an intrinsic factor) and lateral, longitudinal, and angular offsets, as well as reflections (extrinsic factors) as defined in Fig. 49.3:

$$IL = -10\log\left[\left(\frac{16n_f^2 n_g^2}{(n_f+n_g)^4}\right)\frac{4\sigma}{q}\exp\left(-\frac{\rho\mu}{q}\right)\right]\text{dB} \tag{49.1}$$

where

$$\rho = 0.5(Kw_t)^2; \quad q = G^2 + (\sigma+1)^2$$
$$\mu = (\sigma+1)F^2 + 2\sigma FG\sin q + \sigma(G^2+\sigma+1)\sin^2\theta$$
$$F = \frac{2x}{Kw_t^2}; \quad G = \frac{2z}{Kw_t^2}$$
$$\sigma = \left(\frac{w_r}{w_t}\right)^2; \quad K = \frac{2pn_g}{\lambda} \tag{49.1}$$

where

n_f, n_g = refractive indices of the fiber and the medium between the fibers, respectively
w_t, w_r = mode field radii of the transmitting and receiving fibers, respectively
λ = operating wavelength
x, z = lateral and longitudinal offset, respectively
σ = angular misalignment

Although the fields of dispersion-shifted fibers (fibers having minimum dispersion near 1550 nm) are not truly Gaussian, this general formula is still applicable, particularly for the small offsets and misalignments that are required by connectors and splices designed for communication applications (<0.5 dB). In the preceding general formula, terms in the first parentheses, inside the square bracket, account for the loss due to reflection at two fiber/gap interfaces, which can be as high as 0.32 dB for air as a gap material. If fiber core-to-core contact is established in a connector or perfect index matching is used in a mechanical splice, this loss reduces to zero. The formula, however, does not account for the loss variation due to interference effects caused by multiple beam interference that may exist in joints with fiber endface separation [Wagner and Sandahl, 1982]. Even though high-performance connectors are designed for fiber core-to-core contact and therefore should not be subject to reflections caused by an air gap, the lack of workable standards and manufacturing tolerances may result in endface separation. Even if the end separation is less than a micrometer, it can still have severe results on insertion loss. Although index matching material is used in mechanical splices, the refractive index of the material is temperature dependent and, therefore, the resulting index mismatches can also have an impact on insertion loss. The effects from multiple beam interference and its impact on connector and splice performance are discussed in a later section.

The preceding general formula can be reduced to determine the coupling loss due only to a mode field diameter mismatch (intrinsic factor). Assuming no extrinsic factors, and assuming core-to-core contact, the formula reduces to

$$\text{coupling loss} = -10 \log \left[4\left(\frac{w_t}{w_r} + \frac{w_r}{w_t} \right)^{-2} \right] \text{dB} \tag{49.2}$$

Using this expression, the loss due to a 10% mismatch in mode field diameters, typical for today's commercial fibers and with no other factors present, is calculated to be 0.05 dB. Generally, to control the fiber's cutoff wavelength and zero dispersion wavelength, the mode field diameters of communications single-mode fibers are usually within 10%. A typical requirement range for the mode field diameter of **dispersion-unshifted single-mode fiber** is from 8.8 to 9.5 μm. It can also be seen in the preceding expression that, contrary to multimode fiber joints, the loss of a single-mode fiber joint having mode field diameter mismatches is independent of the direction of propagation through the joint. In Fig. 49.4, the coupling loss between two fibers having different mode field radii is plotted as a function ratio of transmitting fiber's radii to the receiving fiber's radii.

Although for a given fiber type and design the mode field diameters are usually held to within 10%, the loss in Fig. 49.4 is plotted for ratios from 0.6 to 1.54. This large range is needed to describe the intrinsic loss for fiber joints between dispersion-shifted and **dispersion-unshifted fibers.** This ensemble of fibers may be experienced when and if dispersion-shifted fibers become prevalent in future fiber deployments. It is common to see references that the mode field diameter of dispersion-shifted fiber is similar to that

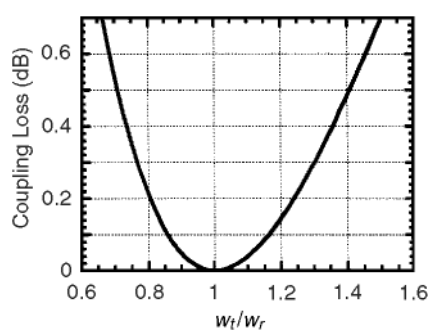


FIGURE 49.4 Plot of coupling loss vs. ratio of mode field radii.

TABLE 49.2 Mode Field Diameters of Different Fiber Types

Fiber Type	Mode Field Diameter (wavelength, 1310 nm)	Mode Field Diameter (wavelength, 1550 nm)
Dispersion-unshifted, mm	9.3	10.5
Dispersion-shifted, mm	6.5	8.1

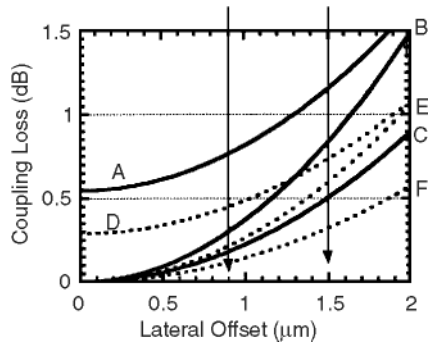


FIGURE 49.5 Plot of coupling loss vs. lateral offset for various fiber type combinations.

of today's fiber of choice, 1310-nm optimized dispersion-unshifted fiber. Although this is true, if one uses the intended wavelength of propagation, 1550-nm for dispersion-shifted and 1310-nm for unshifted fibers, it becomes important to look at the mode field diameters at similar wavelengths. This condition can exist when these different fiber types meet in a connector joint. Furthermore, both wavelengths, 1310 and 1550 nm, can be experienced. Table 49.2 summarizes some typical mode field diameters.

It is noteworthy to look at the intrinsic loss for 1310 nm propagation. The diameter ratios for this wavelength of propagation are about 0.70 or 1.43, depending on which is the receiving fiber. As Fig. 49.4 shows, the intrinsic loss due only to mode field diameter mismatch is 0.55 dB. For the case of 1550-nm propagation, the ratios are 0.77 and 1.30, and the loss due to mode field diameter mismatch is 0.30 dB. Of course, this loss must be added to any extrinsic loss due to existing misalignments.

Continuing, we can also use this general formula to analyze the types of alignments that are necessary for high-performance connectors and splices. As previously stated, angular misalignments and longitudinal offsets (endface separation) are more easily minimized, in connectors and splices, than lateral offsets. Therefore, the main objective in connector and splice design is controlling the lateral offset between the adjacent fibers. For this reason, the preceding general formula can be reduced to the following expression that describes the dependence of insertion loss on lateral offset:

$$\text{insertion loss} = 4.343\left(\frac{x}{w}\right)^2 \text{dB} \tag{49.3}$$

This relationship assumes that identical fibers are in the joint, which is a valid assumption considering the small variation in mode field diameter in the manufacture of today's fibers and its effect on coupling loss. It also assumes no reflections, that is, zero end separation or perfect index matching. This expression has been used to plot insertion loss vs. lateral offset in Fig. 49.5 for both dispersion-shifted (DS) and unshifted (US) fibers at both 1310 and 1550 nm wavelengths. Also plotted (using the preceding general formula) are losses for joints between these two different types of fibers. It should be noted that Eq. (49.3) has been verified in numerous experiments [Kummer and Fleming, 1982].

In Fig. 49.5, coupling loss (insertion loss) is plotted as a function of lateral offset between various combinations of fiber types. Curve A represents a joint between dispersion-shifted and dispersion-unshifted fibers, and curve B represents a joint between two dispersion-shifted fibers; both curves are for a propagation wavelength of 1310 nm. Also plotted is curve C, which represents a joint between two

TABLE 49.3 Insertion Loss for 1.5 μm Lateral Offset

	US/US Fibers (dB)	DS/DS Fibers (dB)	DS/US Fibers (dB)
1310 nm(λ)	0.5	0.8	1.1
1310 nm(λ)	0.3	0.6	0.7

dispersion-unshifted fibers, and curve D, a joint between a dispersion-shifted and dispersion-unshifted fibers; both of these curves are for a propagation wavelength of 1550 nm. Finally, joints between two dispersion-shifted and two dispersion-unshifted fibers, both for a wavelength of 1550 nm, are shown by curves E and F, respectively.

A typical specification for today's connectors is a maximum loss of 0.5 dB and an average loss of 0.25 dB. As can be seen, a lateral alignment of 1.5 μm is required for 0.5 dB and about 1 μm for 0.25 loss, using today's standard dispersion-unshifted fibers and an operating wavelength of 1310 nm (curve C.) The information in Fig. 49.4 can then be used to predict the performance of today's connectors (connectors having 1.5 and 1 μm lateral offsets) with the various combinations of fiber types and operating wavelengths. Table 49.3 summarizes these losses for a 1.5 μm lateral offset.

The preceding discussion is based on the fact that today's insertion loss specifications for high-performance connectors, dominated by lateral offsets, are usually a maximum of 0.5 dB and show averages of 0.25 dB that imply lateral offsets of 1.5 and 1 μm, respectively. In addition, to minimize reflections and insertion loss, the connectors must achieve fiber core-to-core contact throughout the operating range. How this submicrometer control is achieved in these butt-joint ferrule connectors is discussed in a later section.

Reflectance Factors

Reflectance is the ratio of reflected power to incident power at a fiber joint, expressed in decibels. Reflectance can, therefore, be defined by the following expression:

$$\text{reflectance} = 10 \log \left(\frac{\text{reflected power}}{\text{incident power}} \right) \text{dB} \tag{49.4}$$

Historically, insertion loss has been the only performance criterion for connectors and splices. Recently, it has been recognized that a fiber joint's reflectance performance is also important because optical reflections from refractive index discontinuities in the fiber path can adversely affect the performance of direct-detection, coherent-detection, and analog lightwave communications systems, particularly in single-mode fiber applications [Mazurczyk, 1981]. In cases having multiple reflectors in the fiber path (such as multiple connectors and splices), high reflectance may also result in the interferometric conversion of laser phase noise to intensity noise, resulting in system degradations and/or measurement limitations. Optical feedback may also cause detrimental changes in the performance characteristics of active components, resulting in further system degradations and measurement limitations. For example, optical feedback into a laser cavity can increase mode-partition noise and add noise caused by reflection-induced power fluctuations. In optical amplifiers, reflections can cause lasing or multiple reflection noise, thereby severely affecting amplifier characteristics, and drastically restricting the amounts of amplifier gain and/or reflectance that can be allowed in high-speed systems.

Thus, besides insertion loss, the reflectance of optical fiber connectors and splices has become an important optical performance criterion. In this regard, requirements for reflectance are not only routinely specified but are presently becoming more stringent, changing from the minimum of -30 to -40 dB.

Connectors

Since high-performance connectors are usually designed for **physical core-to-core contact** to reduce insertion loss, optically transparent fiber joints might be expected (effectively an optically continuous length of fiber). To understand why physical contact between fiber endfaces does not yet result in reflectances that are equivalent to a continuous fiber, we recall that when bulk silica glass is polished, a

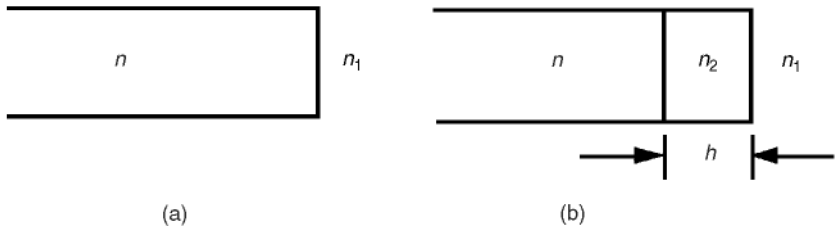


FIGURE 49.6 The model of (a) a cleaved and (b) a polished fiber endface depicting the high-index layer on the polished endface.

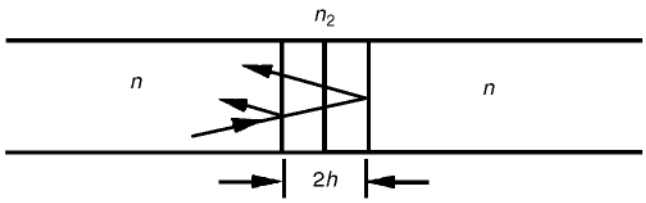


FIGURE 49.7 Schematic representation of two beam interference of a fiber joint having a high index-layer.

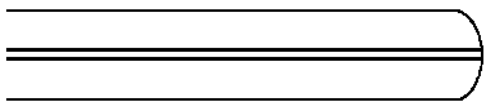


FIGURE 49.8 A nonflat, polished fiber endface.

thin layer with a refractive index higher than the bulk glass is usually formed at the polished surface [Rayleigh, 1937]. This increase in refractive index is the result of compaction of the fiber endface material (increase in density) caused by the mechanics of the polishing process. Almost all connectors use polished endfaces. This polishing-induced increase in the refractive index of the fiber endface creates an interior index discontinuity parallel to the fiber endface. This index discontinuity occurs at a fraction of a micrometer from the fiber endface, and even when fiber endface contact is achieved, these two index discontinuities within the fibers result in increased reflectance. Figure 49.6 shows a model of the endface for the **cleaved fiber** [Fig. 49.6(a)] and for the polished fiber [Fig. 49.6(b)].

The refractive index of the bulk fiber material (here the small difference between the core and cladding index is neglected), the gap material (generally air), and the index of the thin high-index layer are denoted by n, n_1, and n_2, respectively. The thickness of the high-index layer is denoted by h.

When physical contact between two polished fiber endfaces occurs, even though the interface between the fiber endface and the gap material disappears, two index discontinuities separated by a distance $2h$ still remain near the joint (as shown in Fig. 49.7).

These two interfaces can cause a significant degradation in the connector's reflectance performance while only a negligible increase in insertion loss. This discussion assumes that the fiber endfaces are polished flat. Many of the commercially available ferrule-type physical-contact (PC) connectors use nonflat endfaces, as shown in Fig. 49.8.

Usually, the fiber endface is polished spherically convex, as shown in Fig. 49.8. The reason for using a nonflat fiber endface is to ensure fiber core-to-core contact. Because of point or line contact between the two contacting spherical surfaces, the axial compressive spring force in the connector body provides the necessary force to cause **controlled elastic deformation** of the fiber/ferrule endface, thereby achieving fiber endface contact. Thus, the model presented in Fig. 49.7 still holds for such *nonflat* fiber endfaces, and the reflectance performance is still limited by the two interfaces formed by the high-index layers previously described.

Furthermore, since the refractive index n_2 and the thickness h of the high-index layer at the polished fiber endface are dependent on various polishing conditions, such as the polishing material, polishing pressure, grit size of the polishing material, etc., the reflectance of PC connectors depends critically on the polishing technique. For example, polishing with diamond with grit size of 1 or 3 μm or with 1 μm grit size aluminum-oxide coated on a polyester sheet produces a smooth fiber endface having a high-index layer of almost 0.09 μm thick with the refractive index of about 1.54 compared 1.47 for bulk fiber glass. With such a layer present on a fiber endface, a theoretical value of approximately –30 dB is predicted for the reflectance of such a PC connector. Experimentally, –34 dB has been measured for similar connectors, indicating good agreement with theoretical predictions.

When it was recently realized that it was possible to obtain better reflectance performance by using enhanced polishing techniques to significantly minimize the thickness of the compacted endface layer, the performance for some connectors changed from –30 to –40 dB. However, some of these techniques may also produce fiber endfaces that are recessed from the ceramic ferrule endface, therefore making it more difficult to achieve fiber core contact in the connector. The reflectance performance also depends on various environmental conditions. Some conditions may produce an air gap in connectors designed for physical contact, resulting in multiple beam reflections.

Splices

Recall that there are two types of fiber splices: fusion splices and mechanical splices. Fusion of the two fibers together, with respect to reflectance, results in an optically transparent length of fiber. With these types of splices it is very difficult to measure any level of reflectance. On the other hand, because the index matching material used for mechanical splices is temperature dependent, mechanical splices can exhibit wide ranges of reflectance. It is not uncommon for a mechanical splice, cleaved perpendicular fiber endfaces, to have reflectances between –25 and –45 dB over a temperature range of –40∞ and 80∞F. In order to avoid this temperature dependence, cleaved angled fiber endfaces have been recently employed. With endface angles between 5∞ and 10∞, and index matching material, mechanical splices with reflectance performances equal to fusion splices can be easily achieved. Furthermore, it can be theoretically shown that the fibers having angled endfaces can be joined within the splice hardware without regard to ensuring that the endfaces are parallel.

49.3 Multibeam Interference (MBI) Theory

MBI Effects on Transmission

The previous sections discussed factors that cause insertion loss and reflectance in the butt coupling of two single-mode fibers. It was shown that as long as fiber endface contact is achieved with connectors or perfect index matching is achieved with splices, the dominating factor that causes insertion loss in typical high-performance connectors and splices is lateral offset. Furthermore, as the connector's reflectance performance became an important performance criterion, it became a requirement that fiber core-to-core contact be achieved in order to minimize both insertion loss and reflectance. Consequently, it is necessary to ensure fiber endface contact over the expected environmental and climatic conditions. When contact is lost, an air gap is created between the fibers and two glass/air interfaces are formed. The Fresnel reflection at the two interfaces (here, for the sake of clarity, cleaved fiber endfaces are assumed) results in an additional insertion loss of 0.32 dB, as shown by the following formula:

$$\text{insertion loss} = -10\log(1-2R)\,\text{dB}$$

where

$$R = \text{Fresnel reflection} = \left(\frac{n_f - n_g}{n_f + n_g}\right)^2 \tag{49.5}$$

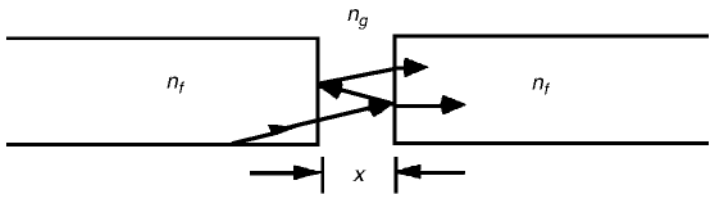


FIGURE 49.9 Multiple beam interference model for an air gap between two single-mode fibers.

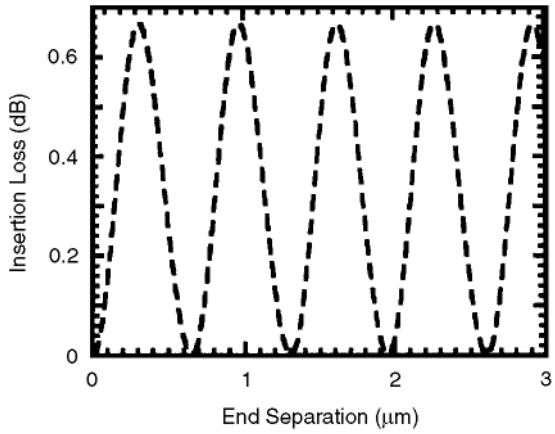


FIGURE 49.10 Plot of insertion loss vs. endface separation of a butt-joint connector.

Although an increase in insertion loss of 0.32 dB nearly doubles the typical performance of high-performance connectors, it is even more important that the two adjacent fibers and the air gap in the connection form a Fabry–Perot etalon. The resulting multiple beam interference, from this gap, can cause additional increases in insertion loss. In analyzing this effect, consider the two beams, as shown in Fig. 49.9.

Figure 49.9 represents an air gap between two fiber endfaces. The two beams shown in Fig. 49.9 will interfere with each other both constructively and destructively, depending on the physical path length difference $2x$. The effect that this two-beam interference can have on the insertion loss performance of the connector can be quite severe. The following expression can be used to evaluate the effect of this multiple beam interference for the small end separations that unfortunately might develop in some high-performance connectors. In cases having larger end separation (not applicable to connectors or splices), it is necessary to include the effects of the coherence length of the optical source, aperturing of the fiber core, and diffraction of the emitted beam.

$$L = -10\log\left[1 - 2R\left(1 - \cos\left(\frac{4\pi n_g s}{l}\right)\right)\right]\text{dB}$$

where

$$R = \left(\frac{n_g - n_f}{n_g + n_f}\right)^2 \tag{49.6}$$

This expression was used to plot insertion loss vs. end separation shown in Fig. 49.10. As shown, the insertion loss is highly dependent on the fiber endface separation x, and as a result, the insertion loss can increase by almost 0.7 dB for just a submicrometer air gap. When the realistic case of a polished connector is considered, the presence of a polishing-induced high-index layer on the fiber endface can increase the peak-to-peak variation in insertion loss from 0.7 dB to as much as 1.1 dB [Shah, Young,

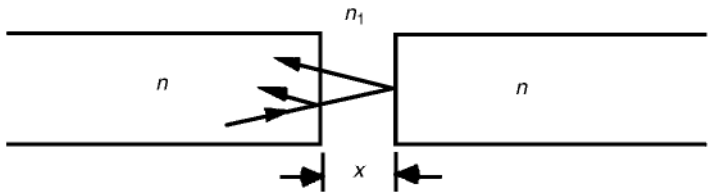


FIGURE 49.11 Schematic representation of two beam interference of a fiber joint having cleaved fiber endfaces.

and Curtis, 1987]. For a performance level of 0.5 dB, such large loss variations are clearly not acceptable and, therefore, must be avoided.

MBI Effects on Reflectance

Next, we consider the effect of multiple beam interference on reflectance performance. In analyzing this effect, consider the two reflected beams shown in Fig. 49.11.

We begin by analyzing the reflectance of a fiber butt joint that has fiber endface separation without index matching, an air gap. This type of fiber joint results in two fiber/air interfaces, and the reflectance is caused by the resulting Fresnel reflections. Again, in cases that have a large air gap, with reference to the coherence length of the optical source, multiple-beam interference effects can be neglected and the following expression can be used to calculate a reflectance of -11.4 dB:

$$\text{reflectance} = 10\log(2R)\text{dB}$$

where

$$R = \left(\frac{n_g - n_f}{n_g + n_f}\right)^2 \tag{49.7}$$

where

n_g = refractive index of gap
n_f = refractive index of fiber

Even though a reflectance level of -11.4 dB is much greater than today's typical performance requirement of <-40 dB, interference effects can occur with air-gap joints that are even more of a concern. Referring to Fig. 49.11, the two reflected beams shown can interfere with each other both constructively and destructively depending on the physical path length difference x, and the effect that this two-beam interference can have on the reflectance performance of the connector can be quite severe. Figure 49.12 illustrates the results of an often verified theoretical analysis.

As Fig. 49.12 shows, a large dependence exists on the fiber endface separation x, and as a result, the reflectance can vary from -8.5 dB to -∞ within a quarter-wavelength change in separation. Since the reflectance at its lowest value is an extremely sensitive function of air gap (asymptotic), its average value tends to be very close to the largest reflectance value, -8.5 dB for the case shown in Fig. 49.11. As previously discussed, a large change in the insertion loss (around 0.7 dB for cleaved endfaces) will simultaneously occur. Thus, when any air gap exists, the connector reflectance performance usually will be severely degraded and approach –8.5 dB, as compared to the extremely low reflectance when fiber endface contact is achieved.

In the case of the polished fiber endfaces, Fig. 49.13 is a schematic of the two-beam reflection model when a high-index layer exists on both fiber endfaces and physical contact between fiber endfaces is achieved. As Fig. 49.13 shows, the two-beam reflection model used in Fig. 49.11 can still be applied to determine the reflectance performance of the connection between two polished fiber endfaces under physical contact.

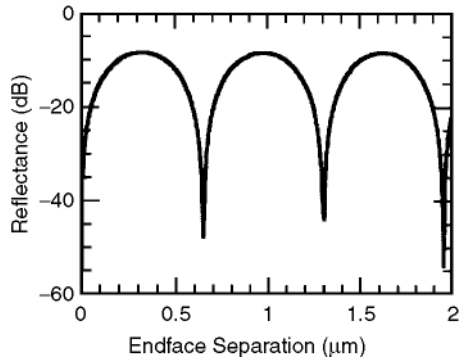


FIGURE 49.12 Theoretical plot of reflectance vs. endface separation for two cleaved fiber endfaces.

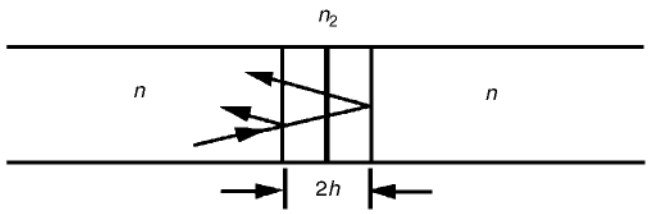


FIGURE 49.13 Schematic representation of two-beam interference of a fiber joint having a high-index layer.

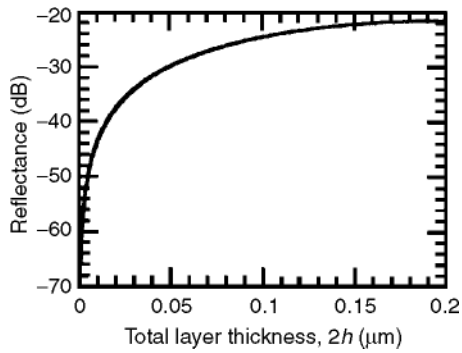


FIGURE 49.14 Reflectance vs. high-index layer thickness for index of 1.54.

The amplitudes and relative phase-shifts of the two beams reflected from the polishing induced index discontinuities (shown in Fig. 49.13) are directly determined by the refractive index of the polishing-induced layer n_2 and the total thickness of this layer $2h$, thereby determining the connector's reflectance performance. Assuming the refractive index of the layer is 1.54, the reflectance performance can be calculated as a function of the total thickness $2h$. As Fig. 49.14 shows, the reflectance increases from -∞ to -40 dB as the layer thickness increases from 0 to 0.0075 μm (total thickness is 0.015 μm).

The reflectance attains its maximum value of about -21 dB for a layer thickness of 0.1 μm (total thickness is 0.2 μm), for a wavelength of 1310 nm, and then decreases with further increase in layer thickness. As Fig. 49.14 shows, reflectances of -30, -40, and -60 dB require polishing techniques that result in a layer thickness of 0.009, 0.0075, and 0.00075 μm, respectively.

Since physical contact may be lost due to environmental factors such as temperature changes and/or due to variations in fiber endface profiles resulting in an air gap at the joint, further analysis of the

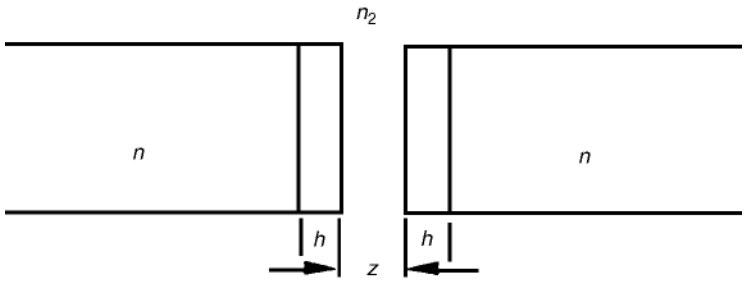


FIGURE 49.15 Air-gap joint with two polished fibers having high-index layers.

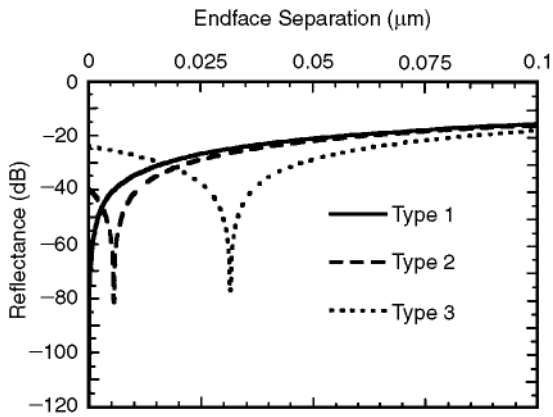


FIGURE 49.16 Reflectance vs. endface separation for three different polishing procedures: –60, –40, and –30 dB; type 1, type 2, and type 3, respectively.

interference effects is required. Using the model in Fig. 49.15, three types of plugs, polished with different polishing procedures (different high-index layer thickness), are considered.

The hypothetical plugs are such that under physical contact condition, the connector reflectance values are -60, -40, and -30 dB, respectively. It is assumed that the index of refraction of the high index layer is 1.54 and that the plugs have different layer thicknesses that yield different reflectance performance.

The reflectances of each of these joints plotted as a function of air gap between the fiber endfaces at a wavelength of 1310 nm are shown in Fig. 49.16.

The plot is an expanded view of the region near zero endface separation showing a rapid asymptotic decrease in reflectance from its value at physical contact to -∞ dB at a small nonzero separation. This effect is not consistent with the cleaved fiber endfaces, where the first reflectance minimum occurs at physical contact. The separations for minimum reflectance for the three plugs are 0.00038, 0.0056, and 0.0385 μm, respectively. Beyond these separations, the reflectance again increases. Thus, for polished fiber endfaces, a situation exists wherein the reflectance performance actually improves within a small range of nonzero endface separation.

In summary, excellent reflectance performance can be achieved with single-mode fiber, physical-contact connectors when appropriate endface polishing techniques are used. In particular, with the enhanced polishing techniques that have been recently developed, reflectance performance of <-40 dB can be routinely obtained, provided fiber endface contact is maintained. However, if the fiber/ferrule endface polishing techniques are not fully under control, the performance of physical contact connectors may not be maintained over all environmental conditions or with random intermating of connectors using different endface polishing treatments.

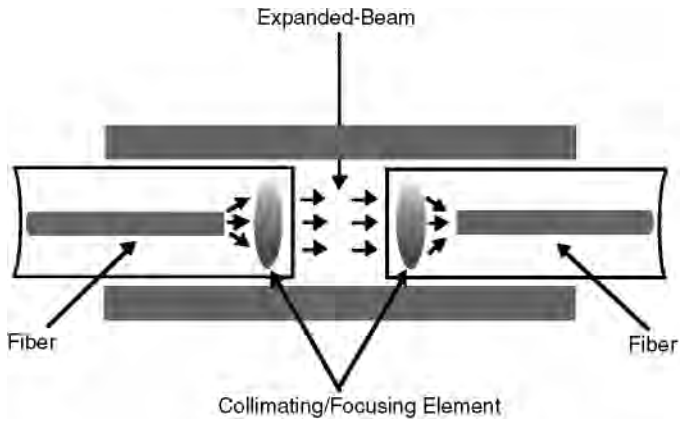


FIGURE 49.17 Schematic of an expanded-beam type optical fiber connector.

49.4 Connector Design Aspects

Introduction (Types of Connectors)

There are two types of connector designs that have received attention: the expanded-beam type and the butt-joint connector. The expanded-beam connectors have not yet found acceptance in communications systems, and are unlikely to find a niche (based on past results); therefore, this chapter does not discuss them in detail. However, for the sake of completeness, we will briefly describe the philosophy of this type of connector design.

The goal of the expanded-beam connector is to make the connector more tolerant to lateral alignment inaccuracies. To accomplish this, the fibers are terminated with a collimating/focusing element (lenses, as shown in Fig. 49.17) that expands and collimates the beam, thereby facilitating connecting and disconnecting [Nica and Tholen, 1981].

Initially, it appears that the expanded-beam connector has a large advantage over the butt coupling of fibers. However, there are aspects of the design that require the same degree of precision alignment. Although it is a fact that the expanded and collimated beam is more tolerant to lateral offsets, angular misalignments of the collimated beam, even as small as a few arc-seconds, can result in large insertion losses. These angular misalignments of the expanded beam result in a lateral displacement of the focus point, resulting in high coupling losses between the fiber and lens. Consequently, the fiber and lens must be aligned to the same submicrometer tolerances that are required in butt coupling of the fibers. Furthermore, to avoid spherical aberrations, and the resulting wavelength dependency of insertion loss, aspheric lenses must be used. Because of these concerns, and the general complexity of the designs, expanded-beam connectors have yet to find general acceptance in communications systems.

Communications systems, both multimode and single-mode fiber systems, have successfully used butt-joint connectors since the very first applications in the mid-1970s. Furthermore, despite the submicrometer alignment tolerances required, more than 30 million butt-joint connectors have been manufactured and deployed. This chapter focuses on the cylindrical-ferrule, physical-contact, butt-joint connectors that are the choice for optical communications systems. Two popular versions of these connectors, bayonet-lock fastened and push/pull snap-lock fastened are shown in Figs. 49.18 and 49.19, respectively.

Factors Contributing to Insertion Loss

Minimizing the lateral offsets between butt-coupled fibers and achieving fiber core-to-core contact are the two most challenging aspects of designing high-performance single-mode fiber connectors. The model in Fig. 49.20 shows the major inaccuracies encountered in minimizing the centering of the fiber's optical field within a single-mode fiber ferrule.

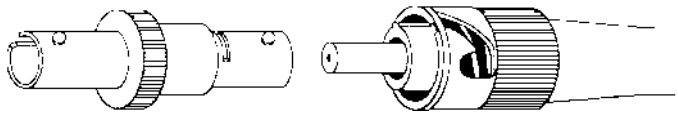


FIGURE 49.18 Bayonet-lock fastened cylindrical ferrule ST type connector.

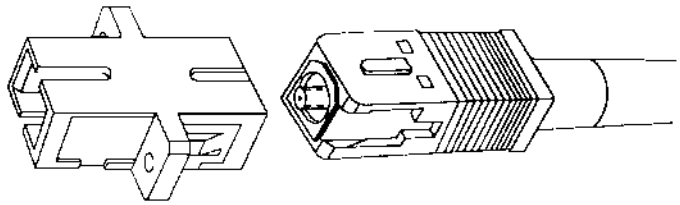


FIGURE 49.19 Push/pull snap-lock fastened cylindrical ferrule SC type connector.

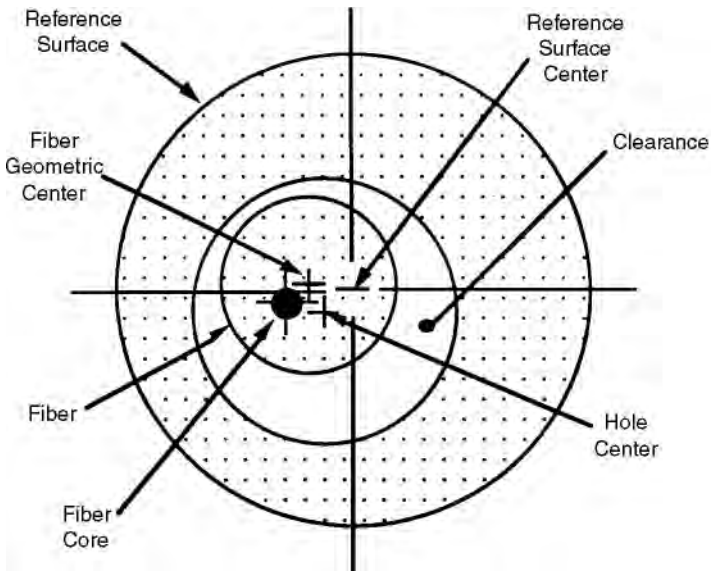


FIGURE 49.20 Model showing inccuracies encountered in assembling a connector ferrule.

In analyzing the model, the first consideration is the quality of the fiber. As Fig. 49.20 shows, the axial offsets between the geometric center of the fiber and the fiber core center (optical center of the mode field diameter), and the clearance between the ferrule's hole and the outside diameter of the fiber, are major contributors to the final quality of the connector ferrule assembly (the lateral offset between the reference surface center of the ferrule and the fiber core).

The contributions from the physical dimensions of the ferrule must be also considered. Similarly to the fiber, fabrication of the connector ferrule results in additional inaccuracies that must be accommodated in the connector design. The outside diameter of the ferrule will vary, as will the hole diameter and the offset between the hole center and the reference surface center. As can be seen in Fig. 49.20, the offset that must be minimized is the offset between the fiber core center and the reference surface center. Recall that when two single-mode fibers are butt coupled, in order to achieve insertion losses <0.25 dB, the total lateral offset should be less than 1 μm.

Statistical analyses of the inaccuracies shown in Fig. 49.20, such as Monte Carlo simulations, have been carried out to quantify the tolerances required for a specific optical performance [Young, 1984].

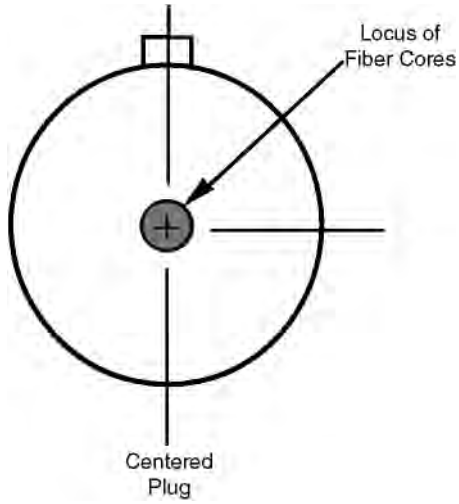


FIGURE 49.21 Ferrule having fiber cores centered to required tolerance.

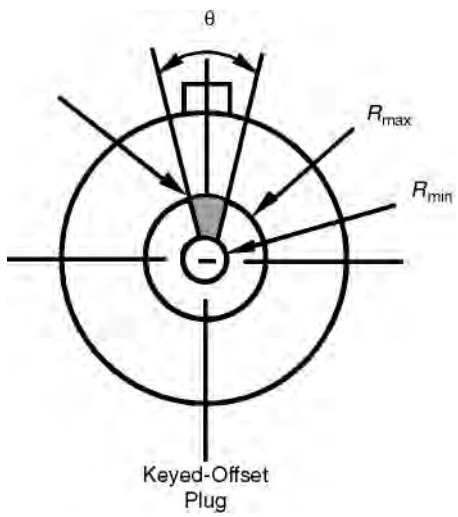


FIGURE 49.22 Model of a performance keyed ferrule.

These statistical analyses have also been experimentally verified. First, to achieve total lateral offsets of <1 μm, the diametral clearance between the hole in the ferrule and the fiber must also be <1 μm.

To more fully understand the difficulty in analyzing the required lateral offsets in the ferrule assembly, two popular ferrule assembly techniques are examined. To begin our analysis, recall that to achieve 1 μm lateral alignment between fiber cores within a connection, which is required for 0.25-dB insertion loss, we need on average 0.5 μm centering of the fiber core within each ferrule. This means that if the ferrules have the same outside diameters, and the ferrule axes are perfectly aligned by the alignment sleeve, then we need to consider only the overlap area that is established by the locus of the core centers in each ferrule assembly. This type of ferrule having the fiber cores centered to within less than 1 μm is shown in Fig. 49.21.

If we consider the popular ceramic cylindrical ferrule connectors, the manufacturing processes do not routinely yield the necessary hole centering. To overcome this aspect of these ferrules, **performance keying** or tuning is frequently used. Figure 49.22 demonstrates how performance keying can be used to

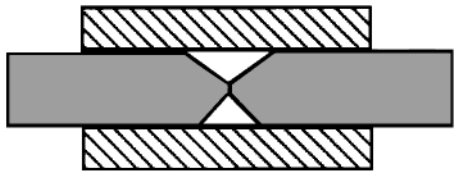


FIGURE 49.23 Schematic of a cylindrical ferrule alignment machanism.

achieve a similar size area for the locus of fiber core centers that is achieved in the directly centered ferrules. In performance keying, the location of the key is determined after the fibers are inserted into the ferrules and the connector endface is finished. Usually, the keying is done by measuring the insertion loss as a function of ferrule rotation against a standard reference ferrule. The use of performance keying is widespread with single-mode fiber cylindrical ferrule connectors.

The final factor is aligning the two ferrules and realizing a connection requires an alignment sleeve. This alignment sleeve is usually allowed to float in a housing that has the appropriate coupling features (screw threads, etc.). This component is referred to as either a coupling or an adaptor. A schematic representation of a cylindrical ferrule connector is shown in Fig. 49.23. At first observation, it might appear that the cylindrical alignment sleeve is a trivial component. However, this is not the case. The materials of the sleeve have ranged from plastic, austenitic stainless steel, copper alloys, and glass to ceramic. It appears there is a tendency today to use **single-slit ceramic sleeves**. The concern over the alignment sleeve designs for the cylindrical ferrule connector is driven by the fact that the sleeve must coaxially align ferrules that have unequal diameters. Even though the diameter of ferrules can be manufactured to a tolerance of 1 μm, an insertion loss performance of <0.25 dB requires lateral alignment of the cores of <1 μm. Additionally, this 1 ㎜ tolerance has to accommodate all inaccuracies, including the fiber core offsets within the ferrules. For these reasons, the alignment sleeve is indeed an important component of an optical fiber connector.

Factors Contributing to Reflectance (Physical Core-to-Core Contact)

As mentioned previously, besides achieving submicrometer lateral alignments, in order to realize low insertion loss and low reflectance, high-performance connectors must achieve and maintain fiber core-to-core contact. Again referring to physical contact cylindrical ferrule connectors, at first glance (Fig. 49.23) it appears trivial to achieve physical core contact. However, one has to investigate the geometry in finer detail.

As stressed in this chapter, the lateral offset between adjacent single-mode fibers must be controlled to less than 1 μm in order to achieve 0.25 dB insertion loss performance. The cylindrical geometry therefore must maintain very small diametral clearances between the ferrules and the alignment sleeve to minimize lateral offsets. These small clearances require precise perpendicularity of the ferrule/fiber endfaces to achieve physical contact of the fiber cores. In practice, this required precision in perpendicularity has not been realized. As a result, the usual practice is to use spherically polished endfaces [Shintaku, Nagase, and Sugita, 1991]. In theory, the spherical endface should result in the fiber core being the highest point on the ferrule endface, thereby always achieving core contact. However, since the material of choice for manufacturing the submicrometer precision ferrule is a ceramic, spherical endface, polishing tends to result in a slight recessing of the fiber endface within the ferrule tip (Fig. 49.24). This recessing is caused by the differential polishing rates of the ceramic ferrules and the glass fibers. This recessing of the fiber endfaces is overcome in the cylindrical ferrule connector by using **controlled elastic deformation** of the ferrules.

Therefore, when two ferrules are forced into contact (Fig. 49.25) by the compression springs within the fastening hardware, line contact is established and large hertz contact forces are exerted on the ferrules' endfaces. If one correctly selects the ferrule's material properties, specifies the radius of the sphere, and controls the amount of fiber recess, then a **controlled elastic deformation** of the ferrule can be used that

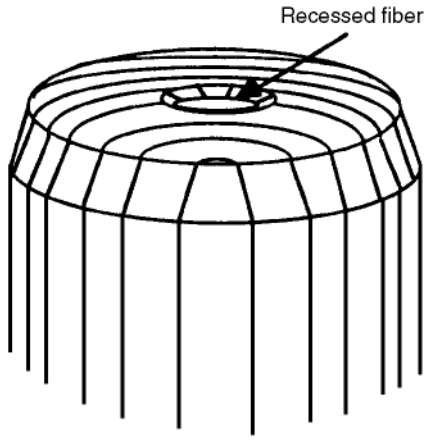


FIGURE 49.24 Ferrule endface with recessed fiber.

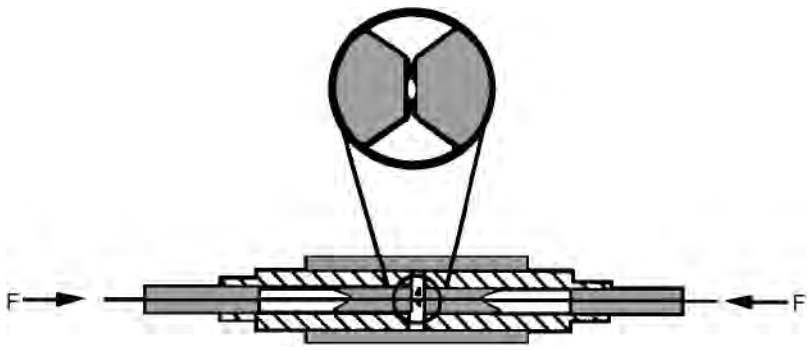


FIGURE 49.25 Ferrule endface contact of spherical ferrule endfaces with fiber recess.

results in fiber core-to-core contact. If these parameters are not controlled, an unstable physical contact connection can result.

To summarize this section on fiber core-to-core contact connectors, the **controlled elastic deformation** technique with type PC connectors is presently the overwhelming choice for low reflectance and low insertion loss connections. There are many details and features that must be controlled to achieve and maintain this core-to-core contact to provide the appropriate elastic deformation within the connectors. Some of these are manufacturing know-how and must be fully developed and socialized within the connector community before appropriate intermateability standards can be realized.

49.5 Splicing Design Aspects

Fiber splices are used in applications where the goal is to achieve a permanent joint between two fibers. Since the splice will be permanent, there are no intermateability concerns as with connectors. Without these concerns, splices are easily designed to have very low insertion losses, <0.1 dB, and very low reflectance, <-55 dB. With the recent improvement in fiber geometry, multiple fiber splices are now becoming feasible as well. As previously mentioned, there are two types of fiber splices: mechanical splices and fusion splices. The recent use of angled endfaces in mechanical splices have made their optical performance equal to that of fusion. A large differentiator between them is the required level of initial capital investment. Fusion splicing is capital intensive, whereas mechanical splicing is more or less a buy-as-needed process.

TABLE 49.4 Typical Steps in Fusion Splicing

- Cable structure removal
- Fiber coating removal
- Cleaning of bare fiber
- Fiber cleaving
- Fusion process
- Installation of strength reinforcement detail

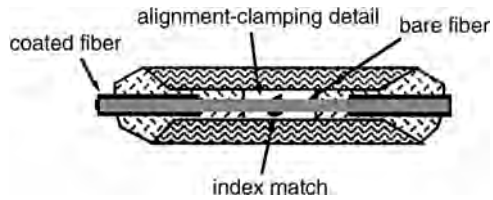


FIGURE 49.26 Schematic of a typical mechanical splice.

Mechanical Splices

Mechanical splicing techniques use butt-joint coupling similar to those used in connectors. However, the fact that the splice is permanent implies there do not need to be any intermateable parts. This fact results in their higher optical performance and lower cost. Essentially, the mechanical splice is a consumable part that mechanically aligns the fibers by referencing on their outside diameters [Patterson, 1988]. The splicing process is quite simple. The plastic coating on the fiber is removed and the fiber endface is cleaved, either perpendicular or with an angle. The two cleaved fibers are inserted or placed in the consumable part (the splice) and are usually held in place by a simple clamping method. The low loss is attributable to the use of a continuous alignment surface similar to a v-groove and the use of an index matching material in the very small gap that exists between the fibers to eliminate reflections. A typical mechanical splice is depicted in Fig. 49.26.

The first mechanical splices and most used today employ perpendicular endfaces. Since the cleaved endfaces are not exactly perpendicular, fiber core contact is not achieved within the splice. Consequently there exists a gap, which is filled with index matching material. Because the refractive index of the available materials varies with temperature, two index mismatches are created, one at each fiber end. These index mismatches result in multiple beam interference as in the case with connectors. These mismatches are not sufficient to cause significant increases in insertion loss (about 0.002 dB) but can severely affect the reflectance performance of the splice. As previously mentioned in this chapter, reflectances ranging between -25 and -40 dB over the expected temperature range can be realized. With the recent reflectance requirements of less than -40 dB, angled endfaces are starting to be used in mechanical splices to avoid this temperature dependence [Young, Shah, and Curtis, 1988]. Simple cleaving tools have been developed that provide endface angles ranging between 5∞ and 10∞. These angled endfaces can be used in existing mechanical splices and thereby yield reflectance performance equal to that of fusion splices.

Fusion Splices

Much progress has been made in the performance of equipment used to perform arc fusion of optical fibers. Essentially, the process is quite simple: the fibers' butt ends are heated to just below their melting point and pressed together to form a continuous fiber [Kato et al. 1984]. After the fusion process is complete, a strength reinforcement member is applied to protect the bare fiber. Table 49.4 lists the typical steps used in the fusion process. In Fig. 49.27, a simple schematic of a holding and alignment mechanism that is part of most splicing machines is shown. In reality, these mechanisms are quite sophisticated, involving computer controlled movements to achieve accurate alignment and positioning. The machines are fully automatic and also provide an estimation of the splice loss.

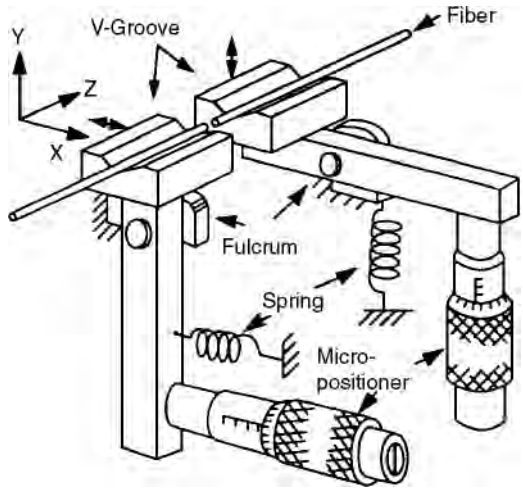


FIGURE 49.27 Schematic of holding and alignment mechanism.

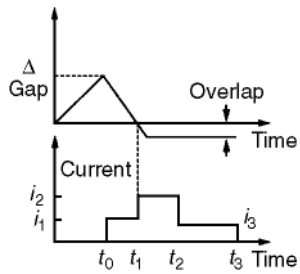


FIGURE 49.28 Example of a fusion splicing cycle.

An example of a fusion splicing cycle is depicted in Fig. 49.28. The cycle is usually fully automatic. The operator loads the cleaved fibers in the v-grooves and grossly positions the ends. The cycle begins by moving the fibers into contact and then separating them a fixed amount. Then the arc is energized with a specific current and the fibers are moved toward each other. This step is used to clean and round the edges of the fiber ends. When the fiber ends reach the contact point, the current is increased to soften the fibers and the fibers are driven past the original contact point (pressed together). Finally, the current is decreased and brought to zero. Although the process seems quite simple, the tolerances for these movements and currents can be quite demanding.

In the fusion splicing of multimode fibers, the surface tension between the fiber endfaces, while softened, can be used to self-align the fibers with reference to their physical geometric axes (outside diameters). This technique makes the fusion splicing of multimode fibers easy and reliable. With the small mode field diameters of single-mode fibers and the existence of fiber core eccentricities, this self-alignment technique has usually needed to be avoided. To avoid this self-alignment requires precise control of the temperature or softness of the end faces during fusion, the distance the fibers move beyond their original contact location (overlap distance), and the fiber cleave angles. Cleave angles of less than 1∞, specific arc duration times within 0.1 s, and overlaps to within 1 μm are usually maintained. Also, since the insertion loss target is typically <0.1 dB, it usually requires that the fibers are aligned with respect to their mode field diameters not their outside diameters. Various techniques have been used to achieve and monitor these alignments. For these reasons, the fusion splicer has developed into a very sophisticated apparatus.

Recently, there has been a marked improvement in the concentricity and diameters of single mode fibers. These improvements have had a positive impact on fusion splicing of these fibers. The self-aligning

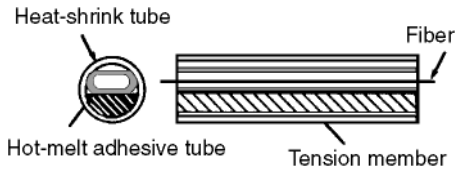


FIGURE 49.29 Fusion splice reinforcement detail.

of the fibers is becoming acceptable. Also, the relaxed tolerances on fiber cleave angles and overlap distances have resulted in simpler and less costly splicers, as well as the multiple fiber fusion splices.

Finally, to complete the splicing operation, the splice must be reinforced. Typically, the tensile strength of an unprotected or reinforced splice is between 10 and 15% of the original coated fiber's strength. The strength of the splice can be improved to between 50 and 70% with the use of a reinforcement detail. An example of a reinforcement detail is shown in Fig. 49.29.

49.6 Conclusions

Major advances have been made in the past 20 years with regard to fiber optic connector and splicing technology. The performance of these components has dropped from 1.5 to 0.25 dB for connectors and to <0.1 dB for splices. Along with impressive improvement transmission performance, reflectance performances have also improved dramatically. Connectors and mechanical splices have been reported and marketed with reflectances < -55 dB. These performances have also been achieved with a drastic reduction in cost. However, in spite of a lot of work over the past 15 years, the intermateability standards for these connectors have not yet been published. Hopefully, intermateability standards will soon be finalized at the international level.

With regard to fusion splices there has also been impressive improvement in recent years. Much of this improvement can be attributed to the improvement in the geometry of the optical fibers themselves. The increase in the precision of the fibers has resulted in simpler and less costly splicing machines, as well as the feasibility of multifiber splicing.

Defining Terms

Cleaved fiber: A fiber having an endface prepared by a combination of a prescribed stress distribution and an induced flaw that follows the Griffith fracture theory.

Controlled elastic deformation: A technique used to achieve core-to-core contact when the fiber endfaces have been recessed below the endface of the connector's ferrule.

Dispersion-unshifted fiber: A single-mode fiber designed to have a minimum dispersion at a wavelength of 1550 nm, the intended operating wavelength window.

Dispersion-unshifted single-mode fiber: A single-mode fiber designed to have minimum dispersion at a wavelength of 1310 nm, the intended operating wavelength window.

Performance keying: A method used to overcome the errors in coaxially aligning the fiber core center with the ferrule's reference surface axis resulting in low insertion loss connections.

Physical core-to-core contact: A connector designed to achieve physical and optical contact between the core areas of adjacent fibers, thereby minimizing insertion loss and reflectance caused by reflections.

Reflectance: Reflectance is analogous with the term return loss, the only difference being that one term is the negative of the other.

Single-slit alignment ceramic sleeves: A ceramic sleeve having a precise inside diameter and a longitudinal slit along its length that increases its compliance to varying ferrule diameters.

References

Chu, T.C. and McCormick, G.R. 1978. Measurement of loss due to offset, end separation and angular misalignment in graded-index fibers. *Bell Sys. Tech. J.,* 57(3):595–602.

Kato, Y., Tanifuji, T., Kashima, N., and Arioka, R. 1984. Arc-fusion splicing of single-mode fibers. *Appl. Opt.,* 23:2654–2658.

Kummer, R.B. and Fleming, S.R. 1982. Monomode optical fiber splice loss. In *Proceedings of Optical Fiber Communications Conference.*

Marcuse, D. 1977. Loss analysis of single-mode fiber splices. *Bell Syst. Tech. J.,* 56:703–117.

Mazurczyk, V.J. 1981. Sensitivity of single-mode buried heterostructure lasers to reflected power at 274 Mbit/s. *Electron. Lett.,* 17:143.

Miller, C.M., Mettler, S.C., and White, L. 1986. *Optical Fiber Splices and Connectors,* Marcel Dekker, New York.

Nemota, S. and Makimoto, T. 1979. Analysis of splice loss in single-mode fibers using a Gaussian field approximation. *Opt. Quantum Elec.,* 11:447.

Nica, A. and Tholen, A. 1981. High-efficient ball-lens connector and related functional devices for single-mode fibers. In *Proceedings of 7th European Conference on Optical Communications.*

Patterson, R.A. 1988. Mechanical optical fiber splice containing and articulated conformable metallic element. In Proceedings of International Wire and Cable Symposium, Reno, NV.

Rayleigh, Lord. 1937. The surface layer of polished silica and glass with further studies on optical contact. *Proc. R. Soc. London, Ser. A,* 160:507–526.

Shah, V., Young, W.C., and Curtis, L. 1987. Large fluctuations in transmitted power at fiber joints with polished endfaces. In *Proceedings of OFC'87,* TUF4, Reno, NV.

Shintaku, T., Nagase, R., and Sugita, E. 1991. Connection mechanism of physical-contact optical fiber connectors with spherical convex polished ends. *Appl. Opt.,* 30(36).

Wagner, R.E. and Sandahl, C.R. 1982. Interference effects in optical fiber connections. *J. Appl. Opt.,* 21(8):1381–1385.

Young, W.C. 1984. Single-mode fiber connectors—Design aspects. In *Proceedings 17th Annual Connector and Interconnection Technology Symposium,* Electronic Connector Study Group, pp. 295–301.

Young, W.C., Shah, V., and Curtis, L. 1988. Optimization of return loss and insertion loss performance of single-mode fiber mechanical splices. In Proceedings of International Wire and Cable Symposium, Reno, NV.

Further Information

Several technical conferences provide the latest trends and results on fiber optic connectors and splices. Two of the major conferences to follow are the Optical Fiber Communications (OFC) Conference, sponsored by the Optical Society of America (OSA) and IEEE (LEOS), and the International Wire and Cable Symposium, sponsored by the U.S. Army Electronics Command, Fort Monmouth, NJ. Both of these conferences are held annually.

Three major journals that report latest work are the *Journal of Lightwave Technology* (OSA), *Applied Optics* (OSA), and *Electronic Letters* (IEE).

Generic requirements and future trends can be followed in the *Bellcore Digest of Technical Information,* Piscataway, NJ.

To follow the standardization of connectors and splices in the U.S., the Telecommunications Industries Association (TIA), Washington, D.C., is recommended. To follow international standardization, refer to the International Electrotechnical Commission in Geneva.

50

Passive Optical Components

Joseph C. Palais
Arizona State University

50.1 Introduction

Passive optical components play a large role in making simple fiber systems (such as point-to-point links) practical and extending their applications to more complicated systems (such as local-area and subscriber-distribution networks). This chapter limits its coverage to a few of the more important bulk passive components not covered elsewhere in this book.

The components to be described are the following optical devices: GRIN rod lens, isolator, directional coupler, star coupler, optical filter, attenuator, circulator, mechanical switch, and polarization controller.

50.2 Losses in a Passive Optical Component

A generic passive optical component is shown in Fig. 50.1. It has several ports (as few as two and, in some cases, over 100), all of which can be input or output ports. The transmission loss generally refers to the coupling efficiency between any two ports. In terms of decibels, the transmission loss is then

$$L = -10 \log(P_{out}/P_{in}) \tag{50.1}$$

where P_{in} and P_{out} are the input and output powers at the two ports being considered. **Insertion loss** is commonly used to refer to the loss in a desired direction of signal travel, and **isolation** or **directionality** is used for the transmission loss between two ports where coupling is undesirable. Typical losses for different devices will be presented in the paragraphs where the devices themselves are described.

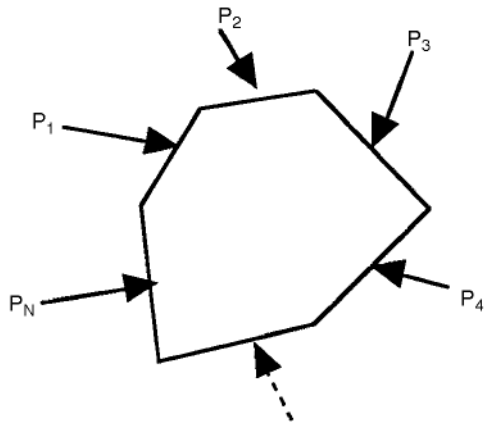


FIGURE 50.1 Passive optical component with *N* ports.

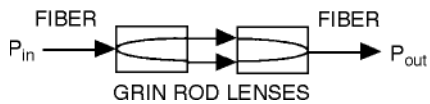


FIGURE 50.2 GRIN rod lens connector.

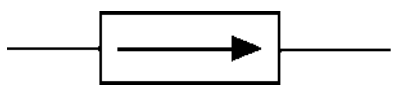


FIGURE 50.3 Isolator showing preferred direction of beam travel.

50.3 Grin Rod Lens

The GRIN rod lens is a short cylindrical length of graded index (GRIN) material that functions as a lens. It is shown schematically in Fig. 50.2 being used to couple between two fibers. An advantage of this connector is the improved tolerance to lateral offsets because of the expanded beam. Another advantage is that significant gaps can exist between the lenses without large losses because of the collimation of the beam. This property also allows the GRIN rod lens to be part of functional devices where gaps are needed to accommodate additional components such as filters, beamsplitters, and gratings. Examples of useful devices employing the GRIN lens will be described later.

50.4 Isolator

An isolator, illustrated in Fig. 50.3, is a one-way transmission line. Its defining characteristics are insertion loss (the loss in the direction of allowed propagation) and isolation (the loss in the undesired direction of propagation). Typical insertion losses are less than 1 dB and isolations are 25 dB or more.

Isolators are used to protect laser diodes from back reflections, which can increase the laser noise level. Isolators are also needed in optical amplifiers to prevent feedback, which could initiate unwanted oscillations of a beam bouncing back and forth through the amplifier due to reflections.

The principle of the isolator is illustrated in Fig. 50.4. A polarizer aligns the beam along one direction. The polarization is rotated 45∞ by a **Faraday rotator** to align it with an output polarizer. For a beam traveling in the opposite direction (to the left in the figure), the beam is rotated such that it aligns at 90∞ to the polarizer. It is thus blocked from further travel.

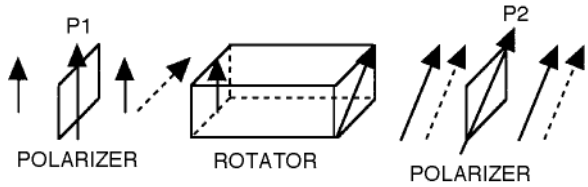


FIGURE 50.4 Principle of the optical isolator. P1 and P2 are polarizers that allow propagation of waves polarized in the directions indicated by the long arrows. The shorter solid arrows refer to waves traveling in the preferred direction (left to right). The dashed arrows refer to waves traveling in the reverse (undesired) direction. Ideally, no wave will get past polarizer P1 (traveling to the left) because the wave will enter P1 at 90∞ with respect to the allowed direction of wave polarization.

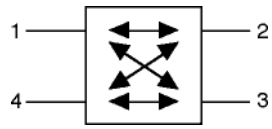


FIGURE 50.5 A four-port directional coupler.

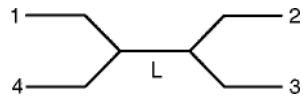


FIGURE 50.6 Fused biconically tapered directional coupler. Two fibers are fused over length *L*.

50.5 Directional Couplers

The four-port directional coupler is shown schematically in Fig. 50.5 [Palais, 1998]. Arrows in the diagram indicate the directions of desired power flow. Power entering port 1 is divided between ports 2 and 3 according to the desired splitting ratio. Any desired split can be obtained, but the most common one is a 50:50 split. Under ideal (lossless) conditions, the insertion loss to ports 2 and 3 would then be 3 dB and the isolation (coupling loss to port 4) would be (in decibels) infinite. In practice, there are **excess losses** associated with the coupler. Excess losses, which can be as low as a few tenths of a decibel for a good coupler, add to the splitting loss. That is, the total insertion loss is the sum of the splitting loss and the excess loss. Typical insertion losses for the 3-dB coupler are, thus, around 3.3 dB. Typical isolations are 40 dB or more.

Applications include local monitoring of the power levels of the transmitting light source, tapping power from a data bus in a local-area network, signal splitting in a subscriber distribution system, power splitting in an optical sensor, and combining signals in a heterodyne (coherent) receiver.

The fused biconically tapered coupler is one popular construction technique. Two fibers are fused together over a short length by heating. During the fusing process, the fibers are stretched, creating a tapered region in the fused section. The fused fiber coupler is shown schematically in Fig. 50.6. The amount of light coupled between the two fibers depends on the coupling length (L) and can be varied from 0 to 100%. A 50% split, creating a 3-dB coupler, is most common.

Directional couplers can also be constructed using GRIN rod lenses. Illustrated in Fig. 50.7 are a prism/beamsplitter and a reflective coating directional coupler.

50.6 Star Coupler

The star coupler is a multiport device (more than four ports) which commonly distributes power entering one port equally among a number of output ports. Referring to Fig. 50.8, power entering any of the N input ports is equally distributed among all of the N output ports. As an example, an ideal 16×16 star coupler would deliver 1/16 of the input power to each output port, resulting in an insertion loss of 12.04 dB. If the coupler were not ideal, for example, having a 2-dB excess loss, the actual insertion loss would be 14.04 dB.

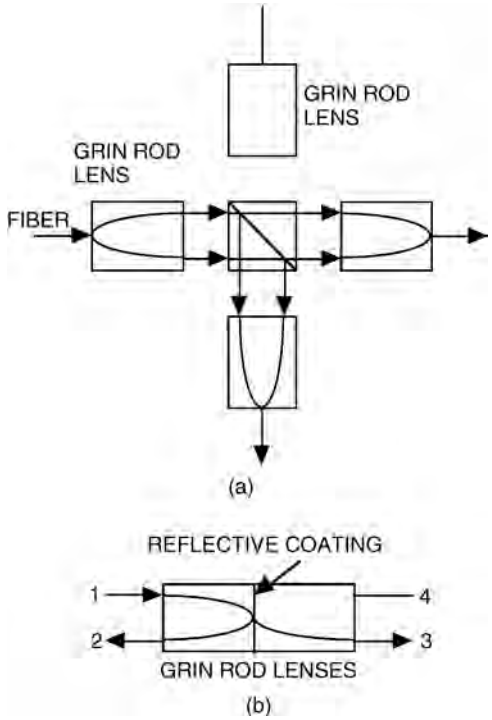


FIGURE 50.7 Directional couplers: (a) prism/beam-splitter coupler and (b) GRIN rod lens reflective coupler.

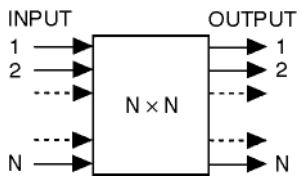


FIGURE 50.8 A star coupler.

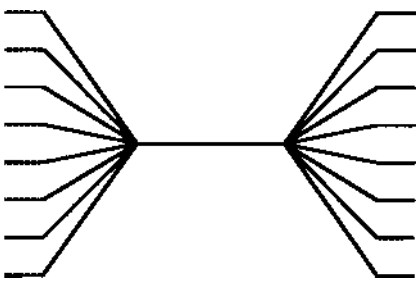


FIGURE 50.9 An 8 × 8 fused star coupler.

The star coupler is heavily used in the passive star local-area network architecture. Using the 16 × 16 coupler just described, 16 terminals (each with its own transmitter and receiver) can be interconnected.

The most common implementation is the fused tapered star coupler, as illustrated in Fig. 50.9, where 8 fibers are fused to produce an 8 × 8 star. The 16 × 16 star would require that 16 fibers be fused together.

50.7 Optical Filter

Optical filters separate (or combine) the power carried by different wavelengths that travel through a fiber. Just like electronic filters, they pass or block different wavelengths. Unlike electronic filters, they often also spatially separate the various wavelengths present, as indicated in Fig. 50.10.

Optical filters are necessary in wavelength-division multiplexing (WDM) systems where multiple optical carriers traverse the same fiber to increase system capacity. The fused directional coupler can be used for a two-channel system because the splitting is wavelength dependent. As an example, referring to the directional coupler in Fig. 50.6, if signals at wavelengths 1300 and 1550 nm are received along the same fiber, they can be separated by a coupler that transmits the 1300-nm beam from port 1 to 2 and the 1550-nm beam from port 1 to 3. At the transmitter, the same coupler can combine the two signals onto a fiber connected to port 1 if the 1300-nm signal enters port 2 and the 1550-nm signal enters port 3. As is commonly the case, the same device can serve as either a **multiplexer** or a **demultiplexer**.

The reflective coupler drawn in Fig. 50.7(b) can also serve as a two-channel multiplexer/demultiplexer. In this case, as illustrated in Fig. 50.11, the coating is reflective at one of the system wavelengths and transmissive at the other.

WDM systems carrying more than two channels require more complex multiplexers/demultiplexers. Two such devices are sketched in Fig. 50.12. One uses a GRIN rod to collimate the beam diverging from the input fiber. The collimated beam strikes the blazed grating, which directs different wavelengths into different directions. The lens then refocuses the collimated (but redirected) beams onto different output fibers leading to the receivers. In the second device, a curved reflector performs the same function as the GRIN rod lens. Both of these components can multiplex/demultiplex 10 or more channels (wavelengths). Insertion losses and isolations are on the order of 1 dB and 25 dB, respectively.

Another filter is based on Bragg reflection by a waveguide grating imbedded in the fiber, as illustrated in Fig. 50.13. This device has a periodic refractive index variation for a short length along the axis of the fiber. The variation is produced by exposing the fiber to a periodic variation in ultraviolet light, which permanently changes the refractive index of the glass fiber, thus producing a grating. The transmission characteristics are determined by the grating periodicity and the depth of the refractive index changes.

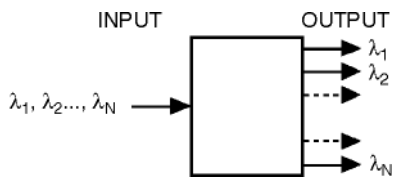


FIGURE 50.10 Optical filter used as a demultiplexer.

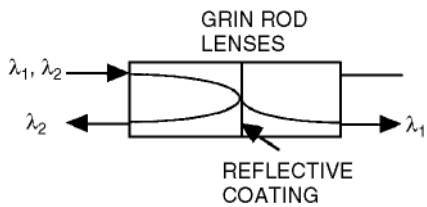


FIGURE 50.11 Two-wavelength demultiplexer. By reversing the direction of the arrows, the device acts as a multiplexer.

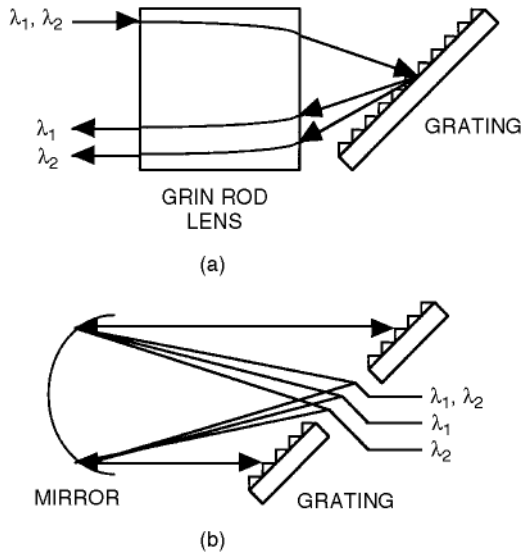


FIGURE 50.12 Two multiplexers/demultiplexers: (a) grating and GRIN rod lens; (b) grating and curved reflector.

FIGURE 50.13 Bragg reflection filter.

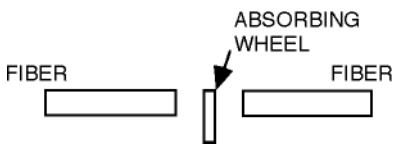


FIGURE 50.14 Fiber variable attenuator using an absorbing wheel.

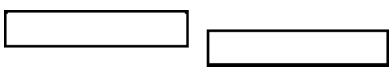


FIGURE 50.15 Fiber variable attenuator using movable offset fibers.

50.8 Attenuator

Optical attenuators make it possible to change the light intensity in a fiber. Applications include testing receivers at various light levels (for example to determine the receiver's dynamic range) and adjusting the light intensity to prevent saturation of the receiver.

An absorbing type attenuator appears in Fig. 50.14. The absorbing wheel has a different amount of absorption at different points so that rotating it provides a variation in the attenuation. Misaligned fibers can also be used as attenuators, as indicated in Fig. 50.15. Varying the gap or the lateral offset provides variable attenuation.

50.9 Circulator

An optical circulator, whose operation is sketched in Fig. 50.16, directs the signal sequentially from one port to the next. That is, an input at port 1 exits at port 2, an input at port 2 exits at port 3, and an input at port 3 exits at port 1. Insertion losses are typically just under 1 dB and isolations are above 25 dB.

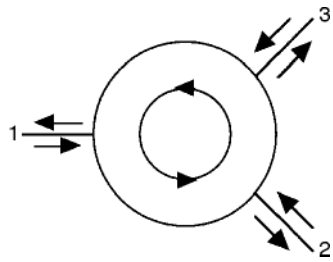


FIGURE 50.16 Fiber circulator.

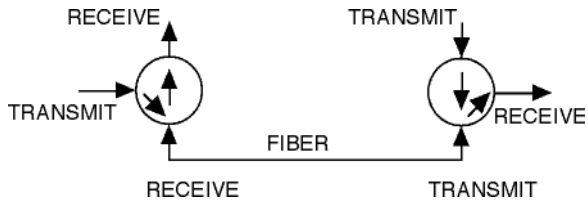


FIGURE 50.17 Full duplex fiber transmission system.

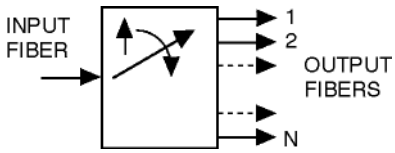


FIGURE 50.18 A 1 ¥ *N* mechanical switch.

The circulator can be used to decouple transmitted and received signals traveling along the same fiber (that is, in a bidirectional communications system such as that sketched in Fig. 50.17).

A circulator uses a Faraday rotator, similar to that found in the optical isolator. A fuller description of the circulator appears in Kashima [1995].

50.10 Mechanical Switch

An optical switch allows an incoming signal to be diverted to any one of two or more different output fibers, as schematically shown in Fig. 50.18. A mechanical switch does this by physical movement of the fibers or some other element of the switch. The insertion loss (coupling to the desired port) should be no more than a few tenths of a decibel and the isolation (undesired coupling to any other port) should be more than 40 dB. Switching speeds are on the order of a few milliseconds, not fast by electronic and optical measures, but quick enough for many fiber applications.

On/off switches allow power to be supplied to a system under test only when needed. Multiple output port switches find applications in measurement systems (for example, to selectively measure the output power from a sequence of fibers). Switches can also redirect signals from a disabled fiber link onto an alternate fiber route. They are used to produce bypass paths directing signals around disabled equipment, such as a node in a local-area network.

In an on/off switch, a block is inserted into a small gap between fibers in the off position and removed in the on position. In a $1 \times N$ switch, the end of the transmitting fiber physically moves into alignment with the desired output fiber.

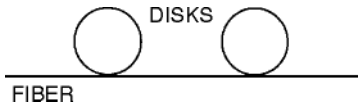


FIGURE 50.19 Fiber polarization controller.

50.11 Polarization Controller

The polarization controller is a fiber device for producing a desired state of polarization. Although most fiber systems operate independently of the state of polarization of the optical beam, coherent systems and certain sensors are very much dependent on the state of polarization of the optical signal.

A simple polarization controller is constructed by winding the fiber around a disk whose diameter is on the order of a centimeter [Chen, 1996]. Usually, two or three of these fiber disks are connected in series (Fig. 50.19). Rotating the disks about the axis of the fiber produces a differential change in the refractive index of the two normal orthogonally polarized beams traversing the fiber. Any desired polarization state can be obtained with the appropriate rotation of the disks.

Defining Terms

Demultiplexer: Separates different wavelengths entering from a single transmission fiber and directs them to separate receivers.

Directionality: The fractional change in power level (expressed in decibels) when a signal passes between two ports which are ideally meant to be uncoupled.

Excess losses: The fractional amount of input power (expressed in decibels) that reaches all the desired output ports of an optical component.

Faraday rotator: A component which rotates the polarization of the electric field passing through it. Often, the rotation of an incident linearly polarized beam is 45∞. The rotation is nonreciprocal.

Insertion loss: The fractional change in power level (expressed in decibels) when a signal passes between two ports which are designed to be coupled.

Isolation: Same as directionality.

Multiplexer: Combines two or more wavelengths from separate light sources onto a single fiber for transmission.

References

Chen, C.L. 1996. *Elements of Optoelectronics and Fiber Optics,* Irwin, Chicago, IL.
Kashima, N. 1995. *Passive Optical Components for Optical Fiber Transmission,* Artech, Boston, MA.
Palais, J.C. 1998. *Fiber Optic Communications*, 4th ed., Prentice-Hall, Englewood Cliffs, NJ.

Further Information

Information on components for optical fiber systems appears in several professional society journals. These include the *IEEE Journal of Lightwave Technology*, the *IEEE Photonics Technology Letters*, and the Optical Society of America's *Applied Optics*. Valuable information is also contained in several trade magazines such as *Lightwave* and *Laser Focus World*.

51

Semiconductor Optical Amplifiers

Daniel J. Blumenthal
University of California

51.1 Introduction

The deployment of optical amplifiers in fiber optic communications systems has received widespread attention over the past several years. Signal repeaters are needed to overcome inherent transmission losses including fiber absorption and scattering, distribution losses, and connector and component losses. Traditional repeaters are based on optoelectronic conversion where the optical signal is periodically converted to an electronic signal, remodulated onto a new optical signal, and then transmitted onto the next fiber section. An alternative approach is to use repeaters based on optical amplification. In optical repeaters, the signal is amplified directly without conversion to electronics. This approach offers distinct advantages including longer repeater spacings, simultaneous multichannel amplification, and a bandwidth commensurate with the transmission window of the optical fiber. The deployment of optical amplifiers in operational networks has reduced maintenance costs and provided a path for upgrading the existing fiberplants since the amplifiers and fibers are transparent to signal bandwidth.

Two classes of optical amplifiers are used in fiber-based systems: active fiber amplifiers and semiconductor optical amplifiers (SOAs). In this chapter we concentrate on semiconductor amplifiers. Fiber-based amplifiers have advanced more rapidly into the deployment phase due to their high-output saturation power, high gain, polarization insensitivity, and long excited state lifetime that reduces crosstalk effects. Fiber amplifiers have been successfully used in the 1.55-μm fiber transmission window but are not ideally suited for the 1.31-μm fiber transmission window. Fiber amplifiers are covered in detail in the companion chapter in this handbook. Recent developments in SOAs have led to dramatic improvements in gain, saturation power, polarization sensitivity, and crosstalk rejection. Semiconductor amplifiers also have certain characteristics that make their use in optical networks very desirable: (1) they have a flat gain bandwidth over a relatively wide wavelength range that allows them to simultaneously amplify

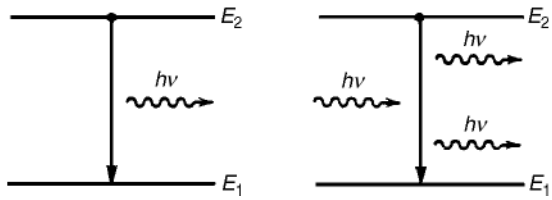


FIGURE 51.1 Spontaneous and stimulated emission occuring between the two energy states of an atom.

signals of different wavelengths, (2) they are simple devices that can be integrated with other semiconductor based circuits, (3) their gain can be switched at high speeds to provide a modulation function, and (4) their current can be monitored to provide simultaneous amplification and detection. Additionally, interest in semiconductor amplifiers has been motivated by their ability to operate in the 1.3-μm fiber transmission window. For these reasons, semiconductor optical amplifiers continue to be studied and offer a complementary optical amplification component to fiber based amplifiers.

In this chapter, we first discuss the principle of operation of SOAs followed by amplifier design considerations. The gain characteristics are described next with discussion on small-signal gain, wavelength dependence, gain saturation, dynamic range, polarization sensitivity, and noise. Amplification of optical pulses and multichannel amplification are treated followed by a brief discussion of applications.

51.2 Principle of Operation

A semiconductor optical amplifier operates on the principle of stimulated emission due to interaction between input photons and excited state electrons and is similar to a laser in its principle of operation. The semiconductor can be treated as a two energy level system with a ground state (valence band) and excited state (conduction band), as shown in Fig. 51.1. Current is injected into the semiconductor to provide an excess of electrons in the conduction band. When a photon is externally introduced into the amplifier, it can cause an electron in the conduction band to recombine with a hole in the valence band, resulting in emission of a second photon identical to the incident photon (**stimulated emission**). As these photons propagate in the semiconductor, the stimulated emission process occurs over and over again, resulting in **stimulated amplification** of the optical input. Since an SOA can amplify input photons, we can assign it a gain. As will be discussed in further detail, the **material gain** is only a function of the basic device composition and operating conditions, whereas the **amplifier gain** defines the relationship between the input and output optical power.

In addition to amplification of the input photons (signal), it is important to consider how noise is generated within the amplifier when an electron in the conduction band spontaneously recombines with a hole without the aid of a photon. This process, known as **spontaneous emission**, results in an emitted photon with energy equal to the energy difference between the electron and hole. Photons created from spontaneous emission will propagate in the amplifier and experience gain through stimulated emission. This process leads to amplification of spontaneous emission and is called **amplified spontaneous emission (ASE)**. ASE is a noise mechanism as it is not related to the input signal and is a random process. Additionally, ASE takes away amplifier gain that would otherwise be available to the signal.

51.3 Types of Semiconductor Optical Amplifiers

SOAs can be classified as either **subthreshold** or **gain clamped**. Subthreshold amplifiers are lasers operated below threshold, and gain-clamped amplifiers are lasers operated above threshold but used as amplifiers. Subthreshold SOAs can be further classified according to whether optical feedback is used. If the SOA amplifies the optical signal in a single pass, it is referred to as a **traveling wave amplifier (TWA)**, as shown in Fig. 51.2. The second type of subthreshold amplifier is a **resonant amplifier**, which contains

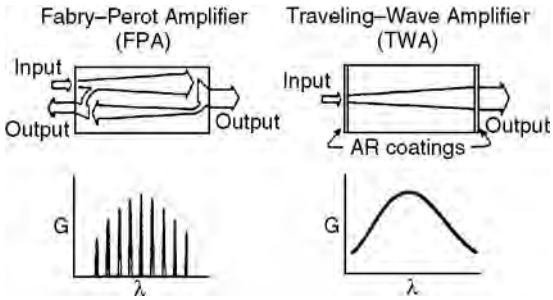


FIGURE 51.2 Classification of semiconductor optical amplifiers. (*Source:* Saitoh, T. and Mukai, T. 1991. In *Coherence, Amplification, and Quantum Effects in Semiconductor Lasers*, Ed. Y. Yamamoto, John Wiley & Sons, New York.)

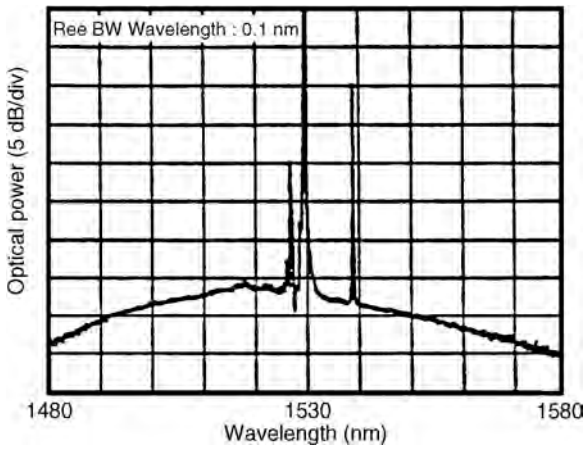


FIGURE 51.3 Output optical spectrum of a gain clamped amplifier with the main lasing mode and the identified amplified signal.

a gain medium and some form of optical feedback. In this case, the gain is resonantly enhanced at the expense of limiting the gain bandwidth to less than that of the TWA case for an equivalent material. An example of a resonant amplifier is the **Fabry–Perot** (FP) **amplifier** shown in Fig. 51.2. The FP amplifier has mirrors at the input and output ends, which form a resonant cavity. The resonant cavity creates an optical **comb filter** that filters the gain profile into uniformly spaced **longitudinal modes** (also see Fig. 51.2). The TWA configuration has a bandwidth limited by the material gain itself but is relatively flat, which is desirable for an optical communications system application. Typical 3-dB bandwidths are on the order of 60–100 nm. Although the FP amplifier exhibits very large gain at wavelengths corresponding to the longitudinal modes, the gain rapidly decreases when the input wavelength is offset from the peak wavelengths. This makes the gain strongly dependent on the input wavelength and sensitive to variations that may occur in an optical communications system.

Gain-clamped semiconductor optical amplifiers operate on the principle that the gain can be held constant by a primary lasing mode and signals can be amplified if their wavelength is located away from the main lasing mode. Laser structures suitable for this approach are **distributed feedback (DFB)** and **distributed Bragg reflector (DBR)** lasers, which lase into a single longitudinal mode. Since only a single mode oscillates, the remainder of the gain profile is available for amplification. Figure 51.3 illustrates the output optical spectra of a gain-clamped amplifier with the main lasing mode and the amplified signal identified. A primary advantage of gain clamped amplifiers is a reduction in crosstalk in multichannel amplification applications discussed in Section 51.7.

51.4 Design Considerations

For optical communications systems, traveling wave SOAs are desirable due to the wide gain bandwidth and relatively small variation in gain over a wide signal wavelength range. Most practical TWAs exhibit some small ripples in the gain spectrum that arise from residual facet reflections. A large effort has been devoted to fabricate amplifiers with low cavity resonances to reduce **gain ripple**. For an amplifier with facet reflectivities R_1 and R_2, the peak-to-valley ratio of the output intensity ripple is given by [Dutta and Simpson, 1993]

$$V = \left[\frac{1 + G_s\sqrt{R_1 R_2}}{1 - G_s\sqrt{R_1 R_2}} \right] \tag{51.1}$$

where G_s is the single-pass gain of the amplifier. For the ideal case $R_1, R_2 \to 0$, $V = 1$, that is, no ripple at cavity mode frequencies in the output spectrum. A practical value of V should be less than 1 dB. Thus, for an amplifier designed to provide gain $G_s = 25$ dB, the facet reflectivities should be such that $\sqrt{R_1 R_2}$ must be less than 3.6×10^{-4}.

Three principle schemes exist for achieving low facet reflectivities. They are: (1) antireflection dielectric coated amplifiers [Olsson, 1989], (2) buried facet amplifiers [Dutta et al., 1990], and (3) tilted facet amplifiers [Zah et al., 1987]. In practice, very low facet reflectivities are obtained by monitoring the amplifier performance during the coating process. The effective reflectivity can be estimated from the peak-to-peak ripple at the FP mode spacings caused by residual reflectivity. Reflectivities less than 10^{-4} over a small range of wavelengths are possible using antireflection coatings [Olsson, 1989]. In buried facet structures, a transparent window region is inserted between the active layer ends and the facets [Dutta et al., 1990]. The optical beam spreads in this window region before arriving at the semiconductor–air interface. The reflected beam spreads even farther and does not couple efficiently into the active layer. Such structures can provide reflectivities as small as 10^{-4} when used in combination with antireflection dielectric coatings. Another way to suppress the FP modes of the cavity is to slant the waveguide (gain region) from the cleaved facet so that the light incident on it does not couple back into the waveguide [Zah et al., 1987]. The process essentially decreases the effective reflectivity. A combination of antireflection coating and the tilted stripe can produce reflectivities less than 10^{-4}. A disadvantage of this structure is that the effective reflectivity of the higher order modes can increase, causing the appearance of higher order modes at the output which may reduce fiber-coupled power significantly.

Another important consideration in amplifier design is choice of semiconductor material composition. The appropriate semiconductor bandgap must be chosen so that light at the wavelength of interest is amplified. For amplification of light centered around 1.55 μm or 1.31 μm, the InGaAsP semiconductor material system must be used with optical gain centered around the wavelength region of interest.

51.5 Gain Characteristics

The fundamental characteristics of the optical amplifiers such as small-signal gain, gain bandwidth, gain saturation, polarization sensitivity of the gain, and noise are critical to amplifier design and use in systems. The component level behavior of the optical amplifier can be treated in a manner similar to electronic amplifiers.

Small-Signal Gain and Bandwidth of Traveling Wave Amplifiers

The optical power gain of an SOA is measured by injecting light into the amplifier at a particular wavelength and measuring the optical output power at that wavelength. The gain depends on many parameters, including the input signal wavelength, the input signal power, material gains and losses, amplifier length, current injection level, etc. For a TWA operated with low-input optical power, the small-signal gain is

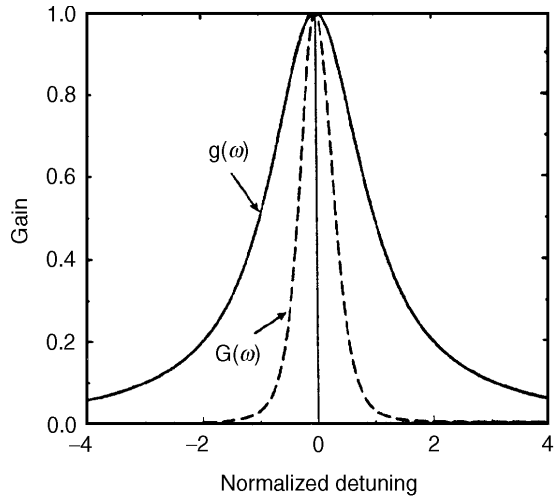


FIGURE 51.4 Lorentzian material gain profile and the corresponding amplifier gain spectrum for an amplifier modeled as a two-level atomic system. (*Source:* Agrawal, G.P. 1995. In *Semiconductor Lasers: Past, Present, and Future*, Ed. G.P. Agrawal, AIP Press.)

given by the unsaturated single-pass gain of the amplifier

$$G_0(\nu) = \frac{P_{\text{out}}(\nu)}{P_{\text{in}}(\nu)} = \exp[(\Gamma g_0(\nu) - \alpha_m)L] \tag{51.2}$$

where Γ is the **optical mode confinement factor**, g_0 is the **unsaturated material gain coefficient**, α_m is the **absorption coefficient**, and L is the amplifier length. The material gain coefficient g_0 has a Lorentzian profile if the amplifier is modeled as a two-level atomic system. An important distinction is made between the material gain bandwidth and the amplifier signal bandwidth. The 3-dB bandwidth (full width at half-maximum) of g_0 is $\Delta\nu_g = 1/\pi T_2$, and the 3-dB bandwidth of the TWA signal gain [Eq. (51.2)], is

$$\Delta\nu_A = \Delta\nu_g\left(\frac{\ell\mathbf{n}2}{\Gamma g_0 L - \ell\mathbf{n}2}\right) \tag{51.3}$$

Figure 51.4 illustrates that the material gain bandwidth is greater than the amplifier small-signal gain bandwidth.

Small-Signal Gain and Bandwidth of FP Amplifiers

The small-signal power transmission of an FP amplifier shows enhancement of the gain at transmission peaks and is given by [Saitoh and Mukai, 1991]

$$G_0^{FP}(\nu) = \frac{(1-R_1)(1-R_2)G_0(\nu)}{(1-\sqrt{R_1R_2}G_0(\nu))^2 + 4\sqrt{R_1R_2}G_0(\nu)\sin^2[\pi(\nu-\nu_0)/\Delta\nu]} \tag{51.4}$$

where ν_0 is the cavity resonant frequency and $\Delta\nu$ is the **free spectral range** (also called the longitudinal mode spacing) of the SOA. The single-pass small-signal gain G_0 is given by Eq. (51.2). Note that for $R_1 = R_2 = 0$, G_0^{FP} reduces to that of a TWA. The 3-dB bandwidth B (full width at half-maximum) of an FP

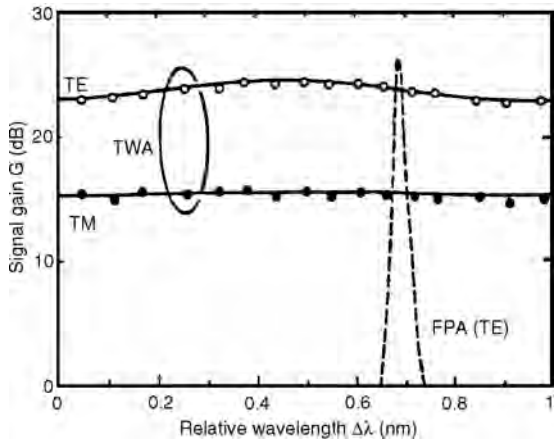


FIGURE 51.5 Unsaturated gain spectra of a TWA and an FP amplifier within one free spectral range. (*Source:* Saitoh, T., Mukai, T., and Noguchi, Y. 1986. In *First Optoelectronics Conference Post-Deadline Papers Technical Digest*, Paper B11-2, The Institute of Electronics and Communications Engineers of Japan, Tokyo.)

amplifier is expressed as [Saitoh and Mukai, 1991]

$$B = \frac{2\Delta n}{\pi}\sin^{-1}\left[\frac{1-\sqrt{R_1R_2}G_0}{(4G_0\sqrt{R_1R_2})^{1/2}}\right] \tag{51.5}$$

whereas the 3-dB bandwidth of a TWA is three orders of magnitude larger than that of the FP amplifier since it is determined by the full gain width of the amplifier medium itself. Figure 51.5 shows the small-signal (unsaturated) gain spectra of a TWA within one free spectral range [Saitoh, Mukai, and Noguchi, 1986]. Solid curves represent theoretical FP curves fitted to the TWA experimental data. Experimental FP amplifier data are also shown (dashed curve). It can be seen that the signal gain fluctuates smoothly over the entire free spectral range of the TWA in contrast to the FP amplifier, where signal gain is obtained only in the vicinity of the resonant frequencies.

Wavelength Dependence of Gain

As seen previously, the amplifier gain varies with the input wavelength for both the TWA and FP cases. The peak and width of the gain can also vary as a function of injection current. The material gain vs. wavelength for a typical amplifier is shown in Fig. 51.6 as a function of injection current. As the injection current increases, the peak gain increases and the location of the peak gain shifts toward shorter wavelengths.

Gain Saturation

The gain of an optical amplifier, much like its electronic counterpart, can be dependent on the input signal level. This condition, known as **gain saturation**, is caused by a reduction in the number of electrons in the conduction band available for stimulated emission and occurs when the rate of input photons is greater than the rate at which electrons used for stimulated emission can be replaced by current injection. The time it takes for the gain to recover is limited by the spontaneous lifetime and can cause intersymbol interference among other effects, as will be described. The **saturated material gain coefficient** is given by

$$g(\nu) = \frac{g_0(\nu)}{1+\frac{P}{P_s}} \tag{51.6}$$

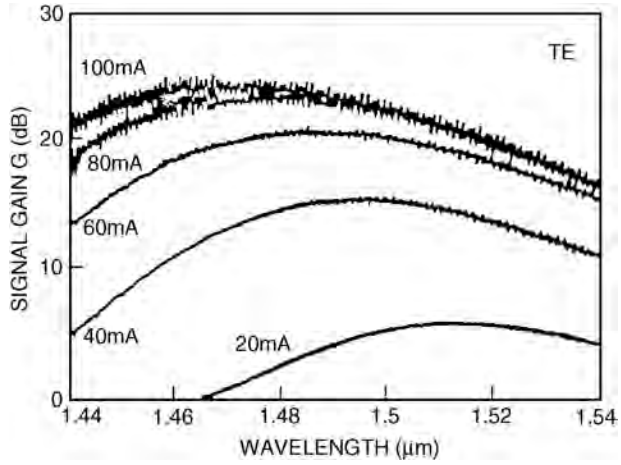


FIGURE 51.6 Gain spectrum of a TWA at several current levels. (*Source:* Saitoh, T. and Mukai, T. 1987. *Electron. Lett.*, 23:218.)

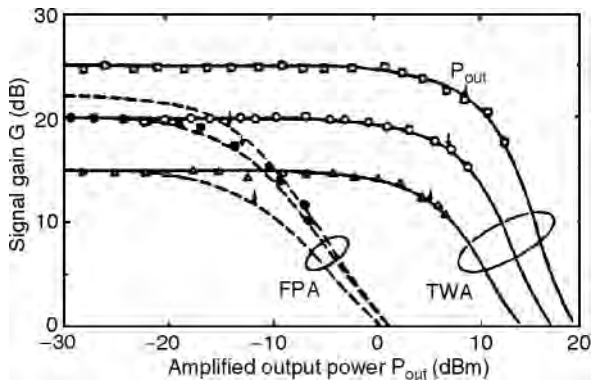


FIGURE 51.7 Theoretical and experimental signal gain of a TWA and FP amplifier as a function of amplified output power. (*Source:* Saitoh, T. and Mukai, T. 1987. *Electron. Lett.*, 23:218.)

where g_0 is the unsaturated material gain coefficient, P is the total optical power in the active layer of the amplifier, and P_s is the saturation power defined as the light intensity that reduces the material gain g to half its value ($g_0/2$). In general, the optical saturation power is related to the carrier lifetime but can be reduced as the optical intensity in the amplifier increases. The single-pass saturated signal gain $G(\nu)$ is given by [Saitoh and Mukai, 1991]

$$G(\nu) = G_0(\nu)\exp\left[-\frac{P_{out}-P_{in}}{P_s}\right] = G_0\exp\left[-\frac{G-1}{G}\frac{P_{out}}{P_s}\right] \tag{51.7}$$

where P_{in} and P_{out} are the input and the output optical powers and G_0 is the unsaturated gain given by Eq. (51.2). Figure 51.7 shows experimental gain saturation characteristics of both FP and TWAs [Saitoh and Mukai, 1987] along with theoretical curves.

Dynamic Range

The dynamic range is defined as the range of input power for which the amplifier gain will remain constant. The gain saturation curves shown in Fig. 51.8 [Adams et al., 1985] show the relationship between input power and gain for various current injection levels. The flat regions are the unsaturated regions.

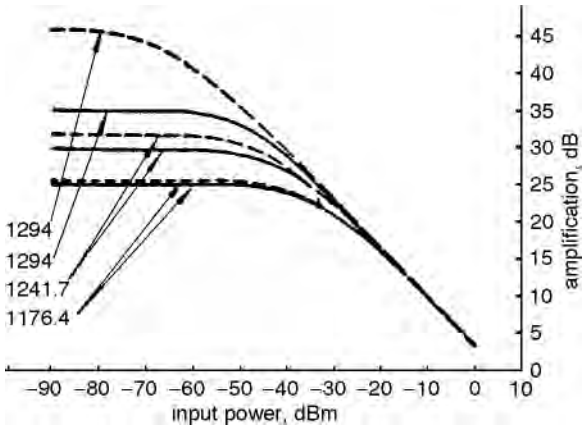


FIGURE 51.8 Effect of amplifier injection current on variation in gain as a function of optical input power. Flat regions are unsaturated gain. Range of unsaturated gain vs. input power changes as a function of unsaturated gain. Values indicated on curves are for injected current density. Dashed curves are obtained without taking into account ASE. (*Source:* Adams, M.J. et al. 1985 *IEEE Proceedings* Part J, 132:58.)

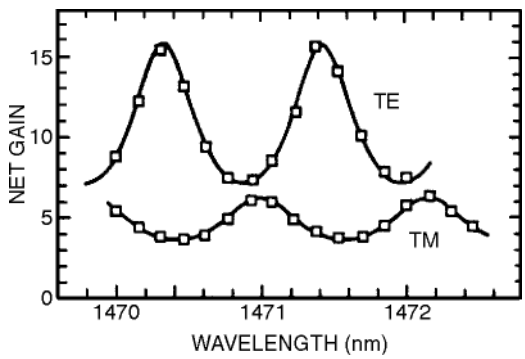


FIGURE 51.9 Theoretical and experimental gain spectra of TE- and TM-polarized input signals. (*Source:* Jopson, R.M. et al. 1986. *Electron. Lett.*, 22:1105.)

As the gain increases, the 3-dB rolloff moves toward lower input power, leading to a decrease in dynamic range.

Polarization Sensitivity

In general, the optical gain of an SOA is **polarization dependent**. It differs for the transverse-electric (TE) and transverse-magnetic (TM) polarizations. Figure 51.9 shows the gain spectra of a TWA for both TE and TM polarization states [Jopson et al., 1986]. The polarization-dependent gain feature of an amplifier is undesirable for lightwave system applications where the polarization state changes with propagation along the fiber.

Several methods of reducing or compensating for the gain difference between polarizations have been demonstrated [Saitoh and Mukai, 1991]. A successful technique that leads to polarization dependent gain on the order of 1 dB involves the use of material strain [Dubovetsky et al., 1994]. The measured gain as a function of injection current for TE and TM polarized light for a tilted-facet amplifier is shown in Fig. 51.10 [Zah et al., 1987]. Figure 51.11 shows the measured optical gain plotted as a function of output power for regular double heterostructure (DH) and multi-quantum well (MQW) amplifiers

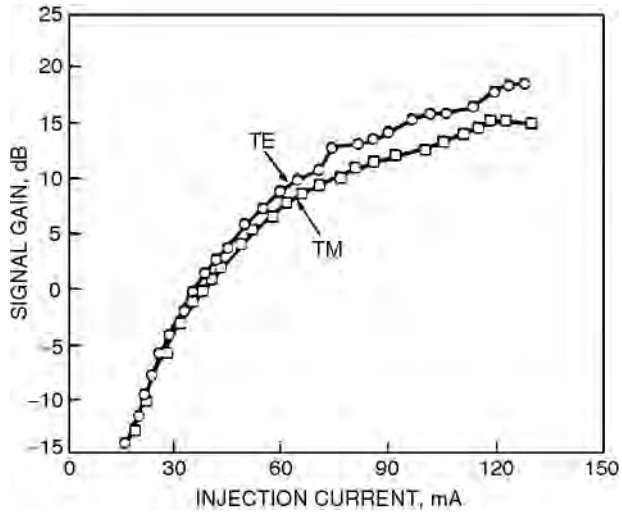


FIGURE 51.10 Measured gain as a function of injection current. (*Source:* Zah, C.E. et al. 1987. *Electron. Lett.*, 23:990.)

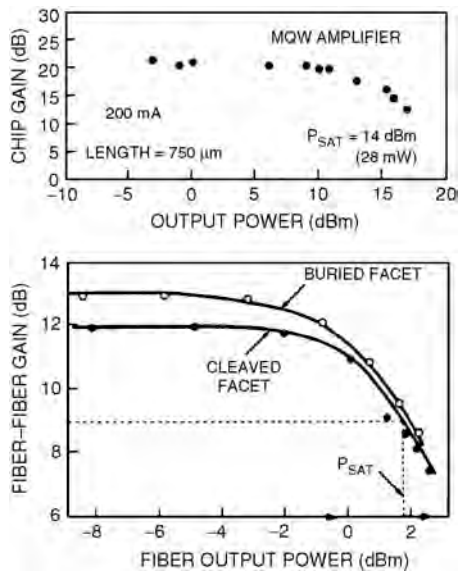


FIGURE 51.11 Measured gain as a function of output power for regular double heterostructure and multiquantum well amplifiers. (*Source:* Dutta, N.K. and Simpson, J.R. 1993. *Progress in Optics*, Ed. E. Wolf, Vol. 31, North-Holland, Amsterdam.)

[Dutta and Simpson, 1993]. The multiquantum well amplifier result is shown for the TE mode, whereas the gain difference between TE and TM modes for double heterostructure amplifiers was less than 1 dB.

Amplifier Noise

Semiconductor optical amplifiers add noise to the amplified optical signal in the form of amplified spontaneous emission. This noise can reduce the signal-to-noise ratio of the detected output and can reduce the overall gain available to the signal through gain saturation. ASE forms a noise floor with a wavelength dependence approximately that of the gain. Figure 51.12 shows a typical optical spectrum at

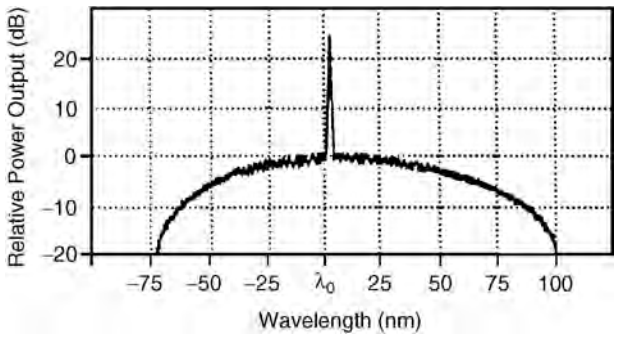


FIGURE 51.12 A typical optical spectrum at the output with the amplified signal and the amplified spontaneous emission. (*Source:* BT&D Technologies, Publication No. DS007, 1992.)

the output with the amplified signal and the ASE. Semiconductor optical amplifiers can be characterized by a noise figure. Typical noise figures are in the range of 5–8 dB and can be calculated from the carrier density due to current injection N, the carrier density required to reach transparency N_0, the saturated gain g, and the internal amplifier losses α_{int} [Saitoh and Mukai, 1991]:

$$F_n = \frac{SNR_{in}}{SNR_{out}} = 2\left(\frac{N}{N-N_0}\right)\left(\frac{g}{g-\alpha_{int}}\right) \tag{51.8}$$

In calculating signal-to-noise ratio, it is useful to consider the total output of an SOA after it has been converted to an electronic signal using a photodetector. Here we consider direct detection where a square law detector is used to convert the optical signal to an electrical signal. In addition to the amplified signal, several noise terms result from interaction between signal and noise in the square law detection process [Mukai, Yamamoto, and Kimura, 1982; Yamamoto, 1980]: amplified signal shot noise, spontaneous emission shot noise, signal-spontaneous beat noise, spontaneous-spontaneous beat noise, and signal excess noise. The first and third types of noise are proportional to the signal power, whereas the second and fourth types of noise are generated from the ASE that exists independently of the input signal. For a single amplifier, the two beat noise terms dominate as they are G times stronger [Saitoh and Mukai, 1991] than the shot noise contributions. The spontaneous-spontaneous beat noise can be reduced by loading a narrowband optical filter matched to the signal frequency, since this noise arises from the beat between the ASE components over a wide gain spectrum [Saitoh and Mukai, 1991]. For a direct detection system with a high-gain optical amplifier prior to the photodetector, the electrical SNR is given by [Jopson and Darcie, 1991]

$$SNR_{dd} = \frac{(P_{in}/h\nu)^2}{[FP_{in}/h\nu + (F/2)^2\delta\nu_{opt}]B_e} \tag{51.9}$$

where the input power to the amplifier is P_{in}, the optical bandwidth of the output of the amplifier is $\delta\nu_{opt}$, the receiver electrical bandwidth is B_e, the amplifier noise figure is F, and the optical frequency is ν.

51.6 Pulse Amplification

The large bandwidth of a TWA suggests that they are capable of amplifying very fast optical pulses without significant pulse distortion. However, the output pulses can be distorted either when the optical power exceeds the saturation power or when the pulse width gets short compared to the carrier lifetime [Agrawal and Olsson, 1989].

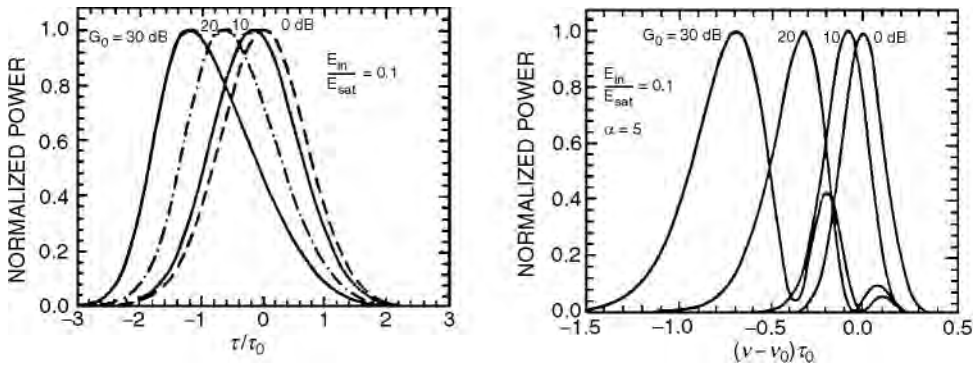


FIGURE 51.13 Output pulse shapes and spectra for several values of the unsaturated gain when the input pulse is Gaussian. The ratio of input pulse energy to the saturation energy is 0.1. (*Source:* Agrawal, G.P. and Olsson, N.A. 1989. *IEEE J. Quant. Electron.*, 25:2297.)

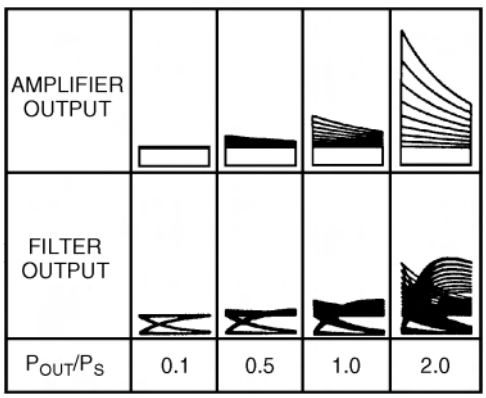


FIGURE 51.14 Effect of gain saturation of binary bit eye diagram. Upper figures show amplifier output as output power varies with saturation power. Decrease in bit height as a function of time is due to time constant of gain saturation. Lower figures show effect of electronic bandpass filtering prior to level discrimination in an optical receiver. Bit period is half the gain saturation time constant. (*Source:* Jopson, R.M. and Darcie, T.E. 1991. In *Coherence, Amplification, and Quantum Effects in Semiconductor Lasers*, Ed. Y. Yamamoto, John Wiley & Sons, New York.)

The Gaussian pulse and its spectrum for various small-signal gains G_0 when the pulse width τ_p is much smaller compared to carrier lifetime τ_c is shown in Fig. 51.13 [Agrawal and Olsson, 1989]. The input pulse energy E_{in} is equal to 0.1 times the saturation energy E_{sat} of the SOA. The pulse gets distorted such that its leading edge becomes sharper compared with the trailing edge. When the pulse width is on the order of the carrier lifetime, saturation can lead to a closing of the eye of a pseudo-random bit stream, as shown in Fig. 51.14 [Jopson and Darcie, 1991]. If a sequence of short pulses (bits) are injected, where the bit duration is less than the carrier lifetime, **bit patterning** can occur.

When the pulse width is comparable to τ_c, the saturated gain has time to recover during the pulse. Figure 51.15(a) shows [Agrawal and Olsson, 1989] the effect of gain recovery on the output pulse shape when the input pulse is Gaussian. As the input pulse width is increased, the output pulse becomes less asymmetric and becomes broader than the input pulse. The gain recovery mainly affects the trailing edge of the pulse. In the presence of partial gain recovery, the output spectrum also becomes less asymmetric and the spectral shift becomes smaller. When $t_p >> t_c$, the gain recovery is complete and the output pulse as well as its spectrum become symmetric. Figure 51.15(b) shows [Agrawal and Olsson, 1989] the output

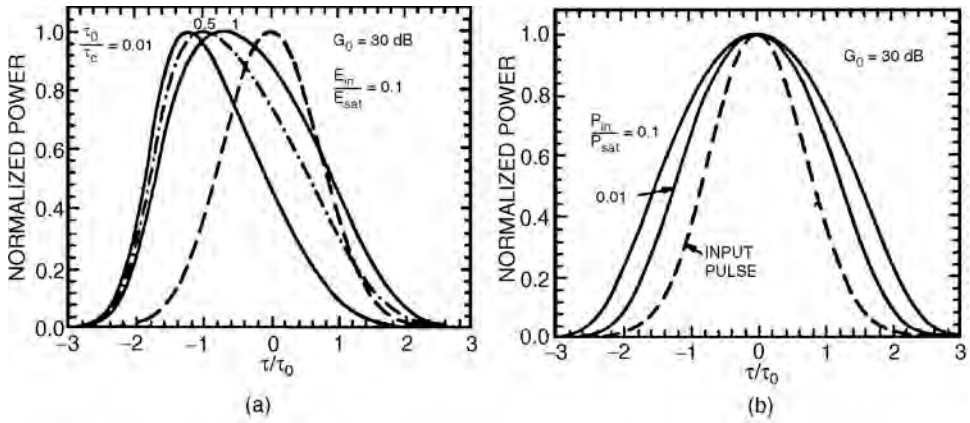


FIGURE 51.15 Output pulse shapes for (a) several values of pulse width in the partial gain recovery regime and (b) several values of input power when the pulse width is much larger than the carrier life time. (*Source:* Agrawal, G.P. and Olsson, N.A. 1989. *IEEE J. Quant. Electron.*, 25:2297.)

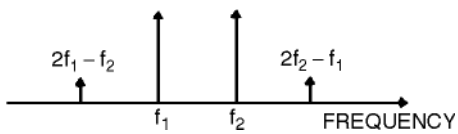


FIGURE 51.16 Generation of new frequencies due to FWM interaction between the signals at frequencies f_1 and f_2.

pulses when $t_p >> t_c$. The pulse is also broadened because the peak experiences less amplification than the wings because of gain saturation.

51.7 Multichannel Amplification

A TWA can greatly simplify optical repeaters used in **wavelength-division multiplexed** (WDM) systems, since its wide gain bandwidth allows a single amplifier to simultaneously amplify signals of different wavelengths. In practice, however, several phenomena in SOAs can induce **interchannel crosstalk: cross-gain modulation (XGM)** and **four-wave mixing (FWM)**. To understand XGM, the optical power in Eq. (51.6) should be replaced by the total power in all channels. Thus, the gain of a specific channel is saturated not only by its own power but also by the power of other channels. If channel powers change depending on bit patterns, as in amplitude-shift keying (ASK), then the signal gain of one channel changes from bit to bit, and the change also depends on the bit pattern of the other channels [Agrawal, 1995]. The amplified signal appears to fluctuate randomly, which degrades the signal-to-noise ratio at the receiver. This crosstalk can be largely avoided by operating SOAs in the unsaturated regime. It is absent for phase-shift keying (PSK) and frequency-shift keying (FSK) systems, since the power in each channel, and therefore the total power, remains constant with time [Agrawal, 1995].

The four-wave mixing causes the generation of new optical frequencies in closely spaced WDM systems. This phenomenon is similar to interharmonic distortion in electronic systems. The presence of multiple wavelengths in the amplifier results in nonlinear amplification. Two wavelengths can generate additional optical frequencies, as shown in Figure 51.16. If these frequencies coincide with existing channels, crosstalk results. This crosstalk can be incoherent [Darcie and Jopson, 1988] or coherent [Blumenthal and Kothari, 1996] in direct detection systems and limitations on the input power can be computed for each case [Darcie and Jopson, 1988; Blumenthal and Kothari, 1996] for a given number of channels,

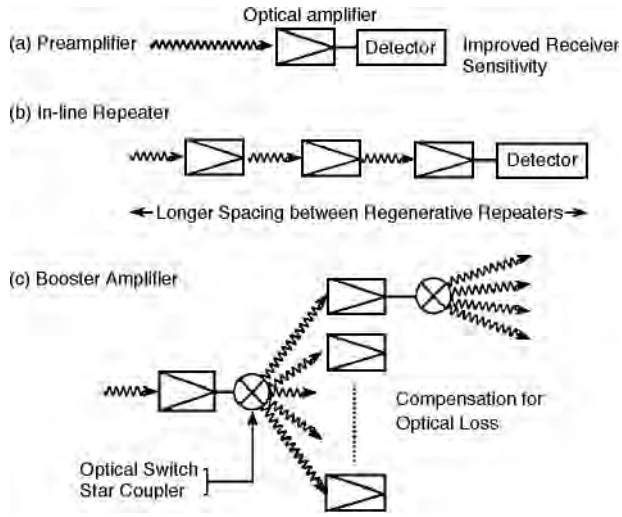


FIGURE 51.17 Configurations for SOA applications. (*Source:* Yamamoto, Y. and Mukai, T. 1989. *Optical Quant. Electron.*, 21:S1.)

channel spacing, and amplifier gain. Unequal channel spacing can be used to minimize crosstalk effects due to FWM [Forghieri et al., 1994].

51.8 Applications

Semiconductor optical amplifiers have multiple potential uses in optical communication systems (see Fig. 51.17). Power boosting, optical preamplification prior to photodetection, compensation of distribution losses, and in-line amplification are examples of transmission applications. In-line amplifiers eliminate the need for electronic regenerators and allow multichannel amplification. As a booster amplifier, the transmission distance can be increased by 10 to 100 km. SQAs can be integrated with a semiconductor laser to obtain a high-power, narrow-linewidth optical source useful for coherent communication systems [Koch and Koren, 1991; Glance et al., 1992]. Transmission distance can also be increased by putting an amplifier to operate the receiver in the shot noise limit.

Another potential application of SOAs is in the area of **dispersion compensation**. Midpoint dispersion compensation in a fiber link is achieved by using the SOA to perform frequency conversion through four-wave mixing [Tatham, Sherlock and Westbrook, 1993]. This process inverts the ordering of the optical spectrum within the pulse, thereby reversing the effect of dispersion in the second section of fiber. An SOA can also be used as a booster amplifier to balance the effect of **frequency chirp** in direct drive semiconductor lasers [Yazaki et al., 1992].

A third class of application, **photonic switching**, may require switching of signals in space and/or wavelength. SOAs can be integrated in **space switches** as transmission gates [Gustavsson et al., 1992] with very high-extinction ratio in the off state. SOAs can also be used to translate signals between wavelengths using wavelength conversion techniques such as cross-gain saturation [Valiente, Simon, and LeLigne, 1993] or four-wave mixing [Zhou et al., 1994].

Acknowledgments

This work was supported in part by the National Science Foundation under a National Young Investigator (NYI) award (contract number ECS-9457148) and the Air Force Office of Scientific Research (AFOSR) (contract number F49620-95-1-0024).

Defining Terms

Absorption coefficient: Amount of optical power absorbed in material per unit length.
Amplified spontaneous emission (ASE): Stimulated replication of photons that were originally generated by spontaneous emission.
Amplifier gain: Ratio of optical power at input of amplifier to optical power at output of amplifier.
Bit patterning: Influence of a previous bit on subsequent bits in a bit stream due to gain saturation effects.
Comb filter: Transfer function of resonant structure that has repeating periodic passband.
Cross-gain modulation (XGM): Modulation of amplifier gain at one wavelength due to variation in signal at another wavelength.
Dispersion compensation: Realignment of phases for different spectral components in a pulse so that pulse broadening can be reduced.
Distributed Bragg reflector laser (DBR): Laser structure with Bragg reflectors at ends of cavity.
Distributed feedback laser (DFB): Laser structure with Bragg reflector in middle of cavity and without mirrors at ends.
Fabry–Perot (FP) amplifier: Resonant amplifier using mirrors at input and output for feedback.
Four-wave mixing (FWM): Generation of new frequencies, or intermodulation products, through gain nonlinearity.
Free spectral range: Spacing between longitudinal modes in resonant cavity.
Frequency chirp: Change in optical frequency as a function of time across a modulated optical signal.
Gain clamped: Operation of gain structure above condition where gain equals cavity losses.
Gain ripple: Periodic variation in gain over amplifier bandwidth.
Gain saturation: Reduction in gain due to increase in optical power.
Interchannel crosstalk: Transfer of information between channels due to a physical crosstalk mechanism.
Longitudinal modes: Single passband of a resonant cavity with comb-filter type transfer function.
Material gain: Number of photons generated per incident photon per unit length of material.
Optical mode confinement factor: Amount of optical signal that passes through amplifying medium.
Photonic switching: Switching or redirection of optical signals without conversion to electronics.
Polarization dependent: Variation in gain for different orientations of optical field.
Resonant amplifier: Amplifier with gain and feedback.
Saturated material gain: Optical signal-level dependent gain.
Space switches: Connection through spatially discrete elements.
Spontaneous emission: Emission of a photon by random transition from a high-energy state to a low-energy state.
Stimulated amplification: Stimulated replication of photons through stimulated emission.
Stimulated emission: Emission of a photon due to transition from a high-energy state to a low-energy state that was caused by an initial photon.
Subthreshold: Operation of gain structure below condition where gain equals cavity losses.
Traveling wave amplifier (TWA): Single-pass amplifier with optical gain.
Unsaturated material gain coefficient: Gain provided to optical signal that is not signal-level dependent.
Wavelength-division multiplexing (WDM): Transmission of simultaneous channels on parallel optical wavelengths.

References

Adams, M.J., Collins, J.V., and Henning, I.D. 1985. *IEEE Proceedings* Part J, 132:58.

Agrawal, G.P. 1995. Semiconductor laser amplifiers. In *Semiconductor Lasers: Past, Present, and Future,* Ed. G.P. Agrawal, American Institute of Physics Press.

Agrawal, G.P. and Olsson, N.A. 1989. *IEEE J. Quant. Electron.,* 25:2297.

Blumenthal, D.J. and Kothari, N.C. 1996. *IEEE Photon. Tech. Lett.,* 8:133.

Darcie, T.E. and Jopson, R.M. 1988. *Electron. Lett.,* 24:638.
Dubovitsky, S., Mathur, A., Steier, W.H., and Dapkus, P.D. 1994. *IEEE Photon. Tech. Lett.,* 6:176.
Dutta, N.K., Lin, M.S., Piccirilli, A.B., Brown, R.L., and Chakrabarti, U.K. 1990. *J. Appl. Phys.,* 67:3943.
Dutta, N.K. and Simpson, J.R. 1993. Optical amplifiers. In *Progress in Optics,* Ed. E. Wolf, Vol. 31, North-Holland, Amsterdam.
Forghieri, F., Tkach, R.W., Chraplyvy, A.R., and Marcuse, D. 1994. *IEEE Photon. Tech. Lett.,* 6:754.
Glance, B., Wiesenfeld, J.M., Koren, U., Gnauck, A.H., Presby, H.M., and Jourdan, A. 1992. *Electron. Lett.,* 28:1714.
Gustavsson, M., Lagerstrom, B., Thylen, L., Jansson, M., Lundgren, L., Morner, A.C., Rask, M., and Stolz, B. 1992. *Electron. Lett.,* 28:2223.
Jopson, R.M. and Darcie, T.E. 1991. Semiconductor laser amplifiers in high-bit-rate and wavelength-division-multiplexed optical communications systems. In *Coherence, Amplification, and Quantum Effects in Semiconductor Lasers,* Ed. Y. Yamamoto, John Wiley & Sons, New York.
Jopson, R.M., Eisenstein, G., Hall, K.L., Raybon, G., Burrus, C.A., and Koren, U. 1986. *Electron. Lett.,* 22:1105.
Koch, T.L. and Koren, U. 1991. *IEEE J. Quant. Electron.,* 27:641.
Mukai, T., Yamamoto, Y., and Kimura, T. 1982. *IEEE J. Quant. Electron.,* QE-18:1560.
Olsson, N.A. 1989. *IEEE J. Lightwave Tech.,* 7:1071.
Saitoh, T. and Mukai, T. 1987. *Electron. Lett.,* 23:218.
Saitoh, T., and Mukai, T. 1991. Traveling-wave semiconductor laser amplifiers. In *Coherence, Amplification, and Quantum Effects in Semiconductor Lasers,* Ed. Y. Yamamoto, John Wiley & Sons, New York.
Saitoh, T., Mukai, T., and Noguchi, Y. 1986. Fabrication and gain characteristics of a 1.5 mm GaInAsP traveling-wave optical amplifier. In *First Optoelectronics Conference Post-Deadline Papers Technical Digest,* The Institute of Electronics and Communications Engineers of Japan, Tokyo, Paper B11-2.
Tatham, M.C., Sherlock, G., and Westbrook, L.D. 1993. *Electron. Lett.,* 29:1851.
Valiente, I., Simon, J.C., and LeLigne, M. 1993. *Electron. Lett.,* 29:502.
Yamamoto, Y. 1980. *IEEE J. Quant. Electron.,* QE-16:1073.
Yazaki, P.A., Komori, K., Arai, S., Endo, A., and Suematsu, Y. 1992. *IEEE J. Lightwave Tech.,* 10:1247.
Zah, C.E., Osinski, J.S., Caneau, C., Menocal, S.G., Reith, L.A., Salzman, J., Shokoohi, F.K., and Lee, T.P. 1987. *Electron. Lett.,* 23:990.
Zhou, J., Park, N., Dawson, J.W, Vahala, K.J., Newkirk, M.A., and Miller, B.I. 1994. *IEEE Photon. Tech. Lett.,* 6:50.

Further Information

Yamamoto, Y., Ed. 1991. *Coherence, Amplification, and Quantum Effects in Semiconductor Lasers,* John Wiley & Sons, New York.
Agrawal, G.P. and Dutta, N.K. 1993. *Semiconductor Lasers,* Van Nostrand Reinhold, New York.

52

Optical Amplifiers

Anders Bjarklev
Technical University of Denmark

52.1 Introduction

From the first demonstration of optical communication, a primary drive in the research activity has been directed toward constant increase in the system capacity. Until very recently, the ultimate capacity limits have been determined by the spectral bandwidth of the signal source and of the fundamental fiber parameters: attenuation and dispersion. On the basis of the limiting term of the transmission link, systems have been denoted as either *loss limited* or *dispersion limited*, and it can be seen that there have been alternations between these two extremes over the past 20 years. In the mid-1980s, the international development had reached a state at which not only dispersion-shifted fibers were available but also spectrally pure signal sources emerged. The long-haul optical communication systems were, therefore, clearly loss limited and their problems had to be overcome by periodic regeneration of the optical signals at repeaters applying conversion to an intermediate electrical signal. Because of the complexity and lack of flexibility of such regenerators, the need for optical amplifiers was obvious.

The technological challenge was to develop a practical way of obtaining the needed gain, that is, to develop amplifiers with a complexity and application flexibility superior to the electrical regenerators. Several means of doing this had been suggested in the 1960s and 1970s, including direct use of the transmission fiber as gain medium through nonlinear effects [Agrawal, 1989], semiconductor amplifiers with common technical basis in the components used for signal sources [Simon, 1983], or doping optical waveguides with an active material (rare-earth ions) that could provide the gain [Koester and Snitzer, 1964]. First, however, with the pronounced technological need for optical amplifiers and the spectacular results on erbium-doped-fiber amplifiers [Mears et al., 1987], an intense worldwide research activity on optical amplifiers was initiated. This resulted in the appearance of commercially available packaged erbium-fiber amplifier modules as early as 1990, and the rapidly expanding technology and transfer of knowledge from research to practical system implementation make it clear that optical amplifiers and, especially, the erbium-doped-fiber amplifier are bringing about a revolution in communication systems.

The following section provides a short review of the general properties and limitations of optical amplifiers. Thereafter, the specific technologies available for the realization of optical amplifiers are reviewed, and their key parameters are compared. Finally, the chapter includes a short review of some of most significant results on optical amplifiers in optical communication systems.

52.2 General Amplifier Concepts

In order to clarify and define the need for optical amplifiers, we discuss here some concepts general to all optical amplifiers. In an ideal optical amplifier, the signal would be amplified by adding, in phase, a well-defined number of photons to each incident photon. This means that a bit sequence or an analog optical signal simply would increase its electromagnetic field strength but not change its shape by passage through the optical amplifier, as illustrated in Fig. 52.1(a). In a perfect amplifier, this process would take place independent of the wavelength, state of polarization, intensity, (bit) sequence, and optical bandwidth of the incident light signal, and no interaction would take place if more than one signal were amplified simultaneously. In practice, however, the optical gain depends not only on the wavelength (or frequency) of the incident signal, but also on the electromagnetic field intensity at any point inside the amplifier. Details of wavelength and intensity dependence of the optical gain depend on the amplifying medium.

Before we turn to describing the different types of amplifiers, however, it is informative to illustrate some general concepts of practical optical amplifiers. To do so, we consider a case in which the gain medium is modeled as a homogeneously broadened two-level system. In such a medium, the gain coefficient (i.e., the gain per unit length) can be written as [Agrawal, 1992]:

$$g(\omega) = \frac{g_0}{1 + (\omega - \omega_0)^2 T_2^2 + P_s/P_{\text{sat}}} \tag{52.1}$$

Here g_0 is the peak value of the gain coefficient determined by the pumping level of the amplifier, ω is the optical angular frequency of the incident signal, and ω_0 is the atomic transition angular frequency. P_s is the optical power of the signal and P_{sat} is the saturation power, which depends on the gain medium parameters such as fluorescence lifetime τ and the transition cross section. The parameter T_2 in Eq. (52.1) is normally denoted as the dipole relaxation time [Agrawal, 1992].

Two important amplifier characteristics are described in Eq. (52.1). First, if the signal power ratio obeys $P_s/P_{\text{sat}} \ll 1$ throughout the amplifier, the amplifier is said to be operated in the unsaturated region. The gain coefficient is then maximum when the incident angular frequency ω coincides with the atomic transition angular frequency ω_0. The gain reduction for angular frequencies different from ω_0 is generally given by a more complex function than the Lorentzian profile, but this simple example allows us to define the general property of the gain bandwidth. This is normally defined as the full-width half-maximum (FWHM) value of the gain coefficient spectrum $g(\omega)$. For the Lorentzian spectrum, the gain bandwidth is given by $\Delta\omega_g = 2/T_2$.

From the communication system point of view, it is more natural to use the related concept of amplifier bandwidth instead of the gain bandwidth, which is related to a point within the gain medium. The difference becomes clear when we consider the amplifier gain G defined as

$$G = P_s^{\text{out}}/P_s^{\text{in}} \tag{52.2}$$

where P_s^{in} is the input power and P_s^{out} is the output power of a continuous wave (CW) signal being amplified. The amplifier gain G may be found by using the relation

$$\frac{dP}{dz} = gP \tag{52.3}$$

where $P(z)$ is the optical power at a distance z from the amplifier input end. If the gain coefficient $g(\omega)$, for simplicity, is considered constant along the amplifier length, the solution of Eq. (52.3) is an exponentially growing signal power $P(z) = P_s^{\text{in}} \exp(gz)$. For an amplifier length L, we then find that

$$G(\omega) = \exp[g(\omega)L] \tag{52.4}$$

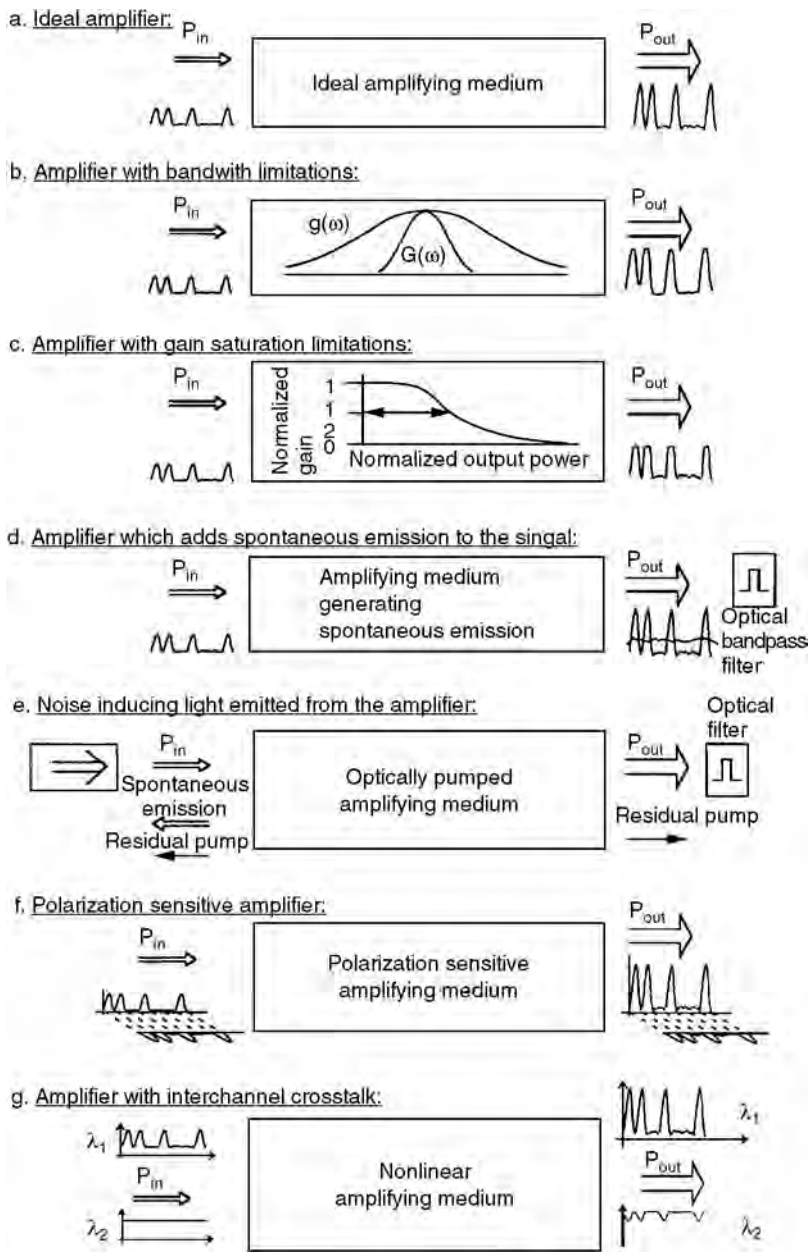


FIGURE 52.1 Optical amplifiers with different limiting properties: (a) ideal optical amplifier, (b) optical amplifier with a Gaussian bandwidth limitation, (c) optical amplifier with saturation limitation, (d) optical amplifier adding noise to the signal, (e) optical amplifier with isolator and filter to limit residual pump and ASE disturbance in the surrounding system, (f) polarization sensitive optical amplifier, and (g) optical amplifier inducing interchannel crosstalk.

Equation (52.4) illustrates the frequency dependence of the amplifier gain G and shows that $G(\omega)$ decreases much faster than $g(\omega)$ with the signal detuning $(\omega - \omega_0)$ because of the exponential dependence of G on g. As shown in Fig. 52.1(b), the amplifier bandwidth $\Delta\omega_A$, which is smaller than the gain bandwidth, results in signal distortion in the case where a broadband optical signal is transmitted through the amplifier.

Another important limitation of the nonideal amplifier is expressed in the power dependence of the gain coefficient. This property, which is known as gain saturation, is included in the example of Eq. (52.1), and it appears in all cases where the term P_s/P_{sat} is nonnegligible. Since the gain coefficient is reduced when the signal power P_s becomes comparable to the saturation power P_{sat}, the amplification factor (or amplifier gain) G will also decrease. This is illustrated in Fig. 52.1(c), where it is shown how G decreases from its unsaturated value G_0.

Gain saturation might be seen as a serious limitation for multiwavelength optical communication systems. However, gain saturated amplifiers have practical use as their self-regulating effect on the signal output power is going to be used in planned transoceanic communication links including many concatenated amplifiers. The principle is as follows: if the signal level in a chain of amplifiers is unexpectedly increased along the chain, due to saturation the following amplifiers will be operated at a lower gain and vice versa for a sudden signal power decrease.

Besides the bandwidth and gain saturation limitations of practical optical amplifiers, another property is inherent for practical amplifiers today. Optical amplifiers, in general, will add spontaneously emitted or scattered photons to the signal during the amplification process, and this will consequently lead to a degradation of the signal-to-noise ratio (SNR). This property is illustrated in Fig. 52.1(d). The SNR degradation is quantified through a parameter F, normally denoted as the amplifier noise figure, which is defined as the SNR ratio between input and output [Yariv, 1990],

$$F = \frac{SNR_{\text{in}}}{SNR_{\text{out}}} \tag{52.5}$$

It should be noted that it is common practice to refer the SNR to the electrical power generated when the optical signal is converted to electrical current by using a photodetector. The noise figure as defined in Eq. (52.5), therefore, in general, would depend on several detector parameters which determine the shot noise and thermal noise associated with the practical detector. However, the influence of detector parameters will not help us clarify the amplifier noise properties, and it is therefore advantageous to consider an ideal detector, whose performance is limited by shot noise only [Yariv, 1990].

In practice, the spontaneous emission is reduced by optical filtering of the received signal. Therefore, the SNR generally also will be dependent on the bandwidth of the optical filters and the spectral power distribution of the spontaneous emission from the amplifier. However, since this filtering is a process independent of the amplifier properties, it is also common practice to eliminate this ambiguity by considering an ideal filter [Agrawal, 1992]. Such an ideal filter is introduced only to allow the signal and the spontaneous emission within the signal bandwidth to pass to the detector. Therefore, it will only be the spontaneous emission spectral power density at the signal wavelength that enters the ideal detector, and the noise figure becomes independent of the spectral shape of the spontaneous emission. Note that since the amplifier deteriorates the signal-to-noise ratio, that is, $SNR_{\text{in}} > SNR_{\text{out}}$, the noise figure will always obey the relation $F > 1$.

For a practical communication system, the amplifier spontaneous emission may have another influence besides that described through the noise figure (i.e., besides the effect that it adds fluctuations to the amplified optical signal power, which are converted to current fluctuations during the photodetection process). Namely, the spontaneous emission, which is emitted from the amplifier input end, may enter the signal source (a semiconductor laser), where it can result in performance disturbances. Therefore, it is often necessary to include isolation between amplifier and light source to avoid additional noise in the system.

Therefore, it must be considered that the optical communication system has to be protected against undesired emission from the amplifier, and it is necessary to include the special properties of optically pumped amplifiers. These may also transmit residual pump power onto the transmitter and/or detector. The actual effect of this will be strongly dependent on the spectral properties of the pump light and whether the signal source or detector is sensitive to such radiation. As shown in Fig. 52.1(e), this often implies the inclusion of separate isolators and filters (if not already included for direct reduction of noise-induced signal power fluctuations).

Another property that has to be evaluated for the optical amplifier is the polarization sensitivity. A high-polarization sensitivity means that the amplifier gain G differs for different polarization directions of the input signal. This is shown in Fig. 52.1(f) for two orthogonal linear polarization states. Since optical communication systems normally do not include polarization control, and because the polarization state is likely to vary due to external factors such as mechanical pressure, acoustic waves, and temperature variations, amplifier polarization sensitivity is naturally an undesired property in most cases. Therefore, for amplifiers that are inherently polarization sensitive, it has been a primary goal to reduce or even eliminate the amplifier output power variation due to changes in the signal input polarization state.

All of the limiting properties, which we have previously discussed, are not surprisingly distinctly different for different types of amplifiers. This is also the case for the crosstalk limitation that relates to multichannel applications of optical amplifiers. In contrast to the ideal case, where all signal channels (or wavelengths) are amplified by the same amount, undesired nonlinear phenomena may introduce interchannel crosstalk (i.e., the modulation of one channel is affected or modified by the presence of another signal channel). The effect of crosstalk is illustrated in Fig. 52.1(g), where one channel is considered to carry a modulated signal, whereas the other carries a continuous wave (CW) signal. By passage of the amplifier, the CW has been modulated by the signal carried in the other channel. We will return to the more specific nature of these nonlinear phenomena in connection with the discussion of the different optical amplifiers.

The final limiting factors that should be mentioned here are closely related to the physical environment in which the amplifier is placed. These may include sensitivity toward vibrations, radioactive radiation, chemicals, pressure, and temperature. However, since, at least, amplifiers for optical communication systems generally are placed in controlled environments, we will not go further into a detailed discussion of such properties at this point.

The relative importance of the different limiting factors as just discussed depends on the actual amplifier application. One application is as replacement of electronic regenerators; in such cases, the amplifiers are placed at a considerable distance from the transmitter and receiver and they are denoted as in-line amplifiers. The optical amplifier may also be used to increase the transmitter power by placing an amplifier just after the transmitter. Such amplifiers are called power amplifiers, or boosters, because their main purpose is to boost the transmitted power. Long-distance systems may also be improved by the inclusion of so-called preamplifiers, which are placed just before the receiver. Finally, optical amplifiers may find use in local area networks in which they can compensate for distribution losses. Thereby, the number of nodes in the networks may be increased.

52.3 Alternative Optical Amplifiers for Lightwave System Applications

The most important factors that limit the performance of practical optical amplifiers have now been identified. The next task then is to investigate the technologies available for obtaining realistic amplifiers. Basically, six different ways of obtaining optical amplification have to be considered: **fiber-Raman** and **fiber-Brillouin amplifiers, semiconductor optical amplifiers (SOA)**, rare-earth-doped fiber amplifiers, **rare-earth-doped waveguide amplifiers**, and **parametric amplifiers**. The aim of the present description is to present the basic physical properties of these candidates as relates to their applicability as practical optical amplifiers.

Fiber-Raman Amplifiers

The fiber-Raman amplifier works through the nonlinear process of **stimulated Raman scattering (SRS)** occurring in the silica fiber itself. SRS is an interaction between light and the vibrational modes of silica molecules. The process appears because the SRS converts a small fraction of the incident power from one optical beam (the pump) to another optical beam (the signal) at a frequency down-shifted by an amount determined by the vibrational modes of the medium.

TABLE 52.1 Comparison of Optical Amplifiers Around 1550 nm

Property	Raman Amplifier	Brillouin Amplifier	SOA	EDFA	RE-Doped Waveguide Amplifiers	Optical Parametric Amplifiers
Small-signal gain	>40 dB	>40 dB	>30 dB	>50 dB	20 dB	16 dB
Efficiency	0.08 dB/mW	5.5 dB/mW	28 dB/mA	11 dB/mW	0.1 dB/mW	10^{-4} dB/mW
Output power	1 W	1 mW	>0.1 W	>0.5 W	2 mW	
Distortion/ Crosstalk	Negligible	Negligible	Significant (in saturation)	Negligible	Negligible	
Dynamic performance	>20 Gb/s	<100 MHz	>25 Gb/s	>100 Gb/s	>100 Gb/s	
Gain bandwidth	A few tens of nm	<100 MHz	60–70 nm	[a]30 nm	A few tens of nm	5000 GHz
Noise figure	[a]3 dB	>15 dB	5–7 dB	[a]3 dB	>3 dB	
Polarization sensitivity	Significant	None	<A few dB	Insignificant (<0.1 dB)	<1 dB	<A few dB
Coupling loss (to fiber)	<1 dB	<1 dB	A few dB	<1 dB		

The light generated by SRS is called the Stokes wave, and Stokes waves are generated both copropagating and counterpropagating with the incident beam. As early as 1962, it was observed that for very intense pump waves a new phenomenon of SRS can occur in which the Stokes wave grows rapidly inside the medium such that most of the pump energy appears in it. If the signal angular frequency ω_s and pump angular frequency ω_p are chosen with a difference of $\Omega_R = \omega_p - \omega_s$ (the Stokes shift) corresponding to the value dictated by the vibrational energy of the molecules, the signal may experience Raman gain.

A significant feature of the Raman gain in silica fibers is that it extends over a large frequency range (up to 40 THz with a typical FWHM width of 6 THz) with a broad dominant peak near 13 THz. This is due to the amorphous nature of fused silica in which the molecular vibrational frequencies spread out into bands overlapping each other and creating a continuum (depending on the fiber core composition). The relatively large bandwidth of fiber-Raman amplifiers and a very large saturation power (typically around 1 W) makes them attractive for optical communication systems. Also, their noise properties are good, since a noise figure close to 3 dB is obtainable. On the other hand, one of the considerable drawbacks of this amplifier type is that to obtain a 20-dB gain in a 100-km-long standard fiber for optical communication, more than 2.5 W of pump power at 1480 nm is needed [Agrawal, 1989]. Another important limitation of fiber-Raman amplifiers is their polarization sensitivity. Experiments show the necessity of matching the polarization directions of pump and signal, since SRS ceases to occur for the case of orthogonal polarizations. Because polarization control is not normally employed in optical communication systems, this property introduces an additional complexity, which, together with the high pump power requirements, makes the Raman amplifier only slightly attractive from the point of view of the optical communication system designer. Table 52.1 lists key parameters for the fiber-Raman amplifier.

Brillouin Amplifiers

Another type of amplifier that makes use of the physical properties of the silica transmission fiber itself is the **stimulated Brillouin scattering (SBS)** amplifier [Agrawal, 1989]. The nonlinear SBS process appears through the generation of a backward propagating Stokes wave that carries most of the input energy and is, in this sense, similar to the SRS process. There are, however, major differences between the SRS and SBS processes. One aspect is that the Stokes wave in the case of SBS propagates backwards and, in contrast to SRS, the input power levels needed to obtain SBS are much lower ([a]1-mW CW pump power is sufficient to obtain SBS). Note also that the SBS depends on the spectral width associated with the pump wave and it nearly ceases to occur for short pump pulses. Finally, the Stokes shift in SBS ([a]10 GHz) is smaller by three orders of magnitude compared with that occurring in SRS. The differences between SRS and SBS may be explained from a single fundamental difference; optical phonons participate

in the Raman process, whereas acoustical phonons (forming acoustic waves) participate in the Brillouin process. In the ideal case Brillouin scattering should not occur in the forward direction in an optical fiber, but due to spontaneous or thermal Brillouin scattering, small amounts of forward scattered light may be measured.

The obtainable SBS gain is nearly three orders of magnitude larger than the Raman gain [Agrawal, 1989]. This means that the fiber-Brillouin amplifier can amplify a signal by 30 to 35 dB with a small pump power in the range of 1 to 10 mW. Also, the saturation power of fiber-Brillouin amplifiers is different from that of the Raman amplifier; in the case of SBS it is in the milliwatt range. However, from the point of view of the optical communication system designer, who in the mid-1980s was seeking an efficient optical amplifier, the central problem of fiber-Brillouin amplifiers was their narrow gain bandwidth (less than 100 MHz). This narrow bandwidth, which, due to fiber inhomogeneities, is broadened even compared to the bulk silica value of around 17 MHz [Agrawal, 1992], makes fiber Brillouin amplifiers less suitable as power amplifiers, preamplifiers, or in-line amplifiers in generally used intensity-modulated direct detection systems. Fiber-Brillouin amplifiers need to be pumped by narrow linewidth semiconductor lasers (<10 MHz), which obviously limit their potential use in optical communication systems since, in practice, a wavelength tuneable narrow bandwidth pump is required. As a consequence of this, SBS has, since its first observation in optical fibers in 1972, primarily been studied as a nonlinear phenomenon limiting transmitter power in fiber optic communication systems. It is, on the other hand, important to note that the narrow bandwidth nature of SBS lately has attracted attention, for example, in connection with channel selection in multichannel lightwave systems.

Another property that adds to reservations about fiber Brillouin amplifiers is their disadvantageous noise properties. Noise appears due to a large population of acoustic phonons at room temperature, and noise figures as large as 15 dB are likely to appear. This makes the Brillouin amplifier highly problematic both as a preamplifier and for in-line applications.

Semiconductor Optical Amplifiers

The semiconductor optical amplifier, in contrast to both Raman and Brillouin fiber amplifiers, is fabricated in a material and structure different than silica transmission fiber. Therefore, light has to be coupled from fiber to SOA and vice versa, which inevitably results in coupling losses. Unless special action is taken, the SOA itself also will experience relatively large feedback due to reflections occurring at the cleaved end facets (32% reflectivity due to refractive index differences between semiconductor and air), resulting in sharp and highly disadvantageous gain reduction between the cavity resonances of the Fabry–Perot resonator. Therefore, it is necessary to design traveling wave (TW) type SOAs by suppressing the reflections from the end facets; a common solution is the inclusion of antireflection coating of the end facets. It turns out that to avoid the amplifier bandwidth being determined by the cavity resonances rather than the gain spectrum itself, it becomes necessary to require that the facet reflectivities satisfy the condition [Agrawal, 1992] $G\sqrt{R_1R_2} < 0.17$. Here, G is the single-pass amplification factor, and R_1 and R_2 are the power reflection coefficients at the input and output facets, respectively. For an SOA designed to provide a 30-dB gain, this condition will mean that $\sqrt{R_1R_2} < 0.17 \lozenge 10^{-4}$, which is difficult to obtain in a predictable and reproducible manner by antireflection coating alone. Additional methods to reduce the reflection feedback include designs in which the active-region stripe is tilted from the facet normal and introduction of a transparent window region (nonguiding) between the active-layer ends and the antireflection coated facets. Reflectivities as small as 10^{-4} may thereby be provided, and the SOA bandwidth will be determined by the gain medium. Typically, 3-dB amplifier bandwidths of 60–70 nm may be obtained.

The gain in an SOA depends on the carrier population, which changes with the signal power and the injection current. Note that, in contrast to the alternative optical amplifiers, the SOA is not optically pumped, and an important property is the very short effective lifetime of the injected carriers (i.e., of the order of 100 ps). This property becomes specifically relevant for multichannel applications of SOAs, where crosstalk limitations are in question. This crosstalk originates from two nonlinear phenomena:

cross saturation and four-wave mixing. The former appears because the signal in one channel through stimulated recombinations affects the carrier population and, thereby, the gain of other channels. This significant problem may only be reduced by operating the amplifier well-below saturation, but this is not an easy task due to the relatively limited saturation output power of 5 to 10 mW. Four-wave mixing also appears because stimulated recombinations affect the carrier number. More specifically, the carrier population may be found to oscillate at the beat frequencies between the different channels, whereby both gain and refractive index are modulated. The multichannel signal therefore creates gain and index gratings, which will introduce interchannel crosstalk by scattering a part of the signal from one channel to another. Finally, it should be mentioned that, due to their large bandwidth, SOAs have the potential of amplifying ultra-short pulses (a few picoseconds). However, the gain dynamics will play an important role and saturation induced self-phase modulation has to be considered. In other words, the SOA will introduce a frequency chirp, which may find positive use in terms of dispersion compensation.

The noise properties of the SOA are determined by two factors. One is the emission due to spontaneous decays and the other is the result of nonresonant internal losses α_{int} (e.g., free carrier absorption or scattering loss), which reduce the available gain from g to (g - α_{int}). Also, residual facet reflectivities (i.e., loss of signal input power) increase the noise figure. Typical values of the noise figure for SOAs are 5 to 7 dB.

An undesirable characteristic of the early SOAs is the polarization sensitivity, which appears because the amplifier gain differs for the transverse-electric (TE) and transverse-magnetic (TM) modes in the semiconductor waveguide structure. However, intense research over the last few years has reduced the problem, and an SOA with much reduced polarization sensitivity has been reported by many groups [Stubkjaer et al., 1993]. For such state-of-the-art SOAs, a net gain of 25 dB may be achieved for a bias current of 100 mA. At this gain level, the polarization sensitivity is ±1 dB, and it becomes insignificant for gain levels lower than ^{a}20 dB. Other methods of using serial or parallel coupled amplifiers or two passes through the same amplifier are also suggested, but such schemes increase complexity, cost, and stability requirements.

To conclude the discussion of the semiconductor optical amplifier, it is important to note that SOAs through the 1980s have shown potential for optical communication applications. However, drawbacks such as polarization sensitivity, interchannel crosstalk, and large coupling losses have to be overcome before their use becomes of practical interest. In favor of the SOA are the large amplifier bandwidth and the possibility of monolithical optoelectronic integration, especially within the receiver, where the input signal powers are weak enough to avoid undesired nonlinearities. Therefore, it is likely that SOAs will find application in future optical communication systems. It should be mentioned that the SOA itself can be used as an amplifier and a detector at the same time, because the voltage across the pn-junction depends on the carrier density, which again interacts with the optical input signal. Promising results for transparent channel drops, channel monitoring, and automatic gain control have been demonstrated [Stubkjaer et al., 1993]. Note that possible future applications of SOAs will not be in their ability as amplifiers alone, but rather as wavelength conversion elements and optical gates [Stubkjaer et al., 1993].

Rare-Earth-Doped Fiber Amplifiers

Parallel to the maturation of semiconductor optical amplifiers, another development has taken place, which in only three years has demonstrated a large impact on optical communication systems. With a point of reference in work on rare-earth- (RE-) doped glass lasers initiated as early as 1963 [Koester and Snitzer, 1964], the first rare-earth-doped fiber amplifiers (as possible useful devices for telecommunication applications) were demonstrated in 1987 [Mears et al., 1987]. Progress since then has multiplied to the extent that amplifiers today offer far-reaching new opportunities in telecommunication networks.

The possible operational wavelengths of rare-earth-doped fiber amplifiers are determined by the emission spectra of the rare-earth ions moderately modified by the host material in which they are embedded. Only a few rare-earth materials become relevant for optical communication purposes, primarily erbium, which may provide amplification in the 1550 nm wavelength band, and neodymium or

praseodymium, which may be operated around the 1300 nm band. The absorption spectrum holds accurate information about the location of possible pump wavelengths and the ability to excite the rare-earth ions to higher energy levels. From this higher energy level the electron can relax, transferring its packet of energy to the ground state either radiatively or nonradiatively. The nonradiative de-excitation involves the creation of phonons (i.e., a quantized vibration of the surrounding medium), whereas the radiative decay to lower energy levels takes one of two forms. These are known as spontaneous and stimulated emission, and in both cases photons are emitted. Spontaneous emission always takes place when the electrons of a collection of atoms are in an excited state; therefore, spontaneously emitted light may not be avoided in a fiber amplifier. Stimulated emission is the process that allows signal amplification to take place and, therefore, is the desired property of the fiber amplifier. The process may be explained as follows. A photon incident on the medium, with an energy equal to the difference in energy of the ground state and an excited state, promotes de-excitation, with the creation of a photon that is in phase with the incident photon. When the ions are pumped to energy levels above the upper laser level, the population cascades down through the intervening levels by nonradiative transitions and eventually leads to a population of the upper laser level. This level is normally described as metastable, since transitions from this to lower manifolds only appear because the crystal field has broken the inversion symmetry of the ion's environment, permitting the electric dipole transitions to occur. Note that for this reason, the involved lifetimes in the upper laser level (typically, 10 to 14 ms for erbium) are many orders of magnitude larger than the nonradiative lifetimes of the higher energy levels.

Focusing on the **erbium-doped-fiber amplifier (EDFA)**, which may be fabricated using a silica glass host, semiconductor pump sources have been successfully used to pump amplifiers in the 800-, 980-, and 1480-nm absorption bands. In addition, pumping at lower wavelengths is also applicable but less interesting due to the lack of practical pump sources [Bjarklev, 1993]. Basic differences exist between the application of the three mentioned pump choices. First, it should be noted that amplification occurs according to a three-level scheme when 800- or 980-nm pumping is applied, but the erbium ion works as a two-level system when 1480-nm pumping is used. Furthermore, 800-nm pumping is much less efficient than 980-nm pumping due to a pronounced excited-state absorption (ESA) of pump photons [Bjarklev, 1993]. Because of these differences, today only 980- and 1480-nm pumping is considered in practical system applications. In cases where very low-noise figures are required, 980-nm pumping is the preferred choice, since the three-level nature of the system makes excitation of all erbium ions (full inversion) possible, resulting in a noise figure very close to 3 dB. For 1480-nm pumping, only 70% of the erbium ions may be excited to the upper laser level, and noise figures down to 5 dB are generally obtainable. However, 1480-nm pumping still provides highly efficient amplifiers, and more reliability testing of the semiconductor pump sources have been performed at this wavelength, making 1480-nm pumped amplifiers preferred for many present applications.

EDFAs readily provide small-signal gain in the order of 30 to 40 dB, but values as high as 54 dB have been demonstrated [Hansen, Dybdal, and Larsen, 1992]. The EDFA also provides high-saturation powers, and signal output powers of more than 500 mW have been demonstrated. In such power amplifier applications, 80% or more of the pump photons may be converted to signal photons, indicating the high efficiency of the EDFA. A 30-nm bandwidth of the EDFA may easily be obtained, and although this is not quite as large as that of the SOA, it will be quite sufficient for most multiwavelength applications of optical amplifiers.

The long lifetime in the upper laser level makes the EDFA an almost ideal energy reservoir, and it is important to note that, with the power levels used in modern optical communication systems, one signal pulse will only interact with very small fraction of the erbium ions within the EDFA. The following pulse is, therefore, very unlikely to interact with the same erbium ions, and it will remain unaffected by other pulses. For this reason, in practice, no crosstalk will be seen in fiber amplifiers. The EDFA is also very insensitive to polarization variations of the signal, and only systems with cascaded amplifiers have lately demonstrated degradation due to polarization effects. A typical gain variation of 0.1 dB is found [Mazurczyk and Zyskind, 1993] when the polarization state of the signal is switched between orthogonal linearly polarized states. Finally, it should be mentioned that since the EDFA is an optical fiber itself, very low

coupling losses (comparable to ordinary splice losses) are involved in the application of EDFAs. The key numbers describing the EDFA are shown in Table 52.1.

In contrast to the EDFA, the 1300-nm fiber amplifiers have to deal with a number of difficulties, and their applicability still remains an unanswered question. The neodymium-doped fiber amplifier operates at a wavelength that is too high (around 1340 nm) to provide a realistic gain, and so the best candidate in the 1300-nm wavelength band is the **praseodymium-doped fiber amplifier (PDFA)**. The primary difficulty is that for 1300-nm applications, the rare-earth dopants have to be placed in a host material other than silica (otherwise no net gain will appear) [Bjarklev, 1993]. Until now the most promising host materials have been different compositions of fluoride glasses, which suffer from low strength and hygroscopy. However, a small-signal gain of 34 dB (fiber-to-fiber) and a saturated output power of 200 mW have been demonstrated for a praseodymium-doped fiber amplifier module. Another problem is that the 1300-nm amplifiers normally have a lower pump efficiency than EDFAs. However, the latest results present values as high as 4 dB/mW for praseodymium-doped amplifiers, which also show noise figures in the 5 to 8 dB range.

Rare-Earth-Doped Waveguide Amplifiers

Rare-earth-doped waveguide amplifiers are only different from fiber amplifiers in the respect that the rare-earth material is embedded in a planar optical waveguide and not a fiber. This means, at first glance, that only another waveguide geometry has to be considered. However, two important factors have to be considered in the design of rare-earth-doped waveguide amplifiers. First, the background loss is several orders of magnitude larger in a planar waveguide (around 0.1 dB/cm) than in a fiber, and second, a much higher rare-earth-dopant concentration is necessary in the relatively short planar waveguides. This raises the problem of energy transfer between the rare-earth ions and a resulting lower amplifier efficiency. Not surprisingly, the focus has been on the erbium-doped planar waveguide amplifiers in silica, and integration with a 980/1530-nm wavelength division multiplexer has been presented [Hattori et al., 1994]. In this component, a maximum gain of 13.5 dB for a 600-mW pump power was shown, but gain values up to 20 dB have been demonstrated in other amplifiers. Of course, these early amplifiers are far from optimized, but the results indicate that the rare-earth-doped planar waveguides have realistic and attractive application possibilities in integrated-optical devices. We may consider other host materials such as lithiumniobat and new applications beside optical communication. On the other hand, it is not realistic to expect that the rare-earth-doped waveguide amplifier will completely replace EDFAs as high-performance amplifiers in optical communication systems.

Optical Parametric Amplifiers

Finally, it should be mentioned that the possibility of using second-order nonlinearity in electro-optic materials for amplification exists. In this case, energy is transferred from a strong optical pump to the signal with the aid of an optical idler in a degenerate four-wave-mixing process. It is referred to as a parametric process as it originates from light-induced modulation of a medium parameter such as the refractive index, and its origin lies in the nonlinear response of bound electrons of the material to an applied optical field. The main difficulty is that phase matching of the three interacting waves has to be maintained. Furthermore, a high-pump-power requirement is a serious limitation, and parametric amplifiers are, at present, inferior to erbium-doped and semiconductor optical amplifiers. However, such phase sensitive amplifiers are suitable for squeezed-state amplification, and they have no lower limit for the noise figure in the limit of high amplification [Kumar, Kath, and Li, 1994].

Comparison of Amplifier Characteristics

Table 52.1 lists the key parameters of all of the different optical amplifiers discussed. In principle, the different types may be operated at any wavelength, except for the erbium-doped amplifiers. However, since the 1550-nm wavelength band is one of the most important for optical communication, Table 52.1 refers to this wavelength band, and, as such, it serves as a source of comparison between the different amplifiers.

52.4 Summary

Through the description of the alternative optical amplifiers for optical communication systems we learn that fiber-Raman and fiber-Brillouin amplifiers have not been able to completely fulfill the demands of optical communication systems, and it was with the demonstration of practical EDFAs that the development of optical amplifiers for communication systems really first started. The tremendous progress provided by the EDFA over the past five years has been complemented by the development of high-performance SOAs, and, lately, rare-earth doped planar waveguide amplifiers and parametric amplifiers have been investigated. However, the last two still have to mature before a practical system application is demonstrated. For these reasons, we end this description by looking at the system performance of the EDFA and, to a minor extent, the SOA.

The main reason for the strong impact that especially the EDFAs have had and will have on data communication systems is to be found in the fact that they are unique in two respects: they are useful for amplifying input signals of different bit rates or formats, and they can be used to achieve simultaneous amplification of multiwavelength optical signals over a wide spectral region. These multiwavelength optical signals can carry different signal formats, including digital data, digital video, analog video, etc., allowing flexible upgrading in a broadband distribution network by adding extra wavelength-division multiplexed (WDM) channels as needed. With optical amplifiers, therefore, the bottleneck of the narrow and fixed bandwidth of electrical repeaters is avoided. The outstanding properties offered by the EDFAs can be summarized as follows: high gain, high-power conversion efficiency, low noise, low crosstalk, high-saturation power, polarization insensitivity, broad spectral bandwidth, very low-coupling losses, and low cost. This short list of fiber amplifier properties underlines the huge potential of these components, which most likely will extend to a large variety of applications in the years to come.

Some of the most significant system capability results achieved over the last few years are plotted in Fig. 52.2 as a function of the publication date. For comparison, the highest performance results of direct detection systems without optical amplifiers are also included. Compared to these, it is obvious that only by including EDFAs at the ends of the transmission link (as either power boosters or preamplifiers) can a significant bit rate distance improvement be obtained. One of the highest capacities achieved with only one EDFA is 2000 Gb/s ◊ km, which was demonstrated by transmitting 20 Gb/s over a distance of 100 km. This is equivalent to the transmission of 300,000 simultaneous telehone channels transmitted over 100 km.

Not surprisingly, most of the later system experiments include one or a multiple of in-line EDFAs, and transmission over transoceanic distances is obtained in this way. The breakthrough for capacity increase

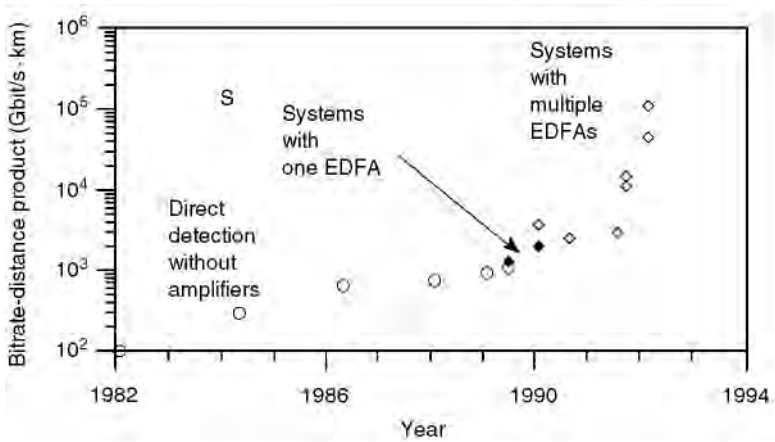


FIGURE 52.2 Capability of optical direct detection systems measured as reported bit rate distance product against data of report.

that happened with the availability of optical amplifiers and especially EDFAs in 1989 is therefore evident. It should also be mentioned that all reported high performance system experiments today use optical amplifiers. The present record is as much as 45 Tb/s · km, with which, according to Fig. 52.2, we have obtained a bit rate distance product improvement of a factor of 100 compared to systems without EDFAs.

However, limits also exist for direct detection systems with EDFAs. In contrast to electronic repeaters, the EDFA cannot regenerate the shape of the optical pulses. Thus, in dispersive fiber links, the transmission length will be dispersion limited, especially at high bit rates. However, methods to avoid this are studied intensively at the present time, and optical amplifiers can also help to overcome the dispersion limitation by exploiting the nonlinear effects.

From the results in Fig. 52.2, it is evident that optical amplifiers have demonstrated their relevance in high-speed, long-distance repeaterless communication systems. The applications here are many, such as undersea and off-shore links, island jumping, and trunk lines between main exchanges. However, it is also relevant to consider that many medium-distance applications exist (i.e., interconnection of points spaced by 100–300 km). In such cases (and especially for submarine applications), it may be of interest to avoid the in-line amplifier, since the system maintenance and reliability problems would then be considerably reduced. Additionally, the EDFA will find many applications in optical signal distribution systems, and splits up to 39 million subpaths have been demonstrated for only two stages of amplification. In short, the EDFA has revolutionized thinking on future optical fiber networks.

Finally, note that the application of SOAs in a 10-Gb/s preamplified receiver today realizes system sensitivities of -28 dB (1 mW) with a fiber-to-fiber gain of the SOA of 18.5 dB. The polarization dependency of the receiver has been brought below 0.5 dB. For comparison, a fiber preamplifier used with the same front-end and the same transmitter resulted in a sensitivity of -34 dB (1 mW) [Stubkjaer et al., 1993].

Defining Terms

Erbium-doped fiber amplifier (EDFA): Fused silica fiber codoped with erbium as the active material providing gain in the 1.55-mm wavelength range. The erbium ions are most commonly pumped at either 0.98 or 1.48 mm using high-power semiconductor pump sources.

Fiber-Brillouin amplifier: Optical fiber (typical of fused silica) providing signal gain through the process of SBS.

Fiber-Raman amplifier: Optical fiber (typical of fused silica) providing signal gain through the process of SRS.

Parametric optical amplifier: Amplifier in which optical power is transferred from a strong optical pump to a signal through a degenerate four-wave mixing process. The term parametric amplifier refers to the property of using light-induced modulation of a medium parameter such as the refractive index.

Praseodymium-doped fiber amplifier (PDFA): Optical fiber (typically manufactured from a fluoride glass composition) codoped with praseodymium in order to obtain gain in the 1.3-mm wavelength range through stimulated emission.

Rare-earth-doped waveguide amplifier: Dielectric waveguide doped with a rare earth material. This material has the effect of providing gain through stimulated emission, and the necessary power for this process appears through absorption of pump photons (propagating within the waveguide structure at a different frequency than the signal).

Semiconductor optical amplifier (SOA): Electrically pumped active semiconductor component providing optical gain. SOAs have a large potential as amplifiers integrated with receivers, for example, and as wavelength conversion elements.

Stimulated Brillouin scattering (SBS): Interaction between light and acoustic phonons (forming acoustic waves) resulting in conversion of light at one frequency (the pump) to another (the signal). SBS couples energy from one mode to another propagating in the opposite direction and is a quite narrowband phenomenon (bandwidth less than 100 MHz).

Stimulated Raman scattering (SRS): Interaction between light and vibrational modes (optical phonons) of the transmission medium that converts a fraction of the incident power from one optical beam (the pump) to another optical beam (the signal) at a frequency downshifted by an amount determined by the vibrational modes of the medium.

References

Agrawal, G.P. 1989. *Nonlinear Fiber Optics,* Academic Press, London.

Agrawal, G.Y. 1992. *Fiber-Optic Communication Systems,* Wiley, New York.

Bjarklev, A. 1993. *Optical Fiber Amplifiers: Design and System Applications,* Artech House, Boston, MA.

Hansen, S.L., Dybdal, K., and Larsen, C.C. 1992. Upper gain limit in Er-doped fiber amplifiers due to internal Rayleigh backscattering. Optical Fiber Communications, OFC'92, San Jose, CA. 1:TuL4.

Hattori, K., Kitagawa, T., Oguma, M., Ohmori, Y., and Horiguchi, M. 1994. Erbium-doped silica-based planar-waveguide amplifier integrated with a 980/1530-nm WDM coupler. Optical Fiber Communications, OFC'94, San Jose, CA. 1:FB2.

Koester, C.J., and Snitzer, E. 1964. Amplification in a fiber laser. *Appl. Opt.,* 3(10):1182.

Kumar, P., Kath, W.L., and Li, R.-D. 1994. Phase-sensitive optical amplifiers. Integrated Photonics Research IPR'94, San Francisco CA. 1:SaB1.

Mazurczyk, V.J., and Zyskind, J.L. 1993. Polarization hole burning in erbium doped fiber amplifiers. Conference on Lasers and Electro-Optics, CLEO'93, Baltimore, MD, May 2–7. CPD26.

Mears, R.J., Reekie, L., Jauncey, I.M., and Payne, D.N. 1987. Low-noise erbium-doped fiber amplifier operating at 1.54 μm. *Electron. Lett.,* 23(19):1026.

Simon, J.C. 1983. Semiconductor laser amplifier for single mode optical fiber communications. *J. Opt. Commun.,* 4(2):51–62.

Stubkjaer, K.E., Mikkelsen, B., Durhuus, T., Joergensen, C.G., Joergensen, C., Nielsen, T.N., Fernier, B., Doussiere, P., Leclerc, D., and Benoit, J. 1993. Semiconductor optical amplifiers as linear amplifiers, gates and wavelength converters. Proceedings of the 19th European Conference on Optical Communication, ECOC'93. 1:TuC5.

Yariv, A. 1990. Signal-to-noise considerations in fiber links with periodic or distributed optical amplification. *Opt. Lett.,* 15(19):1064–1066.

Further Information

A thorough treatment of fiber optical amplifiers and their applications in optical communication systems can be found in *Optical Fiber Amplifiers: Design and System Applications*, by A. Bjarklev, 1993, Artech House, Norwood, MA.

Fiber lasers and amplifiers are treated in detail in *Rare Earth Doped Fiber Lasers and Amplifiers*, by M.J.F. Digonnet, 1993, Marcel Dekker Inc., and in *Optical Fiber Lasers & Amplifiers*, by P.W. France, 1991, Blackie and Son Ltd.

The journals *IEEE Electronics Letters, IEEE Photonics Technology Letters*, and *IEEE Journal of Lightwave Technology* reports advances in optical amplifiers and optical communication systems. Also, Proceedings of the Annual European Conference on Optical Communication (ECOC) and the annual OSA Conference on Optical Fiber Communication (OFC) present state-of-the-art results on optical amplifiers. The annual SPIE conference OE/FIBERS also presents valuable information relevant to the subject.

53

Coherent Systems

Shinji Yamashita
The University of Tokyo

53.1 Introduction

Optical fiber communications have been expected to play a leading role in communication networks and have already been put into commercial operation in domestic and international trunk lines. The present systems employ intensity modulation/direct detection (IM/DD), where the transmitter laser diode (LD) is directly on–off modulated and the signal is directly detected by an avalanche photodiode. Although advantageous in simplicity and low cost, these systems are as primitive as the radio communication systems prior to 1930, and do not fully utilize the potential ultra-high capacity of an optical fiber. On the other hand, in radio communications, the **heterodyne scheme** has become common, and sophisticated coherent modulations such as frequency or phase modulation are also widely used. Immediately after the invention of the laser, some heterodyne-type optical communication systems had been conceived, where they were assumed to use gas lasers and lens waveguides, and were far from practical. After the advent of LDs and optical fibers in 1970, the IM/DD was the only scheme; the heterodyne scheme had been given up because of the poor spectral purity and frequency stability of the LDs and because of the multimode nature of an optical fiber.

Interest in heterodyne-type optical fiber communications, hereafter called coherent optical fiber communications or coherent systems including the **homodyne scheme**, was revived around 1980 [Okoshi and Kikuchi, 1980] when the truly single-frequency LDs such as distributed-feedback (DFB) LDs and single-mode optical fibers began to be available. Since then, research on coherent systems has been expanded, along with the drastic improvement of LDs' spectral purity and the spread of the single-mode optical fibers, because the coherent systems have two main advantages: (1) dramatic improvement of the receiver sensitivity and the resulting elongation of the repeaterless transmission distance and (2) drastic improvement of the frequency selectivity and the resulting possibility of dense frequency-division multiplexing (FDM) [Okoshi and Kikuchi, 1988; Smith, 1987; Linke and Gnauck, 1988; Ryu, 1995]. Many experimental demonstrations of long distance and FDM systems have been reported, and, recently, field tests of a practical 2.5-Gb/s trunk system over 300 km [Hayashi, Ohkawa, and Yanai, 1993] and a 128-channel FDM distribution system [Oda et al., 1994] have been reported.

This chapter describes the fundamentals, practical implementation, and constraints and countermeasures of coherent optical fiber communication systems. The emphasis is on digital systems.

53.2 Fundamentals of Coherent Systems

Basic construction of the coherent optical fiber communication system is shown in Fig. 53.1. Amplitude, frequency, or phase of an optical carrier is modulated at a transmitter directly or by using a modulator. The electric field of the optical signal is expressed as

$$E_S(t) = E_S d(t) \exp[\omega_S t + \phi_S + \theta(t)] \tag{53.1}$$

where E_S, ω_S, and ϕ_S are the field amplitude, frequency, and phase of the carrier; $d(t)$ is the baseband information impressed on the amplitude and $\theta(t)$ is that on the frequency or phase. After passing through a single-mode optical fiber, the optical signal is mixed with a local oscillator (LO) light using a beam splitter (BS) or a fiber coupler, and detected by the pin photodiode (PD). When we denote the LO light field as,

$$E_{LO}(t) = E_{LO} \exp(\omega_{LO} t + \phi_{LO}) \tag{53.2}$$

the detected photocurrent becomes, assuming no mismatch of the **states of polarization** (**SOP**) between the signal and the LO light fields and ignoring the loss,

$$\begin{aligned} i_{photo}(t) &= K|E_S(t) + E_{LO}(t)|^2 \\ &= KE_S^2 d(t)^2 + KE_{LO}^2 + 2KE_S E_{LO} d(t) \cos[\omega_{IF} t + \phi_{IF} + \theta(t)] \end{aligned} \tag{53.3}$$

where $\omega_{IF} = \omega_S - \omega_{LO}$, $\phi_{IF} = \phi_S - \phi_{LO}$ and K is a constant. The third term in Eq. (53.3) is a beat component between the signal and the LO light. It is found to be a down-converted signal from the optical carrier frequency into the microwave beat frequency. This beat frequency is called an intermediate frequency (IF), and this process is called optical heterodyne detection. The obtained IF signal is fed into a heterodyne demodulator, which is basically identical to that used in radio communications, and then the baseband signal is regenerated.

In a special case where the IF is equal to zero, that is, where the optical carrier and LO light have the same frequency, it is called optical homodyne detection. In this case, the baseband signal is obtained directly; however, both the frequencies and phases of the optical carrier and LO light must be matched exactly.

The high receiver sensitivity near the shot-noise limit of a coherent receiver stems from the fact that the shot noise due to the high-power LO light predominates over the thermal noise that a DD receiver suffers from, and the high-frequency selectivity from an electrical filter with sharp cutoff characteristics can be used at the IF stage instead of a dull optical filter.

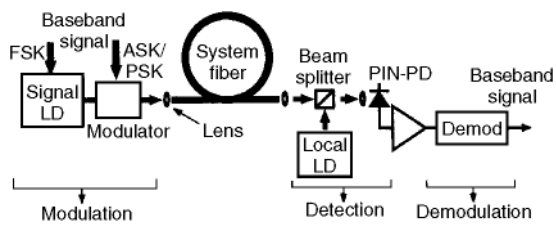


FIGURE 53.1 A basic construction of the coherent optical fiber communication system.

53.3 Modulation Techniques

In digital coherent systems, information is impressed on an optical carrier in one of three ways: amplitude-shift keying (ASK), frequency-shift keying (FSK), and phase-shift keying (PSK). These are basically identical to those of digital radio communication systems. In a binary ASK, $d(t)$ in Eq. (53.1) is equal to 1 for mark data and 0 for space data; similarly, $\theta(t) = +\Delta\omega t/2$ and $-\Delta\omega t/2$ in a binary FSK, and $\theta(t) = 0$ and π in a binary PSK, respectively. PSK signals can also be expressed as $d(t) = 1$ and -1, which means that a binary PSK signal is a carrier-suppressed binary ASK signal.

The FSK systems are classified by the magnitude of the frequency deviation $\Delta\omega$. Normalizing it by the bit rate R_b, the modulation index m is defined as $m = \Delta\omega/2\pi R_b$. When $m \geq 2$, it is referred to as wide-deviation FSK, and when $m \leq 2$, it is referred to as narrow-deviation FSK. The minimum value of m is 0.5, which is referred to as minimum-shift keying (MSK). If the FSK signal has continuous phases between the mark and the space, which is the case of direct FSK modulation of an LD, it is called continuous-phase FSK (CPFSK).

In many cases, the ASK and PSK signals are produced by using $LiNbO_3$ external waveguide modulators, whereas the FSK signal is obtained by directly modulating the drive current of the transmitter DFB-LD. Direct ASK modulation of the LD, that is, the IM used in the conventional systems, is not available because of the accompanying large frequency shift (chirping). In the $LiNbO_3$ crystal, it is possible to change the refractive index by applying an electric field (electrooptic effect). Thus, the $LiNbO_3$ waveguide can be used directly as a phase modulator, as shown in Fig. 53.2(c), or as an amplitude modulator by forming a Mach–Zehnder interferometer, as shown in Fig. 53.2(a). High-speed modulation over 10 Gb/s has been realized by making these modulators traveling-wave types. A drawback of the use of these $LiNbO_3$ modulators is that high voltage around 10 V needs to be applied to get π phase shift; therefore, a high-power, high-speed, and high-cost driver is required. In contrast, the frequency can be modulated over 10 Gb/s by using the LD's chirping characteristics, that is, by modulating its drive current slightly through a bias circuit, as shown in Fig. 53.2(b), which is easy and simplifies the transmitter.

The data can be transmitted more efficiently by using multilevel keying formats, which are common in modern radio systems. In coherent optical fiber communication systems, however, their use has not yet been considered seriously, partly because of the added complication in the system and technical difficulties and partly because the available vast bandwidth in an optical fiber has not yet been fully utilized.

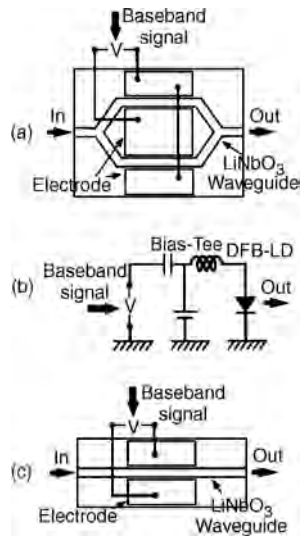


FIGURE 53.2 Modulation schemes: (a) ASK external modulator, (b) FSK direct modulation, and (c) PSK external modulator.

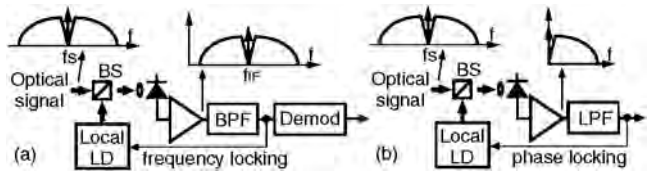


FIGURE 53.3 Detection schemes: (a) heterodyne and (b) homodyne.

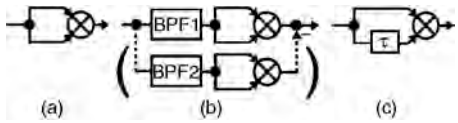


FIGURE 53.4 Demodulation schemes: (a) square-law demodulator for ASK, (b) single- or dual-filter demodulator for wide-deviation FSK, and (c) differential demodulator for narrow-deviation FSK and PSK.

53.4 Detection and Demodulation Techniques

Detection schemes are basically classified into heterodyne and homodyne types by the IF, as previously described. A heterodyne-detected IF signal is a down-converted signal from the optical carrier frequency into the microwave IF, that is, the heterodyne-scheme has spectral transparency. Demodulation must be performed to get the baseband signal after detection, whereas only the IF is to be locked, as shown in Fig. 53.3(a). Required bandwidth of the PD and IF amplifiers is usually 5–6 times larger than the baseband, which makes it difficult to detect a high bit rate signal.

On the other hand, in a homodyne receiver, detection and demodulation are performed simultaneously, and the baseband signal is obtained directly. Only the bandwidth identical to the baseband is required; therefore, it is suitable for detection of a high bit rate signal. However, an optical phase-locked loop (PLL) is indispensable, as shown in Fig. 53.3(b). Though there are some realizations such as the pilot carrier PLL and the Costas PLL [Smith, 1987], the optical PLL is extremely difficult due to the large **phase noise** of the LDs.

Demodulators in heterodyne receivers are also basically identical to those used in digital radio receivers. In most cases, however, asynchronous (square law or differential) demodulators are used because synchronous demodulations are difficult because of the large phase noise of the LDs. An ASK signal is demodulated by a square-law demodulator [Fig. 53.4(a)]. For a wide-deviation FSK signal, a bandpass filter is used to extract the mark or space spectrum and convert it to an ASK signal, which is fed into a square-law demodulator [Fig. 53.4(b)]. The signal-to-noise (SN) ratio is enhanced if both the mark and space spectra are demodulated using dual filters and subtracted. A differential demodulator is used for a narrow-deviation FSK signal as a frequency discriminator [Fig. 53.4(c)]. For a CPFSK signal obtained by the direct modulation of an LD, the optimum delay τ becomes $1/2mR_b$ [Linke and Gnauck, 1988]. Likewise, a PSK signal is demodulated by a differential demodulator having the delay τ of 1-bit period $1/R_b$. In this case, however, the demodulator output is encoded according to $d(t)d \cdot (t - \tau)$ $[d(t) = \pm 1]$, and a decoder is required to get the original signal $d(t)$. This implementation is called differential PSK (DPSK).

53.5 Receiver Sensitivity

In the detection of the light in the classical coherent state, a quantum mechanical noise exists, which is called shot noise. In a DD receiver, the shot-noise power is proportional to the signal power P_S. However, the thermal noise predominates over the shot noise when P_S is small, and it makes the DD receiver much less sensitive.

On the other hand, in a coherent receiver, the shot-noise power is proportional to the LO power P_{LO} (and predominates over the thermal noise when $P_{LO} \gg P_S$), thus, highly sensitive receivers near the shot-noise limit are possible. Since the beat signal power is proportional to $P_S P_{LO}$, the shot-noise-limited SN ratio for the heterodyne scheme is given by

$$\gamma_{\text{het}} = \frac{\eta P_S}{2\hbar\omega B} \tag{53.4}$$

and for the homodyne scheme is

$$\gamma_{\text{hom}} = \frac{\eta P_S}{\hbar\omega B} \tag{53.5}$$

where $\hbar$ is Planck's constant, η is the quantum efficiency of the PD (ideally, $\eta = 1$), and B is the receiver bandwidth. The SN ratio improvement by 3 dB in a homodyne scheme stems from the fact that the required receiver bandwidth becomes half [Okoshi and Kikuchi, 1988]. The expression of the SN ratio is simplified using the number of received photons per bit N. If we assume that the receiver bandwidth is half of the bit rate (Shannon's essential limit), $N = P_S/2\hbar\omega B$. In this case, $\gamma_{\text{het}} = \eta N$, and $\gamma_{\text{hom}} = 2\eta N$.

The bit error rate (BER) is approximately given as

$$\text{BER} = \frac{1}{2}\exp\left(-\frac{g}{n}\right) \tag{53.6}$$

where n is an integer inherent in the modulation formats. For ASK and single-filter FSK, $n = 4$; for dual-filter FSK, $n = 2$; and for differential FSK and PSK, $n = 1$. Equation (53.6) holds for both heterodyne and homodyne schemes. Using Eq. (53.6), we can compute the required N for various schemes to achieve a prescribed BER. Table 53.1 shows the result of a computation assuming that $\eta = 1$ and BER = 10^{-9}. The required beat linewidth is also shown and will be described in the next section. The PSK homodyne scheme has the highest sensitivity ($N_{\text{required}} = 10$), the PSK and differential FSK heterodyne schemes are lower by 3 dB and the next best sensitivity ($N_{\text{required}} = 20$), the ASK homodyne and dual-filter FSK heterodyne schemes have a further 3-dB decrease (with $N_{\text{required}} = 40$), and the ASK and single-filter FSK heterodyne is the lowest ($N_{\text{required}} = 80$). Note that the optical power or the number of photons is defined to be the peak values.

For example, we will consider a 10-Gb/s PSK homodyne system having an ideal performance. If we assume that a signal at 1.55 μm of 10 mW is incident to the system fiber, the number of photons in one bit is calculated as $P_S/\hbar\omega R_b = 7.8 \times 10^6$. As $Ns_{\text{required}} = 10$ for an ideal PSK homodyne receiver, it has the power margin of $7.6 \times 10^5 = 58.9$ dB. The loss in a fiber at 1.55 μm is less than 0.2 dB/km; therefore, unrepeated transmission of about 300 km is possible in this case.

The differences in sensitivity among the schemes can be attributed to the degree of coherence the light utilizes in modulation and demodulation; the PSK scheme is coherent modulation, the ASK, incoherent.

TABLE 53.1 Receiver Sensitivity and Linewidth Requirement

Detection	Modulation, Demodulation	N_{required}	Required Beat Linewidth
Homodyne	PSK	10	$5 \times 10^{-4}\ R_b$
	ASK	40	$5 \times 10^{-4}\ R_b$
Heterodyne	DPSK	20	$3\text{–}7 \times 10^{-3}\ R_b$
	FSK (narrow-dev.), differential	20	$3\text{–}20 \times 10^{-3}\ R_b$
	FSK (wide-dev.), dual-filter	40	$0.1\text{–}0.2\ R_b$
	FSK (wide-dev.), single-filter	80	$0.1\text{–}0.2\ R_b$
	ASK, square-law	80	$0.1\text{–}0.2\ R_b$

The PSK homodyne is the most coherent system. As the degree of coherence of the system is enhanced, its realization becomes more and more difficult.

53.6 Practical Constraints and Countermeasures

The coherent systems promise excellent performance compared with the conventional IM/DD system, as previously mentioned. However, practical systems have many difficulties, such as the phase and intensity noise of the LD, frequency response and bandwidth of the modulators and receivers, fluctuation of the SOP in a fiber, and the dispersion and nonlinearity in a fiber.

Phase and Intensity Noise in the Laser Diode

The prime concern in realizing coherent systems has been the spectral purity of the signal and local LDs. Because the spectral shape of a single-mode laser under continuous wave (CW) operation becomes Lorentzian, the spectral purity is characterized by the linewidth of the full width at half-maximum (FWHM). The origin of the spectral broadening is phase noise generated by spontaneous emission events. Typical linewidth of a commonly used DFB-LD is around 10 MHz, which is in contrast to electrical or microwave oscillators having linewidths less than 1 Hz. The required beat linewidth to hold the sensitivity penalty below 1 dB is shown for various schemes in Table 53.1 [Linke and Gnauck, 1988]. It is proportional to the bit rate for all schemes, which can be understood by the fact that the phase correlation between adjacent bits becomes higher as the bit duration becomes shorter. As mentioned previously, the homodyne schemes require extremely narrow linewidth. It is tolerated in heterodyne schemes, but PSK and narrow-deviation FSK schemes require still narrower linewidth. In contrast, ASK and wide-deviation FSK can accept wider linewidth. As the bit rate has grown higher recently (typically, 2.5 Gb/s), the requirement has been relaxed and PSK or the narrow-deviation FSK schemes have become common. Nevertheless, homodyne schemes are still difficult.

To avoid the extremely severe linewidth requirement of the homodyne schemes, the **phase-diversity receiver** has been proposed, as shown in Fig. 53.5(a), in which two orthogonal in-phase (I) and quadrature (Q) homodyne-detected signals $i_I(t)$ and $i_Q(t)$ are produced using a so-called 90° optical hybrid as

$$i_I(t) = 2KE_S E_{LO} d(t) \cos[\phi_{IF} + \theta(t)] \tag{53.7}$$

$$i_Q(t) = -2KE_S E_{LO} d(t) \sin[\phi_{IF} + \theta(t)] \tag{53.8}$$

By feeding the signals to demodulators (square-law demodulators for ASK and differential demodulators for PSK) and combining, the baseband signal is obtained irrespective of the phase term ϕ_{IF} [Kazovky, 1989]. The 90° optical hybrid uses the properties of SOPs of the signal and LO light; I and Q signals are produced by combining the signal of a 45° linear SOP and the LO of a circular SOP and by splitting them using a polarization beam splitter (PBS).

If the I and Q signals are up-converted by two electrical LO waves having mutually orthogonal phases $\cos \omega_{IF2} t$ and $\sin \omega_{IF2} t$, as shown in Fig. 53.5(b), instead of demodulating them, the combined up-converted

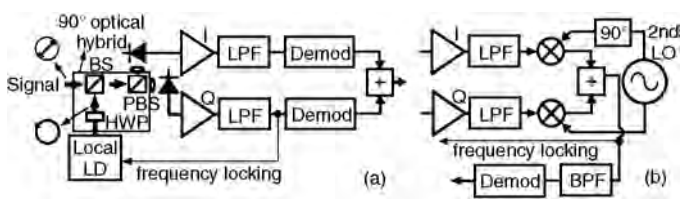


FIGURE 53.5 Phase-diversity receivers: (a) single-stage and (b) double-stage.

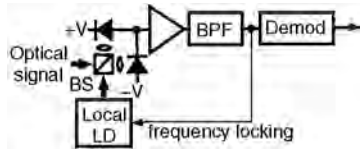


FIGURE 53.6 Balanced heterodyne receiver.

signal becomes

$$i_{\mathrm{IF2}}(t) = 2KE_S E_{\mathrm{LO}} d(t) \cos[\omega_{\mathrm{IF2}} t + \phi_{\mathrm{IF}} + \theta(t)] \tag{53.9}$$

which is identical to a heterodyne-detected IF signal having the IF of ω_{IF2}. The baseband signal is obtained by feeding it to a heterodyne demodulator. It is called the double-stage phase-diversity (DSPD) receiver [Okoshi and Yamashita, 1990]. The receiver sensitivities and the linewidth requirement of the phase diversity and DSPD receivers are both the same with the heterodyne receivers.

An LD not only has phase noise but also **intensity noise**, which is also generated by spontaneous emission events. Specifically, the intensity noise in a local LD produces additional large noise through the second term in Eq. (53.3). It can be suppressed by using a **balanced receiver** as shown in Fig. 53.6. When we use a 50:50 BS or fiber coupler, the two outputs have the same power but the relative phases between the signal and LO fields have a 180° difference, because the phase does not shift on transmission but shifts by 90° on reflection. The two outputs are detected independently and subtracted to be

$$\begin{aligned} i_{\mathrm{bal}}(t) &= K|E_S(t) + E_{\mathrm{LO}}(t)|^2 - K|E_S(t) - E_{\mathrm{LO}}(t)|^2 \\ &= 4KE_S E_{\mathrm{LO}} d(t) \cos[\omega_{\mathrm{IF}} t + \phi_{\mathrm{IF}} + \theta(t)] \end{aligned} \tag{53.10}$$

It is found that terms of the direct-detected signal and LO lights are suppressed. The balanced receiver has another merit in that it can effectively utilize the optical power that is unused in the single-detector receiver. The subtraction is easily performed by combining two PDs in tandem. Therefore, the balanced receiver is implemented in most systems. It can also be combined with phase-diversity receivers.

Frequency Response and Bandwidth of Modulators and Receivers

In most of the recent experiments, CPFSK modulation is employed because there is no need for an external modulator and direct modulation is much easier, as shown in Fig. 53.2(b). However, the normal DFB-LDs have different frequency responses for direct modulation between the low- and high-frequency regions and cause distortion of the signal waveform. This is a result of the synergism of the heat effect and carrier effect on the modulation of the LDs. To avoid this, we can use an equalization circuit before the bias circuit for direct modulation, or we can use the more sophisticated LDs, having flat frequency response such as three-electrode distributed Bragg reflector (DBR) LDs. The $LiNbO_3$ modulators used for ASK/PSK modulations do not have this problem, although a high-power, high-speed driver is necessary and the transmitter becomes more complicated.

In a heterodyne receiver, the bandwidth of the PD and IF amplifiers must be over 5–6 times the baseband, as shown in Fig. 53.3(a), which makes high-speed signal (over 5 Gb/s) detection difficult. In this point, a homodyne receiver is advantageous because the required bandwidth is only the baseband, as shown in Fig. 53.3(b). The homodyne receiver has another advantage in that it has no image interference from adjacent channels when it is used in an FDM system. These are the reasons why the phase-diversity techniques have been studied.

In principle, it is possible to build an FDM system with an arbitrarily large number of channels having channel spacing equal to the bit rate R_b. Actually, this is, of course, not exact since the modulated spectrum

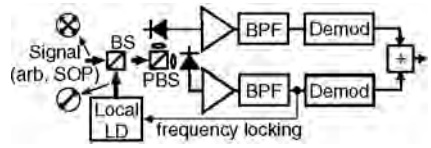


FIGURE 53.7 Polarization-diversity heterodyne receiver.

can be much wider and some guard band is necessary. The narrow-deviation CPFSK signal has a spectrum as narrow as R_b, and the required channel spacing has been found to be about $3R_b$ in a homodyne receiver [Linke and Gnauck, 1988]. It is increased to be $6R_b$ in a heterodyne receiver due to image interference. It should be noted that the balanced receiver is indispensable in the FDM system in order to reject the beat components between channels.

Polarization Fluctuation in the Fiber

In the derivation of Eq. (53.3), we assumed the SOPs of the signal and LO light fields are matched. In practice, however, the SOP of the signal fluctuates in a fiber due to the birefringence change caused by the external stress or the temperature change. There are four countermeasures: (1) use of a polarization-maintaining (PM) fiber such as a PANDA (Polarization-maintaining AND Absorption-reducing) or a bow-tie fiber as the system fiber, (2) use of an active polarization controller at the receiver, (3) use of a polarization-scrambling scheme in which the SOP of the signal is scrambled within one bit at the transmitter, and (4) use of a **polarization-diversity receiver**. Among these countermeasures, the fourth is considered to be superior and is implemented in many receivers. Figure 53.7 shows the polarization-diversity receiver. It uses an optical circuit similar to the 90° optical hybrid. The LO light is set to be a 45° linear SOP and split by the PBS to be x and y linear SOPs having the same amplitude. The PBS splits the signal to be $k{:}1-k$ according to the power ratio of x to y elements. Then, the IF signals detected by the x and y branches have the amplitude ratio of $\sqrt{k}{:}\sqrt{1-k}$. Just like the phase-diversity receiver, they are fed to demodulators having square-law characteristics and combined to be baseband signal independent of the value of k. The polarization diversity scheme can be combined with phase diversity and balanced receivers.

Dispersion and Nonlinearity in the Fiber

As bit rate, incident power, transmission distance, and number of channels are increased, optical fiber properties other than transmission loss become increasingly important. The loss of the single-mode optical fiber (0.2 dB/km at 1.55 μm and 0.35 dB/km at 1.3 μm) is the basic limitation on transmission distance. However, dispersions and various nonlinear effects can dominate under certain conditions [Linke and Gnauck, 1988]. The single-mode optical fiber has **chromatic dispersion** around 17 ps/km · nm in the lowest loss window at 1.55 μm. It is a phenomenon in which the speed of the light is different in frequency and acts on the spectrum spread by the modulation. As the different spectral components arrive at different times, it produces distortion in the demodulated waveform, resulting in intersymbol interference in the received signal and degradation of the receiver sensitivity. Its effect is proportional to $R_b^2 L$, where L is the total system length, and, therefore, becomes more significant as the bit rate becomes higher. A solution is the use of a dispersion-shifted fiber, which has zero dispersion around 1.55 μm, but it costs much more and is inconsistent with the normal fiber already installed for conventional systems. In a heterodyne receiver, we have another solution: the use of an equalizer at the IF stage [Iwashita and Takachio, 1990]. A microstrip line (typically 20 cm long for the compensation of 100-km fiber) has been found to be adequate as the equalizer because it has the reverse characteristics in dispersion. It is possible because the heterodyne scheme is spectrally transparent, that is, the IF spectrum is the down-converted

optical spectrum, as shown in Fig. 53.3(a). In a homodyne or a phase-diversity receiver, the detected spectrum is folded into baseband and, thus, equalization is not possible. The DSPD receiver, in contrast, can equalize at the second IF stage. Thus, the DSPD receiver has both the advantages of the heterodyne and the homodyne schemes.

The birefringence in the fiber causes not only the fluctuation of the SOP but also **polarization mode dispersion (PMD)**. It is a phenomenon where two orthogonal polarization modes propagate at different speeds in a fiber, and the different spectral components arrive at different SOPs. The propagation delay is known to be proportional to $\sqrt{L}$. The PMD in a typical fiber is not very large ($0.1 - 2.0$ ps/$\sqrt{\text{km}}$), but it becomes important in a long-distance high-speed system.

Other problems in the optical fiber are the various nonlinear effects [Agrawal, 1989]. For coherent systems, the **stimulated Brillouin scattering (SBS)**, the **stimulated Raman scattering (SRS)**, and **nonlinear Kerr effect** are the main problems. The SBS and SRS are phenomena where the light incident to a fiber is back-reflected due to the interaction between a photon and a phonon when the power exceeds the threshold. The threshold of the SBS is typically 3–4 mW, whereas the fiber amplifier can amplify the signal to 10–100 mW when used as the postamplifier after the transmitter LD. Fortunately, the bandwidth of the SBS is as narrow as 100 MHz. The SBS can be avoided if the high-speed carrier-suppressed modulations (PSK or CPFSK) or the optical spread-spectrum modulations are employed. The threshold of the SRS is much higher (~1 W) in a single-channel system, although the SRS brings crosstalk at lower power in an FDM system. The Raman induced crosstalk is known to be reduced by closely spacing the channels, which can be done easily in a coherent system.

The nonlinear Kerr effect is a third-order nonlinear effect. In the presence of the Kerr effect, the refractive index of the fiber becomes $n = n_0 + n_2|E|^2$, where n_0 and n_2 are the linear and nonlinear refractive indices. In the ASK system, the modulation in amplitude causes modulation in phase (AM-PM conversion). It is called self-phase modulation (SPM). It causes chirping in the signal spectrum and produces waveform distortion in conjugation with the chromatic dispersion. In an FDM system, the phase of a channel may be modulated by the amplitude of the other channel. This is called cross-phase modulation (XPM). SPM and XPM can be avoided by using the modulations having constant envelope (PSK or FSK). The most severe restriction in FDM systems is four-wave mixing (FWM), in which the two lightwaves $E_1\exp(j\omega_1 t)$ and $E_2\exp(j\omega_2 t)$ generate another two lightwaves $E_3\exp(j\omega_3 t)$ and $E_4\exp(j\omega_4 t)$ through the Kerr effect, where $\omega_3 = 2\omega_1 - \omega_2$ and $\omega_4 = 2\omega_2 - \omega_1$. The generated lights may interfere with the other channels. It has also been pointed out that, in long-distance single-channel systems using optical amplifiers, phase noise is enhanced through the nonlinear interactions between the signal and amplified spontaneous emission (ASE) noise, which makes the long-distance coherent system difficult [Ryu, 1992; Ryu, 1995].

53.7 Summary and Conclusions

Coherent techniques offer the potential of nearly ideal receiver sensitivity together with frequency selectivity improvements. Most of the technical obstacles described here have been overcome, and coherent systems are ready to be practical. Unfortunately, the advent and recent development of the erbium-doped fiber amplifiers has reduced the first advantage and seems to keep the coherent systems away from practical use. Nevertheless, the second advantage still remains very important because it enables us to exploit the vast bandwidth of an optical fiber. Moreover, coherent techniques are found to be effective to enhance the performance of the amplifier systems [Yamashita and Okoshi, 1994]. Coherent techniques will certainly play a leading role in future communication networks.

Defining Terms

Balanced receiver: A coherent receiver having two detectors combined in tandem, which can suppress intensity noise.

Chromatic dispersion: A phenomenon in which the speed of the light is different in optical frequency.

Heterodyne scheme: A coherent detection scheme in which the optical signal is down-converted to the microwave intermediate frequency.

Homodyne scheme: A coherent detection scheme in which the optical signal is down-converted directly to the baseband.

Intensity noise: Random intensity fluctuation of the light emitted from a laser, which is generated by spontaneous emission events.

Nonlinear Kerr effect: The third-order nonlinear effect in a fiber, which changes the refractive index in proportion to the intensity of the light.

Phase-diversity receiver: A homodyne receiver composed of a 90° optical hybrid and two sets of detectors and demodulators, in which the linewidth requirement is very relaxed.

Phase noise: Random phase fluctuation of the light emitted from a laser, which is generated by spontaneous emission events.

Polarization-diversity receiver: A heterodyne receiver comprising an optical hybrid and two sets of detectors and demodulators, which is insensitive to polarization fluctuation.

Polarization mode dispersion (PMD): A phenomenon in a fiber, in which two orthogonal polarization modes propagate at different speeds.

State of polarization (SOP): A spatial trace drawn by the electric field vector of the lightwave, which is characterized by amplitude and phase differences between the vertical and longitudinal wave components.

Stimulated Brillouin scattering (SBS) and stimulated Raman scattering (SRS): Phenomena in a fiber where the incident light is back-reflected due to the interaction between a photon and a phonon when the power exceeds the threshold.

References

Agrawal, G.P. 1989. *Nonlinear Fiber Optics,* Academic Press, New York.

Hayashi, Y., Ohkawa, N., and Yanai, D. 1993. Estimated performance of 2.488 Gb/s CPFSK optical non-repeated transmission system employing high-output power EDFA boosters, *J. Lightwave Tech.,* 11(8):1369–1376.

Iwashita, K. and Takachio, N. 1990. Chromatic dispersion compensation in coherent optical communications, *J. Lightwave Tech.,* 8(3):367–375.

Kazovsky, L.G. 1989. Phase- and polarization-diversity coherent optical techniques, *J. Lightwave Tech.,* 7(2):279–292.

Linke, R.A. and Gnauck, A.H. 1988. High-capacity coherent lightwave systems, *J. Lightwave Tech.,* 6(11):1750–1769.

Oda, K., Taba, H., Nosu, K., Kato, K., and Hibirm, Y. 1994. Long-term error-free operation of a fully-engineered 128 channel optical FDM distribution-system. *Flee. Lett.,* 30(l): 62–63.

Okoshi, T. and Kikuchi, K. 1980. Frequency stabilization of semiconductor lasers for heterodyne communication systems. *Elec. Lett.,* 16(5):179–181.

Okoshi, T. and Kikuchi, K. 1988. *Coherent Optical Fiber Communications,* KTK/Kluwer, Tokyo.

Okoshi, T. and Yamashita, S. 1990. Double-stage phase-diversity (DSPD) optical receiver: analysis and experimental confirmation of the principle, *J. Lightwave Tech.,* 8(3):376–386.

Ryu, S. 1992. Signal linewidth broadening due to nonlinear Kerr effect in long-haul coherent systems using cascaded optical amplifiers. *J. Lightwave Tech.,* 10(10):1450–1457.

Ryu, S. 1995. *Coherent Lightwave Communication Systems,* Artech House Publishers, Boston/London.

Smith, D.W. 1987. Techniques for multigigabit coherent optical transmission. *J. Lightwave Tech.,* 5(10):1466–1478.

Yamashita, S. and Okoshi, T. 1994. Suppression of beat noise from optical amplifiers using coherent receivers. *J. Lightwave Tech.,* 12(6):1029–1035.

Further Information

More details on the general aspects of coherent systems are described in Okoshi and Kikuchi [1988] and Ryu [1995]. Details on practical constraints can be found in Linke and Gnauck [1988] and Smith [1987]. Current research is reported in a number of *Technology Letters*, and *IEEE Electronics Letters*. Annual international conferences of interest include the Conference on Optical Fiber Communications (OFC), the European Conference on Optical Communication (ECOC), the International Conference on Integrated Optics and Optical Fiber Communication (IOOC), the Opto-Electronics Conference (OEC), and many more.

54

Fiber Optic Applications

Chung-Sheng Li
IBM T.J. Watson Research Center

54.1 Introduction

The recent rapid advance of optical technologies includes the following:

- Low-loss single-mode fiber
- Single longitudinal mode distributed feedback (DFB) laser diodes
- Low-threshold-current multiple-quantum-well (MQW) lasers
- Vertical cavity surface emitting lasers (VCSEL)
- High-speed p-i-n and MSM photodetectors and receivers
- Optical amplifiers
- Integrated optics

These enable the deployment of fiber optical components to both computing and communication systems. The rapid commercialization of these components has already had a profound impact in the following areas:

- Optical interconnect within a system or between systems
- Local area networks
- Access networks
- Backbone networks
- Fiber optic sensors

In this chapter, we will examine each individual application in detail.

54.2 Optical Interconnects

Current metal technology has difficulty in supporting bus and backplane bandwidth exceeding several hundred megahertz. Therefore, a serious bandwidth bottleneck exists between processor and memory or between processors. Thus, new technology is required to overcome these obstacles in order to create low-latency, high-bandwidth communication paths whose characteristics match the performance of ever-increasing processor speed. Different metal-based technology standards such as Rambus [Farmwald and Mooring, 1992], Ramlink [Gjessing et al., 1992], and scalable coherent interface (SCI) [IEEE, 1992] have been proposed for parallel metal interconnects to meet the requirements of systems with clock speeds up to 500 MHz and are in the process of being implemented. These systems show that careful engineering such as the use of a controlled-impedance transmission line, fully differential signalling (SCI and Ramlink), and limiting the interconnection structure to point-to-point (SCI and Ramlink) can push the metal interconnection technology to higher bandwidths than could be supported with existing technology. Nevertheless, the cost of the new technology is initially high, and there is no guarantee that the solutions can be scaled to clock speeds higher than their present limits. If any of these interconnection technologies is going to succeed, it must be available at low cost in high volume, and must be able to evolve to support future clock speeds of 1 GHz and beyond.

Optical interconnections provide an alternative technology to solve the interconnection problem. The multigigabit bandwidth allowed by this technology is more than sufficient for applications, such as communications within a multimedia system, for the foreseeable future. Furthermore, it is easier to control reflections in this technology for both point-to-point and multidrop structures than in metal links [Stone and Cook, 1991]. Optical links generally exhibit less noise problems due to ground-loop noise because fibers do not carry currents as do metal links. For these reasons, optical interconnections may be an attractive alternate technology for building high-speed board and backplane interconnections in future multimedia systems.

An optical link can be designed to be an almost one-to-one replacement for metal point-to-point or multidrop connections. The conventional line driver is replaced by a laser driver and an edge-emitting or surface-emitting laser diode/LED, or a laser diode and an external modulator such as a Mach–Zehnder interferometer, directional coupler, total internal reflection (TIR) modulator spatial light modulator (SLM), self-electro-optic device (SEED), or vertical-to-surface transmission electrophotonic device (VSTEP) at the transmission end. The conventional line receiver is replaced by a light sensitive device such as a p-i-n diode or metal-semiconductor-metal (MSM) photodetector and an amplifier at the receiving end. The light can be guided from the transmission end to the receiving end through single-mode or multimode fiber ribbon cable, polyimide or silica-on-silicon channel waveguides, or free-space microlenses and/or holograms.

Figure 54.1 shows the structure of a typical optical interconnect, which consists of a driver array, a laser diode or LED (LD/LED) array, a waveguide or fiber ribbon array, a photodetector array (p-i-n or MSM), and a receiver array. Using optical interconnects for high-bandwidth communication channels between boxes has been demonstrated, for example, in Lockwood et al. [1995]. It is conceivable that optical interconnect can also be used within a box (both at the board and backplane levels), shown in Fig. 54.2.

Interconnect Media

Possible media that can be used for optical interconnects include:

- *Free-space interconnect.* Light travels fastest in free space. In addition, free-space interconnects also offer the highest density and the most sophisticated interconnection patterns. Unfortunately, bulk optical elements, such as lenses, holograms, beam splitters, etc., are usually unavoidable in free-space optical interconnects and, thus, make the alignment of optical beams very difficult and unstable with respect to environmental disturbances.

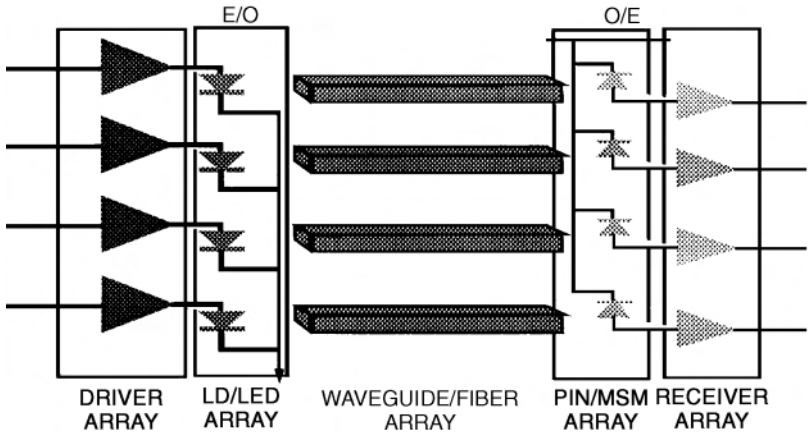


FIGURE 54.1 Structure of an optical interconnect system.

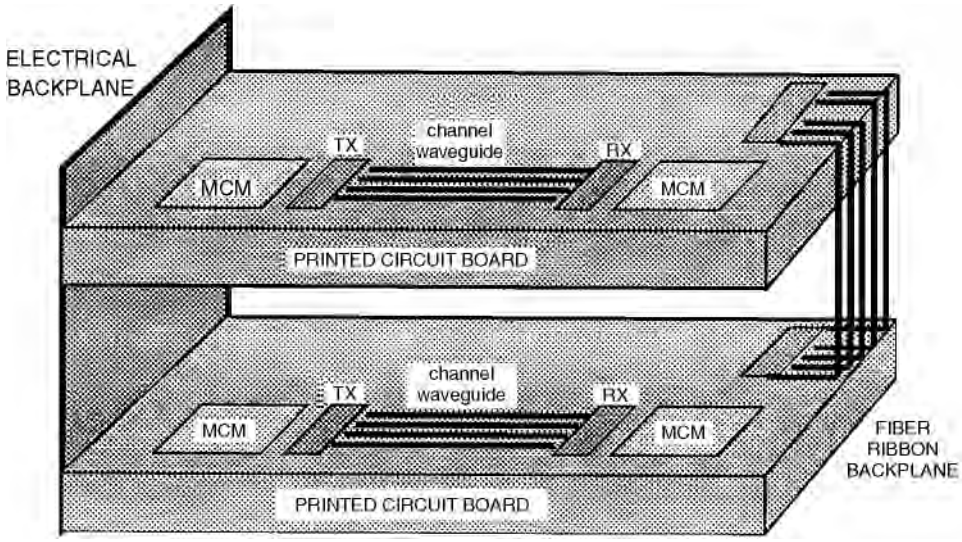


FIGURE 54.2 A packaging structure uses optical interconnect at board and backplane levels.

- *Optical fiber ribbon*. Optical fiber has the least loss compared with the other two media, and most of the technologies used in fabrication are already mature. Fiber ribbon cable also has the potential of providing reasonable interconnect density with regular interconnection pattern. However, fibers are incompatible with the existing packaging technology at the board or multichip module (MCM) level and they are not suitable for interconnects with very short distance or complicated patterns due to the possibly excessive volume occupied by the fiber cable. A lot of research effort has thus been devoted to develop compatible connector and packaging technologies to interconnect fiber ribbon and optical transceiver arrays.
- *Planar optic waveguide.* Passive planar optic waveguides are emerging as a viable alternative to optical fiber for very short distance interconnects. They have a higher propagation loss than optical fiber (0.01 ~ 0.5 dB/cm as compared to 0.2 dB/km) but use technologies that are compatible with existing PCB or MCM technology. Therefore, they are more suitable for short-distance dense interconnect applications. However, coupling of light into and from the waveguides is also difficult and careful alignment cannot be avoided.

Backplane Optical Interconnects

The function of a backplane is to provide a logical bus for all of the boards connected to it. Free space, fibers, and planar waveguides are all suitable for backplane interconnects. An optical backplane can be achieved through the use of star couplers, as shown in Fig. 54.3. Each board in the architecture occupies one input port and one output port from each of the star couplers so that signals input to any of the input ports will be broadcast to all of the output ports. The total number of star couplers required can be greatly reduced by multiplexing several channels into a single waveguide, with each channel using a different wavelength.

On the other hand, a topological bus can also be used to interconnect from one board to another, as shown in Fig. 54.4. The bus is either folded back at the end or two independent buses are used because a unidirectional optical bus structure is usually easier to implement.

Board and Multichip-Module Optical Interconnects

Board-level optical interconnects have to provide interconnects between different single chip modules (SCMs) or MCMs, whereas MCM-level optical interconnects have to provide interconnects between unpackaged wire-bonded or solder-ball-bonded flipped chips. At the board level, the electrical to optical (E/O) and optical to electrical (O/E) conversion can be performed within an SCM/MCM, or through separate special-purpose E/O and O/E chips. Similarly, the E/O and O/E conversion at the MCM level

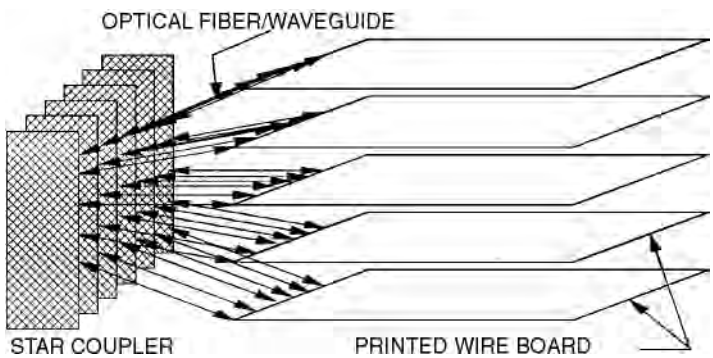


FIGURE 54.3 Optical backplane interconnects: star couplers are used to combine and redistribute the data signals.

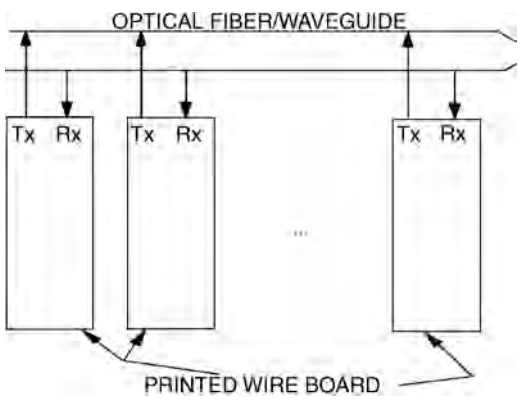


FIGURE 54.4 Optical backplane interconnects: a topological bus is used to provide a communication path between any two boards connected to the backplane.

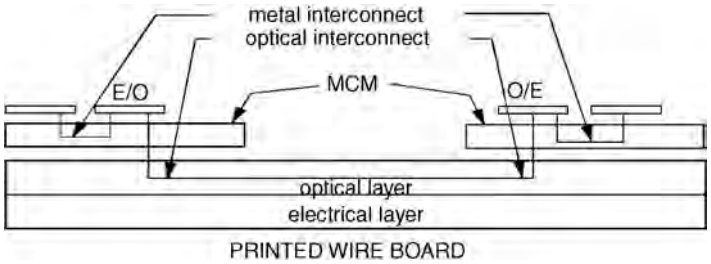


FIGURE 54.5 E/O and O/E conversion: conversion is performed at the same packaging level as the electrical signal is generated.

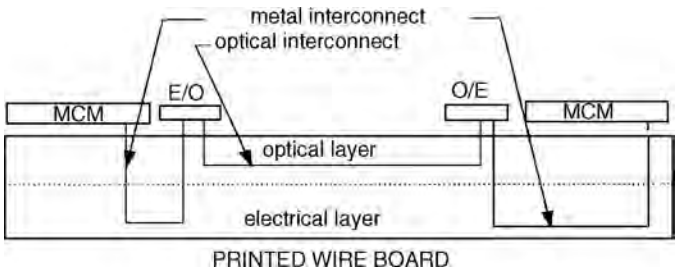


FIGURE 54.6 E/O and O/E conversion: conversion is performed at the next higher packaging level.

can be performed within the chip where the logical signals are generated or via separate special-purpose E/O and O/E chips on an MCM.

If the E/O and O/E conversion is performed before the package is connected to the next higher level, as shown in Fig. 54.5, the electrical discontinuity can be minimized but the optical alignment is more difficult. On the other hand, more electrical discontinuity and, thus, more signal distortion are introduced if the E/O and O/E conversion is performed after the package is connected to the next level, as shown in Fig. 54.6. However, this is acceptable for applications that require only moderate data rates.

In both cases, there already exist multiple layers of metal interconnect to provide signal lines as well as power and ground plane. Optical interconnects can be developed on top of these metal interconnect layers to allow optical signals to propagate from one chip/module to another chip/module. In some cases, more than one optical layer may be necessary in order to provide a sufficient interconnect density (such as at the MCM level) just similar to its metal counterpart.

54.3 Local Area Networks and Input/Output (I/O) Interconnections

Examples of local area networks and input/output (I/O) channels that use fiber optics include:

- Fiber distributed data interface (FDDI), FDDI-II
- Fiber channel
- Gigabit ethernet (802.3z)

Since the protocol of FDDI and ethernet have been described elsewhere, this section will only focus on the fiber channel.

The major application area of fiber channel, which became an American National Standards Institute (ANSI) standard in 1994, is to provide high-bandwidth (up to 100 megabyte/s) connectivity to clusters of supercomputers, mainframes, workstations, disk arrays, and other peripheral devices. Its origin has strong ties with high performance parall interface (HIPPI) and enterprise systems connection (ESCON),

both of which provide circuit switching, and small computer systems interface (SCSI), which is the primary interconnection mechanism for peripheral devices. Fiber channel defines a point-to-point or switched point-to-point link, which includes definitions of the physical layer, transmission code, and higher level functions such as flow control and multiplexing. Figure 54.7 shows a collection of nodes (called *N-port*) interconnected by a fabric through the fabric ports (called *F-port*).

The fiber channel standard includes definitions of the following functions:

- *Physical layer.* The physical layer specifies four data rates: 133, 266, 533, and 1062.5 Mb/s. Because of the adoption of 8B/10B line coding technique, there is a 25% increase in the transmission data rate. The transmission media defined for a fiber channel includes miniature coax, multimode and single-mode fibers. The physical layer also defines bit synchronization, serial-to-parallel conversion, and an open fiber control protocol. This protocol performs automatic shutdown of the laser diode when there is a fiber cut and automatically establishes a full duplex link when the fiber connection is restored.
- *Data link layer.* This layer specifies a line coding technique, the 8B/10B code, to ensure sufficient transitions are present in the data stream. In addition, byte-to-word conversion and word synchronization are also performed in this layer.
- *Arbitrated loop layer.* The arbitrated loop layer specifies the loop control and bypass mechanisms for an end node when such an end node participates in the connection of a loop, as shown in Fig. 54.8.

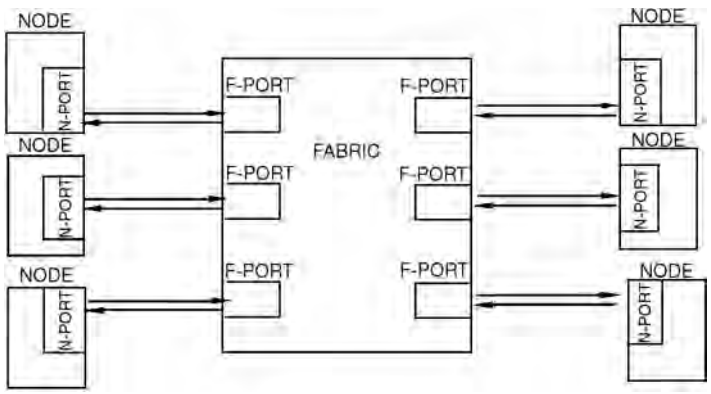


FIGURE 54.7 Several nodes interconnected by a fabric.

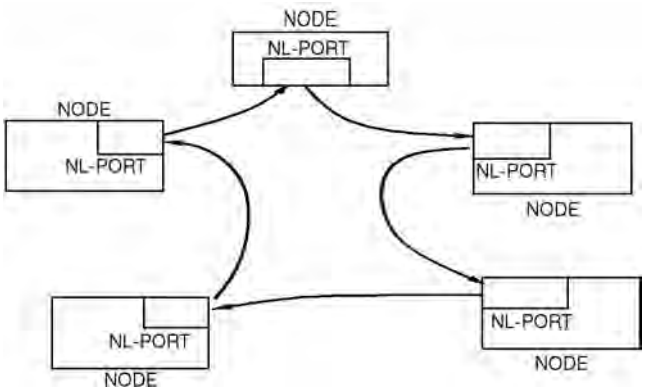


FIGURE 54.8 Several nodes interconnected by a fabric.

This configuration has the advantage of reduced cost at the cost of reduced performance and is particularly attractive for disk storage applications.

- *Frame processing layer.* The processing and validation of the integrity of each received frame including, CRC checking is performed in this layer.

54.4 Access Networks

Access networks provide interconnections between the customer premises equipment (CPE) and the central offices. The average distance of these types of networks is usually less than 3 mi. The components of access networks are extremely cost sensitive. Although the price of fiber has gone down dramatically, the price of a single-mode distributed feedback (DFB) laser has not. The design of the access network is, thus, to minimize the use of active devices such as the laser diodes or confine the use of such devices within central offices. The current fiber deployment strategies are divided into the following two categories: hybrid fiber coaxial (HFC) and fiber to the home (or fiber in the loop). Although fiber to the home is the ultimate answer to broadband services, HFC seems to be the most tangible interim solution.

The hybrid fiber coaxial cable structure deploys both the fiber and the coaxial cable. The basic fiber/coax network is already taking shape as a result of the cable TV delivery system. Signals in most of the existing HFC networks are transmitted using fiber with broadband RF spectrum to an headend. A typical architecture is shown in Fig. 54.9. The signals are then converted back into an electrical RF signal and distributed to the home. The major advantages of this star-configured fiber network with coax bus are as follows [Paff, 1995]:

- Better quality video
- More bandwidth possible with lower beginning investment
- Better reliability
- Dedicated feed into and out of a small area

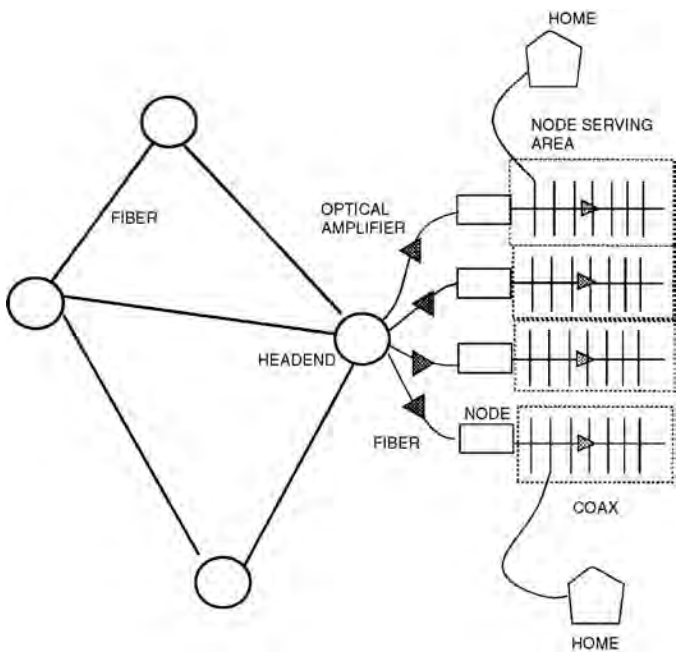


FIGURE 54.9 A broadcast-and-select network.

Although the current HFC network is a broadcast and isolated architecture, a lot of research had been devoted to the next generation HFC network that can provide an interactive and broadband element within the public telecommunication fabric.

54.5 Wavelength-Division Multiplexing-Based All Optical Networks

The advantages of a wavelength-division multiplexing (WDM)-based optical network can be grouped into two aspects. One aspect lies in its potential in utilizing the vast bandwidth offered by the single-mode optical fiber since there are no electronic components in the network. The other equally important aspect is its simple network control. To transmit a packet, the source only needs to identify the destination through a unique wavelength. There is no need to know about the exact network topology for route selection. There is no buffering within the network and, therefore, no need for flow control or congestion control at or below the network layer. Only the associated end nodes need to do handshaking between themselves.

Two types of WDM-based networks are commonly discussed:

- Broadcast-and-select networks
- Wavelength routed networks

Broadcast-and-Select Networks

A broadcast-and-select network consists of nodes interconnected to each other via a star coupler, as shown in Fig. 54.10. An optical fiber link, called the *uplink*, carries signals from each node to the star. The star combines the signals from all of the nodes and distributes the resulting optical signal equally among all its outputs. Another optical fiber link, called the *downlink*, carries the combined signal from an output of the star to each node. Examples of such networks are Lambdanet [Goodman et al., 1990] and Rainbow [Janniello, Ramaswami, and Steinberg, 1993].

Wavelength-Routed Networks

A wavelength-routed network is shown in Fig. 54.11. The network consists of *static* or *reconfigurable* wavelength routers interconnected by fiber links. Static routers provide a fixed, nonreconfigurable routing pattern. A reconfigurable router, on the other hand, allows the routing pattern to be changed dynamically. These routers provide static or reconfigurable *lightpaths* between end nodes. A lightpath is a connection consisting of a path in the network between the two nodes and a wavelength assigned on the path. End nodes are attached to the wavelength routers. One or more controllers that perform the network management functions are attached to the end node(s).

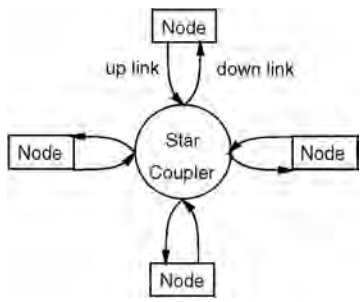


FIGURE 54.10 A broadcast-and-select network.

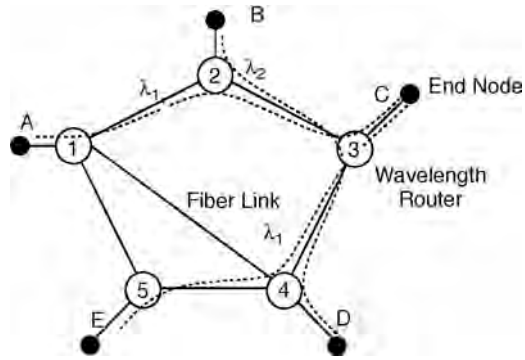


FIGURE 54.11 A wavelength-routed network.

54.6 Fiber Sensors

It is well known that optical techniques, interferometry in particular, offer extremely high precision. However, it is difficult to apply these high-precision techniques outside a laboratory for the following two reasons:

- Environmental perturbations tend to cause rapid misalignment of the optical components.
- Optical components have to be within line of sight and, thus, remote sensing is difficult to achieve.

This situation has recently changed dramatically due to the incorporation of fiber optic waveguides into sensors where the fiber optic waveguides can act as the sensor or serve as a transceiver to transfer light signals to a remote sensor. Several examples of the application of optical techniques are as follows:

- *Biomedical transducer* [Mignani and Baldini, 1995]. The development of optic fiber sensors for medical applications dates back to the 1970s. The current-generation fiber optical sensors, which are mostly composed of silica and plastic fibers, are coupled to sensitive fiber sections or optrodes. These sensors are usually based on intensity modulation interrogation schemes, as shown in Fig. 54.12. Applications of these sensors include cardiovascular and intensive care, angiology, gastroenterology, ophthalmology, oncology, and neurology, to name just a few.
- *Chemical transducer*. Because of their immunity to electromagnetic interference (EMI), electric and intrinsic safety, possibility of miniaturization, and biocompatibility, optical fiber pH sensors are ideal for blood monitoring [Deboux et al., 1995]. These sensors are based on the observation of the changes in the optical properties of colorimetric or fluorimetric pH indicators. The changes are measured by reflection or through interaction with the evanescent field changes in the absorption due to changes in color of colorimetric indicators. Optical sensors based on optical time-domain reflectometry can also be used to detect the water ingress point.
- *Intelligent sensors*. Fiber optical sensors can be embedded in cement mortar bodies for nondestructive monitoring of structures and for deformation measurement [Habel and Polster, 1995].
- *Magnetic/electric field transducer* [Ezekiel et al., 1995]. Magnetic and electric fields can be measured through spectroscopic techniques, as shown in Fig. 54.13. These techniques usually involve the excitation of atoms and molecules in a cell using a tunable light source, such as a laser, and measuring the resulting absorption or fluorescence. In this way, it is possible to identify the type of molecule in the vapor, as well as the concentration. This method is generally referred to as a single-step or single-photon spectroscopy, since the excitation involves a single pair of energy levels. Recently, there are also new techniques that utilize two-step or two-photon spectroscopy, in which two radiation fields simultaneously interact with the atom or molecule in the vapor.

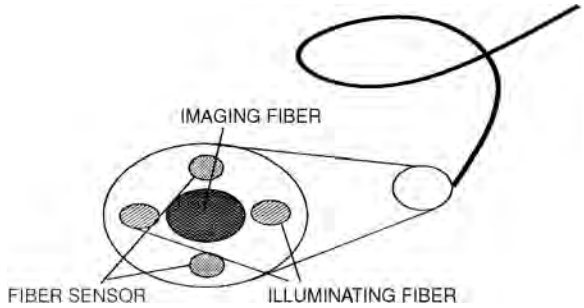


FIGURE 54.12 Configuration of a biosensor.

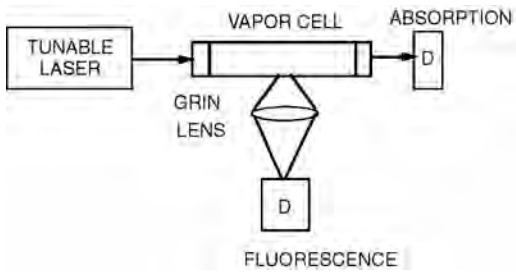


FIGURE 54.13 Configuration of a spectroscope setup for measuring absorption spectrum.

- *Moisture transducer*. Instead of using a large number of electrically passive sensors, moisture can be detected through the use of a single fiber based on optical time-domain reflectometry. This technique enables the presence of water to be detected as a linear function of the fiber length.

References

Deboux, B.J.C., Lewis, E., Scully, P.J., and Edwards, R. 1995. A novel technique for optical fiber pH sensing based on methylene blue absorption. *J. Lightwave Tech.,* 13(7).

Ezekiel, S., Smith, S.P., Shahriar, M.S., and Hemmer, P.R. 1995. New opportunities in fiber-optic sensors. *J. Lightwave Tech.,* 13(7).

Farwald, M.P. and Mooring, D. 1992. A fast path to one memory. *IEEE Spect.,* 29(Oct.):50–51.

Gjessing, S., Gustavson, D.B., James, D.V., Stone, G., and Wiggers, H. 1992. A RAM link for high speed. *IEEE Spect.,* 29(Oct.):52–53.

Goodman, M.S., Kobrinski, H., Vecchi, M., Bulley, R.M., and Gimlett, J.M. 1990. The LAMBDANET multiwavelength network: architecture, applications and demonstrations. *IEEE J. Selec. Areas Commun.,* 8(6):995–1004.

Habel, W.R. and Polster, H. 1995. The influence of cementitious building materials on polymeric surfaces of embedded optical fibers for sensors. *J. Lightwave Tech.,* 13(7).

IEEE. 1992. Scalable coherent interface: Logical, physical, and cache coherence specifications. IEEE P1594 Working Group, IEEE Standard Department.

Janniello, F.J., Ramaswami, R., and Steinberg, D.G. 1993. A prototype circuit-switched multi-wavelength optical metropolitan-area network. *J. Lightwave Tech.,* 11(May/June):777–782.

Lockwood, J.W., Duan, H., Morikuni, J.J., Kang, S.M., Akkineni, S., and Cambell, R.H. 1995. Scalable optoelectronic ATM networks: the iPOINT fully functional testbed. *J. Lightwave Tech.,* 13(6): 1093–1103.

Mignani, A.G. and Baldini, F. 1995. In-vivo biomedical monitoring by fiber optic systems. *J. Lightwave Tech.*, 13(7).

Paff, A. 1995. Hybrid fiber/coax in the public telecommunications infrastructure. *IEEE Commun. Mag.*, 33(4).

Stone, H.S. and Cocke, J. 1991. Computer architecture in the 1990s. *IEEE Comput. Mag.*, 24(9): 30–38.

55

Wavelength-Division Multiplexed Systems and Applications

Mari W. Maeda
IPA

55.1 Introduction

Research in wavelength-division multiplexing (WDM) technology is driven by the desire to exploit the immense bandwidth available for transmission in a single-mode optical fiber. The low-loss wavelength range around 1.55 μm alone can support over 10 THz of bandwidth. However, this bandwidth cannot be directly utilized by time-division multiplexed (TDM) systems due to the speed limitations posed by fiber dispersion. WDM offers a more practicable solution of multiplexing many high-speed channels at different optical carrier frequencies and transmitting them over the same fiber. It is expected that WDM technology will play a greater role in the coming years as the telecommunications networks evolve to meet the growing demand for bandwidth.

In this chapter, we describe the fundamental concepts of WDM systems, beginning with a review of key optical components and their performance requirements. This is followed by a discussion of system design issues and a review of several promising applications for telecommunications networks.

55.2 Optical Components for Wavelength-Division Multiplexed Systems

WDM systems are commonly classified into three general groups based on the choice of channel spacing and the aggregate transmission capacity supported. The simplest configuration, often known as the *coarse* WDM architecture, supports transmission of two channels in the broad low-loss windows of the fiber at 1.3 μm and at 1.55 μm. A greater capacity can be supported by *dense* WDM systems, where several

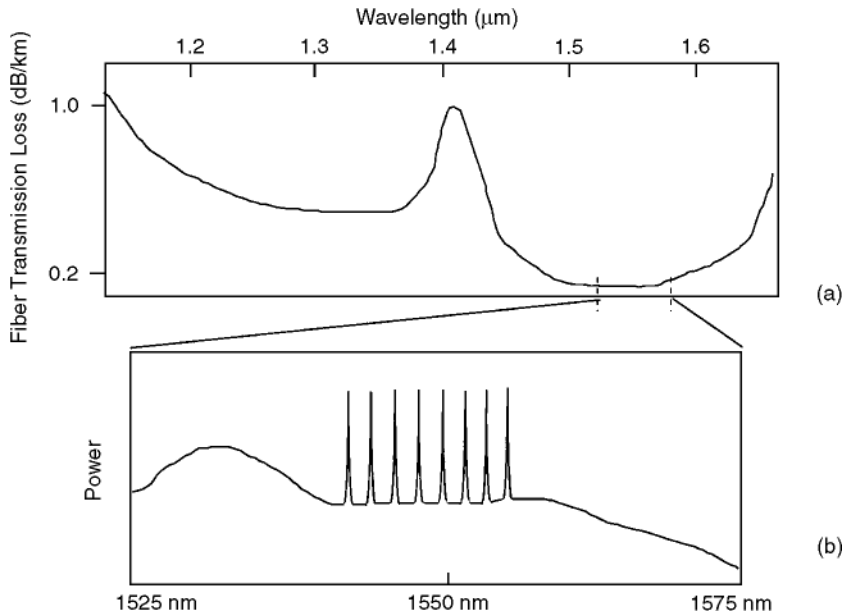


FIGURE 55.1 (a) Fiber transmission loss showing the low-loss windows at 1.3 and 1.5 μm. (b) Transmission spectrum for an eight-channel WDM system using an erbium-doped fiber amplifier (EDFA). The background detail shows the spontaneous emission spectrum from EDFA.

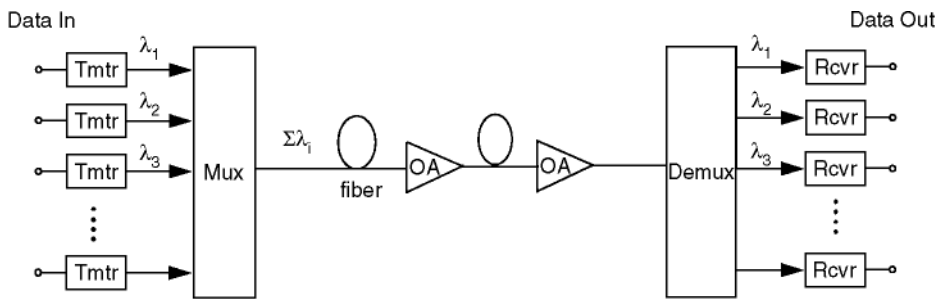


FIGURE 55.2 Basic WDM system configuration, optical amplifier (OA).

wavelengths are transmitted in the 1.55-μm region with a typical channel spacing on the order of one to several nanometers, as shown in Fig. 55.1. Optical frequency-division multiplexed (OFDM) systems support even greater density of channels with the spacing in the range of 0.1 nm (~10 GHz). Although both dense WDM and OFDM systems can support aggregate transmission capacity of over 100 Gb/s, dense WDM is becoming widely accepted as offering a more practical, near-term solution in various applications and will be the primary focus of this chapter.

Figure 55.2 shows a basic WDM system for point-to-point high-capacity transmission applications. The signals at different wavelengths are optically multiplexed so that N wavelengths are simultaneously carried on a common transmission fiber. Optical amplifiers may be used to compensate for fiber attenuation and component loss. At the receiving end, the signals are demultiplexed and detected by optoelectronic receivers. The key optical components for WDM systems are now reviewed.

Transmitter

Lasers for high-speed transmission systems must emit single wavelengths (i.e., a single longitudinal mode) and support high output power (>1-mW coupled into fiber pigtail) over a large modulation bandwidth with a specified spectral stability. Output spectral characteristic is one of the key considerations in designing high-speed systems where the transmission distances can be limited by fiber chromatic dispersion. Well-confined optical spectrum is especially important for WDM since the channel spacing must be selected to minimize interchannel crosstalk. Optical sources that are in widest use today are distributed feedback (DFB) lasers. DFB lasers can support multigigabit per second rate intensity modulation through direct injection current modulation and produce an output signal with a stable spectrum over the entire modulation bandwidth. For applications requiring ~10-Gb/s modulation rates, external modulation can be used to reduce the extent of wavelength chirping that normally accompanies direct modulation. It should be noted that incoherent sources such as light emitting diodes (LEDs) or super luminescent diodes together with wavelength filters can also be used to generate quasi-monochromatic signals [Iannone, 1995]. This so-called *spectral-slicing* technique is applicable for systems that emphasize simplicity and cost effectiveness of the transmitter design.

Multiplexer and Demultiplexer

Key considerations in selecting the multiplexing and demultiplexing technologies are the passband characteristics, insertion loss, polarization sensitivity, and thermal stability. For systems with wavelength spacing in the range of 1–10 nm, commonly used passive de/multiplexers are based on diffraction gratings and multilayer interference filters. For systems employing narrower channel spacing in the range of 0.1 nm, cascaded Mach–Zehnder or Fabry–Perot (FP) interferometer devices may be used for filtering the desired channel. A commonly used alternative to using a wavelength-selective coupling device is to use a passive N by N star coupler, which evenly splits each input signal between N outgoing ports. The intrinsic splitting loss of a star coupler may be compensated by the use of an optical amplifier. Star couplers are often employed in a *broadcast-and-select* configuration together with tunable filters on the receiving end. Examples of tunable filters include the index-tuned liquid crystal FP filter, acousto-optic filter, and DFB filter amplifier. For further details on tunable filter technology and its application in multiaccess networks, readers may consult Mestdagh [1995].

Optical Amplifier

Recent technological advances have shown optical amplification to be practical in the form of erbium-doped fiber amplifiers (EDFA). EDFAs possess the desirable features of high fiber-to-fiber gain (20 to 40 dB), high output power [>10 mW], and quantum limited noise figure, and offer approximately 30 nm of useful gain bandwidth in the 1535 to 1565 nm range [Desurvire, 1994]. For WDM systems employing cascaded EDFAs, cumulative gain nonuniformity over the gain spectrum can be significant and channel equalization is necessary (for example, by using filters or a signal pre-emphasis technique). The high signal powers made possible with EDFAs also raise the issue of signal degradations due to fiber nonlinearity, which must be carefully addressed especially in long-haul, high-speed system design.

55.3 Wavelength-Division Multiplexed System Design

Successful WDM system design entails a tradeoff of channel spacing, per channel bit rate, and feasible transmission distances based on various physical and practical constraints. Table 55.1 lists key factors limiting system performance along with possible solutions for relaxing the constraints.

The maximum transmission distances that can be attained without optoelectronic regeneration are limited by three characteristics of the fiber: linear loss, chromatic dispersion, and nonlinearity. Linear loss in the fiber, due primarily to Rayleigh scattering and molecular absorption, can be largely compensated by the use of optical amplifiers. For systems employing multigigabit per second channel rates, viable

TABLE 55.1 Sources of Limitations in WDM Systems, Their Effects, and Possible Solutions

Limitation Sources	Effect	Solutions
Fiber dispersion	Waveform distortion at high bit rates	Reduce signal chirp (e.g., use external modulator) Use dispersion-shifted fiber Use dispersion compensation
SPM/XPM	Waveform distortion at high bit rates and high launch powers	Use dispersion-shifted fiber Use dispersion compensation Increase channel spacing Reduce power
Four-wave mixing	Nonlinear crosstalk and loss at high launch powers	Increase channel spacing Avoid dispersion-shifted fiber Reduce per-channel power
Amplifier gain nonuniformity	Uneven gain for different λ	Use equilization or pre-emphasis Use narrow spacing to confine the channel spread
Wavelength drift	Crosstalk and signal fading	Require active λ stabilization Increase channel spacing

transmission distances become limited by the fiber dispersion and optical spectral characteristics. For example, with a *conventional* (nondispersion shifted) single-mode fiber that is widely deployed in the field, the dispersion level of ~17 ps/km-nm at 1550 nm limits the transmission distances to a mere 11 km when launching a 10-Gb/s signal with a spectral width of 1.0 nm. However, when the spectral width is reduced to 0.1 nm, the transmission distance is extended to 65 km. Such reduction in the optical spectral width can be attained by replacing direct laser modulation with the use of an external modulator. The chromatic dispersion problem can be also alleviated with the use of dispersion-shifted fiber whose zero-dispersion wavelength is in the 1550-nm range instead of the 1300 nm for conventional fiber. However, when such redeployment of fiber is not an option, as the case may be when upgrading a network using already installed fibers, dispersion management techniques can be used to extend the transmission span. For example, a dispersion compensating device or a length of negative-dispersion fiber can be inserted periodically into the fiber span to correct the signal distortions [Lin et al., 1980; Onaka et al., 1996]. In laboratory experiments, several different dispersion management techniques have been successfully employed to extend the regeneratorless transmission spans into the thousands of kilometers range.

When the transmitted signal powers are high, as they are in systems employing optical amplifiers, the effects of fiber nonlinearities must be taken into consideration during system design [Chraplyvy, 1990]. Fiber nonlinearities cause energy transfer between propagating wavelengths and lead to nonlinear loss and crosstalk if the system parameters such as launch powers and channel spacing are not carefully selected. One such effect is nonlinear phase modulation, which results from the power dependence of the refractive index. For single-wavelength systems, this nonlinearity takes the form of self-phase modulation (SPM), whereas in WDM systems, cross-phase modulation (XPM) plays a greater role as the phase of one wavelength is influenced by the intensity of all other wavelengths [Marcuse, 1994]. SPM and XPM broadens the signal spectrum, causing pulse distortions due to fiber chromatic dispersion. Degradations can be alleviated by reducing the launched signal power, increasing the channel spacing, and using a passive dispersion compensation technique mentioned earlier.

Another nonlinear effect that has a deleterious effect on WDM systems is four-wave mixing (FWM), where nonlinear interaction between three wavelengths generates a fourth wavelength, analogous to third-order intermodulation in radio systems. FWM interaction between different combinations of multiplexed wavelengths results in an appearance of interfering light at a signal channel, causing a crosstalk penalty.

Because FWM requires optical phase matching between the interacting signals, the effect can be reduced by using unequal or large channel spacings and by locally avoiding dispersion of magnitude less than about 1 to 2 ps/nm-km in the transmission fiber. The latter can be accomplished by using the dispersion compensating technique whereby the dispersion level in an individual fiber segment is kept greater than the required 1 to 2 ps/nm-km but the overall dispersion is minimized.

There are other factors, in addition to fiber nonlinearity, that influence the choice of WDM channel spacing. When the wavelength spacing is very small, it is possible to pack more channels into the wavelength regime where the EDFA gain is relatively uniform. However, when the spacing is reduced to a fraction of a nanometer, the design requirements for various optical components become quite stringent. For example, stabilization of optical sources and multiplexers becomes necessary so that wavelength drifts do not cause signal fading and crosstalk. Optical demultiplexers and filters also require more careful design to minimize interchannel crosstalk. Although there have been demonstrations of OFDM systems using wavelength spacing in the range of 0.1 nm (~10 GHz) [Nosu et al., 1993], a typically selected spacing for most applications is in the range from 0.6 nm to several nanometers. Indeed, a preliminary agreement in the International Telecommunication Union (ITU) standards endorses the usage of a 100-GHz (1.1 nm) frequency grid referenced to an absolute frequency of 193.100 THz (1552.52 nm).

55.4 Trunk Capacity Enhancement Applications

Near-term applications for high-speed WDM systems can be found in fiber trunk capacity enhancements. Currently available upgrade options are: (1) use of higher speed TDM systems, (2) use of additional fibers, and (3) use of wavelength multiplexing. Although a mix of three approaches will most likely be used in different parts of the network depending on the need and the economics, a transmission technique of combining high-speed TDM with WDM is especially attractive for long-haul applications since cost saving is made possible through fiber and amplifier sharing. Also, in those parts of the network where existing capacity is nearing exhaustion, the use of WDM offers a lower cost alternative to installing new cables or constructing new ducts.

Figure 55.3 shows the growing transmission capacity demonstrated in laboratory experiments from the mid-1980s to 1996. A two to three orders of magnitude increase in per-fiber capacity has been achieved

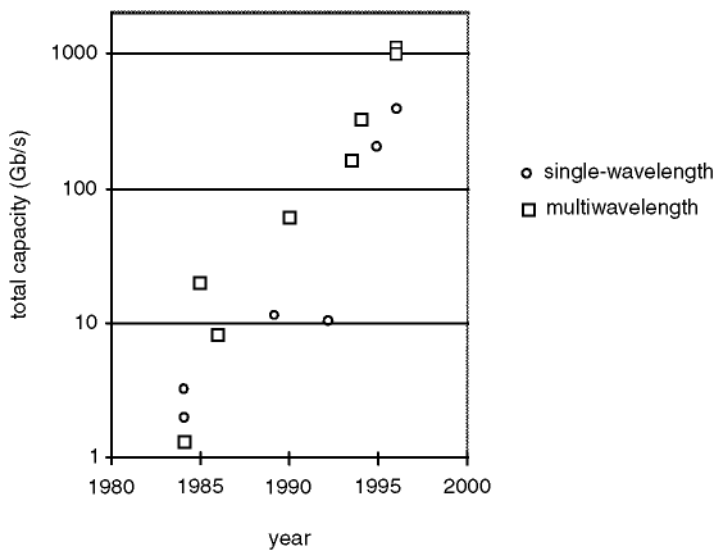


FIGURE 55.3 Trend in experimental demonstrations: total transmission capacity over one single-mode fiber for single-wave systems and for WDM systems.

with improvements in device speeds and through the use of WDM. A record 1.1-Tb/s transmission over 150 km was demonstrated in early 1996 with the use of 55 wavelength sources, each operating at 20 Gb/s [Onaka et al., 1996]. An equally remarkable trend has been witnessed in the transmission distances attained with WDM systems. In a recent experiment, eight 5-Gb/s channels were transmitted over 4500 km of fiber using 69 optical amplifiers [Suyama et al., 1996]. A separate experiment has demonstrated the feasibility of transmitting 20 5-Gb/s WDM channels over 9100 km of fiber in a circulating-loop configuration [Bergano et al., 1996].

Although it will be a while before terabit per second systems become a commercial reality, deployment of WDM systems has begun in terrestrial and undersea networks. Most terrestrial long distance networks in the U.S. are currently based on single-wavelength 2.5-Gb/s [optical carrier-OC-48] SONET systems. Next generation TDM systems operating at the speed of 10 Gb/s (OC-192) are nearing deployment while, simultaneously, WDM-based upgrade solutions are becoming commercially available. Network architects are also investigating the use of WDM in global undersea communication networks. For example, the Africa Optical Network, which will encircle the African continent with a 40,000-km trunk and branch network, will utilize a WDM technique [Marra et al., 1996]. Of eight wavelengths to be transported over the main ring fibers, four wavelengths will carry interregional traffic while the remaining four wavelengths will carry intraregional traffic. Over the trans-Atlantic and trans-Pacific spans, single-wavelength systems incorporating optical amplifiers are currently deployed. However, the feasibility of upgrading the existing cable capacity using WDM has been experimentally verified [Feggeler et al., 1996] and WDM technology will no doubt play an important role in future generations of transoceanic networks.

55.5 Wavelength-Division Multiplexed Networking and Reconfigurable Optical Transport Layer

Whereas point-to-point WDM systems are becoming a commercial reality, the role of WDM in a reconfigurable optical transport network is being researched by a number of laboratories around the world [Hill et al., 1993; Sato, 1994; Brackett et al., 1993; Alexander et al., 1993]. Such WDM networks can be dynamically reconfigured to carry existing transport signals and are especially suited for supporting high-capacity demands that vary with time, a situation that may become likely as broadband services become widespread [Berthold, 1996]. WDM networks may consist of nodes or network elements that can multiplex, switch, and terminate wavelengths. For example, an N by N wavelength cross-connect routes wavelengths from input WDM ports to output WDM ports. A wavelength add–drop multiplexer (W-ADM) performs the function of assigning an input wavelength and multiplexing this signal to be routed over the network. A chain of W-ADMs can be deployed in a linear or ring configuration, similar to SONET ADMs.

In addition to reconfigurability and capacity enhancement, the WDM network can offer an advantage of built-in network survivability. For example, in an event of a fiber failure, the entire multiplexed signal traversing the failed link can be switched and routed on a protection fiber using a loopback mechanism similar to that used in SONET rings. Alternatively, a wavelength path can be protected on an end-to-end basis using a pre-configured protection path that becomes activated only after the primary path failure.

Figure 55.4 shows an example design for a W-ADM with single and multiwavelength ports. An incoming wavelength at a single-wavelength port is wavelength adapted through a receiver/laser pair or multiplexed directly through the optical cross connect fabric. Incoming multiplexed signals are demultiplexed and dropped or transparently routed through the optical cross connect fabric. In contrast to SONET ADMs, electronic signal processing and regeneration are kept to a minimum with most signal processing being done in the optical domain. The key active elements are optical cross connects for routing signals and optical amplifiers for compensating signal loss. The integrity of the signals is monitored at various points internal to the node, allowing dynamic adjustments to be made in the event of signal degradations.

Research is currently being pursued to define various aspects of the network architecture and the nodal design, including the performance requirements, protection switching mechanisms, and methods to support embedded management channels. Because the viability of WDM networks hinges strongly on

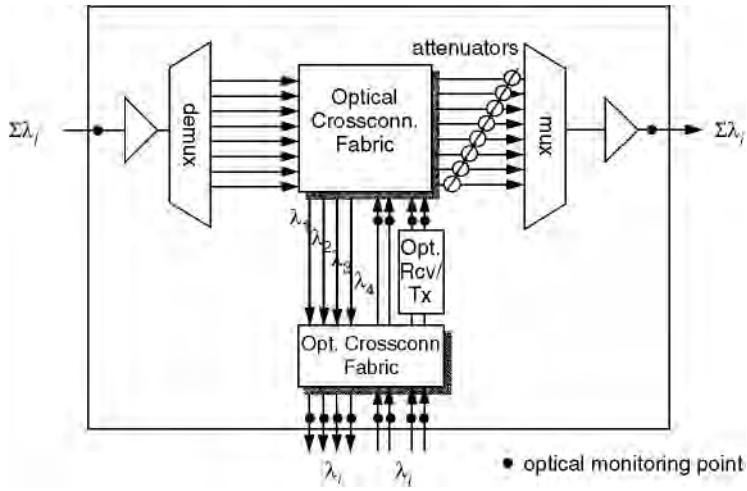


FIGURE 55.4 WDM add–drop multiplexer. Part of the design supporting WDM signals in only one direction is shown. Signal levels, noise levels, and wavelength readings are tracked at various monitor points.

the economics, both cost effectiveness and manageability of the network are being closely examined by a number of research consortia.

55.6 Summary

There have been rapid advances in the lightwave technology over the past decade, especially in the areas of single-frequency sources, optical amplifiers, wavelength multiplexers, demultiplexers, and optical receivers. These advances have made possible a tremendous increase in the per-fiber transmission capacity, as exemplified by the terabit per second WDM transmission experiments demonstrated in early 1996. WDM systems are beginning to be deployed in the terrestrial networks, as the as well as in undersea global networks and are expected to play a greater role in the future as the demand for bandwidth continues to grow and the need for economical capacity upgrades becomes more compelling.

References

Alexander, S.B. et al. 1993. A precompetitive consortium on wide-band all-optical networks. *J. Lightwave Tech.,* 11(May/June):714–735.

Bergano, N. et al. 1996. 100 Gb/s Error free transmission over 9100 km using twenty 5 Gb/s WDM data channels, Proceedings OFC'96, Postdeadline Paper 23-1, Feb.

Berthold, J. 1996, Application of WDM in local exchange networks, Proceedings OFC'96, Paper ThN1, Feb.

Brackett, C.A. et al. 1993. A scaleable multiwavelength multihop optical network: a proposal for research on all-optical networks. *J. Lightwave Tech.,* 11(May/June):736–753.

Chraplyvy, A.R. 1990. Limitation on lightwave communications imposed by optical-fiber nonlinearities. *J. Lightwave Tech.,* 8:1548.

Desurvire, E. 1994. *Erbium-Doped Fiber Amplifiers,* John Wiley & Sons, New York.

Feggeler, J. et al. 1996. WDM transmission measurements on installed optical amplifier undersea cable systems, Proceedings OFC'96, Paper TuN3, Feb.

Gnauck, A. et al. 1996. One terabit/s transmission experiment, Proceedings OFC'96, Postdeadline Paper 20-2, Feb.

Hill, G.R. et al. 1993. A transport network layer based on optical network elements. *J. Lightwave Tech.,* 11(May/June):667–679.

Iannone, P. 1995. A WDM PON architecture with bi-directional optical spectral slicing, Proceedings OFC'95, Paper TuK2, Feb.

Lin, C. et al. 1980. Optical-pulse equalization of low-dispersion transmission in single-mode fibers in the 1.3–1.7 μm spectral region. *Opt. Lett.,* 5(Nov.):476–478.

Marcuse, D. 1994. Dependence of cross-phase modulation on channel number in fiber WDM systems. *J. Lightwave Tech.,* 12(May):885–898.

Marra, W. et al. 1996. Africa ONE: the Africa optical network. *IEEE Commun. Mag.,* 34(Feb):50–57.

Mestdagh, D.J.G. 1995. *Fundamentals of Multi-Access Optical Fiber Networks,* Artech House, Norwood, MA.

Nosu, K. et al. 1993. 100 Channel optical FDM technology and its applications to optical FDM channel-based networks. *J. Lightwave Tech.,* 11(May/June):764–776.

Onaka, H. et al. 1996. 1.1 Tb/s WDM transmission over a 150 km 1.3 mm zero-dispersion single-mode fiber, Proceedings OFC'96, Postdeadline Paper 19-1, Feb.

Sato, K. 1994. Network reliability enhancement with optical path layer technologies. *IEEE J-SAC,* 12:159–170.

Suyama, M. et al. 1996. Improvement of WDM transmission performance by non-soliton RZ coding—a demonstration using 5 Gb/s 8-channel 4500 km straight line testbed, Proceedings OFC'96, Postdeadline Paper 26-2, Feb.

Further Information

For futher information on WDM technology and networking applications, see *Fiber-Optic Networks*, by P.E. Green, and *Special Issue on Broadband Optical Networks, IEEE Journal of Lightwave Technology*, Vol. 11, May 1993. Readers interested in access network applications may consult the article, "Technology and System Issues for a WDM-Based Fiber Loop Architecture," by S. Wagner and H. Lemberg, in *IEEE Journal of Lightwave Technology*, vol. 7, pp. 1759–1768, Nov. 1989. The latest developments in WDM components, systems, and applications are published each year in *Technical Digest of the Optical Fiber Communication Conference.*

V

Satellite

56

Geostationary Communications Satellites and Applications[1]

Bruce R. Elbert

Application Technology Strategy, Inc.

56.1 Introduction

Communication satellites, whether in geostationary Earth orbit (GEO) or low Earth orbit (LEO), provide an effective platform to relay radio signals between points on the ground. The users who employ these signals enjoy a broad spectrum of telecommunication services on the ground, the sea, and in the air. In recent years, such systems have become practical to the point where a typical household can have its own direct-to-home (DTH) satellite dish. Whether or not we reach the stage of broad acceptance depends on the competition with the more established broadcasting media, including over-the-air TV and cable TV. The success of British Sky Broadcasting in the U.K. market shows this is entirely possible. In any case, GEO and non-GEO satellites will continue to offer unique benefits for other services such as mobile communications and multimedia, but it is up to us to exploit them.

Technology and applications have progressed significantly since the late 1980s; however, the basic principles remain the same. Satellite communication applications extend throughout human activity, both occupational and recreational. Many large companies have built their foundations on TV, data communications, information distribution, maritime communications, and remote monitoring. For many others, satellites have become a hidden asset. In the public sector, satellite applications are extremely effective in emergency situations where terrestrial lines and portable radios are not available or ineffective for a variety of reasons.

[1]Adapted with permission from *The Satellite Communication Applications Handbook,* Bruce Elbert, Artech House, Norwood, MA, 1997.

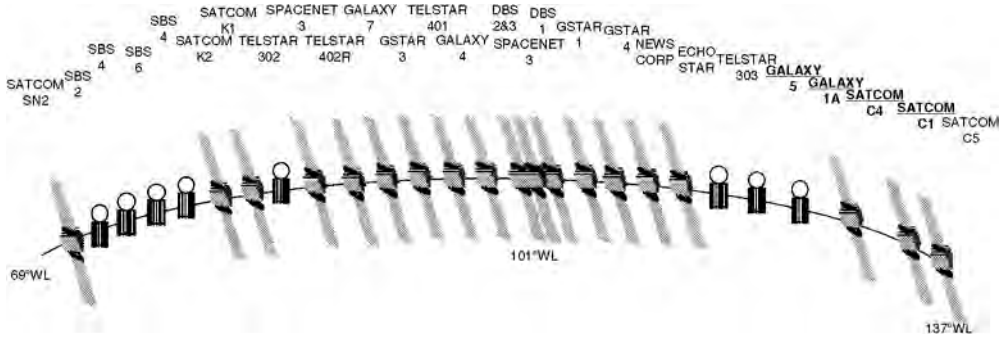


FIGURE 56.1 An illustration of the U.S. GEO orbital arc as of 1996, highlighting the most popular cable TV satellites.

From this experience, those of us who have offered satellite services to large user communities know that the orbit position is paramount to achieving a solid business. Consider as an example the U.S. geostationary orbital service arc shown in Fig. 56. 1. Highlighted are the most popular cable TV satellites: Galaxy I-R, Satcom C-1, Galaxy 5, and Satcom C-3. These four satellites collect revenue comparable to the remaining 25 satellites. Users of these *hot birds* pay a premium for access to the ground infrastructure of cable TV receiving antennas, much like tenants in a premium shopping mall pay to be in an outstanding location and in proximity to the most attractive department stores in the city. In the case of cable TV, access is everything because the ground antenna is, in turn, connected to households where cable services are consumed and paid for. For new satellite operators to get into such an established market often requires them to subsidize users by providing new antennas or paying for repointing. This then reduces the potential profitability as these switching costs must come out of the forthcoming revenues.

Satellite operators, who invest in the satellites and make capacity available to their customers, generally prefer that users own their own Earth stations. This is because installing antennas and associated indoor electronics is costly for satellite service providers. Once working, this investment must be maintained and upgraded to meet evolving needs. The owner/user benefits because: (1) there is complete control of the service facility and (2) advances in microcircuitry and satellite sophistication have reduced the cost and complexity of ownership and operation. A typical small Earth station is no more complex than a cellular telephone or video cassette recorder (VCR). Larger Earth stations such as TV uplinks and international telephone gateways are certainly not a consumer item, but the advantages of control typically motivate leading TV networks and communication companies to take these on.

Several U.S. corporations attempted to introduce DTH service in the mid-1980s at a time when cable TV was still establishing itself. The first entrants experienced great difficulties with the limited performance of low- and medium-power Ku-band satellites, hampering the capacity of the networks and affordability of the home receiving equipment. Europe and Japan had problems of their own creating viable DTH systems with the technologies at hand, choosing first to launch high-power Ku-band satellites with only a few operating channels. It was not until BSkyB and NHK were able to bring attractive programming to the public on their respective satellites that consumers adopted the services.

In the U.S., DTH first emerged in the 1980s through the backyard C-band satellite TV business that allowed consumers to pull existing cable TV programming from cable TV hot birds like Galaxy I and Satcom 3R. This clearly demonstrated the principle that people would vote with their money for a wide range of attractive programming, gaining access to services that were either not available or priced out of reach. Early adopters of these dishes paid a lofty price per installation because the programming was essentially free since the signals were unscrambled. Later, HBO and other cable networks scrambled their programming using the Videocipher 2 system, resulting in a halt to expansion of backyard dishes. This market settled back into the doldrums for several years.

In 1994, Hughes Electronics introduced its DIRECTV™ service through a high-power BSS satellite (labeled DBS-1 in Fig. 56.1). With a combined total of over 150 channels of programming, DIRECTV, Inc., and another service provider, USSB, demonstrated that DTH could be both a consumer product and a viable alternative to cable. An older competing service, Primestar, was first introduced by TCI and other cable operators as a means to serve users who were beyond the reach of their cable systems. This subsequently provided the platform for a new and improved Primestar, rebuilt along lines very similar to DIRECTV. By converting from an analog to a digital system and expanding programming, the Primestar service has seen much greater demand than when it was first introduced in 1992.

Satellite communication applications can establish a solid business for companies that know how to work the details to satisfy customer needs. Another example is the IDB Communications Group. From a very modest start from distributing radio programs by satellite, IDB branched out into every conceivable area of the telecommunication industry. Whether the requirement is to cover a media event in Antarctica or to connect Moscow to the rest of the world with voice and video links, IDB proved its worth. They capitalized on the versatility of satellite communication in the pursuit of a growing company and a demanding base of investors. Sadly, IDB ran into financial difficulties in 1994 and had to merge with LDDS, the fourth largest long distance carrier in the U.S. IDB's early success points out that needs can be met and money is to be made in the field of satellite communication.

56.2 Satellite Network Fundamentals

A commercial satellite communication system is composed of the space segment and the ground segment, illustrated in Fig. 56.2. In this simple representation of a GEO system, the space segment includes the satellites operating in orbit and a tracking, telemetry, and command (TT&C) facility that controls and manages their operation. This part of the system represents a substantial investment, where the satellite operating lifetime is typically of the order of 12 years based on fuel availability. Satellites tend to operate peacefully after they have survived launch and initiation of service. Failures of on-board equipment are generally transparent to users because the satellite operator can activate spare electronic equipment, particularly traveling wave tubes and solid state power amplifiers (the critical elements in the communications subsystem). From a user's perspective, the more important portion of the system is the ground segment, which is the focus of this chapter. It contains Earth stations of various types: large hubs or gateways and much smaller stations that render services directly to individual users.

Every satellite application achieves its effectiveness by building on the strengths of the satellite link. A satellite is capable of performing as a microwave repeater for Earth stations that are located within its coverage area, determined by the altitude of the satellite and the design of its antenna system [Elbert, 1987]. The arrangement of three basic orbit configurations is shown in Fig. 56.3. A GEO satellite can cover nearly one-third of the Earth's surface, with the exception of the polar regions. This includes more than 99% of the world's population and essentially all of its economic activity.

Satellite Orbits and Constellations

The LEO and medium Earth orbit (MEO) approaches require more satellites to achieve this level of coverage. Because of the fact that the satellites move in relation to the surface of the Earth, a full complement of satellites (called a constellation) must be operating to provide continuous, unbroken service. The tradeoff here is that the GEO satellites, being more distant, incur a longer path length to Earth stations, whereas the LEO systems promise short paths not unlike those of terrestrial systems. The path length introduces a propagation delay since radio signals travel at the speed of light. Depending on the nature of the service, the increased delay of MEO and GEO orbits may impose some degradation on quality or throughput. The extent to which this materially affects the acceptability of the service depends on many factors.

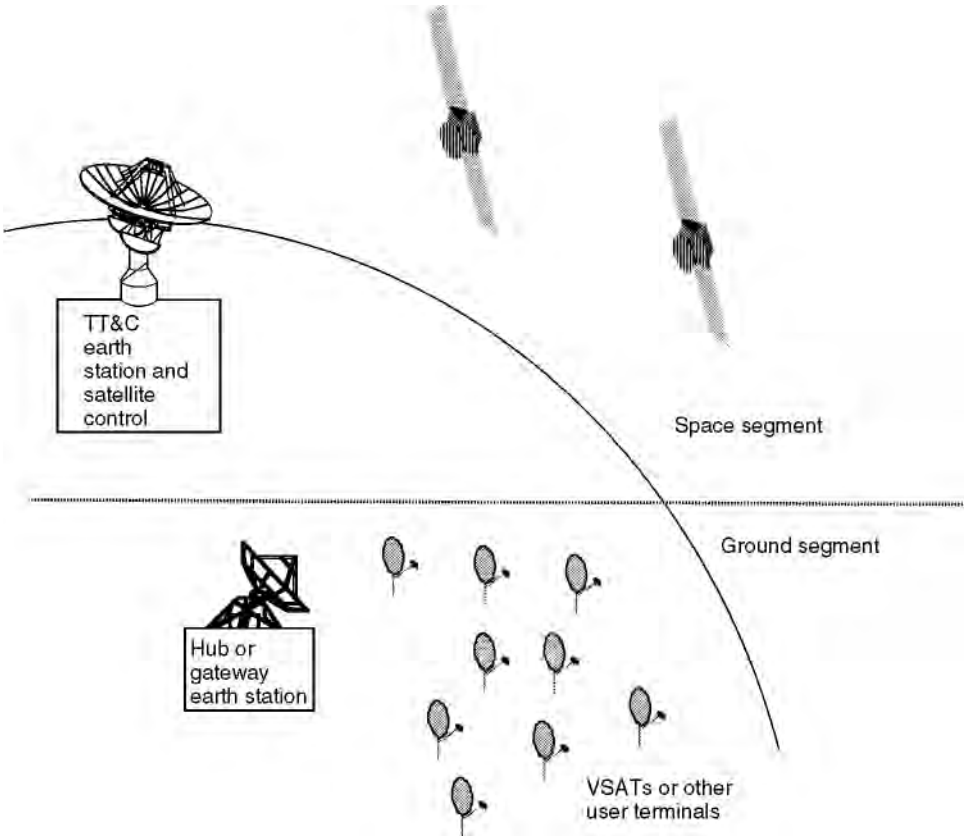


FIGURE 56.2 Elements of a satellite system, including the space and ground segments.

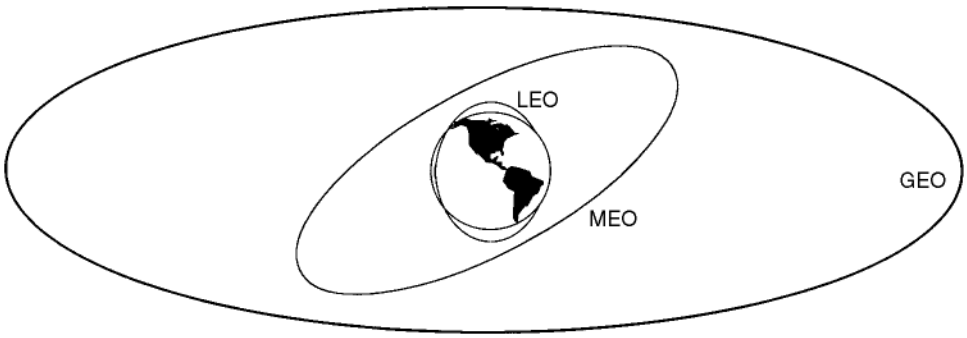


FIGURE 56.3 The three most popular orbits for communication satellites are low Earth orbit (LEO), medium Earth orbit (MEO), and geostationary Earth orbit (GEO). The respective altitude ranges are LEO (500–900 km), MEO (5,000–12,000 km), and GEO (36,000 km). Only one orbit per altitude is illustrated, even though there is a requirement for constellations of LEO and MEO satellites to provide continuous service.

At the time of this writing, the only non-GEO system, Orbcom, has entered commercial service with its first two satellites. The applications of Orbcom include paging and low-speed burst data. Systems capable of providing telephone service, such as Iridium and I-CO Communications, will achieve operation in the near future. On the other hand, GEO satellite systems continue to be launched for literally all services contemplated for satellite, and the ones currently operating have lives (based on fuel reserves for stationkeeping) that will carry them through the year 2010.

Frequency Spectrum

Satellite communication services are provided through GEO satellites operating between 1 and 30 GHz. As microwave frequencies, transmissions to and from the satellite propagate along a line-of-sight path and experience free space loss that increases as the square of the distance. As indicated in Fig. 56.4, the spectrum is allocated in blocks which are designated L-, C-, X-, Ku-, and Ka-bands. The scale along the x axis is logarithmic in order to show all of the satellite bands; however, observe that the bandwidth available for applications increases in real terms as one moves toward the right (i.e., frequencies above 4 GHz). Also, the precise amount of spectrum that is available for services in a given region or country is usually less than Fig. 56.4 indicates. Precise definition of the services and their allocated bandwidths can be found in the current edition of the Radio Regulations, published by the International Telecommunication Union in Geneva, Switzerland.

As a general rule, the lower the band is in frequency, the better the propagation characteristics. This is countered by the second general principle, which is that the higher the band, the more bandwidth that is available. The mobile satellite service (MSS) is allocated to the L and S bands, where there is a greater degree of diffraction bending and penetration of physical media like foliage and non-metallic buildings. This tends to improve service in a mobile environment. Yet, the bandwidth available between 1 and 2.5 GHz, where MSS applications are authorized, must be shared not only among MSS applications, but with all kinds of radio broadcast and point-to-point services as well. Therefore, the competition is keen for this spectrum due to its excellent space and terrestrial propagation characteristics. The rapid rollout of wireless services such as cellular radiotelephone and personal communications services (PCS) will probably conflict with advancing GEO and non-GEO MSS systems.

Fixed satellite service (FSS) refers to applications between fixed Earth stations on the ground. These satellites deliver a wide range of services for video, telephone, and data communication. Wideband services such as DTH and broadband integrated services digital network (ISDN) using asynchronous transfer mode (ATM) can be accommodated at frequencies above 3 GHz, where there is more than 10 times the bandwidth available. Add to this the benefit of using directional ground antennas that effectively multiply the number of orbit positions. Some wideband services have begun their migration from the

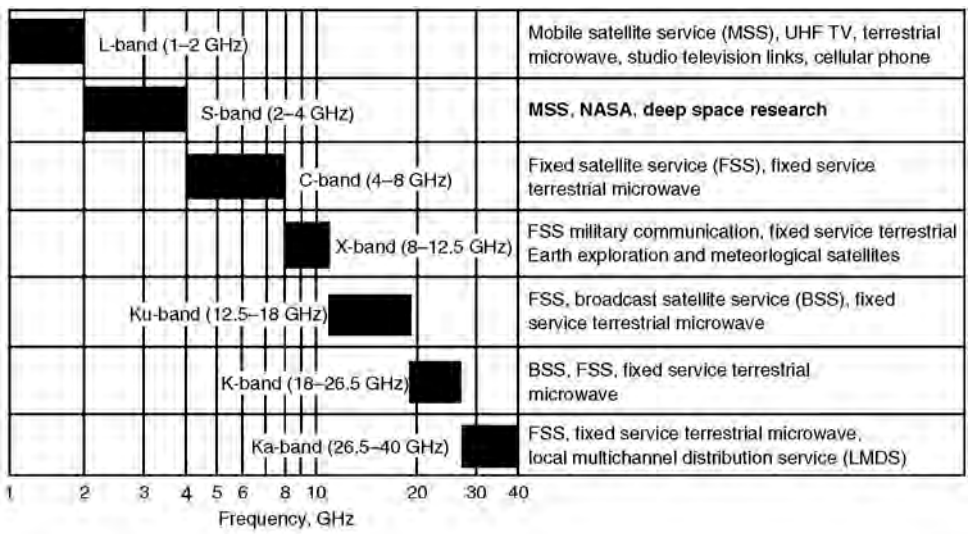


FIGURE 56.4 The general arrangement of the frequency spectrum that is applied to satellite communications and other radiocommunication services. Indicated are the short-hand letter designations along with an explanation of the typical applications. Note that the frequency ranges indicated are the general ranges and do not correspond exactly to the ITU frequency allocations and allotments.

well-established world of C band to Ku band and Ka band. The X-band allocation for FSS has been reserved by governments for their internal uses, notably for military fixed communications.

Higher satellite effective isotropic radiated power (EIRP) used at Ku and Ka bands allows the use of relatively small Earth station antennas. An important factor in the use of these higher bands is the additional attenuation produced by heavy rain. Additional link margin is demanded in cases where the service must achieve an availability of greater than about 99%. The bands below Ku (e.g., C, S, and L) experience relatively minor amounts of rain attenuation and, therefore, can use their power to improve other aspects of service. C band, in particular, should maintain its strength for video distribution to cable systems and TV stations, particularly because of the extensive investment in uplink and downlink antennas and electronic equipment. One needs to identify and evaluate all of the critical factors before making a selection.

Network Topology

Applications are delivered through a network architecture that falls into one of three categories: point-to-point (mesh), point-to-multipoint (broadcast), and multipoint interactive [very small aperture terminal (VSAT)]. Mesh-type networks mirror the telephone network. They allow Earth stations to communicate directly with each other on a one-to-one basis. To make this possible, each Earth station in the network must have sufficient transmit and receive performance to exchange information with its least-effective partner. Generally, all such Earth stations have similar antennas and transmitter systems, so that their network is completely balanced. Links between pairs of stations can be operated on a full-time basis for the transfer of broadband information such as TV or multiplexed voice and data. Alternatively, links can be established only when needed to transfer information, either by user scheduling (reservation system) or on demand (demand assignment system).

By taking advantage of the broadcast capability of a GEO satellite, the point-to-multipoint network supports the distribution of information from a source (the hub Earth station) to a potentially very large number of users of that information (the remote Earth stations). A broadcast of information by the satellite is more efficient than terrestrial arrangements using copper wires, fiber optic cables, or multiple wireless stations. Therefore, any application that uses this basic feature will tend to provide value to the business.

Many applications employ two-way links, which may or may not use the broadcast feature. The application of the VSAT to interactive data communication applications has proven very successful in many lines of business. A hub and spoke network using VSATs competes very effectively with most terrestrial network topologies that are designed to accomplish the same result. This is because the satellite provides the common point of connection for the network, eliminating the requirement for a separate physical link between the hub and each remote point. Other interactive applications can employ point-to-point links to mimic the telephone network, although this tends to be favored for rural and mobile services. The coming generation of Ka-band satellite networks that support very low-cost VSATs are expected to reduce barriers to entry into satellite communication.

Satellite vs. Terrestrial Communication

The degree to which satellite communication is superior to terrestrial alternatives depends on many interrelated factors, many of which are listed in Table 56.1. Although satellite communication will probably never overtake terrestrial telecommunications on a major scale, these strengths can produce very effective niches in the marketplace. Once the satellite operator has placed the satellite into service, a network can easily be installed and managed by a single organization. This is possible on a national or regional basis (including global in the case of non-GEO systems). The satellite delivers the same consistent set of services independently of the ground infrastructure, and the costs of doing this are potentially much lower as well. For the long term, the ability to serve mobile stations and to provide communications instantly are features that will gain strength in a changing world.

TABLE 56.1 The Primary Strengths of Satellite Communications

Feature of Satellite Service	Application
Wide-area coverage	Domestic, regional, global
Wide bandwidth	Up to 1 GHz per coverage
Independent of land-based networks	Does not require connection to terrestrial infrastructure
Rapid installation	Individual sites can be installed and activated in one day for VSAT or two months of major hub
Low cost per added site	Depends on type of service; can be as low as $600 for DTH
Uniform service characteristics	Determined by coverage and type of transmission system
Total service from a single provider	By the satellite operator or a separate organization that leases transponder capacity
Mobile Communication	Requires line-of-sight path over the coverage area

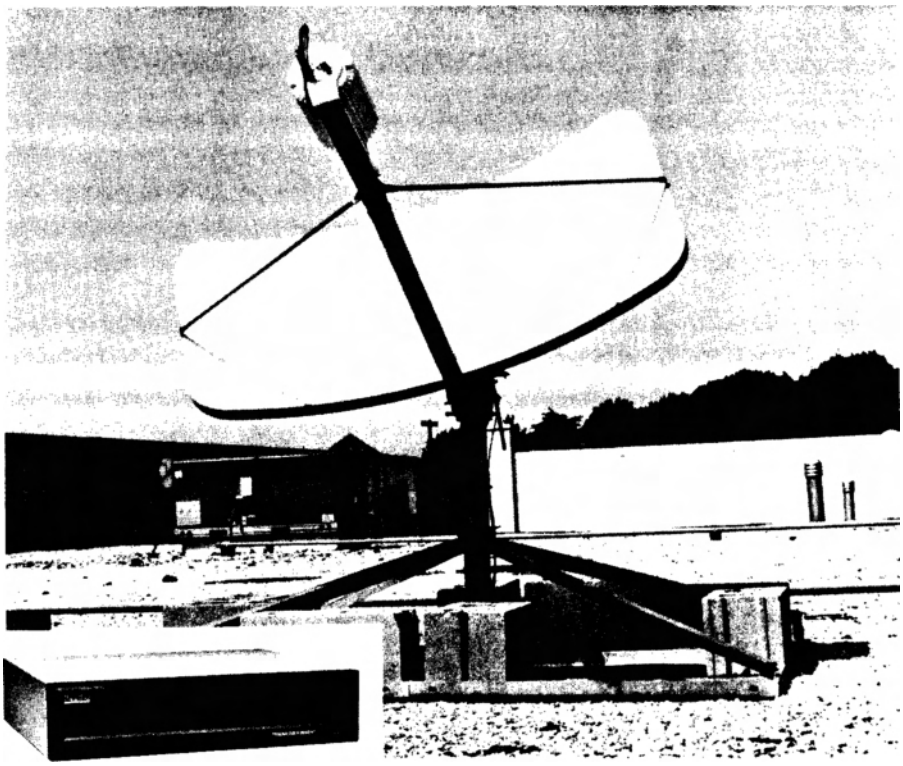


FIGURE 56.5 An example of a small customer premise VSAT used in data communication. (Photograph courtesy of Hughes Network Systems.)

Satellite and Earth Station Designs

Originally, Earth stations were large, expensive, and located in rural areas so as not to interfere with terrestrial microwave systems that operate in the same frequency bands. These monsters had to use wideband terrestrial links to reach the closest city. Current emphasis is on customer premise Earth stations: simple, reliable, and low cost. An example of a Ku-band VSAT is illustrated in Fig. 56.5. Home receiving systems for DTH service are also low in cost and quite inconspicuous. Expectations are that the coming generation of Ka-band VSATs will reduce the cost and thereby increase their use in smaller businesses and,

FIGURE 56.6 Measat, first launched in 1996, is a small satellite of the Hughes HS-376 class. With 1100 W of payload power, it provides both low-power C-band and medium-power Ku-band transponders to Malaysia and surrounding countries in the Asia–Pacific region. (Photograph courtesy of Hughes Space and Communications Company.)

ultimately, the home. There is the prospect of satellites delivering improved Internet access ahead of the terrestrial networks, which have yet to deliver on the promise of ubiquitous wide digital bandwidth.

The other side of the coin is that as terminals have shrunk in size, satellites have grown in power and sophistication. There are three general classes of satellites used in commercial service, each designed for a particular mission and capital budget. Smaller satellites provide a basic number of transponders, usually in a single-frequency band. One such satellite can be launched by the old and reliable Delta II rocket or dual-launched on the Ariane 4 or 5. Satellite operators in the U.S., Canada, Indonesia, and China have established themselves in business through this class of satellite. The Measat satellite, illustrated in Fig. 56.6,

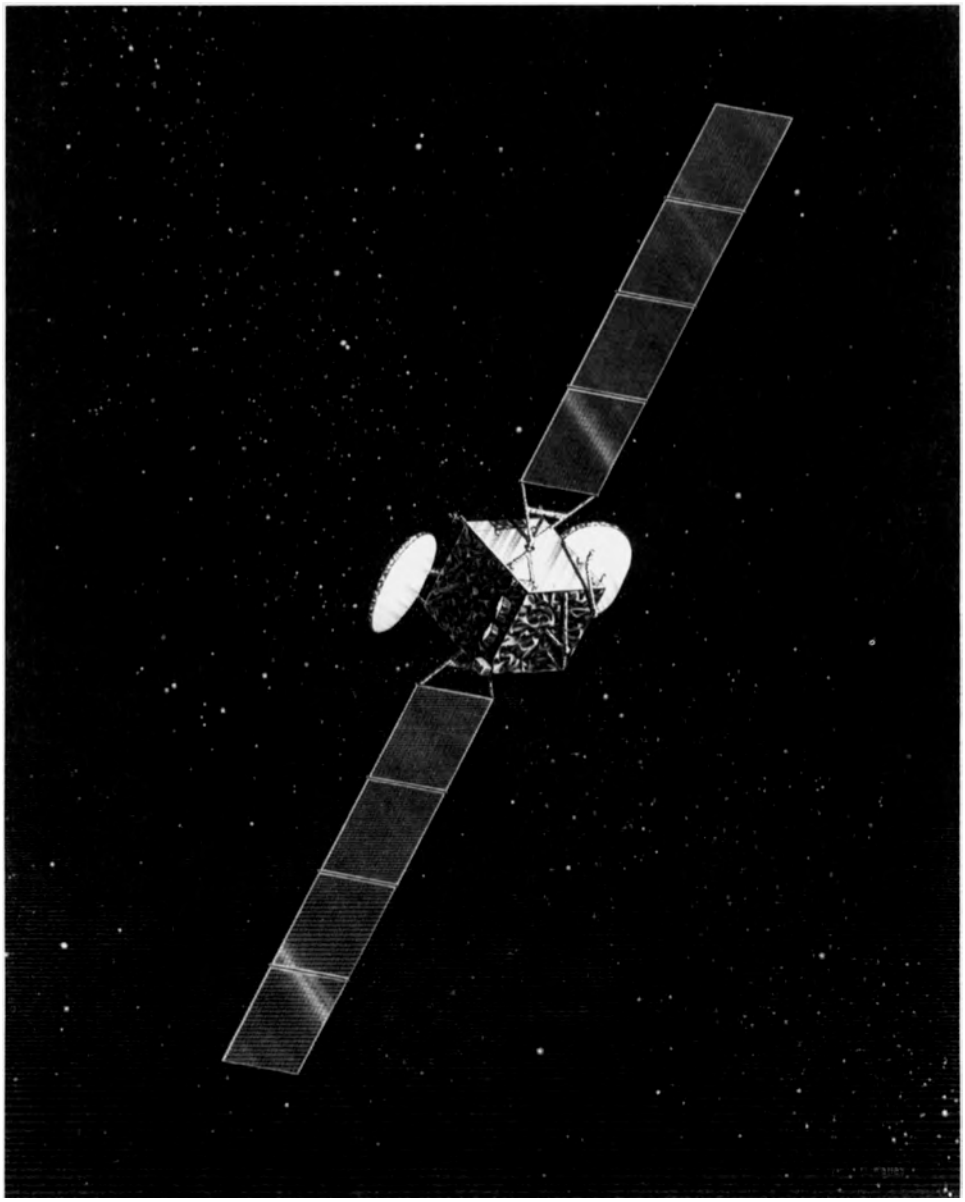


FIGURE 56.7 Galaxy VII, launched in 1993, is a medium-power Hughes HS-601 satellite that provides 24 low-power C-band and 24 medium-power Ku-band transponders to the U.S. market. It is used by cable TV and over-the-air broadcasting networks to distributed programming to affiliates. (Photograph courtesy of Hughes Space and Communications Company.)

is an example of this class of vehicle which delivers a total of 18 transponders to Malaysia and neighboring countries in Asia. Introduction of mobile service in the LEO will involve satellites of this class as well.

Moving up to the middle range of spacecraft, we find designs capable of operating in two frequency bands simultaneously. Galaxy VII, shown in Fig. 56.7, provides 24 C-band and 24 Ku-band transponders to the U.S. market. A dual payload of this type increases capacity and decreases the cost per transponder. First generation DTH and MSS applications are also provided by these spacecraft. Because of the increased weight of this size of vehicle, launch service must be provided by larger rockets such as the Atlas 2,

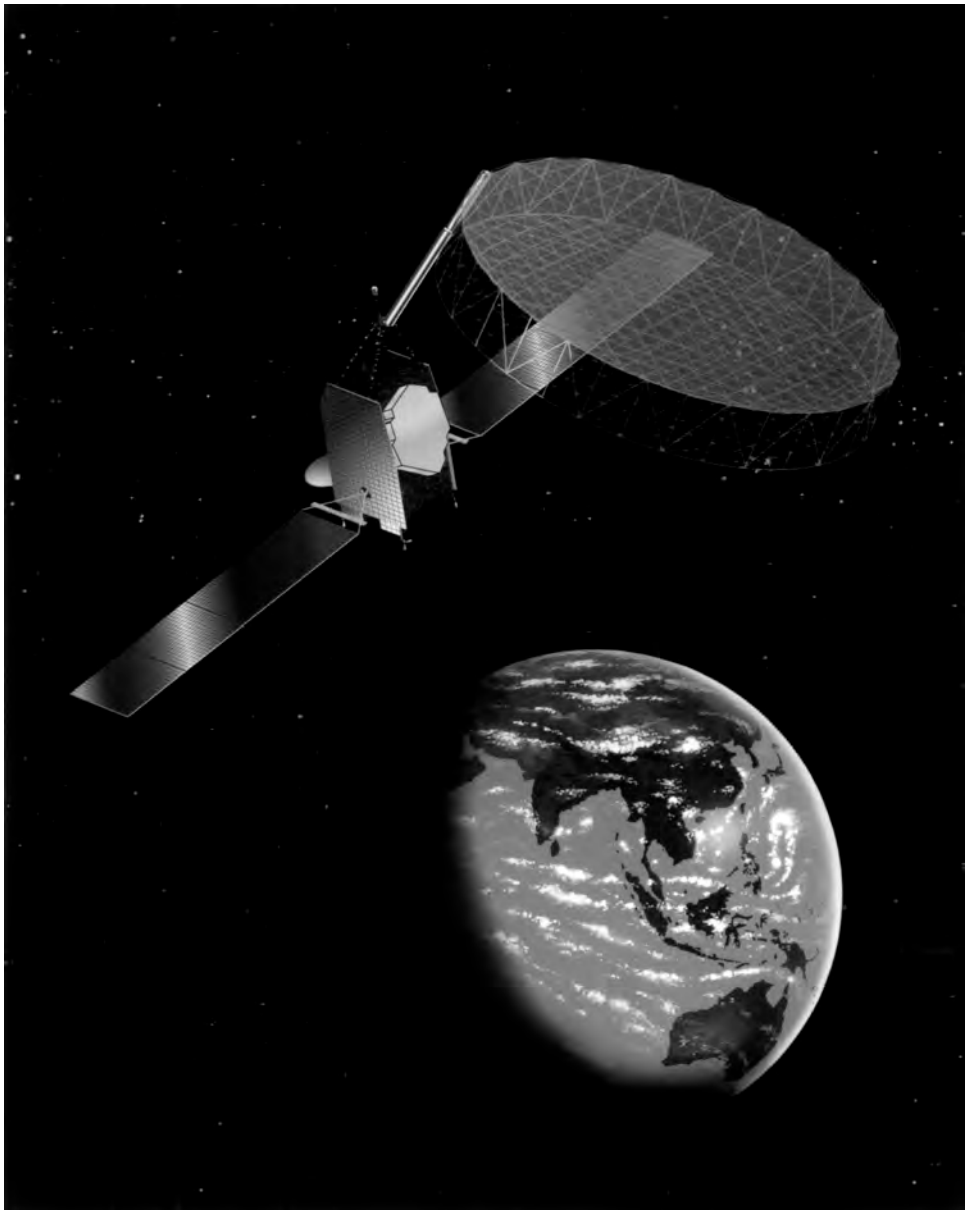


FIGURE 56.8 APMT is a regional GEO mobile communications satellite of the largest HS-601 HP class produced by Hughes. It serves the Asia Pacific region and employs a 12-m reflector to serve mobile and fixed users, many of which will use handheld user terminals. (Photograph courtesy of Hughes Space and Communications Company.)

Ariane 4 or 5, the Russian Proton, or the Chinese Long March 2E or 3B (an improved Delta capable of launching such spacecraft was under development at the time of this writing).

Finally, the largest satellites serve specialized markets where the highest possible power or payload weight are required. The generation of GEO mobile satellites that are capable of serving handheld phones fall into this classification. An example of one of these satellites, APMT, is shown with its 12-m antenna deployed in Fig. 56.8. Also, the trends to use the smallest possible DTH home receiving antenna and to cover the largest service area combine to demand the largest possible spacecraft. The total payload power

of such satellites reaches 8 kW, which is roughly eight times that of Measat. Designs are on drawing boards at the time of this writing for satellites that can support payload powers of up to 16 kW.

These general principles lead to a certain set of applications that serve telecommunication users. Satellite manufacture is a strong and viable business, as three American and a like number of European firms compete in the global market. Earth stations are produced by specialists in the U.S., Canada, Europe, Japan, and Korea. However, world-class consumer electronics and telecommunication equipment firms are lining up for the opportunity to address the megamarkets in DTH and MSS systems. In the next section, we review the most popular applications.

56.3 Satellite Application Types

Applications in satellite communication have evolved over the years to adapt to competitive markets. Evolutionary development, described in Elbert [1987], is a natural facet of the technology because satellite communication is extremely versatile. This is important to its extension to new applications in the coming century.

Television and Video Services

Commercial TV is the largest segment of the entertainment industry; it also represents the most financially rewarding user group to satellite operators. It may have taken 10 or more years for the leading networks in the U.S. and Europe to adopt satellites for distribution of their signals, but by 1985, it was the mainstay. Prior to this date, pioneering efforts in Indonesia and India allowed those countries to introduce nationwide TV distribution via satellite even before the U.S. made the conversion from terrestrial microwave. European TV providers pooled their resources through the European Broadcasting Union (EBU) and the EUTELSAT regional satellite system. Very quickly, the leading nations of Asia and Latin America adopted satellite TV delivery, rapidly expanding this popular medium to global levels.

There are four fundamental ways that GEO satellites are used to serve the TV industry:

- Point-to-point transmission of video information from a program source to the studio (also called *backhaul*)
- Point-to-multipoint distribution of TV network programming from the studio to the local broadcast station
- Point-to-multipoint distribution of cable TV programming from the studio to the local cable TV network
- Point-to-multipoint distribution of TV network and/or cable TV programming from the studio directly to the subscriber (i.e., DTH)

The first technique is a part of the three that follow, since programming material must be acquired from a variety of sources. A simplified view of the latter three is given in Fig. 56.9. In each case, the video channels are transmitted to the satellite by an uplink Earth station with the capability to obtain and assemble the programming. This can involve fixed antennas to backhaul programming from other satellites, tape play machines and video servers to play back prerecorded material, and fiber optic connections to distant studios. DTH systems with their smaller antennas require higher satellite power than cable TV and TV distribution. Hence, a common satellite for all three classes is rarely the case.

Over-the-Air TV Broadcasting

Video distribution by satellite is now standard for TV broadcasting in the very high-frequency (VHF) and ultra high-frequency (UHF) bands, which use local TV transmitters to cover a city or area. The satellite is used to carry the network signal from a central studio to multiple receive Earth stations, each connected to a local TV transmitter. This has been called TV distribution or TV rebroadcast. When equipped

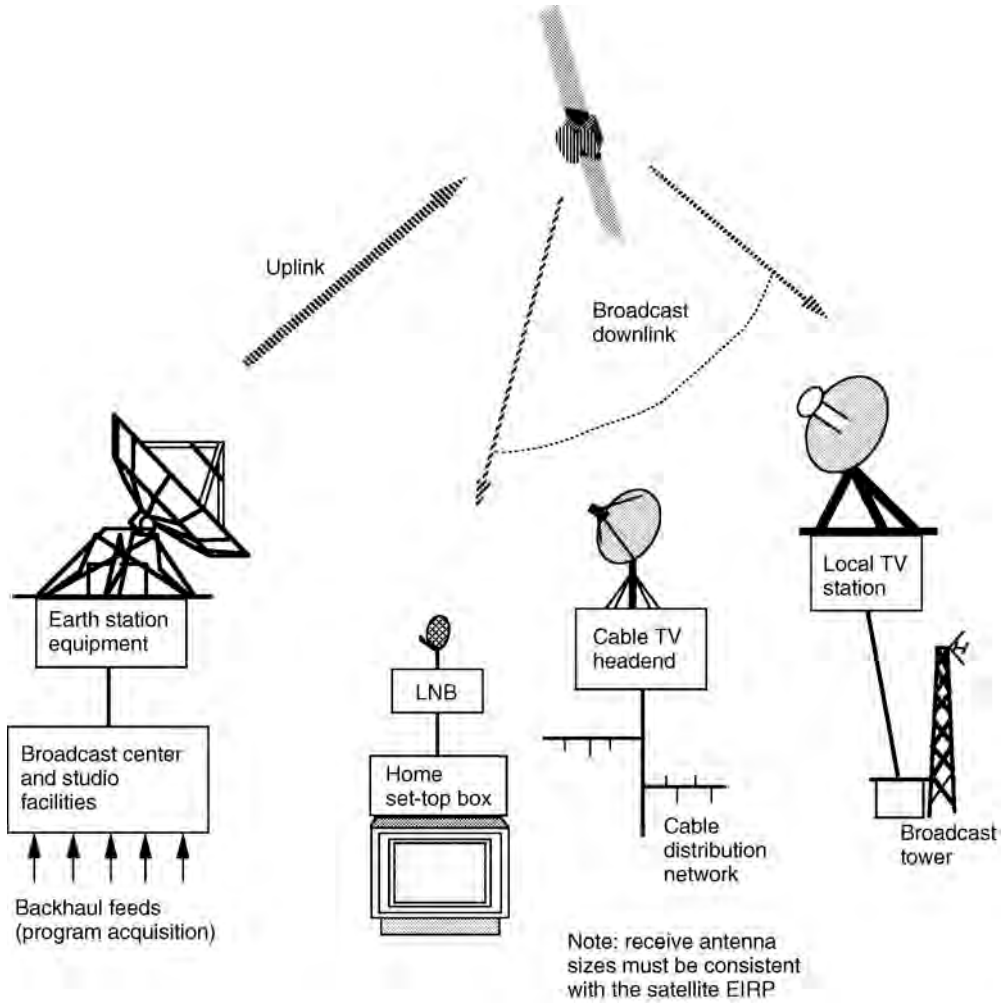


FIGURE 56.9 Framework for satellite communication applications in local broadcasting, cable TV, and direct-to-home satellite broadcast.

with transmit uplink equipment, the remote Earth station can also transmit a signal back to the central studio to allow the station to originate programming for the entire network. U.S. TV networks such as CBS employ these reverse point-to-point links for on-location news reports. The remote TV uplink provides a transmission point for local sporting and entertainment events in the same city. This is popular in the U.S., for example, to allow baseball and football fans to see their home team play an away-from-home game in a remote city. More recently, TV networks employ fiber optic transmission services from a national fiber network, but satellites continue as the most flexible routing mechanism.

Revenue for local broadcast operations are available from two potential sources: advertisers and public taxes. From its beginnings in the U.S., TV provided an excellent medium to influence consumers on their purchase behavior. In exchange for watching commercials for soap, airlines, and automobiles, the consumer is entertained for virtually nothing. This has produced a large industry in the U.S. as stations address local advertisers and the networks promote nationwide advertising.

An alternative approach was taken in many European countries and Japan, where government operated networks were the first to appear. In this case, the consumer is taxed on each TV set in operation. These revenues are then used to operate the network and to produce the programming. The BBC in the U.K. and NHK in Japan are powerhouses in terms of their programming efforts and broadcast resources.

However, with the rapid introduction of truly commercial networks, cable TV, and DTH, these tax-supported networks are experiencing funding difficulties.

Public TV in the U.S. developed after commercial TV was well established. Originally called educational TV, this service existed in a fragmented way until a nonprofit organization called the Public Broadcasting Service (PBS) began serving the nation by satellite in 1978. The individual stations are supported by the local communities through various types of donations. Some are attached to universities; others get their support by donations from individuals and corporations. PBS itself acquires programming from the member stations and from outside sources such as the BBC. Programs are distributed to the members using satellite transponders purchased by the U.S. government. Therefore, it must compete with other government agencies for congressional support. PBS programming is arguably of better quality than some of the popular shows on the commercial networks. Even though PBS addresses a relatively narrow segment, its markets are now under attack by even more targeted cable TV networks such as A&E, the Discovery Channel, the Learning Channel, and a more recent entry, the History Channel. All of these competitors built their businesses on satellite delivery to cable systems and DTH subscribers.

Cable TV

Begun as a way to improve local reception in the fringe areas and thinly populated sections of the U.S., cable TV has established itself as the dominant entry point of TV programming. Its popularity grew after pioneering cable TV networks such as HBO and Turner Broadcasting (which are not available over the air) were added by cable operators seeking more subscribers. These organizations, joined by Viacom, Disney, ABC, and NBC, used C-band satellite transmission to bypass the conventional TV networks to deliver programming to cable head ends and, thence, to the home. By 1980, 40% of urban-U.S. homes were using cable to receive local TV stations (because cable provided a more reliable signal); at the same time, the first nationwide cable networks were included as a sweetener and additional revenue source. During the 1980s, cable TV became an eight billion dollar industry and the prototype for this medium in Europe, Latin America, and the developed parts of Asia.

Cable TV networks offer programming as a subscriber service to be paid for on a monthly basis, or as an almost free service such as commercial TV broadcasting. HBO and the Disney Channel are examples of premium (pay) services, whereas the Discovery Channel and TBS are examples of commercial channels that receive most of their revenue from advertisers. The leading premium channels in North America and Europe are immensely successful in financial terms, but the business has yet to be firmly established in economies with low-income levels.

Cable TV became the first to offer a wide range of programming options that are under the direct control of the service provider. This ensures that subscription fees and service charges can be collected from subscribers. If the fees are not paid, the service is terminated. Wireless cable, a contradiction in terms but nevertheless a viable alternative to wired cable, is gaining popularity in Latin America and Asia. The principle here is that the cable channels are broadcast locally in the microwave band to small dishes located at subscribers. Just as in the case of DTH, wireless cable depends on some form of conditional access control that allows the operator to electronically disconnect a nonpaying user.

Direct-to-Home Satellite Broadcasting

The last step in the evolution of the satellite TV network is DTH. After a number of ill-fated ventures during the early 1980s, DTH appears to be ready to take off. B Sky B in the U.K., NHK in Japan, DIRECTV in the U.S., and STAR TV in Asia are now established businesses, with other broadcasters following suit. Through its wide area broadcast capability, a GEO satellite is uniquely situated to deliver the same signal throughout a country or region at an attractive cost per user. The particular economics of this delivery depend on the following factors:

- The size of the receiving antennas: Smaller antennas are easier to install and maintain and are cheaper to purchase in the first place. They are also less noticeable (something which is desirable

in some cultures but not necessarily in all others). However, the diameter of the receiving antenna is inversely proportional to the square root of the effective radiated power of the satellite.

- The design of the equipment: It is simple to install and operate (this author's RCA™ brand Digital Satellite System (DSS) installation, needed to receive DIRECTV, took only 2 h: 105 min to run the cables and 15 min to install and point the dish).
- Several users can share the same antenna: This is sensible if the antenna is relatively expensive, say in excess of $1000; otherwise, users can afford their own. A separate receiver is needed for each independent TV watcher (i.e., for private channel surfing).
- The number of transponders that can be accessed through each antenna: The trend is now toward locating more than one satellite in the same orbit position, to increase the available number of transponders. The more channels that are available at the same slot, the more programming choices the user will experience.
- The number of TV channels that can be carried by each transponder: Capacity is multiplied through digital compression techniques such as Moving Pictures Expert Group (MPEG) 2. This multiplies the quantity of simultaneous programs by five or more.

The ideal satellite video network delivers its programming to the smallest practical antenna on the ground, has a large number of channels available (greater than 100), and permits some means for users to interact with the source of programming. A simple connection to the public-switched telephone network (PSTN) allows services to be ordered directly by the subscriber. Thus, the system emulates a direct point-to-point connection between user and programming supplier, something that broadband fiber optic networks promise but have yet to deliver on a wide-scale basis.

Trends in Satellite Voice Communications

Voice communications are fundamentally based on the interaction between two people. It was recognized very early in the development of satellite networks that the propagation delay imposed by the GEO tends to degrade the quality of interactive voice communications, at least for some percentage of the population. However, voice communications represent a significant satellite application due to the other advantages of the medium. For example, many developing countries and lightly inhabited regions of developed countries continue to use satellite links as an integral part of the voice network infrastructure. Furthermore, one area where satellite links are essential for voice communications is the mobile field.

The public-switched telephone network within and between countries is primarily based on the requirements of voice communications, representing something in the range of 60–70% of all interactive traffic. The remainder consists of facsimile (fax) transmissions, low- and medium-speed data (both for private networks and access to public network services such as the Internet), and various systems for monitoring and controlling remote facilities. The principal benefit of the PSTN is that it is truly universal. If you can do your business within the limits of 3000 Hz of bandwidth and can tolerate the time needed to establish a connection through its dialup facility, the PSTN is your best bet.

Propagation delay has become an issue when competitively priced digital fiber optic networks are introduced. Prior to 1985 in the U.S., AT&T, MCI, and others were using a significant amount of analog telephone channels both on terrestrial and satellite links. An aggressive competitor in the form of U.S. Sprint invested in an all-digital network that employed fiber optic transmission. Sprint expanded their network without microwave or satellite links and introduced an all-digital service at a time when competition in long distance was heating up. Their advertising claimed that calls over their network were so quiet you could "hear a pin drop." This campaign was so successful that both MCI and AT&T quickly shifted their calls to fiber, resulting in rapid turn down of both satellite voice channels and analog microwave systems.

Satellite-based telephone service is acceptable in cases where it is the best alternative. For example, wired networks cannot provide mobile communications on the oceans and in very remote areas. Here, digital voice over satellite is widely accepted. A particular application, called mesh telephony, allows

separate VSATs to operate as independent connection points for users or small village. Gateways in major cities connect calls between these remote VSATs and subscribers on the PSTN. Major telephone Earth stations are part of the international telecommunication infrastructure, connecting to and between countries that are not yet served by modern fiber optic systems. The basic voice transmission quality is the same on satellite or terrestrial fiber because it is determined by the process of converting from analog to digital format. All long-distance transmission systems, satellite and cable, require echo cancellers to maintain quality. GEO satellite links experience a delay of about 260 ms and so are different from their terrestrial counterparts. However, they can be quite acceptable given that the application favors this medium for economic or practical reasons.

The economics of satellite voice communications are substantially different from those of the terrestrial PSTN, even given the use of digital technology on both approaches. With low-cost VSAT technology and high-powered satellites at Ku and Ka bands, satellite voice is the cheapest and quickest way to reach remote areas where terrestrial facilities are not available. It will be more attractive to install a VSAT than to extend a fiber optic cable over a distance greater than several hundred meters. A critical variable in this case is the cost of the VSAT, which is expected to drop from the $10,000 level to perhaps as low as $1500 before 2000.

Fiber optic transmission is not the only terrestrial technology that can address the voice communication needs of subscribers. There is now a new class of terrestrial wireless system called fixed cellular. Low-cost cordless phones or simple radio terminals are placed in homes or offices, providing access to the PSTN through a central base station. The base stations are concentrating points for traffic and can be connected to the PSTN by fiber or even satellite links. The cost of the base station and network control is kept low by not incorporating the automatic hands-off feature of cellular mobile radio. Instead, user terminals of different types make the connection through the closest base station, which remains in the same operating mode throughout the call. Fixed cellular can grow in cost beyond VSATs as more cells are added.

New classes of public network services are expected in coming years, under the general category of broadband communications. The underlying technology is asynchronous transfer mode, a flexible high-speed packet-switched architecture that integrates all forms of communications. ATM services can be delivered through fiber optic bandwidths and advanced digital switching systems and include the following, among others:

- High-speed data on demand (384 kb/s to 155 Mb/s, and greater)
- Multichannel voice
- Video services on demand
- High resolution fax (300 dots per inch and better)
- Integrated voice/data/video for enhanced Internet services

Because of the high cost of upgrading the terrestrial telephone plant to provide ATM services on a widespread basis, many of these services will not appear in many places for some time. However, they represent the capability of the coming generation of public networks being implemented around the globe.

Satellite communication currently addresses these services as well and, in fact, has been the only means available to provide them on a consistent basis. In private network applications, satellite communication has a big advantage for delivering broadband services to a large quantity of sites. Fiber optic networks are attractive for intra- and intercity public networks and can offer broadband point-to-point transmission, which is potentially lower in cost per user. Yet, this is easier to do with a satellite because it provides a common traffic concentration point in the sky (a principle of traffic engineering). The bandwidth is used more effectively and, therefore, the network can carry more telephone conversations and generate more revenue.

Satellite networks are very expandable because all points are independent and local terrain does not influence the performance of link. Consider the example of the largest German bank, Deutsche Bank, which needed to offer banking services in the new states of the former East Germany. The telecom

infrastructure in East Germany in 1990, although the best in eastern Europe, was very backward by western European standards. Deutsche Bank installed medium-sized Earth stations at new bank locations and was then able to offer banking services that were identical to those of their existing branches in the west. Another excellent example is Wal*Mart Department Stores, reviewed later in this chapter.

Digital Information Services

Satellite networks are able to meet a wide variety of data communication needs of businesses and noncommercial users such as government agencies. The wide-area coverage feature combined with the ability to deliver relatively wide bandwidths with a consistent level of service make satellite links attractive in the developing world as well as in the geographically larger developed countries and regions.

The data that is contained within the satellite transmission can take many forms over a wide range of digital capacities. The standard 36-MHz transponder, familiar to users of C-band worldwide, can easily transfer over 60 Mb/s, suitable for supercomputer applications and multimedia. Most applications do not need this type of bandwidth; therefore, this maximum is often divided up among multiple users who employ a multiple access system of some type. In fact, the multiple access techniques used on the satellite mirror the approaches used in local-area networks (LANs) and wide-area networks (WANs) over terrestrial links. As in any multiple access scheme, the usable capacity decreases as the number of independent users increases. Satellite data networks employing VSATs offer an alternative to terrestrial networks composed of fiber optics and microwave radio. There is even a synergy between VSATs and the various forms of terrestrial networks, as both can multiply the effectiveness of their counterpart.

Some of the must successful users of VSATs are familiar names in North American consumer markets. Wal*Mart, the largest U.S. department store chain, was an early adopter of the technology and has pushed its competitors to use VSATs in much the same manner that they pioneered. Chevron Oil also was first among the gasoline retailers to install VSATs at all of their company-owned filling stations to speed customer service and gain a better systemwide understanding of purchasing trends.

Whereas voice networks are very standardized, data networks cover an almost infinite range of needs, requirements, and implementations. In business, information technology (IT) functions are often an element of competitive strategy [Elbert, 1992]. In other words, a company that can use information and communications more effectively than its competitors could enjoy a stronger position in the ultimate market for its products or services. A data communication system is really only one element of an architecture that is intended to perform business automation functions.

A given IT application using client/server computing networks or broadband multimedia will demand efficient transfer of data among various users. Satellite communication introduces a relatively large propagation delay, but this is only one factor in the overall response time and throughput. These include: the input data rate (also in bits per second), the propagation delay through the link, the processing and queuing delay in data communication equipment, any contention for a terrestrial data line, and the final processing time in the end computer system (client or server). This is shown in Fig. 56.10. Each of these contributors must be considered carefully when comparing terrestrial and satellite links.

Propagation delay from a satellite link could reduce the throughput if there is significant interaction between the two ends of the link. The worst-case condition occurs where the end devices need to exchange control information to establish a connection or confirm receipt. Modern data communication standards, such as transmission control protocol/internet protocol (TCP/IP) and systems network architecture/synchronous data link control (SNA/SDLC), guard against the loss of throughput by only requesting retransmission of blocks of data which have errors detected in them by the receiver. However, to optimize throughput, there is a need to test and tune the circuit for the expected delay, type of information, and error rate.

The main deterrent to wider usage of digital satellite links is the cost of the terminal equipment, which today is much higher than that of a personal computer (PC). True low-cost VSATs have yet to appear at the time of this writing. Once the $2000 U.S. barrier is broken, many residential and small business users will be able to afford to purchase their own systems. This is not unlike the situation with backyard TVRO

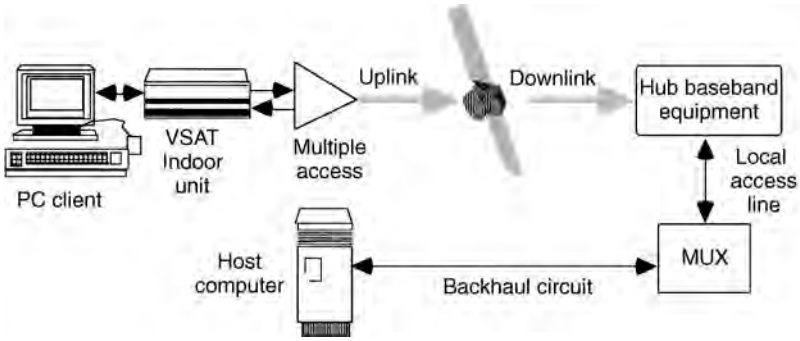


FIGURE 56.10 Contributors to the throughput and end-to-end delay of a data communication link involving a VSAT and a hub.

dishes, which, during the 1980s, were selling above $2000 U.S. Through innovations such as DIRECTV, Inc.'s digital satellite system, which sells for around $600 U.S., the home dish has become as attainable as the VCR. This also has benefits for data applications. If the bulk of the data transfer is outbound only (so that the user installation is receive only from the satellite), the DSS type of installation easily can deliver over 20 Mb/s of downlink capacity. A key example is the DirectPC™ service from Hughes Network Systems that provides turbospeed Internet access service in the U.S.

Mobile and Personal Communications

The world has experienced an explosion of demand for mobile telephone and data communications. The basis for this is the technology of cellular radio systems, which allow handheld mobile phones to connect to the public network as if they had access by wire. Availability of cellular radio at a reasonable cost has allowed a much larger subscriber base to develop than was possible during earlier generations of mobile phone technology. Now, we are prepared for the advance of cellular like systems that connect users directly to orbiting satellites, thus eliminating the terrestrial base station (a much smaller quantity of gateway Earth stations are still needed to connect mobile calls to PSTN subscribers).

Satellite-based mobile radio was first introduced to the commercial market in the late 1970s in the form of maritime mobile satellite service (MMSS) by COMSAT. Within a 10-year period, MMSS overtook the radio telegraph as the primary means of allowing ships to communicate. In the 1990s, ships ceased carrying conventional radio stations, which have been completely replaced by MMSS terminals. Today, the International Maritime Satellite (Inmarsat) system serves commercial ships and has been extended to commercial airlines as well.

Like INTELSAT, Inmarsat serves the needs of its owners and their respective clients. Within the last five years, Inmarsat introduced land-based communications in a service called land mobile satellite service (LMSS). The first LMSS terminals are portable so that they can be set up quickly from a stationary location.

In 1994, Optus Communications of Australia began providing the Mobilesat LMSS service to its domestic market. Mobilesat follows the model of the second generation of LMSS phones, allowing subscribers to communicate from moving vehicles. Optus is the pioneer in this type of system, having directed the development of the ground infrastructure and mobile terminals. The Optus B series of satellites includes the necessary L-band payload for Mobilesat and also carries Ku-band transponders to connect to gateways to the PSTN. The American Mobile Satellite Corp (AMSC) and Telesat Mobile Inc. extended this concept to the North American market in 1996. In this case, a satellite is dedicated to providing mobile telephone services.

A new generation of LMSS is expected before the year 2000, capable of supporting handheld phones similar to those used in cellular. In fact, the promoters of this advanced class of satellite network show

models of modified cellular phones to promote their concept. The two basic approaches being taken are either the use of GEO, with satellites that serve specific countries or regions, or the LEO or MEO approach, where an orbiting constellation of satellites covers the entire globe. Figure 56.3 illustrates only one orbit with a single LEO satellite, whereas in the typical LEO constellation, the globe is covered by multiple orbits, each containing several satellites. The number of required orbits and satellites decreases as the orbit altitude is increased. The minimum number of three satellites to provide global coverage (except the polar regions) is obtained at GEO. The advantage of the GEO approach is its lower initial cost and possibly lower risk, whereas the advantage of the LEO approach is the lower propagation delay and ability to connect users to each other anywhere in the world.

Many of these wireless networks have value to business as a vehicle for extensive data communications and paging services. Mobile data services have been around in the U.S. for over a decade, first introduced for private purposes. Federal Express equips their drivers with radio-based data communications terminals, and today a majority of competing delivery companies have also adopted the technology. An explosion may occur before the year 2000 due to the creation of several satellite mobile data networks in the U.S. and other countries. There is now the prospect of providing a public data network service to all users who wish to access information databases and other facilities while on the move. This is something that the cellular network has had difficulty demonstrating, due to technical issues of maintaining a solid circuit connection during the entire data call.

All together, the MSS is poised to become a substantial contributor to the future of the satellite industry. Spacecraft manufacturers are developing technical designs which even a few years ago looked like science fiction (see Fig. 56.7). Once again, today's science fiction becomes tomorrow's science fact.

References

Adamson, S. et al. 1995. Advanced satellite applications—potential markets. Booz Allen & Hamilton, Noyes Data Corp., Park Ridge, NJ.

Elbert, B.R. 1987. *Introduction to Satellite Communication,* Artech House, Norwood, MA.

Elbert, B.R. 1992. *Networking Strategies for Information Technology,* Artech House, Norwood, MA.

Elbert, B.R. and Martyna, B. 1994. *Client/Server Computing—Architecture, Applications and Distributed Systems Management,* Artech House, Norwood, MA.

Elbert, B.R. 1997. *The Satellite Communication Applications Handbook,* Artech House, Norwood, MA.

Macario, R.C.V. 1991. On behalf of the Institution of Electrical Engineers. In *Personal and Mobile Radio Systems,* Peter Peregrims, Ltd., London, U.K.

Morgan, W.L. and Gordon, G.D. 1989. *Communications Satellite Handbook,* John Wiley & Sons, New York.

Schwartz, M. 1987. *Telecommunication Networks,* Addison–Wesley, Reading, MA, p. 522.

57

Satellite Systems

Robert L. Douglas
Harding University

57.1 Introduction

Communication satellites in **geostationary** orbits have had a tremendous impact on our lives by providing almost instantaneous worldwide electronic communications. The unique capabilities of communication systems that use geostationary satellites have revolutionized the design of telecommunication and television systems.

The purpose of this chapter is to provide an explanation of the specialized techniques that are used to communicate electronically using satellite repeaters or **transponders**. This chapter will be useful to anyone interested in learning about satellite communication techniques. Engineers and technicians, who are already familiar with communication concepts, can use this chapter to learn about the unique characteristics of communicating by satellite. Teachers and students will find this chapter to be an excellent reference in communication concepts and techniques. Managers and nontechnical people can use this chapter as a reference to obtain an understanding of the concepts and technical terms used by communication engineers. A glossary of satellite communication terms used in this chapter is provided in the Defining Terms section at the end of the chapter.

A communication satellite, often called a *COMSAT*, is a spacecraft that receives electrical signals from a transmitter on the Earth, amplifies these signals, changes the carrier frequencies, and then retransmits the amplified signals back to receivers on the Earth [Douglas, 1988]. Since communication satellites simply amplify and retransmit signals, they are often called repeaters or relays. Communication satellites are usually placed in geostationary Earth orbits, which allows the satellites to appear stationary to the transmitters and receivers on the Earth. Placing communication satellites in this special orbit prevents problems caused by the Doppler shift. One **Earth station** transmitter and one satellite can provide signals to receivers in an area that covers almost one-third of the Earth's surface. Satellites have revolutionized electronic communications since the first **commercial satellite** was launched in 1965. There are now more than 200 communication satellites in orbit around the world, and new satellites are launched regularly. Communication satellites have made it possible to routinely receive live television coverage of important news and sporting events from anywhere in the world. Satellites have also improved worldwide telephone communications, navigation, and weather information.

Arthur C. Clarke suggested the amazingly simple concept of communicating by satellites in stationary orbit in 1945. He called the satellite repeaters "extra-terrestrial relays" [Clarke, 1945]. At that time, television was being developed commercially, and Clarke maintained that it would be much cheaper to develop and use communication satellites than to build a ground-based (**terrestrial**) network of transmitters and receivers with repeaters every 30–50 mi to amplify and retransmit the signals. He was correct, of course, but the technology needed to build the satellites and place them in geostationary orbit was not developed until the 1960s. By that time, many countries had developed extensive terrestrial systems for the distribution of television and other wideband communication signals. Communication satellites are rapidly replacing these terrestrial systems.

Communication satellites have several advantages when compared to other forms of electronic communications. One of these advantages is the wide area covered by a single transmitter. The transmitting antennas on a communication satellite can be designed to provide signals to service areas as small as a city or as large as a country. The maximum area that can be covered by a geostationary satellite is approximately one-third of the Earth's surface. With satellites, it is just as easy to receive a signal in a remote location such as a farm or an island as it is to receive the signal in a city. One satellite can easily provide coverage for the entire U.S. Another advantage of a communication satellite is the low average operating cost. Manufacturing and launching a satellite is very expensive, but a typical satellite will operate continuously for 8–12 years. This extended operating lifetime makes satellite communication systems very competitive with terrestrial microwave systems and transoceanic cables.

57.2 Satellite Systems

A typical geostationary satellite, which is used to receive and retransmit signals from Earth, is launched from Earth or a spacecraft and placed in orbit at a particular longitude. After a satellite has been placed in geostationary orbit, a *station keeping subsystem* in the satellite is activated from a control station on Earth to maintain the satellite's position within plus or minus 0.1° of longitude and **equatorial inclination**. An *electrical power subsystem* in the satellite is activated to generate and regulate the electrical power needed by the transponders and other equipment in the satellite. An *antenna subsystem* provides the receiving and transmitting antennas needed for the transponder and other special purpose antennas such as an omnidirectional antenna that is used as a beacon. The antenna subsystem on each satellite provides the radiation pattern needed to transmit and receive signals to and from specific regions on Earth. A *command and control telemetry subsystem* is provided on each satellite to transmit operating information about the satellite's systems to Earth and to receive and process control signals from Earth to the satellite. The *transponder subsystem* on each satellite provides the **low-noise amplifiers** (**LNAs**), the **downconverters**, and the **high-power amplifiers** (**HPAs**) needed to receive signals from the Earth station transmitters and to retransmit signals to the Earth station receivers. A typical example of each of these subsystems is further described later in this chapter. Communication satellites have a typical lifetime of 8–12 years. The amount of fuel (gas) in the station keeping subsystem is one of the factors that limits the useful lifetime of a satellite. The gradual deterioration of the solar cells that are used in the electrical power subsystem also limits the useful lifetime.

57.3 Transponder Systems

The primary purpose of communication satellites in geostationary orbit is to receive and retransmit signals from the Earth. Transponders provide the electronic equipment needed to accomplish this objective. A communication satellite is capable of receiving and retransmitting one or more independent signals. Each **uplink** signal is amplified and translated to a lower **downlink** carrier frequency by electronic circuits in the satellite. The electronic equipment used to amplify and shift the frequency of each signal is called a transponder. Typically, each satellite has 16–24 active transponders plus 6–8 spares.

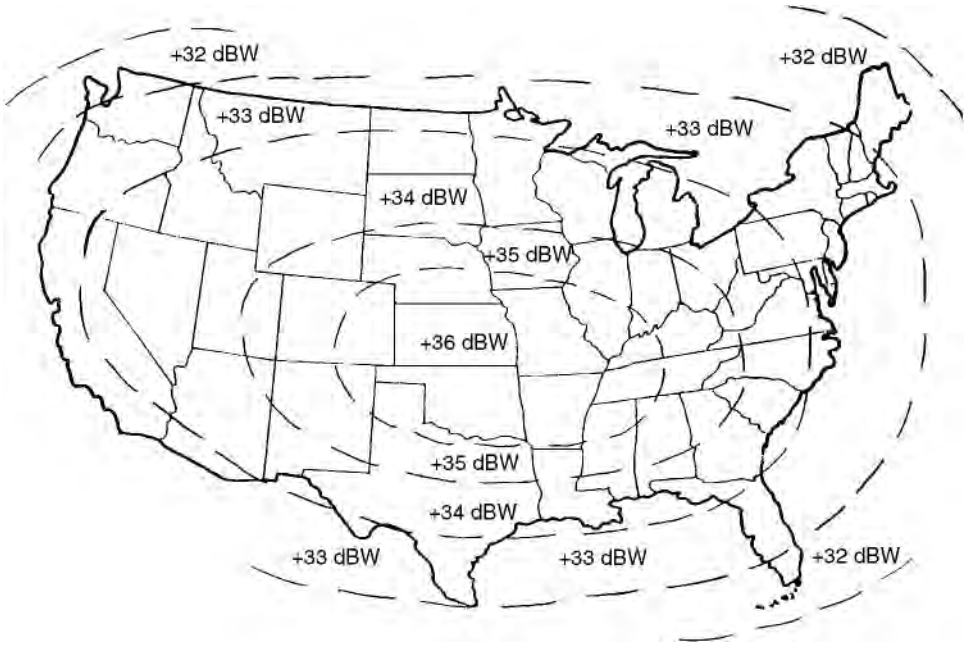


FIGURE 57.1 Satellite footprint.

Satellite Footprints

The radiation pattern of a satellite transmitting antenna determines the region on Earth where the downlink signal can be received. The radiation pattern for a transmitting antenna of any satellite can be superimposed onto a map of the region covered to display the **effective isotropic radiated power (EIRP)** radiated by the satellite to each area of the region. Maps that give the EIRP for a satellite are called satellite **footprints**. Vertical and horizontal radiation patterns for a satellite are usually combined on one footprint. Figure 57.1 shows a typical satellite footprint. The footprint given in Fig. 57.1 can be used to determine the EIRP for any location in the continental U.S. For example, the EIRP for Memphis, Tennessee, is +36 **dBW** (decibel referenced to 1 W) and the EIRP for Boston, Massachusetts, is +33 dBW. The actual power of a signal received from the satellite represented by this footprint is determined by subtracting the **space loss** to any Earth location from the EIRP.

Transponder Analysis

Figure 57.2 provides a functional block diagram of a simplified transponder which has only one input signal. The uplink signal received from an Earth station transmitter is amplified by the combination of the receiving antenna of the satellite and a low-noise amplifier. Tunnel diode amplifiers, operated at reduced temperatures, are often used in the LNAs in satellite receivers. After amplification, the output signal from the LNA is mixed in a **translator** with a signal from a fixed frequency oscillator. The output signal from the mixer contains the sum and difference frequencies of the inputs. A filter passes the difference frequencies and blocks the sum frequencies. The output of the downconverter contains the same information as the input, but the center frequency of the carrier has been shifted down by 2225 MHz. The output of the downconverter is amplified in a high-power amplifier to increase the power of the downlink signal. Very few electronic devices provide the power output needed (10–50 W) and the lifetime required (up to 12 years without maintenance) for the 4–20 GHz signals amplified in the HPAs. Most of the HPAs used in satellite transponders are either **traveling wave tube amplifiers (TWTA)** or **gallium-arsenide field-effect transistor (GaAsFET) amplifiers**. GaAsFETs have replaced TWTAs at 4 GHz as new satellites have been

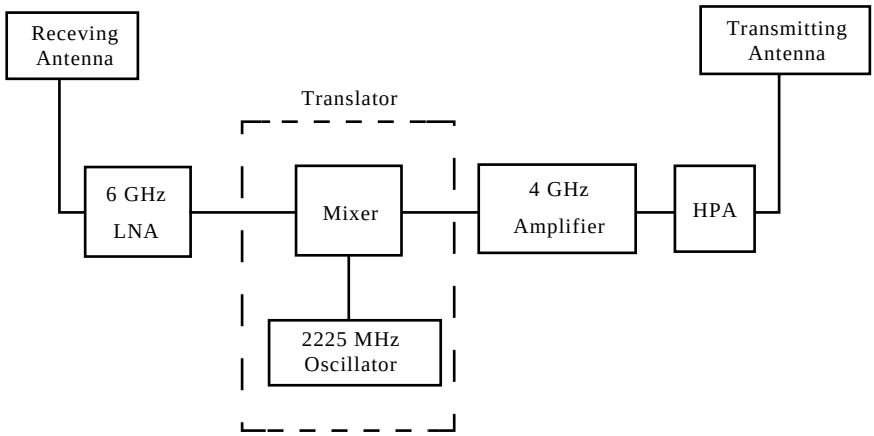


FIGURE 57.2 Transponder block diagram.

launched to replace older satellites. GaAsFETs HPA are called **solid-state power amplifiers** (**SSPAs**). TWTAs are still being used at 12 GHz and 20 GHz. The output of the HPA is amplified by the transmitting antenna and radiated back to Earth.

As an example of a simplified satellite transponder, assume that the uplink carrier signal has a frequency of 5945 MHz ± 18 MHz. This signal is amplified by the receiving antenna and the LNA and then mixed with a 2225-MHz signal. The output of the mixer is 3720 MHz ± 18 MHz (5945 - 2225) and 8170 MHz ± 18 MHz (5945 + 2225). The 3720 MHz ± 18 MHz signals pass through a filter to the HPA. The HPA and the transmitting antenna amplify and concentrate the output signal from the filter to provide the effective power needed to radiate this 3720 MHz ± 18 MHz downlink signal back to Earth. Notice that the uplink signal is not demodulated in the satellite.

Keeping the Pipeline Full

Satellite transponder systems provide a bandwidth of 500 MHz to over 2 GHz to the signals being relayed. **Frequency**- and/or **time-division multiplexing** is used to take full advantage of this available bandwidth. The design of a specific transponder system depends on the type of **multiplexing** used. For example, analog television video and audio signals can be frequency modulated onto transponder channels, which have a bandwidth of 40 MHZ. Telecommunication systems use different multiplexing techniques and, therefore, require different transponder designs to fully utilize the available bandwidth.

Television Transponder Systems

Most of the **C-band** communication satellites use two 12-channel transponders to receive, translate, and retransmit 24 independent channels simultaneously. One of these transponders receives and transmits signals using vertically polarized antennas while the other transponder receives and transmits signals using horizontally polarized antennas. Each LNA amplifies one of these uplink signals and then shifts the uplink frequencies in all 12 channels to corresponding downlink frequencies. The output of the translator is split into 12 separate signals and each signal is amplified by a separate HPA. The outputs of the HPAs are recombined and transmitted back to the Earth. Figure 57.3 provides a functional block diagram of a typical 4/6-GHz transponder system that has 24 channels. Two receiving antennas are used to separate the uplink signal into one vertically polarized and one horizontally polarized signal. The vertically polarized signal contains all of the even-numbered channels from 2 to 24, and the horizontally polarized signal contains all of the odd-numbered channels 1 through 23. Each channel has a bandwidth of 36 MHz, and there is a 4-MHz guardband between each channel. Two separate LNA-translator circuits are required because the horizontally polarized and vertically polarized signals **reuse** the same frequencies.

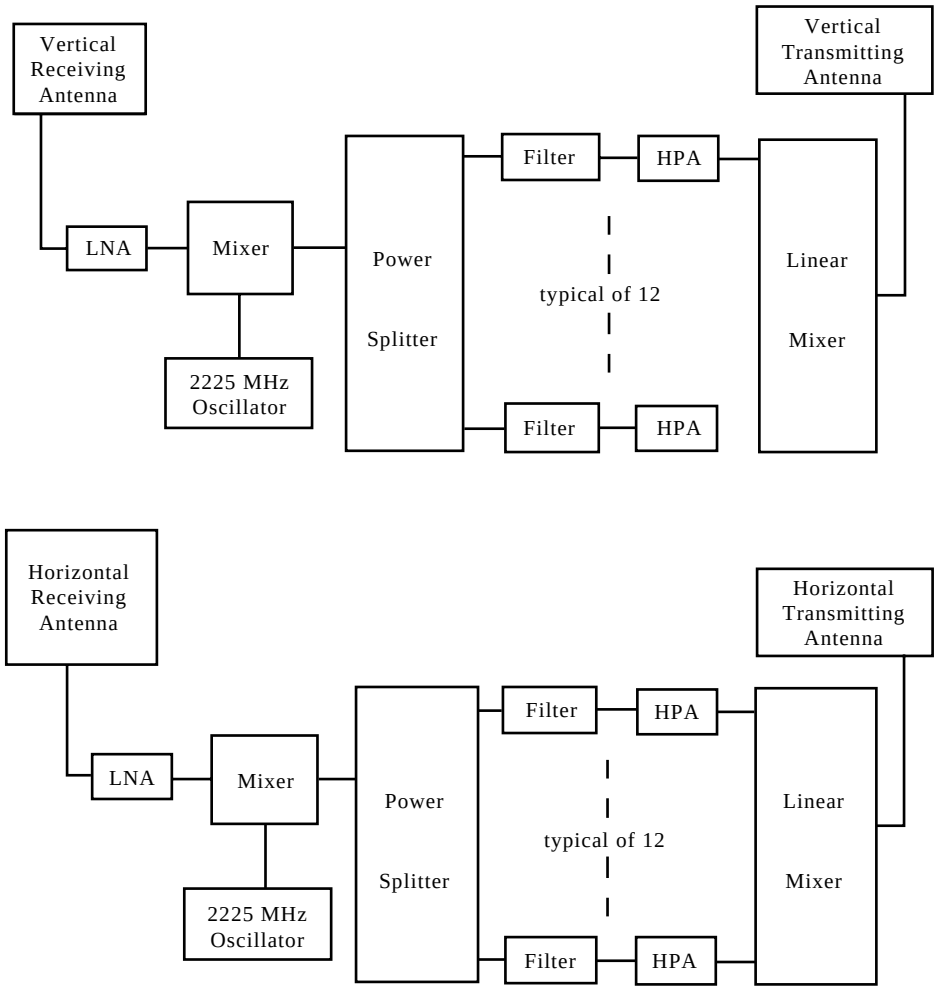


FIGURE 57.3 Simplified 4/6-GHz transponder system.

After amplification in the HPAs, the even-numbered channels are recombined and radiated to the Earth using a vertically polarized transmitting antenna and the odd-numbered channels are recombined and radiated to the Earth using a horizontally polarized transmitting antenna.

Redundancy is used in most transponder systems to improve reliability. For example, most satellites usually have 4–6 spare transponders and one or more spare LNA translators that are used to replace defective circuits. The command and telemetry subsystem is used by a controller on Earth to switch spare circuits in and out of the system. An input multiplexer uses bandpass filters to separate the uplink signal into individual channels between the downconverter and the HPA. There is a separate HPA for each channel. An output multiplexer (diplexer) is used to recombine the outputs of the HPAs into a composite downlink signal.

Antenna Subsystem

A satellite usually has an omnidirectional antenna that is used to receive command signals from Earth and transmit telemetry signals to Earth. These command signals are used to control all of the equipment in the satellite. For example, a command signal is used to start the **apogee** motor to transfer a satellite from elliptical orbit to geostationary orbit. Command signals are also used to start or stop the spinning

of a satellite and to position the attitude of a satellite so that the other satellite antennas are pointing toward the Earth. Telemetry signals are used to provide information about a satellite to the Earth control station. An omnidirectional antenna must be used for the telemetry and control subsystem during the launching and positioning of the satellite because the other antennas on the satellite may not be pointing toward the Earth. The directional antennas can be used for the command and telemetry signals after a satellite has been properly positioned and oriented.

Some satellites have Earth coverage antennas. Earth coverage antennas are horn (cone) shaped with radiation patterns that cover all of the Earth's surface visible from the geostationary orbit. Earth coverage antennas can be used to transmit or receive signals to or from any place on the Earth. A beamwidth of 17.2∞ is required for Earth coverage from the geostationary orbit.

Parabolic antennas are used on satellites to generate spot-beams or radiation patterns that cover only a portion of Earth's surface. In general, the diameter of parabolic antennas and the diameter of a radiation pattern on Earth are inversely related. Large parabolic antennas are used to transmit or receive signals from a small specific region on Earth. Smaller parabolic antennas are used for larger areas.

The radiation pattern of a parabolic reflector can be tailored to cover a particular region by using multiple feed horns that are positioned to provide the coverage desired. For example, the radiation pattern of an antenna can be tailored to cover the continental U.S. (**CONUS**), Alaska, or portions of Canada. Tailoring a radiation pattern to cover a specific region on Earth is called beam shaping. One particular communication satellite uses two 60-in.-diameter parabolic reflectors with multiple feed horns to obtain a beamwidth of 7° by 3.2^{o} for CONUS coverage. Additional feed horns are used to provide a 4° by 2° beam for Alaska, a 3° beam for Hawaii, and a 3° beam for Puerto Rico [Gould and Lum]. Satellites that use vertical and horizontal **polarization** to reuse the available bandwidth often have two transmit parabolic antennas (one for vertical polarization and one for horizontal polarization) and two receive parabolic antennas (one for vertical polarization and one for horizontal polarization). Many satellites have two receiving and two transmitting antennas. One antenna is used to receive vertically polarized uplink signals, and the other receiving antenna is used for horizontally polarized uplink signals. Similarly, one transmitting antenna is used for vertically polarized downlink signals and the other for horizontally polarized downlink signals.

57.4 Launching Satellites

A communication satellite is launched into a geostationary orbit using either an unmanned launch vehicle that is expendable or a manned space shuttle that is reusable. The launch vehicle has one or more rocket stages that provide the acceleration needed to place a satellite into Earth orbit. The satellite is normally placed in a circular or elliptical Earth orbit by the launch vehicle and then transferred into a geostationary orbit.

An unmanned launch vehicle usually places the satellite directly into an elliptical Earth orbit that is called a transfer orbit. The highest point in this elliptical orbit is called the apogee and the lowest point is called the **perigee**. The launch is designed to provide an apogee that is the same height as the geostationary orbit. A special rocket in the satellite provides the acceleration needed to transfer the satellite from an elliptical orbit into a geostationary orbit. This special rocket in a satellite is called an apogee motor or an **apogee kick motor** (**AKM**). Elliptical orbits have an angle of inclination that corresponds to the latitude of the launch site. For example, satellites launched from Cape Canaveral, Florida, have an elliptical orbit with an inclination of approximately 27.5^{o} . The apogee motor must also provide the acceleration to change the angle of inclination to zero as a satellite's orbit is changed from elliptical to geostationary.

Space shuttles provide an alternate method to launch communication satellites. One or more satellites are stored in the payload bay of a space shuttle. The space shuttle, which is manned, is then launched into a circular orbit that is approximately 200 mi above Earth's surface. Astronauts unload the satellites from the payload bay and place them into an elliptical Earth orbit. The satellites are then transferred from this elliptical orbit into a geostationary orbit. Launching satellites with a space shuttle is less expensive

than using unmanned launch vehicles because space shuttles can be reused. Space shuttles have also been used to retrieve satellites from space.

57.5 Station Keeping and Stabilization

Ideally, a satellite in geostationary orbit should remain in a fixed position because the gravitational and centrifugal forces on it are equal. Unfortunately, the gravitational forces on a geostationary satellite are not constant. The gravitational forces of the sun and moon on a geostationary satellite cause a change in the inclination angle of the satellite. The station keeping subsystem in a satellite provides small thruster rockets that are periodically used to move the satellite back to an inclination angle of zero degrees. Orbital calculations also assume that the mass of the Earth is uniformly distributed in a perfect sphere. Actually, the Earth is slightly egg shaped. This error causes a change in the longitude of geostationary satellites. The station keeping subsystem is also used to keep geostationary satellites in their assigned longitudes. The thruster rockets used for station keeping use gas that is stored in tanks in the satellites. Hydrazine gas is often used for the thruster rockets on satellites. The amount of gas stored in the tanks for the thruster rockets is one of the limits on the effective lifetime of a geostationary satellite.

The attitude of a geostationary satellite refers to the orientation of the satellite with respect to Earth. For example, the antennas on a satellite must be positioned so that their radiation pattern covers a specific region on Earth. Once the attitude of a satellite has been established, it is maintained by a stabilization subsystem. There are two different types of stabilization subsystems used in communication satellites.

Spinner-type satellites provide stabilization by causing the body of the satellite to rotate at several hundred revolutions per minute. The angular momentum of the spinning satellite acts like a gyroscope to stabilize the attitude of the satellite. Antennas on spinner satellites are **despun** because they must maintain a fixed pointing position and polarization. The bodies of spinner satellites are usually cylindrical in shape, and these satellites have a pendulumlike device called a **nutation damper** to damp out any periodic wobbling motion that may occur in the spin axis.

Another type of stabilization system uses three inertia wheels to create three-dimensional gyroscopic forces. Satellites that use **three-dimensional stabilization** have one inertia wheel rotating in each dimension. Each inertia wheel is a completely enclosed rotating sphere or cylinder that provides stabilization in one dimension. The three dimensions are referred to as pitch, roll, and yaw. The body of a three-dimensional satellite is usually square or rectangular in shape since it does not spin. The velocities of the inertia wheels can be changed to provide notation damping.

57.6 Electrical Power Subsystem

Solar cells which convert sunlight into electrical energy are used to generate the electrical power used in the satellites. Satellites require a large number of solar cells to generate the electrical power needed. The solar cells are mounted on satellites using two different methods. Spinner satellites have their solar cells mounted on the rotating cylindrical body of the satellite so that a portion of the solar cell array is always exposed to sunlight as the body of the satellite rotates. This method is inefficient because few of the solar cells are continuously exposed to direct sunlight. The electrical power generated by the solar cells on a spinner satellite must be transferred by slip rings from the body of the satellite to the equipment that is despun. Three-dimensional stabilized satellites (that do not spin) use large flat panels or arrays of solar cells. These arrays are positioned so that all of the solar cells on the panel are exposed to direct sunlight simultaneously. These arrays of solar cells are often called **solar sails**. Solar sails are usually folded against the body of three-dimensional stabilized satellites during launch so that the satellite will fit inside of the launch vehicle. The solar sails are deployed after a satellite has been placed in geostationary orbit and stabilized. Solar sails must be rotated by stepper-motors once every 24 h to keep the solar cells pointing toward the sun. DC-to-DC converters are used to change the electrical output of the solar cells into the voltages needed throughout a satellite.

The relative positions of the sun, the Earth, and a geostationary satellite change continuously as the Earth rotates on its axis and revolves around the sun. A satellite eclipse occurs when the Earth moves between the sun and a satellite. Since the solar cells on a satellite cannot generate electrical power during a satellite eclipse, satellites must have storage batteries to provide electrical power during this time. These batteries are recharged by the solar cells between eclipses.

The use of solar cells limits the lifetime of a satellite. The electrical output of a solar cell gradually decreases with age; after 8 to 10 years, the electrical output from a solar-cell will decrease by about 20%. Eventually, there is not enough electrical power available to keep all of the transponders operating. However, the lifetime of a satellite can be extended by turning one or more of the transponders off. Communication satellites are usually replaced after about 10 to 12 years of continuous service.

Defining Terms

Apogee: The place in an elliptical orbit that is farthest from Earth.

Apogee kick motor (AKM): A special rocket in a satellite that is used to provide the acceleration needed to transfer the satellite from an elliptical orbit into a geostationary orbit.

C-band: A 4/6-GHz bind of microwave frequencies used in many satellite communication systems.

Commercial satellites: Satellites owned and operated by private companies.

CONUS: The contiguous states in the U.S.

Demand assignment: A communication system that assigns channels to users upon demand.

Despun: The parabolic antennas on a spinner satellite do not spin because they must maintain a fixed pointing position and polarization.

Direct broadcast satellite (DBS): A communication system that broadcasts signals directly from a satellite into homes and businesses.

Domestic satellite (DOMSAT): A satellite that provides communication service within a country.

Downconverter: An electronic circuit that is used to translate a modulated signal to a lower band of frequencies.

Downlink: The path from a satellite to an Earth station.

Earth station: The site of the transmitting and receiving equipment used to communicate with satellites.

Effective isotropic radiated power (EIRP): The effective power radiated from a transmitting antenna when an **isotropic radiator** is used to determine the gain of the antenna.

Elevation: The angle at which a satellite is viewed from a site on Earth.

Equatorial inclination: The angle that the orbital plane of a satellite makes with the equator.

Equatorial orbit: A satellite is in an equatorial orbit when the orbital plane includes the equator. The angle of inclination for an equatorial orbit is zero degrees.

F(noise factor): A method used to specify noise in electronic circuits and systems.

Footprint: The radiation pattern of the transmitting antenna of a satellite superimposed onto a map of the region covered.

Frequency-division multiplexing: The combining of two or more input signals which are separated by frequency into a composite output signal.

Frequency reuse: The use of two different antenna polarizations to transmit and receive two independent signals that use the same frequency bands.

Gain/temperature figure of merit (G/T): The ratio of antenna gain to system noise temperature in decibels per keluin. G/T is used to specify the quality of a **satcom receiving system**.

Gallium-arsenide field-effect transistor (GaAsFET): A special type of transistor that is used in low-noise amplifiers and high-power amplifiers.

Geostationary: A special Earth orbit that allows a satellite to remain in a fixed position above the equator.

Geosynchronous: Any Earth orbit in which the time required for one revolution of a satellite is an integral portion of a sidereal day.

High-power amplifier (HPA): A linear power amplifier that is used to amplify modulated signals at microwave frequencies.
Hydrazine: A gas that is used in the thruster rockets on satellites.
Inclinometers: A tool that is used to measure the angle of elevation of an antenna.
Isotropic radiator: A reference antenna that radiates energy in all directions from a point source.
Ku-band: A 12/14-GHz band of microwave frequencies used in many satellite communication systems.
LNA: Low-noise amplifier.
Look angles: The azimuth and elevation angles used to point or aim an Earth station antenna toward a satellite.
Multiple access: Several users have access to the same satellite channels.
Multiplexing: The combining of two or more input signals into a composite output signal.
Nutation damper: A pendulumlike device that is used to damp out any periodic wobbling motion that occurs in the spin axis of a satellite.
Perigee: The place on an elliptical orbit that is nearest to the Earth.
Polarization: The direction of the electric field radiated from an antenna.
Polar orbit: A satellite is in a polar orbit when the orbital plane includes the North and South Poles. The angle of inclination for a polar orbit is 90° .
Satcom: A satellite communication system.
SATCOM: A series of satellites owned by RCA.
Solar sails: A large flat array of solar cells.
SSPA: Solid-state power amplifier.
Space loss: The attenuation that occurs as an electromagnetic wave propagates through space.
Spinner: A satellite that spins at several hundred revolutions per minute to stabilize the attitude of the satellite.
Terrestrial: Satcom equipment on the Earth.
Three-dimensional stabilization: Three inertia wheels are used to provide roll, pitch, and yaw stabilization of a satellite.
Time-division multiplexing: The combining of two or more input signals that are separated by time into a composite output signal.
Translator: The shifting of an AM or FM signal up or down in frequency without changing the frequency spacing or relative amplitudes of the carrier and sideband components.
Transponder: The equipment in a satellite that is used to receive the uplink signals and transmit the downlink signals.
Traveling wave tube amplifier (TWTA): A microwave amplifier that is used as the high-power amplifiers in many satellites.
Uplink: The path from an Earth station to a satellite.

References

Clarke, A.C. 1945. Extraterrestrial relays. *Wireless World,* 51 (Oct.):305–308.
Douglas, R.L. 1988. *Satellite Communications Technology,* Prentice-Hall, Englewood Cliffs, NJ.
Gould, R.G. and Lum, Y.F. *Communication Satellite Systems,* IEEE Press, New York.

Further Information

For further information on satellite systems and transponders, see:

- *Modern Electronic Communication Techniques*, Harold B. Killen, Macmillan, New York, 1985.
- *Electronic Communication Techniques*, Paul H. Young, Macmillan, New York, 1994.

58

The Earth Station

David R. Pritchard
Norlight Telecommunications

58.1 Introduction

The term **Earth station** is generally used to refer to a ground based communications facility made up of the transmitters, receivers, and antennas used to relay signals to or from a communications satellite. By definition, the Earth station also includes the equipment required to interconnect the satellite signals to the end user, whether the user is located at the Earth station facility or at a remote location. Information may be transmitted or received in the form of analog or digital carriers for voice, data, or video transmission. The signal is transmitted, or uplinked, from the ground station to a satellite **transponder** (Fig. 58.1). The satellite transponder relays the signal back to Earth where it can be received, or downlinked, anywhere within the coverage area known as the **footprint**. The majority of commercial Earth stations are designed to operate with satellites in geosynchronous orbit operated by AT&T, GE, Hughes, and other domestic and international carriers. Geosynchronous satellites are positioned approximately 22,300 mi. above the equator and rotate at the same speed as the Earth. From a fixed point on the ground, a satellite in geosynchronous orbit appears stationary and can provide continuous service to the end user. Figure 58.2 is an aerial view of Teleport Chicago showing multiple earth station antennas used for video and audio transmission. There are approximately 35 commercial satellites that can be accessed by locations in North America, which provide domestic C- and Ku-band satellite service (not including the Direct Broadcast Satellites).

58.2 Components of the Earth Station

An earth station will consist of one or more antennas, either stationary or motorized, equipped to receive and/or transmit signals to a communications satellite. Commercial satellites for voice, data, and video services operate at C- and Ku-band frequencies. C-Band satellites operate with uplink frequencies of 5.9 to 6.4 GHz and downlink frequencies from 3.7 to 4.2 GHz. Ku-band satellites operate with uplink frequencies from 14.0 to 14.5 GHz and downlink frequencies from 11.7 to 12.2 GHz. The Ku-band satellites providing international service operate with slightly different frequency plans. Because of the shorter wavelength of the higher frequencies used, Ku-band satellite systems experience attenuation

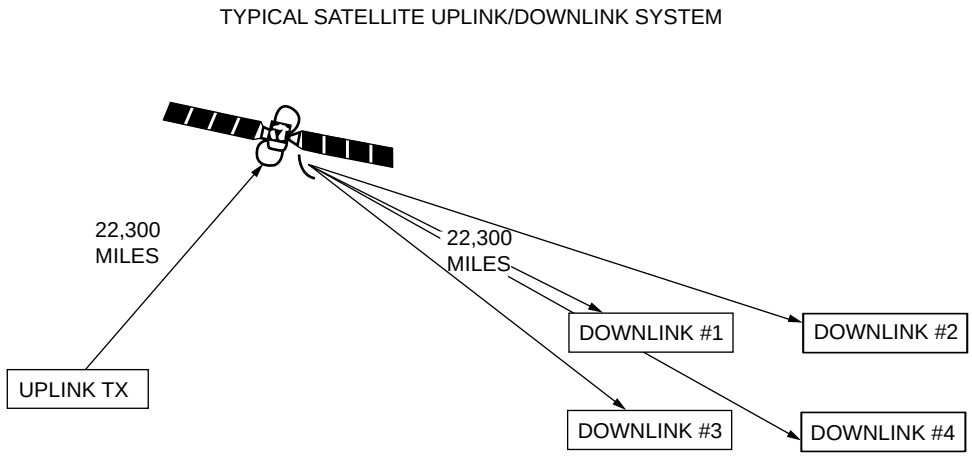


FIGURE 58.1 Typical satellite uplink/downlink system.

FIGURE 58.2 Teleport Chicago Earth station.

during periods of heavy rain. Overall link availability of Ku systems will be decreased depending on the annual rainfall amount in the geographic area where the earth station is located. C-band transmissions are not appreciably affected by rainfall.

Some of the newer generation satellites have both C- and Ku-band capability. These satellites are referred to as hybrids. Some of their transponders can actually be configured by the satellite operator to receive a ground transmission on C-band and relay it back to earth on Ku-band. Most C-band satellites operate with 24 channels or transponders, each using a bandwidth of 40 MHz although only 36 MHz is actually used to provide for a guard band at each end of the transponder. The 36 MHz bandwidth will support one standard National Television Standards Committee (NTSC) television signal, including audio channels, using FM modulation. It will also support multiple analog or digital carriers for voice, data, or compressed video transmission. Ku-band satellites do not follow the same transponder plan as C-band. Many Ku satellite

FIGURE 58.3 Prime-focus feed system.

transponders have wider bandwidths available. In many cases, two FM video carriers may operate side by side on one Ku-band transponder. Both C- and Ku-band satellites receive and transmit signals on alternate polarization to reduce the chance of interference from adjacent transponders in the frequency plan.

The antenna itself can vary in design, but it is usually a parabolic dish with a feedhorn assembly. The feedhorn is the point at which satellite signals are collected from the antenna reflector surface to be fed to receive equipment. It is also the point where transmit signals are sent out to the reflector surface of the antenna to be focused into a beam aimed toward the satellite. The diameter of the reflective surface determines the gain and beamwidth of the antenna. The gain of the antenna becomes an important part of calculating a system link budget to determine the quality and availability of the signal.

Parabolic dish antennas can use many types of feedhorn assemblies. The most common are the prime-focus feed (Fig. 58.4) or the **Cassegrain** (Fig. 58.5) feed system. The prime focus system uses a feedhorn placed at the focal point of the reflector. Transmit or receive signals are directed from the reflector to the feed point. In the Cassegrain feed system, a subreflector is used to direct the transmit or receive signals to a feedhorn extending from the hub of the antenna. The actual feedhorn is hollow and the opening is covered with a protective window. The material used for the window is transparent to the satellite signal and keeps moisture or foreign objects from entering the assembly. Cassegrain feed systems are usually pressurized with nitrogen or dry air to help prevent moisture from degrading the signal. Pressurization systems should be monitored for any sign of leakage since damage to the feed window can occur from hail, birds, or vandalism. Even a small amount of moisture can affect transmit cross polarization, attenuate the transmit and receive signals, and generally degrade system performance. It is advisable to have an alarm tied into the pressurization system so that corrective action can quickly be taken if a feed window is damaged.

Feedhorn assemblies can be designed to accommodate up to four interface ports for transmitters and receivers. Four-port feedhorns are the most complicated and are sometimes referred to as frequency reuse feeds. They contain a diplexing assembly at the base of the feedhorn to separate both transmit and receive frequencies. The diplexer also divides signals into two orthogonal ports for transmit and two for receive. This provides signals for both polarizations used by the satellite. Part of the diplexer assembly also contains a transmit reject filter to prevent signals that are being transmitted to the satellite from

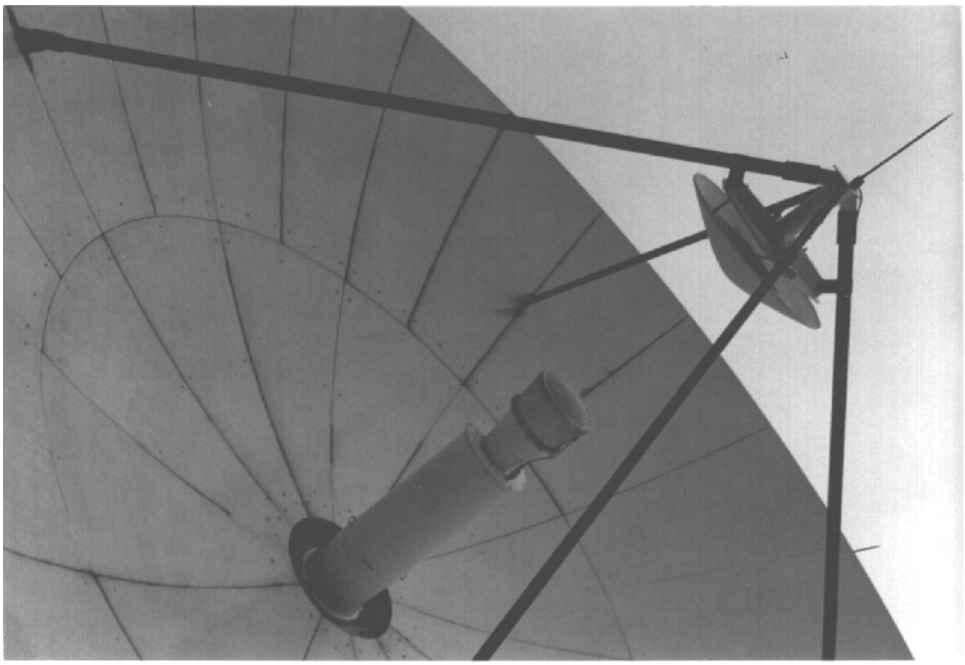


FIGURE 58.4 Cassegrain feed system.

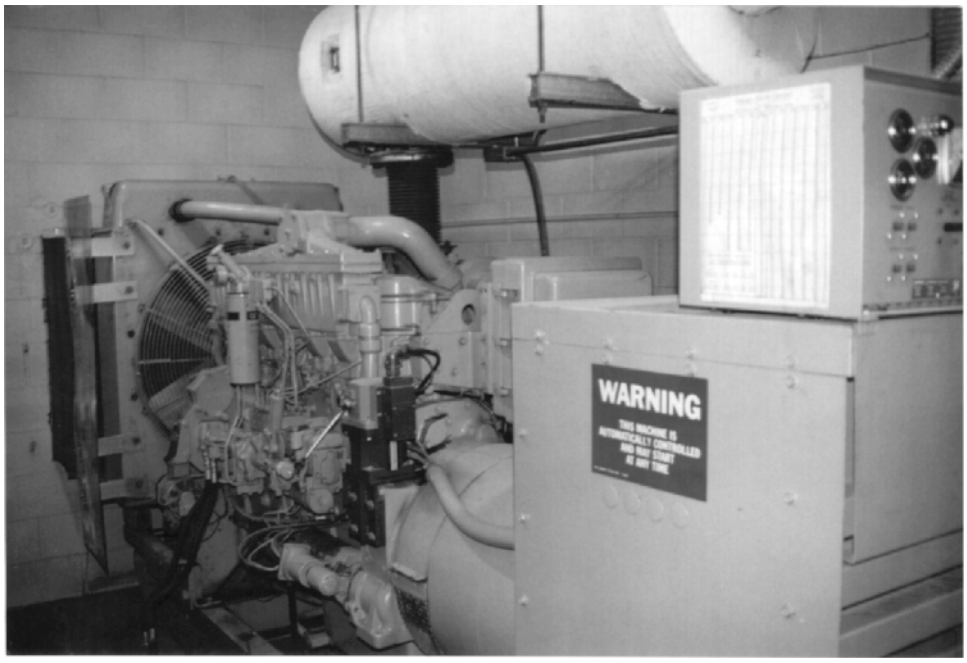


FIGURE 58.5 Diesel engine generator.

interfering with the signals being received within the feedhorn assembly. Some international and non-commercial satellite applications use circular or linear polarization for transmission, which requires a different diplexer design. The Federal Communications Commission (FCC) requires that satellite transmit antennas be able to maintain a radiation pattern that limits the amount of energy radiated in the sidelobes. This is to reduce the potential for interference to adjacent satellites in orbit. Antennas meeting this requirement are advertised as being 2° compliant and usually require adjustments of the subreflector (or feedhorn, in prime-focus systems) at the time of installation to meet the FCC requirement. To certify the antenna is 2° compliant, pattern testing is performed. The received signal level of an unmodulated carrier (transmitted from a different antenna) is observed as the antenna under test is moved in azimuth and elevation. The level of the sidelobes in relation to the main carrier is observed on the spectrum analyzer and a plot is made in both the azimuth and elevation axes to determine if 2° performance is being achieved. Transmit pattern testing may also be performed on the antenna in the same manner, but it is difficult to perform without causing interference to adjacent satellites as the antenna is moved in azimuth. Receive-only antennas are not required to be 2° compliant by law. However, they may suffer from some adjacent satellite interference if they are not compliant because of the closer spacing used by satellites in orbit today. Specific information with regard to FCC antenna performance requirements can be found in Part 25 of The Code of Federal Regulations.

Depending on the Earth station location, snow and ice can be a factor. The antenna reflector will collect ice and snow, reducing gain and affecting the transmit pattern and antenna gain. Earth station antennas can be equipped with deicing systems that keep the reflector clear during storm conditions. Deicing systems use either electrical heating elements or gas-fired furnaces to supply heat to the rear of the reflector surface. Cassegrain antenna designs usually include electric heaters in the feedhorn and subreflector surfaces to keep those areas free of snow and ice. Automatic sensors can be mounted at the antenna to detect snow and ice and place the deicing system into operation.

Received satellite signals are extremely low in level at the antenna feed point and require amplification to a level acceptable to be passed on to the satellite receiver or downconvertor. The amplifier must be capable of considerable gain with low noise, and it is usually mounted at the feedhorn to achieve the best results. This portion of the system is referred to as the **low-noise amplifier (LNA)** or **low-noise block down-converter (LNB)**. Two LNAs or LNBs are required if both polarizations are used on the receive system. The LNA and LNB both use a gallium-arsenide field effect transistor (GAsFET) design to achieve the low-noise performance. The LNA amplifies the incoming receive signal and passes it to the receiver at the original operating frequency. The LNB amplifies the received signal and converts it to an L-band frequency of 950–1450 MHz. Noise performance of an LNA or LNB is measured in degrees Kelvin. Today it is not uncommon to find LNAs or LNBs with noise temperatures of 30° K or better. The gain of the LNA or LNB along with its noise temperature are used when calculating total receive system gain.

Older system designs use an LNA as the heart of the receive system. Because the signal travels from the antenna to the receiver at the original operating frequency, low-loss feedline has to be used.

Most newer satellite receivers are designed to operate with an L-band frequency input of 950–1450 MHz so an LNB can be used at the antenna. The L-band signal can be more easily fed from the LNB directly to the receiver or downconvertor using a conventional shielded coax. Both the LNA or LNB receive their DC power through the signal cable. The LNA/LNB power supply is located internally on some receivers or it can be a separate component in the Earth station. Many of today's video and data applications are using signal compression to allow more carriers to be placed on a single satellite transponder. To receive these signals, a **digital integrated receiver-decoder (IRD)** is used. The IRD contains satellite receive circuitry as well as on-board digital processing to decode compressed or encrypted signals. IRDs are available for different digital transmission formats such as **MPEG** and **DVB**. For these applications, a phase-stable LNB should be used. A phase-stable LNB will minimize frequency drifting that can produce bit errors, which will degrade the video signal.

U.S. television broadcasters are required to begin using a new digital television (DTV) standard by 2006. Some of this programming may be in high-definition format, or broadcasters may elect to transmit several standard definition signals over one television channel. Earth stations involved in the delivery of

these signals to broadcasters will most likely utilize digital IRDs. Therefore, new Earth station facilities should consider initially installing phase-stable LNBs, and older facilities may need to upgrade in order to be able to provide reliable digital services.

A newer style analog video receiver with an LNB input can still be used with an LNA equipped antenna. A device called a block downconvertor can be inserted between the LNA output and the receiver. The block downconvertor will convert the LNA signal to the L-band frequencies required by the newer receiver.

For analog FM video transmission, a device called an **exciter** is used as the first stage in the uplink chain. The exciter performs baseband processing of the video signal and includes separate subcarrier modulators for the audio portion of the transmission. FCC regulations require that all video earth stations include an **automatic transmitter identifier signal (ATIS)** in their carrier. This signal consists of an additional subcarrier modulated with Morse code, which contains the Earth station FCC callsign and telephone number. The ATIS signal is used to help protect against unintentional interference by a video uplink. It is used by the satellite controller to identify stray carriers which have inadvertently been left transmitting to the satellite. Newer video exciters include an onboard ATIS generator. There are outboard models available, which may be added on to older units.

The processed video, audio, and ATIS subcarriers are applied to the modulator stage of the exciter. The intermediate frequency (IF) output of the modulator is then applied to the upconvertor, which is the last stage of the exciter. The upconvertor takes the IF signal and multiplies it to the desired uplink frequency, which is selected by the operator. The exciter output is then sent to the **high power amplifier (HPA)**.

For compressed video signals, an encoder or **codec** is used instead of an exciter. Video compression allows several video signals to be placed onto a single satellite transponder in order to save space and cost. The level of compression used will dictate picture quality since some artifacts due to the compression process may be visible. Another advantage to digital video transmission is that many encoders provide for remote control of the IRDs receiving their signal. This allows the uplink operator to select the receive sites that will be authorized to receive a transmission.

For data or other high-speed traffic such as Internet service, satellite modems are used in place of an exciter. Modems can be configured to match the bandwidth requirements for a given application. Satellite modems generally contain forward error correction circuitry to increase the reliability of the satellite link. In the case of Internet traffic, a satellite link is used primarily in areas where terrestrial connections are unavailable, especially internationally. A router is usually placed between the satellite modem and the Internet customer.

The modulator, satellite modem, or codec produces an IF output, which is fed to an outboard upconvertor. This **upconvertor** performs the same function as the upconvertor in a video exciter. The operator must tune the upconvertor to the correct frequency for the desired satellite transponder. The output is then fed to the HPA. The HPA is the last device in the transmit chain prior to the antenna. This amplifier will develop the power necessary to send the signal from the Earth station antenna to the satellite. The exact amount of power required will vary depending on the type of modulation used, the transmission line loss, the gain of the antenna, and the characteristics of the satellite transponder. Most HPAs used for fixed Earth station applications use a **travelling wave tube (TWT)** or **klystron** as the main amplifying component. TWTs are wide bandwidth devices and do not require any tuning before transmission. Klystrons are manufactured in single- or multiple-channel configurations. The multiple-channel Klystrons have internal mechanical tuning cavities, which must be set to the desired transponder frequency before transmitting. Some Earth stations and some mobile uplink trucks will use a solid-state HPA mounted at the antenna feed point. This minimizes transmission line loss and reduces the output power requirement of the amplifier. Solid state HPAs are not suited for all applications, as they cannot develop the power levels that a TWT or Klystron amplifier can.

HPAs will usually have an input attenuator to control the level of drive from the exciter or upconvertor. This attenuator is used to set the power output level of the HPA. The attenuator may operate manually but some amplifiers use a motorized system for remote control of output power. The HPA output is connected to the antenna using a waveguide. Provision must be included in the system design for routing the HPA output to the appropriate antenna and polarization if the system is used to provide occasional

service to multiple users. Provision should also be included for routing the HPA into a dummy test load. This helps facilitate testing of the HPA and reduces the chance that the system may be inadvertently left transmitting to the satellite. In redundant transmit systems, the backup HPA is usually connected to a dummy load unless a failure occurs on the primary unit.

58.3 Earth Station Site Selection

Earth stations in the United States are required to be licensed by the Federal Communications Commission before they can be placed into service. Part 25 of the Code of Federal Regulations should be referred to for the information needed to make an application for station authorization. Part of the procedure will include performing a frequency coordination analysis. This will assess the impact of a new Earth station to existing common carrier facilities. Placement of a receive-only Earth station is limited by several factors. One is the potential for interference from ground-based communication systems. C-band downlink frequencies are also used by common carriers for point-to-point microwave communications. However, common carrier use of these frequencies has diminished in many areas as higher capacity fiber optic systems replace microwave radio as the means for transmission. This has made it less likely that interference from these carriers will occur to Earth stations. Nonetheless, a spectrum analyzer should be used with a low-noise amplifier and handheld horn antenna to survey any potential Earth station site for external interference. There are companies that will perform this service for a fee during the initial site selection process. Ku-band downlink frequencies are also shared by common carriers for temporary local television transmission. Common carriers are secondary users of this band and are required by the FCC to coordinate transmissions with all satellite users in their service area. It is not uncommon, however, to experience interference from these terrestrial systems, especially in metropolitan areas where signals may be reflected into downlink antennas from nearby structures. Antennas can be equipped with mechanical shielding devices, which help reduce received interference. These devices may reduce overall system gain, but when used in combination with outboard filters in either the RF or IF stages of a receiver, interference can be reduced to a tolerable level.

In North America, the domestic satellite arc extends from about 63° –139° west longitude. A potential Earth station site should have a clear view of the entire satellite arc. Depending on geographic location, North American Earth stations can also see some of the international satellites. This will extend the arc clearance required.

Potential Earth station locations should be studied for adjacent trees or buildings that may cause obstructions. It is advisable to check with local zoning authorities to determine any future building plans near the proposed site location when making the initial visual inspection. An inclinometer or transit can be used to calculate elevation clearance. There are many shareware and commercially available computer programs that will calculate each satellite's azimuth and elevation angle as viewed from the geographic coordinates of the proposed site.

Local authorities will also be able to provide information regarding zoning restrictions that may limit the placement of Earth station facilities. In 1996, the FCC issued a report and order (FCC 96-326) adopting new guidelines for evaluating the environmental effects of radio frequency (RF) radiation from FCC licensed transmitters. As part of this report, the FCC has adopted RF exposure limits for both controlled and uncontrolled environments at Earth station facilities. An uncontrolled environment includes areas around an Earth station site where individuals might be exposed to RF radiation without their knowledge. The Earth station operator will need to include proof that a proposed antenna installation will meet RF exposure limits when submitting a license application to the FCC. In planning a new Earth station installation, a **link budget** is normally prepared, which determines the maximum HPA power output that will be required. The link budget also takes into account the proposed transmit antenna size, satellite transponder characteristics, and atmospheric losses and determines overall quality of service. Calculations can be performed using this data to determine the level of RF exposure that the Earth station will present to surrounding areas. In most cases, the services of a professional engineering firm should be contracted to perform the calculations.

Earth stations may be collocated with the end user of the satellite service they provide. This is true for many video applications where programming originates or terminates at a television studio. In other cases, provision must be made for relaying the signal to the end user. This can be done through fiber optic cable, microwave radio, or a combination of both when route diversity is required. It is important to note that if the Earth station will use terrestrial microwave radio channels to receive or transmit satellite signals to the end user, the site selection process will need to include the feasibility of erecting a tower to support the microwave antennas. A microwave path study will then need to be performed to determine the tower height required. Path length and the frequency band used in a microwave interconnection will determine the potential impact of rain attenuation.

The frequency used for the interconnecting microwave facilities should be carefully planned. C-band earth stations transmit in the range of 5.9–6.4 GHz, which is still shared by some telephone common carriers. Most Earth stations will not use microwave radio in this frequency band because it is likely to receive interference from the higher power signals transmitted from the Earth station to the satellite. Earth stations operated by common carriers usually use microwave channels in the 12-, 18- or 23-GHz common carrier bands when fiber optic cable is not practical. Although this solves the problem of possible interference to the microwave channels from C-band transmissions, the 12-GHz common carrier band falls within some international Ku-band satellite downlink frequency plans. Therefore, these microwave channels may interfere with international satellite downklinks.

Local telephone and alternate access carriers should be contacted during the site selection process to determine the availability and cost of providing telephone and fiber optic interconnections to the Earth station. Plan ahead for future growth when contracting for these facilities to be brought into the site. At the time of installation, it is advisable to provide surge protection on all incoming telephone circuits to protect equipment from transients occurring elsewhere on the line due to lightning strikes, etc.

The local electric utility should be contacted to determine any special construction that will be required to bring sufficient power to the site. If the Earth station is to be manned, gas, water, and sewer utilities should be contacted to determine their installation costs into the site. If a hot-air deicing system is going to be used on the antenna, it is advisable to plan a route for the installation of the gas line prior to the beginning of construction.

58.4 Power Distribution

Reliability of the power is another major concern for the Earth station. Depending on the application, different levels of primary and backup power systems may be required. To increase reliability, some Earth stations use two different sources of commercial power from the utility company that are fed from different grids. The electrical service entrance must be properly sized by calculating the load for all anticipated equipment. The calculated load should include ancillary systems such as antenna deicing and heating, ventilation, and air-conditioning (HVAC) systems. The electric service entrance should also be equipped with surge protectors to help guard against equipment damage from lightning strikes. The Earth station equipment building or shelter should be designed so that sufficient room is provided for the power distribution equipment required. It is sometimes helpful to provide separate distribution for critical and noncritical loads. When backup systems are used, there are sometimes capacity limitations that require noncritical systems to remain off-line during power outages.

One or more outboard engine generators may be used to supply primary or standby power. If a generator is planned, the required capacity should be calculated by determining the initial total site load and potential for growth. The maximum rate of fuel consumption per hour can then be calculated and a fuel storage tank can be sized accordingly. Consideration must be given to the placement of a fuel tank so that it remains accessible at all times for refueling. Local ordinances may stipulate whether the tank can be above or below ground and they will also dictate if special spill containment basins will be required. Newly installed underground tanks will require automatic leak detection systems and double-walled piping to comply with recent federal Environmental Protection Agency (EPA) rulings. As part of a standby generator system, an electrical or mechanical transfer switch will need to be installed. The transfer switch

will sense a failure of the incoming utility power feed and automatically place the generator on-line. Because of the initial current demand placed on a generator at the moment of transfer, time delay relays can be installed to stagger the start up time of high current devices such as air-conditioner compressors. This arrangement can help keep the line frequency stable, especially on generators that use a mechanical governor system. Electrical governors tend to provide better line regulation under varying loads.

Earth stations which utilize electric heaters on the antennas for deicing may wish to include an option for using either commercial AC power or the standby generator for power. Using the generator to power deicing systems is a good way to exercise the unit under an actual load condition.

It is a good idea to include remote alarms on the generator to monitor critical parameters such as oil pressure, water temperature, and fuel level. Most generators have built-in alarms that will initiate shutdown if a dangerous situation occurs. However, few generators have factory-installed prealarms to alert the user to a potential shutdown condition. It is also a good idea to have the generator vendor install prealarms so that the Earth station operator can take corrective action before a shutdown occurs.

Critical earth station components can be powered from an **uninterruptable power system** (**UPS**). A UPS system will keep electrical service from being interrupted while backup generators come on-line during a power failure. As with a standby generator, the UPS system should have the capacity to accommodate all of the existing and anticipated site load. It may not be practical or economically feasible to run the entire Earth station from UPS power. In this case, the electrical distribution should be planned to route the UPS system to critical loads only.

Satellite transmission equipment is often designed with redundant power supplies. The UPS system can be used as one source of power to feed the equipment, and the utility power can be used as the other source. This will prevent an interruption to transmission if one power source fails.

An important part of any installation will be the grounding system. Grounding must conform to the National Electrical Code. A poor ground system becomes a safety hazard for personnel working at the site and also increases the risk of damage from a lightning strike. All antenna systems should be grounded, and if a tower is used for microwave interconnection, it should also be grounded. Equipment racks in the Earth station shelter or building should be tied into the same ground system as all of the outdoor equipment.

58.5 Batteries

In some applications, storage batteries are used to power equipment in the event of commercial AC failure. Storage batteries will also be part of a UPS system. The size and number of battery cells required will be determined by the power requirements of the Earth station equipment. The desired battery reserve time will also be a factor, especially in facilities that do not utilize a backup engine generator. Once the equipment load is calculated, the battery manufacturer should be consulted to determine the type of cell and charging system that will work best. Battery cells must be properly housed and maintained to ensure safe, reliable operation. Instructions provided by the battery manufacturer should always be followed regarding the storage, installation, and maintenance of batteries. Although many storage batteries are of a lead-acid design, newer cells are available using chemical compositions that require minimal maintenance and greatly reduce the chance of a chemical spill.

The temperature of the batteries will affect their ability to deliver power under load and can also limit their overall life expectancy. In general, the batteries should be located in an area that will be kept between 72° and 78° F. Earth station design should include a separate room for batteries to provide a proper environment and safe operation. Local ordinances will address any special requirements for battery room designs such as ventilation, spill containment, or explosion-proof fixtures for battery cells that contain acid. Entrance to battery rooms should be restricted to maintenance personnel, and no one should be permitted to enter the room without proper safety clothing and eye protection. Special safety equipment including eye wash stations and clean-up equipment should be located in every battery room.

58.6 Antenna Foundation Requirements

The antenna manufacturer will provide specific information regarding the foundation requirements for each particular model. Because of the complexity of foundation design, it is recommended that a professional engineer be consulted. The information provided by the manufacturer will include wind loading data and anchor bolt patterns. These should be provided to the foundation engineer before the design process begins.

For antennas requiring a reinforced concrete foundation or "pad," the engineer must take several items into consideration. Soil bearing capacity should be measured from samples obtained at the site. Once an excavation is made, additional samples may be taken from the base to determine if any special soil reinforcement is required prior to the construction of the foundation. The engineer will specify the minimum strength the concrete mix must achieve after curing. This will be based on the analysis of wind load requirements and soil condition. The design of the pad should specify any steel reinforcements required.

The installation contractor should be familiar with local zoning and building ordinances and obtain all of the necessary permits before work begins. The contractor should also be provided with a survey of the site to insure proper placement of the foundation. It is essential that the foundation is oriented correctly so that the desired portion of the satellite arc can be reached. Placement of each antenna foundation should be planned so that any future antenna installations will not have a portion of their desired arc coverage obstructed by existing antenna installations. As construction progresses, the foundation will be framed (Fig. 58.6) and a template can be laid out for the setting of anchor bolts before the concrete is poured. Most antenna manufacturers will provide a template to speed the layout of the anchor bolts. Test cylinders should be taken by the contractor while the concrete is being poured per the recommendation of the foundation engineer. These cylinders are normally tested by an independent laboratory after the concrete is cured to determine the strength it has achieved. The concrete strength needs to meet or exceed the figure specified by the design engineer before the antenna can be erected.

FIGURE 58.6 Antenna pad construction.

58.7 Site Heating, Ventilation, and Air Conditioning

Adequate heating and cooling must be provided at the Earth station site to prevent equipment failures from occurring. In a small Earth station shelter, redundant window style air conditioners may be sufficient for cooling. If heating is required, a small wall-mounted electric furnace will usually work well. In larger facilities, designing the heating and cooling systems becomes more complicated. In many cases, it is wise to contract the services of an HVAC engineer to assist with the project. Some level of cooling redundancy is required in large facilities since the loss of air conditioning during the summer months will almost certainly cause equipment to overheat in a short amount of time. In many cases, heating a large facility in the winter months is not even required as the heat given off by the equipment may be utilized. The heat load may be such that the Earth station will require air conditioning all year long. When calculating the site heat load, be sure to include the heat given off by the HPA dummy loads.

Generally, the HPAs used in Earth stations are air-cooled devices. This requires that a minimum amount of airflow be maintained through the cabinet to adequately cool the TWT or klystron. Although many of the newer domestic satellites require less operating power than earlier designs, high power amplifiers tend to produce the same amount of heat output whether they are operating at 30 or 3000 W of power. In large Earth station installations, multiple amplifiers should be ducted directly outdoors and may require the use of an external exhaust fan. This type of system will most likely use an external air intake fan as well.

HVAC contractors can also include ductwork in the system design to take a portion of the HPA exhaust air and heat other parts of the building. A concern with large air handler systems is that any failure of an intake or exhaust motor will result in rapid heat buildup in the amplifier and increases the risk failure. Intake louvers should contain air filters, which must be serviced regularly. If filters become clogged, air flow can drop significantly, which also increases the risk of failure. An air flow alarm should be a part of air handling system design to warn the operator of a problem.

58.8 Safety Considerations

There are many safety concerns when operating an Earth station. The Earth station operator should ensure that the facility is compliant with current Occupational Safety and Health Administration (OSHA) regulations. It is worth pointing out some special concerns that apply to Earth station safety. Fire suppression equipment must be provided. Local ordinances may dictate the type of fire suppressant to be used, and special care must be taken to select a suppressant that can safely be used on live electrical equipment. Unmanned sites should have remote alarms for smoke or fire detection which are monitored 24 h a day. The local fire department personnel should visit the Earth station so that they can document any special access requirements and be trained on the special requirements if a fire occurs in the transmission equipment. Operating personnel should be trained on the use of fire extinguishers and any other suppressant systems on site.

Earth station facilities present an RF radiation exposure hazard to operation and maintenance personnel. As stated previously, Earth station operators are responsible for complying with the FCC's guidelines for RF exposure limits. The FCC considers the Earth station facility to be a controlled environment. This means that employees and other personnel within the facility must never be exposed to RF energy at a level higher than the specified limit. RF hazard meters are available that can make the required RF measurements. Qualified personnel should perform these measurements at regular intervals to ensure that the building and grounds remain in compliance with FCC limits.

HPAs are provided with interlocks to protect maintenance personnel from the risk of high voltage or accidental exposure to RF radiation. These interlocks should be tested regularly. Warning lights can be mounted at antenna foundations, which automatically illuminate during transmission. This will help alert maintenance personnel working in the vicinity of the antenna that the system is active.

Klystron tubes contain beryllium oxide (BeO). If the klystron is physically damaged, a risk of exposure to toxic fumes exists. Maintenance personnel should always follow the manufacturers' recommended handling procedures to minimize the risk of damage to the klystron.

The HPAs contain voltages that can be lethal upon contact. During maintenance, it is sometimes necessary to make adjustments to an HPA while the high voltage is present. Maintenance personnel should never perform work on an energized HPA alone.

Motorized antennas can be remotely controlled and may move without warning. Many antenna control systems provide for local overrides at the foundation to be used by maintenance personnel. The override prevents the antenna motors from engaging while work is being performed. These overrides should be used in addition to standard OSHA lockout/tagout procedures to fully protect maintenance personnel from possible injury.

Finally, site security should be considered. Most facilities will have a fence to control access to the site and use special lighting for added security. If the site is unmanned, it is wise to have remote intrusion alarms report to a 24-h monitoring center to help prevent theft or vandalism.

58.9 Operation and Maintenance

With limited exception, anytime a satellite transponder is accessed by an Earth station, the satellite controller must be notified by the operator performing the uplink. For Earth stations that carry occasional video services, the Earth station may need to contact the satellite controller many times per day. Each satellite carrier has specific transponder access procedures. In general, the following guidelines should be followed by Earth stations when accessing a video transponder:

1. Ensure the antenna is properly oriented on the correct satellite.
2. Ensure that the exciter and HPA are properly tuned and routed to a dummy load or disabled.
3. Set the HPA for a power level of approximately 10 watts.
4. Verify the correct transmit polarization setting.
5. Set the exciter or upconvertor to produce an unmodulated carrier.
6. Contact the satellite controller and request permission to access the transponder. Verify the service order and time of service with the controller.
7. Begin transmission at low power when authorized by the controller. Be prepared to immediately cease transmission if commanded to do so by the controller.
8. The controller will verify the proper frequency and polarization setting. Be prepared to rotate the antenna polarization, if asked to do so.
9. Gradually increase output power to operating level when permitted by the controller.
10. Apply modulation and ATIS carrier when permitted to do so by the controller.
11. Provide the Earth station telephone number where you may be reached at any time during the transmission.
12. Contact the satellite controller with the good-night time when service is terminated.

The majority of unintentional satellite interference occurs from operator error. The chance of interference is greatly reduced when Earth station operators are properly trained and follow the access procedures for every transmission.

Maintenance of the Earth station is largely a matter of following equipment manufacturers' recommended service intervals. The major source of downtime results from equipment failure due to poor maintenance practices. Antenna drive system lubrication should be checked at regular intervals, especially when an antenna is not moved for an extended period of time. HVAC and equipment air filters should be checked weekly and changed as necessary. Many equipment failures have resulted from overheating due to poor air circulation. Equipment fans or blowers should be regularly checked for the same reason. Storage batteries need to be serviced at the manufacturers' suggested interval. Backup power systems including engine generators should be regularly tested under load.

58.10 Conclusion

New spacecraft are being launched and the demand for transmission service from Earth station facilities continues to increase. New digital video, data, and Internet applications are using satellite technology to deliver reliable service for broadcast, business, and educational applications. Proper planning during the design and construction of an Earth station facility is the key to starting a successful uplink or downlink operation. A well-planned and maintained Earth station staffed by trained operators will continue to be part of a vital link to global telecommunications.

Defining Terms

Automatic transmitter identification system (ATIS): This is a Morse code modulated subcarrier that is required by the FCC to be transmitted with all video carriers.

Cassegrain: This is a style of satellite antenna that uses two reflectors to direct energy to the feedhorn. One reflector is the main body of the antenna, which directs energy to a smaller reflector mounted at the focal point of the main reflector. The energy is then directed into the feedhorn itself.

Codec: A codec is an acronym for coder/decoder. The coder takes an analog voice, data, or video source and converts it to a digital output for transmission. Conversely, a decoder takes a received digital bit stream and converts it back into an analog signal.

Digital integrated receiver—decoder (IRD): A satellite receiver that contains on-board circuitry to process and decode digital video signals. The processor may also contain circuitry to control conditional access and restrict programming.

DVB: DVB stands for digital video broadcasting, which is a standard developed for the transmission of digital video and data signals over satellite and terrestrial systems.

Earth station: An earth station is a ground-based communications facility that is made up of the transmitters, receivers, and antennas used to relay signals to or from a communications satellite.

Exciter: An exciter is a device that takes a voice, data, or video signal and processes it for uplink. The exciter contains a modulator to convert the incoming signal to an intermediate frequency (IF). It also contains an upconvertor to take the processed IF and multiply it to the radio frequency (RF) desired. A video exciter usually contains audio subcarrier modulators to be combined with the video signal.

Footprint: The footprint is a satellite transponder's coverage area as seen from space. The size of the footprint and the level of the signal received in different areas of the footprint will vary depending on the design of the satellite and its orbital location.

Geosynchronous: An orbit at approximately 22,300 mi. above the Earth's equator. This allows satellites to rotate at the same speed as the Earth itself. The satellites then appear stationary from a point on the ground.

High-power amplifier (HPA): The last device before the antenna in most installations. This amplifier develops the power necessary to send a signal to the satellite.

Klystron: A transmitting tube used at ultra high frequency (UHF) and microwave frequencies to produce relatively high-power output over a wide bandwidth.

Link budget: A series of calculations used to determine the quality of service that can be expected over a given satellite link. Link budget calculations include transmit power, antenna gain, path (atmospheric) loss, transponder characteristics, and receive system performance.

Low-noise amplifier (LNA): A high-gain amplifier, usually of GAsFET design, that amplifies the incoming satellite signal to a level suitable to be fed to the receiver. It is designed to add as little noise as possible to the signal during the amplification process.

Low-noise block downconverter (LNB): The same as an LNA except it converts the incoming satellite signal to an L-band frequency (950–1450 MHz) for use by newer style receivers.

MPEG: MPEG is a standard developed by the Moving Pictures Experts Group that is used when compressing video (and corresponding audio) signals into a digital format for transmission or storage.

Transponder: The device internal to the satellite that receives a ground-based uplink transmission and retransmits it back to Earth on the corresponding downlink frequency.

Traveling-wave tube (TWT): A medium-to-high power transmitting tube used for microwave and satellite transmission. It operates in a linear mode with wide bandwidth.

Upconvertor: A device used to multiply an IF signal to the desired satellite operating frequency. It can be internal to the exciter or a separate outboard device. The upconvertor output is fed to the HPA.

Uninterruptable power system (UPS): This is a device that supplies continuous AC power to the load during a failure of the normal power source. It utilizes a large standby battery bank and inverter system to reproduce AC power when the normal source fails.

Further Information

Michael J., Downey, M.J., and Edington, B.G. 1993. *Engineers' Handbook for Earth Station Design and Implementation,* M/R Communications.

Weinhouse, N. 1992. *Satellite Uplink Operator Training Course,* Norman Weinhouse Associates.

Federal Communications Commission, Part 25, Code of the Federal Regulations, Office of the Federal Register, National Archives and Records Administration, U.S. Government Printing Office.

The first two publications cover the Earth station in greater detail and the third contains the rules and regulations for licensing the operating an Earth station.

59

Satellite Transmission Impairments

Louis J. Ippolito, Jr.
ITT Industries

59.1 Introduction

The effects of the Earth's atmosphere on radiowaves propagating between Earth and space is a constant concern in the design and performance of space communications systems. The problems become more acute for systems operating in the bands above 10 GHz, where radiowave links can be adversely affected by atmospheric gases (primarily oxygen and water vapor), rain, clouds, fog, and scintillation. These conditions, when present alone or in combination on the Earth–space radiowave link, can cause uncontrolled variations in signal amplitude, phase, polarization, and angle of arrival, which result in a reduction in the quality of analog transmissions and an increase in the error rate of digital transmissions.

Propagation impairments on radiowaves above 10 GHz are primarily caused by constituents in the troposphere, which extends from the Earth's surface to about 20 km. Degradations induced in the ionosphere generally affect frequencies well below 10 GHz, and the ionosphere is essentially transparent to the radiowave at frequencies above 10 GHz.

Even apparent clear sky conditions can produce propagation effects, which can degrade or change characteristics of the transmitted radiowave. Gases present in the Earth's atmosphere, particularly oxygen and water vapor, interact with the radiowave and reduce the signal amplitude by an absorption process. Turbulence or rapid temperature variations in the transmission path can cause amplitude and phase scintillation or depolarize the wave. Clouds, fog, dirt, sand, and even severe air pollution can cause observable propagation effects. Background sky noise will always be present and contributes directly to the noise performance of the communications receiver system.

The degrading effect of precipitation in the transmission path is a major concern associated with satellite communications, particularly for those systems operating at frequencies above 10 GHz. At those frequencies, absorption and scattering caused by rain (and to a lesser extent hail, ice crystals, or wet snow) can cause a reduction in transmitted signal amplitude (attenuation), which can reduce the reliability and/or performance of the communications link. Other effects can be generated by precipitation in the

Earth–space path. They include depolarization, rapid amplitude and phase fluctuations (scintillation), antenna gain degradation, and bandwidth coherence reduction.

The major factors affecting Earth–space paths in the bands above 10 GHz are as follows:

- Gaseous attenuation
- Hydrometeor attenuation
- Depolarization
- Radio noise
- Scintillation

Extensive experimental research has been performed on the direct measurement of propagation effects on Earth–space paths, beginning in the late 1960s, with the availability of propagation beacons on geostationary satellites [Ippolito, 1986]. Satellites involved included the NASA Applications Technology Satellites (ATS-5 and ATS-6); the Canadian/U.S. Communications Technology Satellite (CTS); the domestic U.S. COMSTAR communications satellites; the ETS-II, CS, and BSE satellites of Japan; SIRIO and ITALSAT of (Italy); and the OTS (European Space Agency) Olympus Satellite. The NASA Advanced Communications Technology Satellite (ACTS) recently completed a very successful five-year campaign of measurements in the 20 and 30 GHz bands at seven North American locations [Ippolito and Bauer, 2000].

Concise modeling and prediction procedures for the propagation factors, based on data and measurements from the satellite beacons, have evolved over several iterations and have been successfully applied in the design and performance of satellite telecommunications systems for several years. [Rogers, Ippolito and Davarian, 1997]. The following sections provide a review and assessment of the available propagation models and their application to satellite communications systems.

59.2 Attenuation on Slant Paths

A radiowave propagating through Earth's atmosphere can experience a reduction in signal level from one or more of several constituents that may be present in the path. The primary constituents that can cause radiowave attenuation on a slant path are atmospheric gases and hydrometeors.

Gaseous atmospheric attenuation is caused primarily by oxygen and water vapor and is an absorption process. Hydrometeor attenuation experienced by a radiowave involves both absorption and scattering processes. Rain attenuation can produce major impairments in Earth–space communications, particularly in the frequency bands above 10 GHz. Cloud and fog attenuation are much less severe than rain attenuation; however, they must be considered in link calculations, particularly for frequencies above 15 GHz. Dry snow or ice particle attenuation is usually so low that it is unobservable on space communications links operating below 30 GHz.

Gaseous Attenuation

Gaseous attenuation is the reduction in signal amplitude caused by the gaseous constituents of the atmosphere that are present in the transmission path. Only oxygen and water vapor have observable resonance frequencies in the radiowave bands used for space communications. Oxygen has a series of very close absorption lines near 60 GHz and an isolated absorption line at 118.74 GHz. Water vapor has lines at 22.3, 183.3, and 323.8 GHz. Oxygen absorption involves magnetic dipole changes, whereas water vapor absorption consists of electric dipole transitions between rotational states.

The Radiocommunication Sector of the International Telecommunications Union (ITU-R) provides a prediction method in Recommendation P.676 to calculate the median gaseous attenuation expected on a slant path for a given set of parameters [ITU-R, P.676-5, 2001]. The ITU-R method is based on the specific attenuation and temperature dependence algorithms developed by Gibbons [1986]. The ITU-R Recommendation provides procedures for frequenciies up to 350 GHz, including the high oxygen absorption bands around 60 GHz. Here we present the procedures for the space applications bands below about 100 GHz.

The input parameters required for the calculation are: frequency f in GHz; path elevation angle θ in degrees; height above mean sea level of Earth terminal, h_s in km; surface temperature T in C; surface pressure p in hPa; and water vapor density at the surface, for the location of interest, ρ_w in g/m^3. If ρ_w is not available from local weather services, representative median values can be obtained from ITU-R Recommendation P.836 [ITU-R P.836-2 2001] which presents global maps of average annual and seasonal values of ρ_w for regions of the world.

Three steps are required to calculate the total gaseous absorption with the ITU-R model:

1. Determine the *specific attenuation* values for oxygen and water vapor.
2. Determine the *equivalent heights* for oxygen and water vapor.
3. Determine the *total slant path gaseous attenuation* through the atmosphere.

Two additional parameters required for the calculations are r_p and r_t, given by

$$r_p \equiv \frac{p}{1013} \qquad r_t \equiv \frac{288}{(273+T)} \tag{59.1}$$

The specific attenuation for dry air, γ_o (dB/Km), is determined from one of the following algorithms, based on the frequency of operation. For $f \leq 54$ **GHz**,

$$\gamma_o = \left[\frac{7.34 r_p^2 r_t^3}{f^2 + 0.36 r_p^2 r_t^2} + \frac{0.3429 b \gamma_o'(54)}{(54-f)^a + b}\right] f^2 \times 10^{-3} \tag{59.2}$$

For **66 GHz** $\leq f <$ **120 GHz**,

$$\gamma_o = \left[\frac{0.2296\, d \gamma_o'(66)}{(f-66)^c + d} + \frac{0.286 r_p^2 r_t^{3.8}}{(f-118.75)^2 + 2.97 r_p^2 r_t^{1.6}}\right] f^2 \times 10^{-3} \tag{59.3}$$

where

$$\gamma_o'(54) = 2.128 r_p^{1.4954} r_t^{-1.6032} \exp[-2.5280(1-r_t)] \tag{59.4}$$

$$\gamma_o'(66) = 1.935 r_p^{1.6657} r_t^{-3.3714} \exp[-4.1643(1-r_t)] \tag{59.5}$$

$$a = \ln(\eta_2/\eta_1)/\ln 3.5 \tag{59.6}$$

$$b = 4^a/\eta_1 \tag{59.7}$$

$$\eta_1 = 6.7665 r_p^{-0.5050} r_t^{0.5106} \exp[1.5663(1-r_t)] - 1 \tag{59.8}$$

$$\eta_2 = 27.8843 r_p^{-0.4908} r_t^{-0.8491} \exp[0.5496(1-r_t)] - 1 \tag{59.9}$$

$$c = \ln(\xi_2/\xi_1)/\ln 3.5 \tag{59.10}$$

$$d = 4^c/\xi_1 \tag{59.11}$$

$$\xi_1 = 6.9575 r_p^{-0.3461} r_t^{0.2535} \exp[1.3766(1-r_t)] - 1 \tag{59.12}$$

$$\xi_2 = 42.1309 r_p^{-0.3068} r_t^{1.2023} \exp[2.5147(1-r_t)] - 1 \tag{59.13}$$

The specific attenuation for water vapor, γ_w (dB/Km), for frequencies from 1 to 350 GHz, is determined from:

$$\gamma_w = \left\{ 3.13 \times 10^{-2} r_p r_t^2 + 1.76 \times 10^{-3} \varrho r_t^{8.5} + r_t^{2.5} \left[\frac{3.84 \xi_{w1} g_{22} \exp(2.23(1 - r_t))}{(f - 22.235)^2 + 9.42 \xi_{w1}^2} \right.\right.$$
$$+ \frac{10.48 \xi_{w2} \exp(0.7(1 - r_t))}{(f - 183.31)^2 + 9.48 \xi_{w2}^2} + \frac{0.078 \xi_{w3} \exp(6.4385(1 - r_t))}{(f - 321.226)^2 + 6.29 \xi_{w3}^2}$$
$$+ \frac{3.76 \xi_{w4} \exp(1.6(1 - r_t))}{(f - 325.153)^2 + 9.22 \xi_{w4}^2} + \frac{26.36 \xi_{w5} \exp(1.09(1 - r_t))}{(f - 380)^2}$$
$$+ \frac{17.87 \xi_{w5} \exp(1.46(1 - r_t))}{(f - 448)^2} + \frac{883.7 \xi_{w5} g_{557} \exp(0.17(1 - r_t))}{(f - 557)^2}$$
$$\left.\left. + \frac{302.6 \xi_{w5} g_{752} \exp(0.41(1 - r_t))}{(f - 752)^2} \right] \right\} f^2 \rho \times 10^{-4} \tag{59.14}$$

where

$$\xi_{w1} = 0.9544 r_p r_t^{0.69} + 0.0061 \varrho \tag{59.15}$$

$$\xi_{w2} = 0.95 r_p r_t^{0.64} + 0.0067 \varrho \tag{59.16}$$

$$\xi_{w3} = 0.9561 r_p r_t^{0.67} + 0.0059 \varrho \tag{59.17}$$

$$\xi_{w4} = 0.9543 r_p r_t^{0.68} + 0.0061 \varrho \tag{59.18}$$

$$\xi_{w5} = 0.955 r_p r_t^{0.68} + 0.006 \varrho \tag{59.20}$$

$$g_{22} = 1 + (f - 22.235)^2 / (f + 22.235)^2 \tag{59.21}$$

$$g_{557} = 1 + (f - 557)^2 / (f + 557)^2 \tag{59.22}$$

$$g_{752} = 1 + (f - 752)^2 / (f + 752)^2 \tag{59.23}$$

The total slant path gaseous attenuation through the atmosphere, A_g, is then found from

$$A_g = \frac{\gamma_0 h_0 + \gamma_w h_w}{\sin \theta} \text{ dB} \tag{59.24}$$

where the equivalent heights for oxygen and water vapor, h_o and h_w, in km, are determined from the following, based on the frequency of operation. For **1 GHz ≤ f ≤ 56.7 GHz**,

$$h_o = 5.386 - 3.32734 \times 10^{-2} f + 1.87185 \times 10^{-3} f^2 - 3.52087 \times 10^{-5} f^3 + \frac{83.26}{(f - 60)^2 + 1.2} \text{ km} \tag{59.25}$$

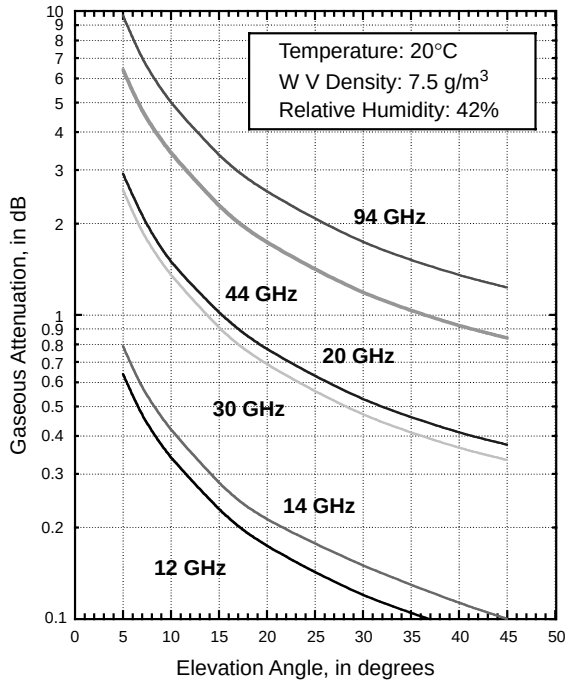


FIGURE 59.1 Total gaseous attenuation for a satellite link under moderate humidity conditions.

For **56.7 GHz < f < 63.3 GHz,**

$$h_o = 10 \text{ km} \tag{59.26}$$

For **63.3 GHz ≤ f < 98.5 GHz,**

$$h_o = f\left\{\frac{0.039581 - 1.19751 \times 10^{-3} f + 9.14810 \times 10^{-6} f^2}{1 - 0.028687 f + 2.07858 \times 10^{-4} f^2}\right\} + \frac{90.6}{(f-60)^2} \text{ km} \tag{59.27}$$

Figure 59.1 shows the total gaseous attenuation calculated from the ITU-R gaseous attenuation model for a range of frequencies and for elevation angles from 5° to 45°. A temperature of 20°C and a water vapor density of 7.5 g/m^3 were assumed. These values correspond to a moderate (42% relative humidity), clear day. The elevation above sea level was set at 0 km.

Rain Attenuation

The attenuation of a radiowave of wavelength λ propagating through a volume of rain of extent L is given by [Ippolito, 1986]

$$a(\text{dB}) = 4.343 \int_0^L [N_0 \int Q_t(r, \lambda, m) e^{-\Lambda r} \mathrm{d}r] \mathrm{d}\ell \tag{59.28}$$

where Q_t is the total attenuation cross-section for a water drop of radius r and refractive index m; N_0 and Λ are empirical constants from the rain drop size distribution, $n(r) = N_0 e^{-\Lambda r}$. The integration over ℓ is taken over the extent of the rain volume in the direction of propagation. Both Q_t and the rain drop size distribution will vary along the path, and these variabilities must be included in the integration process.

A determination of the variations along the propagation path is very difficult to obtain, particularly for slant paths to an orbiting satellite. These variations must be approximated or treated statistically for the development of useful rain attenuation prediction models.

Over the past several years, extensive efforts have been undertaken to develop reliable techniques for the prediction of path rain attenuation for a given location and frequency, and the availability of satellite beacon measurements has provided a database for the validation and refinement of the prediction models. Virtually all of the prediction techniques use surface measured rain rate as the statistical variable and assume a kR^{α} approximation between rain rate R and rain induced attenuation. This approximation shows excellent agreement with direct Mie calculations of specific attenuation for both spherical and nonspherical raindrops. The k and α coefficients for nonspherical drops, that are more representative of a true rain drop, can be calculated by applying techniques similar to the Mie calculations for a spherical drop. An extensive tabulation of k and α coefficients for nonspherical (oblate spheroidal) drops is available from ITU-R Recommendation P.838 [ITU-R P.838-1, 1999].

ITU-R Rain Attenuation Model

The International Telecommunications Union (ITU) first adopted a procedure for the prediction of attenuation caused by rain at its 25th Plenary Assembly in Geneva in 1982. The ITU-R rain attenuation model is the most widely accepted international method for the prediction of rain effects on communications systems. The model is continuously updated as rain attenuation modeling is better understood. The procedure provides the basis for rain attenuation calculations required for international planning and coordination meetings and regional and world administrative radio conferences.

This section describes the ITU-R model as presented in the latest version of Recommendation ITU-R P.618-7, (2001). The ITU-R states that the modeling procedure estimates annual statistics of path attenuation at a given location for frequencies up to 55 GHz.

The input parameters required for the ITU-R Rain Model are:

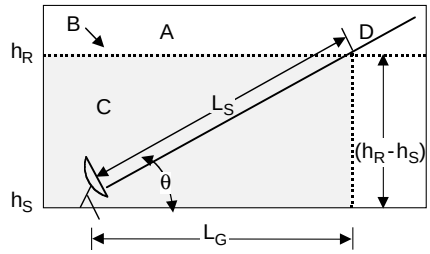
f: the frequency of operation, in GHz
θ: the elevation angle to the satellite, in degrees
φ: the latitude of the ground station, in degrees N or S
τ: the polarization tilt angle with respect to the horizontal, in degrees
h_s: the altitude of the ground station above sea level, in km
$R_{0.01}$: point rainfall rate for the location of interest for 0.01% of an average year, in mm/h

The step by step procedure follows:

1. *Calculate the rain height h_R at the ground station of interest.*
   The freezing rain height (or 0°C isotherm), h_o, is the upper atmosphere altitude at which rain is in the transition state between rain and ice. The rain height is defined in km above sea level. The rain height, latitude and elevation angle are then used to calculate the slant path L_S through a rain cell and the horizontal component of the slant path L_G. The ITU-R rain height assumes that the rain height is a function of latitude (φ). The rain height h_R is found as follows:

$$h_R(\varphi) = h_o + 0.36 \text{ km} \tag{59.29}$$

   where φ is the latitude of the ground station receiver. For locations where h_o is not available from local data, the ITU-R provides a data base for h_o for the entire surface of the earth, in a 1.5° × 1.5° grid of latitude and longitude [ITU-R P.839-3 2001].
2. *Calculate the slant-path length, L_S, and horizontal projection, L_G, from the rain height, elevation angle, and the altitude of the ground receiver site.*
   Rain attenuation is directly proportional to the slant path length. Figure 59.2 shows the path geometry and parameters used in the ITU-R model. The slant path length L_s is defined as the length of the satellite-to-ground path that is affected by a rain cell. The slant path length L_S,



A: frozen precipitation
B: rain height
C: liquid precipitation
D: earth-space path

FIGURE 59.2 Slant path through rain.

expressed in km, is determined from

$$L_S(\theta) = \begin{cases} \dfrac{(h_R - h_S)}{\sin\theta} & \text{for } \theta \geq 5° \\ \dfrac{2(h_R - h_S)}{\left[\sin^2\theta + \dfrac{2(h_R - h_S)}{R_e}\right]^{1/2} + \sin\theta} & \text{for } \theta < 5° \end{cases} \tag{59.30}$$

where

h_R = the rain height (km), from step 1
h_S = the altitude of the ground receiver site from sea level (km)
θ = the elevation angle
R_E = 8500 km (effective earth radius)

For angles greater than or equal to 5°, the slant path length is derived from the geometry from Fig. 59.2. This equation for L_S can result in negative values when the rain height is smaller than the altitude of the ground receiver site. If a negative value occurs, L_S is set to zero.

The horizontal projection is calculated as

$$L_G = L_S \cos\theta \tag{59.31}$$

where L_S and L_G are in kilometers.

3. *Determine the rain rate $R_{0.01}$ (for 0.01% of an average year).*
   Obtain the rainfall rate, $R_{0.01}$, exceeded for 0.01% of an average year (with an integration time of one minute) for the ground location of interest. If this rainfall rate is not available from long-term statistics of local data, the ITU-R provides global maps and a datafile of iso-values of rain intensity, in mm/h, exceeded for 0.01% of the average year for land and ocean regions [ITU-R P.837-3, 2001].
4. *Calculate the specific attenuation γ_R.*
   The specific attenuation is based on the relationship:

$$\gamma_R = kR_{0.01}^{\alpha} \tag{59.32}$$

where γ_R is the specific attenuation in (dB/km) and $\{k, \alpha\}$ are frequency dependent variables calculated using the regression coefficients k_H, k_V, α_H, and α_V from ITU-R Recommendation P.838 [ITU-R P.838-1 1999]. The values have been tested for frequencies up to 40 GHz.

5. *Calculate the horizontal reduction factor, $r_{0.01}$, from the rain rate $R_{0.01}$.*

$$r_{0.01} = \frac{1}{1 + 0.78\sqrt{\frac{L_G \gamma_R}{f}} - 0.38(1 - e^{-2L_G})} \tag{59.33}$$

where L_G is the horizontal projection calculated in step 2.

6. *Calculate the vertical adjustment factor, $v_{0.01}$, for 0.01% of the time.*

$$v_{0.01} = \frac{1}{1 + \sqrt{\sin\theta}\left[31(1 - e^{-(\theta/1+\chi)})\frac{\sqrt{L_R \gamma_R}}{f^2} - 0.45\right]} \tag{59.34}$$

where

$$L_R = \begin{cases} \frac{L_G r_{0.01}}{\cos\theta} \text{ km} & \text{for } \zeta > \theta \\ \frac{(h_R - h_S)}{\sin\theta} \text{ km} & \text{for } \zeta \le \theta \end{cases} \tag{59.35}$$

and

$$\zeta = \tan^{-1}\left(\frac{h_R - h_S}{L_G r_{0.01}}\right) \text{ deg} \tag{59.36}$$

$$\begin{aligned} \chi &= 36 - |\varphi| \text{ deg} && \text{for } |\varphi| < 36 \\ &= 0 && \text{for } |\varphi| \ge 36 \end{aligned} \tag{59.37}$$

7. *Determine the effective path length, L_E.*

$$L_E = L_R v_{0.01} \text{ km} \tag{59.38}$$

8. *Calculate the predicted attenuation $A_{0.01}$ exceeded for 0.01% of an average year.*

$$A_{0.01} = \gamma_R L_E \text{ dB} \tag{59.39}$$

9. *Calculate the attenuation A_p exceeded for other percentages, p, of an average year.*

The attenuation A_p exceeded for other percentages, p, of an average year in the range $p = 0.001\%$ to 5%, is determined from:

$$A_p = A_{0.01}\left(\frac{p}{0.01}\right)^{-[0.655 + 0.033\ln(p) - 0.045\ln(A_{0.01}) - \beta(1-p)\sin\theta]} \text{ dB} \tag{59.40}$$

where

$$\beta = \begin{cases} 0 & \text{if } p \ge 1\% \quad \text{or} \quad |\varphi| \ge 36° \\ -0.005(|\varphi| - 36) & \text{if } p < 1\% \quad \text{and} \quad |\varphi| < 36° \quad \text{and} \quad \theta \ge 25° \\ -0.005(|\varphi| - 36) + 1.8 - 4.25\sin\theta & \text{otherwise} \end{cases} \tag{59.41}$$

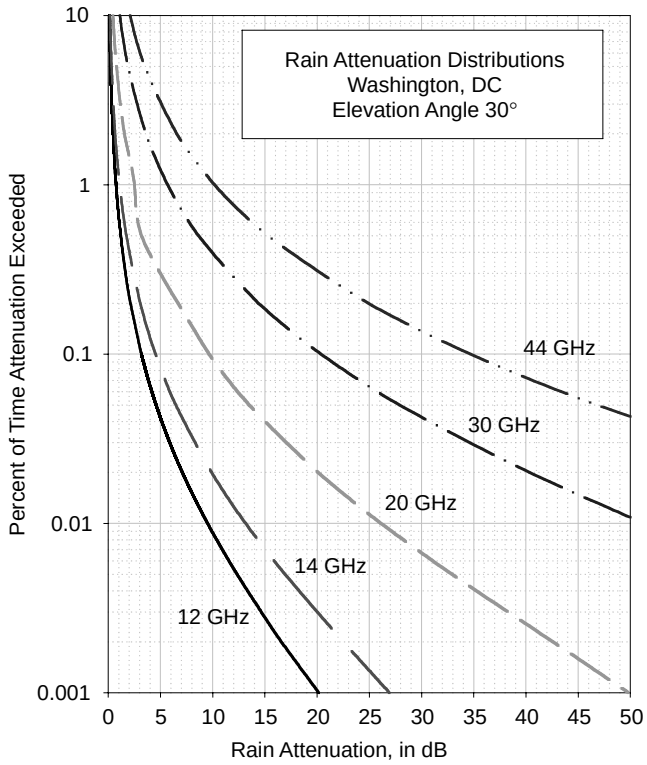


FIGURE 59.3 Rain attenuation distributions from the ITU-R rain attenuation prediction model for Washington, D.C.

This method provides an estimate of long-term statistics due to rain. Large year-to-year variability in rainfall statistics can be expected when comparing the predicted results with measured statistics.

Figure 59.3 shows rain attenuation distributions determined from the ITU-R prediction model for a ground terminal in Washington, D.C., at an elevation angle of 30°, for frequencies from 12 to 44 GHz. The distributions are representative of the prediction curves that can be developed from the ITU-R rain attenuation model.

Cloud and Fog Attenuation

Although rain is the most significant hydrometeor affecting radiowave propagation in the frequency bands above 10 GHz, the influence of clouds and fog can also be present on an Earth–space path. Clouds and fog generally consist of water droplets of less than 0.1 mm in diameter, whereas raindrops typically range from 0.1 to 10 mm in diameter. Clouds are water droplets, not water vapor; however, the relative humidity is usually near 100% within the cloud. High-level clouds, such as cirrus, are composed of ice crystals, which do not contribute substantially to radiowave attenuation but can cause depolarization effects. The average liquid water content of clouds varies widely, ranging from 0.05 to over 2 g/m^3. Peak values exceeding 5 g/m^3 have been observed in large cumulus clouds associated with thunderstorms; however, peak values for fair weather cumulus are generally less than 1 g/m^3.

A detailed study of the radiowave propagation effects of clouds at various locations in the U.S. by Slobin resulted in the development of a cloud model that provides estimates of cloud attenuation and noise temperature on radiowave links [Slobin, 1982]. Extensive data on cloud characteristics, such as type, thickness, and coverage, were gathered from twice-daily radiosonde measurements and hourly

temperature and relative humidity profiles at stations in the contiguous U.S., Alaska, and Hawaii. Regions of statistically consistent clouds were determined, resulting in the definition of 15 cloud regions. The Slobin model defines 12 cloud types based on liquid water content, cloud thickness, and base heights above the surface. Several of the more intense cloud types include two cloud layers, and the combined effects of both are included in the model. Several other cloud attenuation models have been published, and they provide results similar to the Slobin model for frequencies up to about 40 GHz [Gerace and Smith, 1990].

Fog results from the condensation of atmospheric water vapor into droplets that remain suspended in air. There are two types of fog: advection fog, which forms in coastal areas when warm, moist air moves over colder water, and radiation fog, which forms at night, usually in valleys, low areas, and along rivers. Typical fog layers are only 50–100 m in height, and, hence, the extent of fog on a slant path is very small, even for low elevation angles. Fog attenuations are much less than 1 dB for frequencies up to 100 GHz. A prediction model, developed by Altshuler, is available for the estimation of fog attenuation for frequencies above 30 GHz [Altshuler, 1984 rev. 1986].

59.3 Depolarization

Depolarization involves a change in the polarization characteristics of a radiowave caused by (1) hydrometeors, primarily rain or ice particles, or (2) multipath propagation. A depolarized radiowave will have its polarization state altered such that power is transferred from the desired polarization state to an undesired orthogonally polarized state, resulting in interference or crosstalk between the two orthogonally polarized channels. Rain and ice depolarization can be a problem in the frequency bands above about 12 GHz, particularly for frequency reuse communications links, which employ dual independent orthogonal polarized channels in the same frequency band to increase channel capacity. Multipath depolarization is generally limited to very low elevation angle space communications and will be dependent on the polarization characteristics of the receiving antenna.

The cross-polarization discrimination XPD is defined for linearly polarized waves as

$$\mathrm{XPD} = 20 \log \frac{|E_{11}|}{|E_{12}|} \tag{59.42}$$

where E_{11} is the received electric field in the copolarized (desired) direction and E_{12} is the electric field converted to the orthogonal cross-polarized (undesired) direction.

Rain Depolarization

Rain induced depolarization is produced from a differential attenuation and phase shift caused by non-spherical raindrops. As the size of rain drops increase, their shape tends to change from spherical (the preferred shape because of surface tension forces) to oblate spheroids with an increasingly pronounced flat or concave base produced from aerodynamic forces acting upward on the drops [Pruppacher and Pitter, 1971]. Furthermore, raindrops may also be inclined to the horizontal (canted) because of vertical wind gradients. Therefore, the depolarization characteristics of a linearly polarized radiowave will depend significantly on the transmitted polarization angle. When measurements of depolarization observed on a radiowave path were compared with rain attenuation measurements coincidently observed on the same path, it was noted that the relationship between the measured statistics of XPD and copolarized attenuation A for a given probability p could be well approximated by the relationship

$$\mathrm{XPD}_p(\mathrm{dB}) = U - V \log[A_p(\mathrm{dB})] \tag{59.43}$$

where U and V are empirically determined coefficients, which depend on frequency, polarization angle, elevation angle, canting angle, and other link parameters. This discovery was similar to the aR^b relationship observed between rain attenuation and rain rate. For most practical applications, semiempirical relations can be used for the U and V coefficients. The coefficients, as specified in ITU-R Recommendation P.618-7 (2001), are valid for frequencies from 8 to 35 GHz and are determined from

$$\begin{aligned} U &= 30\log f - 10\log[1 - 0.484(1+\cos 4\tau)] - 40\log(\cos\theta) + 0.0052\sigma^2 \\ V &= 12.8 f^{0.19} \quad \text{for } 8 \le f \le 206 \text{ Hz} \\ &= 22.6 \quad \text{for } 20 \le f \le 356 \text{ Hz} \end{aligned} \tag{59.44}$$

where f(GHz) is the frequency, τ is the tilt angle, in degrees, of the polarization with respect to the horizontal (for circular polarization $\tau = 45°$). θ is the elevation angle of the path, in degrees. The elevation angle must be ≤60° for this relationship to be valid. σ is the standard deviation of the raindrop canting angle, in degrees. σ takes the values 0°, 5°, 10°, and 15° for 1%, 0.1%, .01%, and .001% of the time, respectively.

Ice Depolarization

Depolarization due to ice crystals has been observed on Earth–space paths from measurements with satellite beacons. This condition is characterized by a strong depolarization accompanied by very low copolarized attenuation. Also, abrupt changes in the XPD have been observed to coincide with lightning discharges in the area of the slant path, suggesting a change in the alignment of the ice crystals. Ice depolarization occurs when the ice particles are not randomly oriented but have a preferred alignment or direction. Ice crystals most often appear in the shape of needles or plates, and their major axes can abruptly be aligned by an event such as a lightning discharge. When this occurs, differential phase changes in the radiowave can cause an increase in the level of the cross polarized signal level, with little or no change to the copolarized signal. Ice crystals produce nearly pure differential phase shift, without any accompanying differential attenuation, which accounts for the depolarization effects observed in the absence of copolarized attenuation. Ice depolarization effects have been observed in the U.S., Europe, and Japan, at frequencies from 4 to 30 GHz. Ice depolarization often precedes the onset of severe thunderstorm conditions, and XPD degradations of 5–8 dB have been observed. Statistically, however, it occurs for only a very small percent of the time that XPD degradations are observed (1–3%), and so it is generally not a severe impediment to system performance.

The contribution of ice to the XPD distribution experienced on a link can be estimated by a procedure presented in ITU-R Recommendation P.618-7 (2001). The total XPD, caused by rain and ice, for p% of the time, XPD_p, is determined from

$$\text{XPD}_p = \text{XPD}_{\text{rain}} - \left[\frac{0.3 + 0.1\log p}{2}\text{XPD}_{\text{rain}}\right], \text{ dB} \tag{59.45}$$

where XPD_{rain} is the XPD distribution caused by rain, in decibels, either measured or determined from the procedure in the previous section. The term in brackets is the contribution due to the presence of ice on the path.

Figure 59.4 shows the XPD distributions for rain, XPD_{rain}, and rain plus ice, XPD_p, calculated from the rain attenuation distribution for Washington, D.C., at a frequency of 20 GHz. The contribution from ice is seen to be very small, except for high percent values, that is, for 0.1% or greater. At these percentage values the XPD is already high, that is, >30 dB, so system performance will not be significantly degraded by a 1 or 2 dB decrease in XPD.

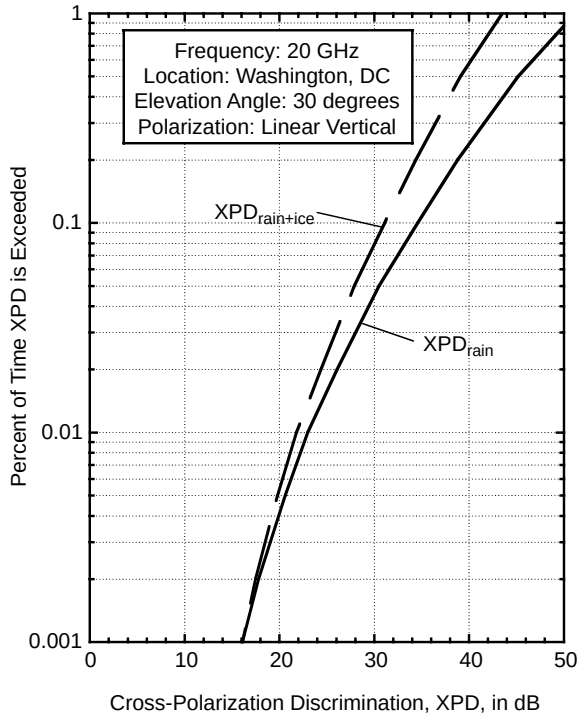


FIGURE 59.4 Rain and ice depolarization distribution for a satellite link to Washington, D.C.

59.4 Radio Noise

The same constituents of the Earth's atmosphere (oxygen, water vapor, rain, clouds, etc.) that cause absorptive attenuation on a radiowave will also produce thermal noise emissions, which are a function of the intensity of the attenuation. This effect follows directly from Kirchoff's law, which specifies that noise emission from a medium in thermal equilibrium must equal its absorption. This result is valid at any frequency. The noise temperature t_S ssproduced by an attenuating medium can be quantitatively described by application of the equation of radiative transfer,

$$t_S = \int_0^\infty t(r)\gamma(r)e^{-\int_0^r \gamma(r')\mathrm{d}r'}\mathrm{d}r \tag{59.46}$$

where $t(r)$ is the physical temperature of the medium in K, $\gamma(r)$ is the absorption coefficient of the medium in km^{-1}, and r is the distance along the path from the antenna in km. If $t(r)$ is replaced by a mean path temperature t_m, the preceding integral equation simplifies to

$$t_S = t_m\left(1 - 10^{-\frac{A(\mathrm{dB})}{10}}\right) \tag{59.47}$$

where $A(\mathrm{dB})$ is the attenuation of the absorbing medium. The mean path temperature t_m is difficult to measure directly for a satellite slant path. Wulfsburg [1964], using a model atmosphere, developed an expression for estimating the mean path temperature from the surface temperature,

$$t_m = 1.12t_g - 50 \tag{59.48}$$

where t_g is the ground surface temperature in K. Using this result, t_m ranges from about 255 to 290 K for the surface temperature range of 0 to 30°C.

Simultaneous measurements of rain attenuation and noise temperature on a slant path using satellite beacons have shown that the best overall statistical correlation of the noise temperature and attenuation measurements occurs for t_m values between 270–280 K for the vast majority of the reported measurments [Ippolito, 1971; Strickland, 1974; Hogg and Chu, 1975]. Similar values for gaseous attenuation or cloud attenuation have also been observed. The noise temperature introduced by the attenuating medium will add directly to the receiver system noise temperature and will degrade the overall performance of the link. The noise power *increase* occurs coincident with the signal power *decrease* due to path attenuation; both effects are additive and contribute to a reduction in link carrier-to-noise ratio.

59.5 Scintillation

Scintillation describes the condition of rapid fluctuations of the signal parameters of a transmitted radiowave caused by time-dependent irregularities in the transmission path. Signal parameters affected include amplitude, phase, angle of arrival, and polarization. Scintillation can be induced in the ionosphere and in the troposphere. The mechanisms and characteristics of ionospheric and tropospheric scintillation differ, and they generally impact different ranges of frequencies. Electron density irregularities occurring in the ionosphere can affect frequencies up to about 6 GHz. Tropospheric scintillation is typically produced by refractive index fluctuations in the first few kilometers of altitude and is caused by high-humidity gradients and temperature inversion layers. The effects are seasonally dependent, vary day to day, and vary with the local climate. Tropospheric scintillation has been observed on line-of-sight links up through 10 GHz and on Earth–space paths at frequencies to above 30 GHz. To a first approximation, the refractive index structure in the troposphere can be considered horizontally stratified, and variations appear as thin layers, which change with altitude. Slant paths at low elevation angles, that is, highly oblique to the layer structure, thus tend to be affected most significantly by scintillation conditions.

Amplitude Scintillation Prediction

Quantitative estimates of the level of amplitude scintillation produced by a turbulent layer in the troposphere are determined by assuming small fluctuations on a thin turbulent layer and applying turbulence theory considerations [Tatarski, 1961]. Karasawa, Yamada, and Allnutt [1988] developed a model to calculate the amplitude fluctuations caused by tropospheric scintillation, based on monthly averages of temperature and relative humidity. The model utilizes low elevation angle measurements in Japan at 11 and 14 GHz. The procedure includes estimates of signal fading as a time percentage as well as standard deviation of signal fluctuations.

The rms amplitude fluctuation, in dB, is determined from

$$\sigma_x = 0.0228\,\sigma_{x\mathrm{REF}} f^{0.45} (\csc\theta)^{1.3} [G(D)]^{0.5}, \quad \theta \geq 5° \tag{59.49a}$$

$$\sigma_x = 0.0228\,\sigma_{x\mathrm{REF}} f^{0.45} \left(\frac{2}{\sqrt{\sin^2\theta + (4.71 \times 10^{-4})} + \sin\theta} \right)^{1.3} [G(D)]^{0.5}, \quad \theta < 50° \tag{59.49b}$$

with

$$\sigma_{x\mathrm{REF}} = 0.15 + \frac{118.51}{(T+273)^2} U e^{\frac{19.7T}{(T+273)}} \tag{59.50}$$

where f is frequency in GHz, θ is the elevation angle in degrees, D is the antenna diameter in meters, T is the average *monthly* temperature in C, and U is the percent average monthly relative humidity. $G(D)$ is the antenna averaging factor of Tatarski [1961].

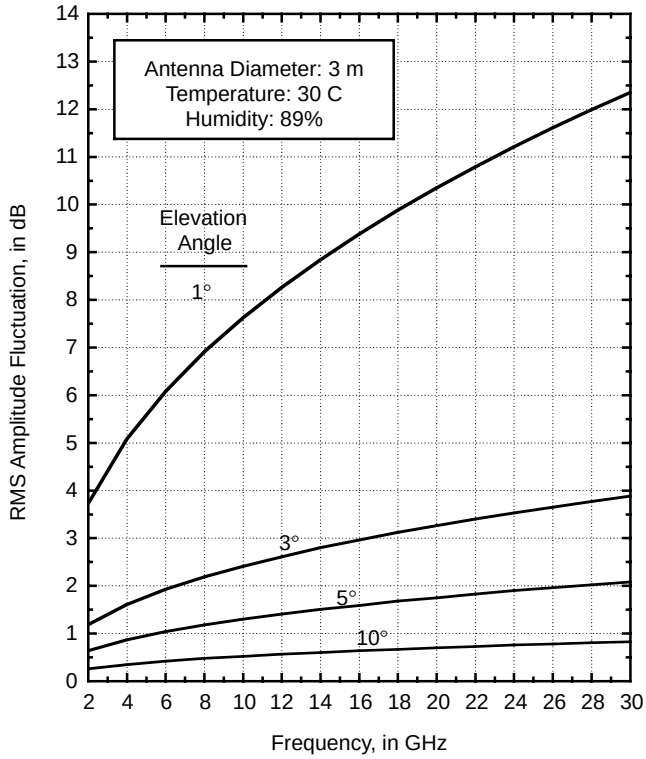


FIGURE 59.5 RMS amplitude fluctuation caused by scintillation as a function of frequency and elevation angle.

Figure 59.5 shows the rms amplitude fluctuation calculated from the Karasawa et al. prediction model as a function of frequency, for elevation angles from 1 to 10°. A 3-m antenna diameter was assumed, with an average monthly temperature of 30°C and an average monthly humidity of 80%. Significant amplitude fluctuations are observed for all frequencies at a 1° elevation angle. The amplitude fluctuations drop quickly as the elevation angle is increased, and at 10° the effect is very small, even at the higher frequencies. The model predicts a slight increase of σ_x with increasing average temperature and a stronger increase in σ_x with increasing average humidity.

Estimates of the time statistics of signal fading due to scintillation can also be determined from the above prediction model. The signal fade level $x(p)$, corresponding to p percent of the cumulative time distribution, is given by

$$x(p) = [3.0 - 0.061 \log^3 p - 0.072 \log^2 p - 1.71 \log p]\sigma_x \tag{59.51}$$

where σ_x(dB) is the rms amplitude fluctuation, determined from Eqs. (59.49a) and (59.49b), and $0.01\% \leq p \leq 50\%$. Figure 59.6 shows the cumulative distribution of signal fading due to scintillation for a 12 GHz link, as calculated from the model. A 3-m antenna diameter was assumed, with a T of 30° C and a U of 80%. The plots again show severe effects at low elevation angles, but at 10° or higher the fade level drops off to only 2 to 3 dB.

The Karasawa et al. scintillation model supports the observation that scintillation is more pronounced under high-temperature and high-humidity conditions, such as in tropical regions or in temperate regions in summer. Significant fading, in excess of 10 dB at low elevation angles, can be expected. At low elevation angles, fading caused by scintillation could exceed attenuation caused by rain, particularly at percentages above about 1%, since the cumulative distributions for scintillation show high fading at the higher time percentages, that is, 1 to 10% and higher.

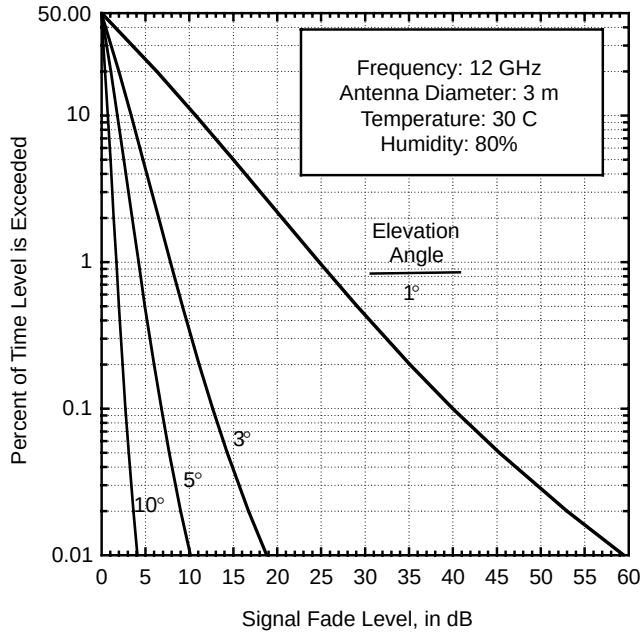


FIGURE 59.6 Cumulative distribution of signal fading due to scintillation on a 12-GHz link.

59.6 Summary

This chapter has reviewed the major propagation impairments affecting Earth–space radiowave links and has provided the latest validated modeling and prediction techniques available for the determination of these effects. Modeling and prediction procedures for gaseous attenuation, rain attenuation, cloud and fog attenuation, rain and ice depolarization, radio noise, and amplitude scintillation are included. The models are validated with extensive satellite beacon measurements obtained throughout the world and are applicable to radiowave links operating up to 100 GHz and a wide range of elevation angles and link conditions.

References

Altshuler, E.A. 1984. A simple expression for estimating attenuation by fog at millimeter wavelengths. *IEEE Trans. Ant. Prop.,* AP-32(7); rev. 1986. AP-34(8).

Crane, R.K. 1980. Prediction of attenuation by rain. *IEEE Trans. Commun.,* COM-28(Sept.):1717–1733.

Crane, R.K. 1982. A two component rain model for the prediction of attenuation statistics. *Radio Sci.,* 17(6):1371–1388.

Gerace, G.C. and Smith, E.K. 1990. A comparison of cloud models. *IEEE Ant. Prop. Mag.,* 32(5):32–37.

Gibbins, C.J. 1986. Improved algorithms for the determination of specific attenuation at sea level by dry air and water vapor, in the frequency range 1–350 GHz. *Radio Sci.,* 21 (Nov.–Dec.):949–954.

Hogg, D.C. and Chu, T.S. 1975. The role of rain in satellite communications. *Proc. IEEE,* 63(9):1308–1331.

Ippolito, L.J. 1971. Effects of precipitation on 15.3 and 31.65 GHz Earth-space transmissions with the ATS-V satellite. *Proc. IEEE,* 59(Feb.):189–205.

Ippolito, L.J. 1986. *Radiowave Propagation In Satellite Communications,* Van Nostrand Reinhold, New York.

Ippolito, L.J. and Bauer, R. 2000, Summary of ACTS Propagation Experiments, Proceedings of Sixth Ka-band Utilization Conference, Cleveland, Ohio.

ITU-R P.618-7 2001. Propagation data and prediction methods required for the design of Earth-space telecommunications systems. International Telecommunications Union, Radiocommunications Sector, Geneva.

ITU-R P.676-5 2001. Attenuation by atmospheric gases. International Telecommunications Union, Radiocommunications Sector, Geneva.

ITU-R P.838-1 1999. Specific attenuation model for rain for use in precipitation methods. International Telecommunications Union, Radiocommunications Sector, Geneva.

ITU-R. P.839-2 2001. Rain height model for prediction methods. International Telecommunications Union, Radiocommunications Sector, Geneva.

Karasawa, Y., Yamada, M., and Allnutt, J.E. 1988. A new prediction method for tropospheric scintillation on Earth-space paths. *IEEE Trans. Ant. Prop.,* 36(11):1608–1614.

Pruppacher, H.R. and Pitter, R.L. 1971. A semi-empirical determination of the shape of cloud and rain drops. *J. Atmos. Sci.,* 28(Jan):86–94.

Rogers, D.V., Ippolito, L.J., and Davarian, F. 1997. System requirements for Ka-band Earth-satellite propagation data, *Proc. IEEE,* 85(June):810–820.

Slobin, S.D. 1982. Microwave noise temperature and attenuation of clouds: statistics of these effects at various sites in the United States, Alaska, and Hawaii. *Radio Sci.,* 17(6):1443–1454.

Strickland, J.I. 1974. The measurement of slant path attenuation using radar, radiometers and a satellite beacon. *J. Res. Atmos.,* 8:347–358.

Tatarski, V.I. 1961. *Wave Propagation in a Turbulent Medium,* McGraw-Hill, New York.

Wulfsberg, K.N. 1964. Sky noise measurements at millimeter wavelengths. *Proc. IEEE,* 52(March): 321–322.

Further Information

For further information on propagation impairments on satellite communications links, see *Radiowave Propagation In Satellite Communications,* L.J. Ippolito, Van Nostrand Reinhold, 1986, and *Satellite-to-Ground Radiowave Propagation,* J.E. Allnutt, Peter Peregrinus, Ltd., London, 1989. For more details on modeling development and propagation databases, see the Series P publications of the International Telecommunications Union Radiocommunications Sector (ITU-R), at www.itu.int.

60

Satellite Link Design

Peter P. Nuspl*
W. L. Pritchard & Co., L.C.

Jahangir A. Tehranil*
INTELSAT, Inc.

60.1 Introduction

This chapter presents a unified approach to satellite link designs for geosynchronous satellites with wideband **transponders** for telecommunications services; it presents and compares **transparent** and **regenerative** link designs. A transponder, as the term is commonly used in telecommunications, is a wideband frequency-translating power amplifier; for example, a C-band transponder could have a useable bandwidth of 72 MHz with an uplink center frequency at 6.25 GHz and a downlink center frequency at about 4.0 GHz. Input and output multiplex filters in the satellite payload provide necessary channelization. As just defined, without other signal processing, this is a transparent transponder, often referred to as a *bent pipe*. A regenerative payload has, in addition, appropriate demultiplexing filters with selected transmission signals as outputs, demodulators (digital), and possibly forward error correcting (FEC) decoders for each of these signals, followed by suitable encoding and modulation. Such onboard processing adds complexity, power consumption, and mass to payloads, but there are many substantial benefits which are being quantified [Nuspl et al., 1987; Mathews, 1988]. In the following sections, general link equations are described and then specialized to digital links of transparent and regenerative types.

In the tables, there are side-by-side comparisons of the transparent and regenerative forms; terminology and units are given in the central column or in the notes to each table. With a few notable exceptions, the forms of the equations are identical, but transparent and regenerative values must be used as appropriate. Where the notation is not explicit or for emphasis, regenerative terms are given in bold. A fundamental

[1] The views expressed herein are those of the authors and do not necessarily represent the views of W.L. Pritchard & Co., or INTELSAT.

attribute of transparent links is that uplink noise, interference, and intermodulation add directly (*noise power addition*); a key feature of regenerative links is that uplink and downlink are separated, and the *errors are additive* but not the noise. As numerical examples, very small aperture terminal (VSAT) parameters are given in some tables, and star VSAT designs for transparent and regenerative links are presented as examples.

The notations used subsequently and in the tables are as follows. For A and B in units such as decibel Hertz,

$$A \oplus B = -10 \log_{10}[10^{-A/10} + 10^{-B/10}] \quad \text{(noise) power addition}$$

$$A \ominus B = -10 \log_{10}[10^{-A/10} \ominus 10^{-B/10}] \quad \text{(noise) power subtraction}$$

The operator $\oplus$ is used to denote multiple summations: $\oplus_1^n$. Also, the identities $[A \oplus B] \pm X \equiv [A \pm X] \oplus [B \pm X]$ are useful.

For A, B, and C with consistent units, including decibel losses, the root sum square (RSS) operator & is defined as follows:

$$A \;\&\; BC = \text{RSS}(A, B, C) = \sqrt{A^2 + B^2 + C^2} \quad \text{root sum squared}$$

60.2 Uplink and Downlink Equations

It is useful to state the uplink and downlink equations separately, for **clear-sky conditions** and with thermal noise contributions only, as in Table 60.1. The saturating flux density ϕ_{sat} is a significant uplink parameter and is based on single-carrier measurements; it is usually selectable from a range of settings. Since the transponder amplifier [a traveling wave tube amplifier (TWTA), solid state power amplifier (SSPA), or a linearized TWTA] is nonlinear, for multicarrier operation it is necessary to operate it backed

TABLE 60.1 Uplink and Downlink Equations, Thermal Noise, Clear Sky

Transparent[a]	Name	Unit	**Regenerative**[b]
$\phi = \phi_{\text{sat}} + IBO - APA_u$	Operating flux density	dBw/m	$\phi = \phi_{\text{sat}} + \boldsymbol{IBO_u} - APA_u$
$\left.\frac{C}{N_0}\right\|_u = \phi - G_{1m2} + \left.\frac{G}{T}\right\|_S - k$	Uplink carrier power-to-noise density	dBHz	$\left.\frac{\boldsymbol{C}}{\boldsymbol{N_0}}\right\|_{\boldsymbol{u}} = \phi - G_{1m2} + \left.\frac{G}{T}\right\|_S - k$
$P_d = EIRP_{\text{sat}} + OBO$	Downlink EIRP	dBW	$\boldsymbol{P_d} = EIRP_{\text{sat}} + \boldsymbol{OBO_d}$
$\left.\frac{C}{N_0}\right\|_d = P_d + APA_d + \left.\frac{G}{T}\right\|_e - k - Lsp_d - La_d$	Downlink carrier power-to-noise density	dBHz	$\left.\frac{\boldsymbol{C}}{\boldsymbol{N_0}}\right\|_{\boldsymbol{d}} = \boldsymbol{P_d} + APA_d + \left.\frac{G}{T}\right\|_e - k - Lsp_d - La_d$

APA = antenna pattern advantage with respect to edge of coverage
$EIRP$ = equivalent isotropically radiated power at the edge of coverage
IBO = input backoff (negative)
OBO = output backoff (negative)
u = uplink
d = downlink
ϕ_{sat} = saturating flux density at the edge of coverage, dBW/m^2

G_{1m2} = gain of 1 square meter at f_u, dBi
G/T = gain-to-temperature, dB K^{-1} [dB/K]
s = satellite receiver
e = Earth station receiver
k = −228.6 Boltzmann's constant, dBW/K-Hz
Lsp_d = spreading loss at f_d, dB
La_d = atmospheric loss at f_d, dB

[a]In a transparent link, OBO is determined by IBO, through the I/O transfer characteristic of the transponder amplifier.
[b]In a regenerative link, $\boldsymbol{IBO_u}$ is a design parameter for the uplink and $\boldsymbol{OBO_d}$ is the key design parameter for the downlink; through the I/O transfer characteristic of the transponder amplifier, the corresponding $\boldsymbol{IBO_d}$ is found and *set onboard*.

off from saturation; even then, intermodulation products are produced, however, at reduced levels. The G/T of a (satellite or Earth station) receiver is often used as a figure of merit. The effective system temperature T is in kelvin and is calculated in Chapter 76. The saturated equivalent isotropically radiated power (e.i.r.p.) of the transponder is the major downlink parameter, whereas the input backoff (IBO) is the only link variable amenable to design once the satellite has been launched.

60.3 Interference Equations

More and more, nonthermal contributions to links are becoming very significant. Table 60.2 is a concise form of the various carrier-to-interference densities C/I_0, due to **interference** and **intermodulation**; note that co-channel interference (CCI) and adjacent satellite interference (ASI) are treated differently.

Cochannel interference arises from reuse of frequency bands by spatial or polarization isolations of the signals; CCI is regarded as being under the designer's control. The carriers in the cochannel transponders can operate at the same frequency; however, they are assigned to the transponders that operate with beams that are spatially isolated. In cross polarization, the signals have opposite polarization. The amount of cochannel interference depends on the modulation characteristics of the carriers, the relative levels and frequency separations between the carriers, and the isolations among the beams. Assuming equal powers in the beams, the beam isolations contribution for the uplink and downlink can be calculated as follows:

1. Spatially isolated beams:

$$\Psi_{\text{up}} = G_{nw} - G_{ni} \qquad \text{uplink isolation at } f_u\text{, dB}$$

$$\Psi_{\text{dn}} = G_{nw} - G_{nw} \qquad \text{downlink isolation at } f_d\text{, dB}$$

TABLE 60.2 Interference and Intermodulation Equations[a]

Transparent	Name	Unit	**Regenerative**
$\left.\frac{C}{I_0\text{cci}}\right\|_u = \oplus_1^n[\phi - \phi_i + \Psi_{i,up}] + B$	Carrier-to-CCI density up	dBHz	$\left.\boldsymbol{\frac{C}{I_{0\text{cci}}}}\right\|_{\boldsymbol{u}} = \oplus_1^n[\phi - \phi_i + \Psi_{i,up}] + B$
$\left.\frac{C}{I_0\text{cci}}\right\|_d = \oplus_1^m[P_d - P_j + \Psi_{j,dn}] + B$	Carrier-to-CCI density down	dBHz	$\left.\boldsymbol{\frac{C}{I_{0\text{cci}}}}\right\|_{\boldsymbol{d}} = \oplus_1^m[P_d - P_j + \Psi_{j,dn}] + B$
$EIRPt_u = \phi - G_{1m2} + Lsp_u + La_u$	Total EIRP up	dBW	$\boldsymbol{EIRPt_u} = \phi - G_{1m2} + Lsp_u + La_u$
$\left.\frac{C}{I_{0\text{asi}}}\right\|_u = EIRPt_u - EIRP_{0,u}$	Carrier-to-ASI density up	dBHz	$\left.\boldsymbol{\frac{C}{I_{0\,\text{asi}}}}\right\|_{\boldsymbol{u}} = \boldsymbol{EIRPt_u} - EIRP_{0,u}$
$\left.\frac{C}{I_{0\text{asi}}}\right\|_d = P_d + APA_d + \Gamma_{\text{asi},d} - EIRP_{0,\text{d}}$	Carrier-to-ASI density down	dBHz	$\left.\boldsymbol{\frac{C}{I_{0\,\text{asi}}}}\right\|_{\boldsymbol{d}} = \boldsymbol{P_d} + APA_d + \Gamma_{\text{asi},d} - EIRP_{0,d}$
$C/IM_{0\,\text{hpa}} = EIRPt_u - HPA_{IM0L} + 10\log 4000$	Carrier-to-HPA density	dBHz	$\boldsymbol{C/IM_{0\,\text{hpa}}} = \boldsymbol{EIRPt_u} - HPA_{IM0L} + 10\log 4000$
$C/IM_0 = C/IM + B$	Carrier-to-IM density	dBHz	$\boldsymbol{C/IM_0} = \boldsymbol{C/IM} + B$

CCI = cochannel interference
ASI = adjacent satellite interference
IM = Intermodulation noise power in the carrier bandwidth
ϕ_i = interfering flux density, dBW/m^2
P_j = interfering EIRP, dBW
B = useable bandwidth, dBHz
Ψ = isolation, dB

Lsp_u = spreading loss at f_u, dB
La_u = atmospheric loss at f_u, dB
$\Gamma_{\text{asi},d}$ = E/S antenna discrimination, dB
$EIRP_{0,u}$ = EIRP density ASI toward wanted satellite, dBW/Hz
$EIRP_{0,d}$ = EIRP density ASI toward wanted E/S, dBW/Hz
HAP_{IM0L} = HPA IM density limit, dBW/4Hz

[a]In multicarrier operation, intermodulation products in the band B depend on the transponder operating point, and C/IM is obtained from simulations or measurements.

where

w = wanted (desired)

i = interfering

G_{nw} = spacecraft antenna gain in the desired sense of polarization of beam n (or m) toward the wanted (w) Earth station, dBi[2]

G_{ni} = spacecraft antenna gain in the desired sense of polarization of beam n (or m) toward interfering (i) Earth station, dBi

2. Oppositely polarized beams:

$$\Psi_{\text{up}} = G_{nw} - 10\log_{10}[10^{Gnix/10} + 10^{(Gni-\rho)/10}]\,\text{dB}$$

$$\Psi_{\text{dn}} = G_{mw} - 10\log_{10}[10^{Gmwx/10} + 10^{(Gmw-\rho)/10}]\,\text{dB}$$

where

G_{nix} = spacecraft antenna gain in the opposite sense of polarization of beam n (or m) toward interfering (i) Earth station, dBi

G_{mwx} = spacecraft antenna gain in the opposite sense of polarization of beam n (or m) toward wanted (w) Earth station

ρ = Earth station polarization purity (receive or transmit), dB

These equations for the oppositely polarized signals assume circular polarization with a relative tilt angle of 45° between the major axis of the polarization ellipses associated with the Earth station and spacecraft antennas. The tilt angle is unknown and can be anywhere between 0° and 90°. An average value of 45° results in factors of 10 (power addition) as shown in the preceding equations. However, the worst case corresponds to a relative tilt angle of 0°, which results in factors of 20 to be used in the preceding equations (voltage addition).

Assuming that the cochannel carriers (wanted and interfering) are similar (i.e., same type of modulation and comparable bandwidth, but not of equal powers), and ignoring antenna pattern advantages, then the ratios of the wanted (ω) to the interfering (i-up, j down) carrier powers can be calculated as follows.

For n uplink interferers, $i = 1,\ldots,n$,

$$(C/I)_{\text{cc},u} = \oplus_1^n[\phi_w - \phi_i + \Psi_{i,\text{up}}]\text{dB} \tag{60.1a}$$

For the same ϕ_{sat} in all cochannel transponders,

$$= IBO_w + \oplus_1^n[-IBO_{\text{cci}} + \Psi_{i,\text{up}}]\,\text{dB}$$

For a constant IBO_I in the interfering channels,

$$= IBO_w - IBO_I + \Psi_{\text{net},u}\,\text{dB} \tag{60.1b}$$

For m downlink interferers, $j = 1,\ldots,m$,

$$(C/I)_{\text{cc},d} = \oplus_1^m[EIRP_{\text{sat},w} + OBO_w - (EIRP_{\text{sat},j} + OBO_j) + \Psi_{j,\text{dn}}]\text{dB} \tag{60.2a}$$

and for the same e.i.r.p.$_{\text{sat}}$ in all co-channels and the corresponding constant output backoff OBO_J,

$$= OBO_w - OBO_J + \Psi_{\text{net},d}\,\text{dB} \tag{60.2b}$$

where ϕ, e.i.r.p., and backoff terms are defined in the tables and

$$\Psi_{\text{net},u} = \oplus_1^n[\Psi_{i,\text{up}}]\text{dB}$$

$$\Psi_{\text{net},d} = \oplus_1^m[\Psi_{j,\text{dn}}]\text{dB}$$

Form (a) of these C/I is reported in Table 60.2, whereas Table 60.7.2 applies form (b).

[2]Here, dBi means antenna gain with respect to isotropic, expressed in decibel.

In the case of cross-polarization interference, the effective isolation is calculated as power addition of the polarization isolations of the spacecraft and the Earth station antennas:

$$\Psi_{\text{eff}} = \Psi_{\text{sc},x} \oplus \Psi_{\text{es},x} \text{ dB}$$

When a number of co-channel transponders exists, the net **isolation** is the power addition of the isolation values for each co-channel transponder, as explained previously and detailed in the following example.

Example: Calculations of net isolations assuming similar co-channel carriers with equal input/output back offs and beam edge conditions.

Assumptions: INTELSAT VI satellite [Brown et al., 1990]

- Desired beam is West Hemi.
- Interfering beams are East Hemi, North West Zone, North East Zone, South West Zone, and South East Zone.

Spacecraft beam isolation:

1. W. Hemi to E. Hemi isolation = 27.0 dB (spatial)
2. W. Hemi to N.E. zone isolation = 30.0 dB (spatial ⊕ cross polarization)
3. W. Hemi to S.E. zone isolation = 30.0 dB (spatial ⊕ cross polarization)
4. W. Hemi to N.W. zone isolation = 27.0 dB (cross polarization)
5. W. Hemi to S.W. zone isolation = 27.0 dB (cross polarization)
6. Spacecraft beam isolation = $\oplus_1^5$ = 21.0 dB (noise power addition)

Uplink isolation:

7. Interfering Earth station-1 cross-polarization isolation = 27.3 dB
8. Interfering Earth station-2 cross-polarization isolation = 27.3 dB
9. Total Earth station cross-polarization isolation = 27.3 ⊕ 27.3 = 24.3 dB
10. Uplink net isolation, $\Psi_{\text{net},u} = \Psi_{\text{eff}} = 21.0 \oplus 24.3 = 19.3$ dB

Downlink isolation:

11. Spacecraft beam isolation (same as in uplink) = 21.0 dB
12. Desired Earth station cross-polarization isolation = 27.3 dB
13. Downlink net isolation, $\Psi_{\text{net},d} = 21.0 \oplus 27.3 = 20.1$ dB

Adjacent satellite interference is not under the designer's control (unless the neighboring satellite is in the same system and the links can be codesigned), but ASI is subject to coordination. Hence, the e.i.r.p. density from the adjacent satellite is a given external parameter; $I_{0,\text{asi}}$ does not change as signal power C is designed. The **discrimination** $\Gamma_{\text{asi},d}$ of the user's own antenna is the key parameter to choose in the design of the $(C/I_0)_{\text{asi},d}$

As examples of typical ASI, for Intermediate Data Rate (IDR) [quaternary phase-shift keying (QPSK) modulation with rate-3/4 FEC] transmission and reception by Revised Standard A INTELSAT [1992] Earth stations, INTELSAT VII hemispheric transponders, the e.i.r.p. densities (e.i.r.p.$_0$) have a nominal value of –32.8 dB(W/Hz) in the uplink, and a corresponding –43.7 dB(W/Hz) in the downlink.[3] Table 60.2 shows the equations for $C/I_{0\text{asi}}$ in the links.

[3]These are averages of maximum and minimum values contained in the INTELSAT system [1992].

60.4 Intermodulation Equations

These brief link designs can hardly do justice to intermodulation issues. The fundamental point is that all amplifiers exhibit nonlinear behavior; this results in intermodulation products, some of which fall within the design band. The only major exceptions are single-carrier operations, as in full-transponder time-division multiple access (TDMA) or in multiplexed time-division multiplexing (TDM) downlinks: there are no intermodulation products, and C/IM is then assigned a very high value. CI/M can be modeled and simulated, or actually measured as input backoff is changed. Table 60.2 assumes C/IM characteristics are known for multicarrier operations. For designs with only a few carriers, C/IM_0 should be calculated for the specific carrier of interest.

60.5 Allowances, Losses, and Margins

Table 60.3 is a summary example of applicable **allowances**, propagation (and other) losses, and (equipment) **margins**.

In order to meet the yearly link **availability** requirement, the amounts of uplink and downlink rain margins need to be determined. The rain statistics of the uplink and the downlink sites need to be studied and the total yearly allocation for link availability needs to be apportioned between the uplink and the downlink portions of the link. In the uplink, the amount of rain attenuation corresponding to the uplink allocation (in percent time of the year) is determined from the statistics; this is called the *uplink attenuation loss* (i.e., the required uplink margin in transparent links). In the downlink, the amount of rain attenuation plus the increase in the receiving Earth station system noise associated with the downlink allocation (in percent time of the year) is determined; this is called the *downlink rain degradation* (i.e., the required downlink margin). A particular link needs to satisfy both the uplink and the downlink margins. This assumes the effect of rain depolarization is negligible.

For a transparent link, the uplink margin is satisfied by allotting the required uplink attenuation loss to the (up)link. Assuming that it does not simultaneously rain on both the uplink and the downlink, with the allotted uplink margin, the available downlink (rain) margin $M_{a,d}$ is calculated for the transparent link. If the available downlink margin is less than the required downlink margin, then the uplink allotment is increased until the downlink margin is satisfied. The system margin M_{sys} for transparent links is the maximum of these two uplink margins.

For a regenerative link, the allotments to each portion of the link are the required margins in that portion.

The numerical values in Table 60.3 are for 99.6% (up and down) link availability of C-band VSAT services. Modem margins in Table 60.3 apply at the design bit error ratio (BER) of 10^{-6} for the **degraded sky**. Onboard decoding would require an *additional* implementation margin of 0.4 dB in the uplink.

60.6 Sums of Link Equations

Table 60.4 presents sums of the C/I_0 equations for the uplink, downlink, and transparent link, including interference (CCI and ASI), intermodulation, allowances, losses, and margins, so that these apply for the degraded sky (also called *threshold*, below which performance is not acceptable for the specified percentage of time).

For a transparent link, note the noise-power combining of uplink and downlink terms but the exclusion of M_{sys}, which is taken into account in the required C/N_0 in Table 60.5.

Regeneration separates the available uplink and downlink C/N_0 sum equivalents; allowances could be equal to or less than allowances for the transparent case. For the regenerative case, the losses and allowances are applied to the uplink and downlink separately, as illustrated in the right column.

In some situations, links are dominated by a particular contribution, and there is not much scope for design. But in designing most multicarrier transparent links, the total available C/N_0 can be optimized; of course, the selection of IBO/OBO is the ultimate design for the desired links. Figure 60.1 illustrates

TABLE 60.3 Allowances, Losses, and Margins[a] (Example for VSAT Networks)

Transparent		Name	Unit	**Regenerative**	Up/Down
% $TERN$ = 10% (of total, clear sky)	10	Terrestrial noise	%		10/10
$TERN = -10 \log_{10} [1 - \%TERN/100]$	0.5	Allowance	dB	$TERN_u/TERN_d$	0.5/0.5
		Other losses			
HPA instabilities and uplink tracking loss	0.5		dB		0.5/—
Downlink tracking loss	0.8		dB		—/0.8
E/S equipment noise (5% total)	0.2		dB		0.1/0.1
OL	RSS 0.9	Allowance	dB	OL_u/OL_d	RSS 0.5/0.8
$AOL = TERN + OL$	1.4	$TERN$ and OL	dB	AOL_u/AOL_d	1.0/1.3
		Propagation[b]			
$L_{r,u}$	1.0	Rain att'n loss, up	dB	$L_{r,u}$	1.0/—
$L_{s,u}$	1.2	Tropo scint, up	dB	$L_{s,u}$	1.2/—
$L_{p,u} = L_{r,u}$ & $L_{s,u}$	1.6	Prop losses, up	dB	$L_{p,u} = L_{r,u}$ & $L_{s,u}$	1.6/—
$L_{D,d}$	1.0	Degraded, down	dB	$L_{D,d}$	—/1.0
$L_{s,d}$	1.2	Tropo scint, down	dB	$L_{s,d}$	—/1.2
$L_{p,d} = L_{D,d}$ & $L_{s,d}$	1.6	Prop losses, down	dB	$L_{p,d} = L_{D,d}$ & $L_{s,d}$	—/1.6
M_m	1.7	Coded modem margins	dB	$M_{m,u}/M_{m,d}$	1.5/1.7
$(C/N_0)_{\text{other};\Delta} = [(C/N_0) \ominus (C/N_0)_d] + \Delta + L_{p,u}$		Other than thermal down[c]	dBHz	n.a.	
$(C/N_0)_{d;D} = (C/N_0)_d + \Delta - L_{p,d} + L_{p,u}$ see Tables 60.1, 60.4		Degraded C/N_0, down	dBHz	n.a.	
$M_{d;D} = (C/N_0)_{ta} + \Delta + L_{p,u} - [(C/N_0)_{\text{other};\Delta} \oplus (C/N_0)_{d;D}]$		Degradation margin, down	dB	n.a.	
$M_{\text{sys}} = \max\{L_{p,u}, L_{p,u} + \Delta; M_{d;D} \le L_{p,d}\}$		Link margins	dB	$M_u = L_{p,u}/Md = L_{p,d}$	1.6/1.6

AOL = allowance for $TERN$ and OL
HPA = high-power amplifier at E/S
OL = other losses
$TERN$ = terrestrial noise
& = root sum squared
$\oplus$, $\ominus$ = noise power addition, subtraction

L_r, L_s = rain, scintillation losses, dB
$L_{D,d}$ = loss in Degraded downlink: rain attenuation plus degradation due to increased system noise temperature
L_p = propagation losses: up, down, dB
Δ = additional uplink margin, dB

[a]See discussion and sample calculation of $M_{a,d}$ and M_{sys} (system margin) for transparent links.
[b]VSAT link availability is typically 99.6%, equally apportioned to the uplink and downlink.
[c]$(C/N_0)_{\text{other}}$ is the portion of $(C/N_0)_{ta}$ which is not affected by downlink fading.
[d]Adjacent channel iinterference effects are assumed negligible.

typical TWTA transfer characteristics and intermodulation characteristics. Having thermal and interference components which increase with less backoff (more carrier power), and another component due to intermodulation products which decreases with less backoff (more nonlinear operation), a maximum value for available C/N_0 can be found.

Example: Optimum backoffs. Figure 60.2 illustrates a typical design, which selects the best input/output backoffs for transparent and regenerative links. In multicarrier mode, regenerative links can apply much more uplink backoff and use slightly more (negative) downlink backoff.

TABLE 60.4 Sums of Link Equations

Transparent, Clear Sky[a]	Name	Unit	**Regenerative**, Degraded Sky[a]
$(C/N_0)_{su} = (C/N_0)_u \oplus (C/I_{0\ \mathrm{cci}})_u \oplus (C/I_{0\ \mathrm{asi}})_u \oplus C/IM_{0\ \mathrm{hpa}}$[b]	Sum uplink C-to-N_0	dBHz	$\boldsymbol{(C/N_0)_{su} = [(C/N_0)_u \oplus (C/I_{0\ cci})_u \oplus (C/I_{0\ asi})_u \oplus C/IM_{0\ hpa}] - M_u}$
n.a.	Available, up C-to-N_0	dBHz	$\boldsymbol{(C/N_0)_{a,u} = (C/N_0)_{su} - AOL_u}$
$(C/N_0)_{sd} = (C/N_0)_d \oplus (C/I_{0\ \mathrm{cci}})_d \oplus (C/I_{0\ \mathrm{asi}})_d \oplus C/IM_0$	Sum downlink C-to-N_0	dBHz	$\boldsymbol{(C/N_0)_{sd} = [(C/N_0)_d - M_d] \oplus (C/I_{0\ cci})_d \oplus (C/I_{0\ asi})_d \oplus C/IM_0}$
$(C/N_0)ta = [(C/N_0)_{su} \oplus (C/N_0)_{sd}] - AOL\ (excluding\ M_{\mathrm{sys}})$[c]	(total) Available C-to-N_0	dBHz	$\boldsymbol{(C/N_0)_{a,d} = (C/N_0)_{sd} - AOL_d}$

[a]Compare the different forms for the transparent and regenerative links. For regenerative links, the C/N_0 sums are shown in bold.

[b]$C/IM_{0\mathrm{hpa}}$ accounts for intermodulation products from all other Earth stations (primarily from within the design network) radiating into the band B.

[c]For transparent links, $M_{d:D}$ is the degradation margin due to downlink; M_{sys} accounts for uplink and downlink rain losses, as in Table 60.5; see the discussion.

TABLE 60.5 Designed BER and Required C/N_0

Transparent	Name	Unit	**Regenerative**
BER[a]	Designed bit error ratio		$\boldsymbol{BER} = BER_{\mathrm{up}} + BER_{\mathrm{dn}}$
Designed BER determines E_b/N_0	Energy-per-bit to noise-density	dB	(Decoded) BER_{up} determines (E_b/N_0) up (Decoded) BER_{dn} determines (E_b/N_0) down
n.a.	Required C/N_0 up	dBHz	$\left.\frac{C}{N_0}\right\|_{r,u} = \left.\frac{E_b}{N_0}\right\|_u + R_b + M_{m,u}$
$\left.\frac{C}{N_0}\right\|_r = \left.\frac{E_b}{N_0}\right\|_r + R_b + M_m + M_{\mathrm{sys}}$	Required C/N_0 (total), down	dBHz	$\left.\frac{C}{N_0}\right\|_{r,d} = \left.\frac{E_b}{N_0}\right\|_d + R_b + M_{m,d}$

E_b = energy per *information* bit, W-s
M_m = modem margin, dB
$M_{m,u}$, $M_{m,d}$ = modem margins, dB
M_{sys} = system margin, dB (transparent)
R_b = information rate, mrR_{sym}, b/s
m = $\log_2 M$, M-ary alphabet of channel symbols ($M = 2, 4, 8$)
r = input bits/output bits, **code rate**
R_{sym} = channel symbol rate, symbol/s

[a]Curves of *BER* vs E_b/N_0 *for coded modems* are used; in principle, such curves are independent of the information rate.

60.7 Designed Bit Error Ratio and Required C/N_0

For digital links, bit error ratio is the most common form of expressing performance. A service BER requirement is in conjunction with a needed availability, usually expressed in percentage of time annually; for VSAT networks operating in C-bands, Tables 60.3 and 60.7.3 assume 99.6% link availability. Given the design values of (uplink and downlink) E_b/N_0 for each carrier, the corresponding required C/N_0 are readily calculated for the three links, as in Table 60.5; Table 60.7.5 gives numerical examples. The use of system and modem margins should be noted.

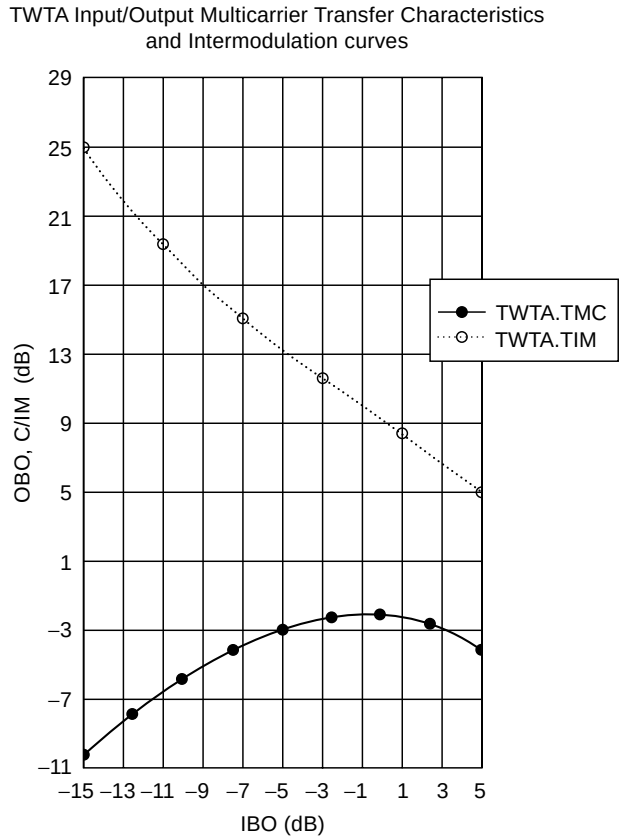


FIGURE 60.1 Output backoff and *C/IM* vs. input backoff.

A classic form of comparing transparent and regenerative links is the **iso-BER curve**, as uplink and downlink values are varied [Mathews, 1988], illustrated in Fig. 60.3. These are required values, so less costly links are lower and to the left on these curves. Point A (4.9, 24.8) off the transparent curve represents a split in noise contributions, with almost all in the downlink. As an example on the regenerative curve, for a downlink-conserving design such as at point B (4.8, 12.8), BER_{up} would be set at a negligible fraction (<1%) of the total *BER*, so that the highest BER_{dn} can be tolerated, to get the lowest required $(E_b/N_0)_{\mathrm{dn}}$. Between transparent and regenerative points A and B, compare the uplink difference of 12 dB and the downlink difference of only 0.1 dB, in favor of using regeneration.

As another example, a link needing an uplink-conserving design is digital satellite news gathering (SNG), where transportability and low-power consumption are important in the uplinking Earth station, and the receive station is usually large. Another link design is required and would show a tradeoff between transparent and regenerative operation.

60.8 Numbers of Carriers, e.i.r.p., and Power per Carrier

Table 60.6 shows the calculations of the numbers of carriers that can be supported by the links; of course, the lesser of n_u and n_d is selected. The power per carrier at an Earth station can then be calculated; with regeneration, the uplink generally requires much less power. The power per carrier in the downlink is also calculated for both types of links.

Example: Hub to VSAT stations. Tables 60.1–60.6 have documented some simplified link design equations; as another set of illustrative examples, the corresponding Tables 60.7.0–60.7.6 provide numerical values of a typical star VSAT link design in the direction hub to VSAT. In Table 60.7.0 are the salient

TABLE 60.6 Number of Carriers, EIRP and Power per Carrier

Transparent[a]	Name	Unit	**Regenerative**[b]
n.a.	Uplink carriers		$\boldsymbol{n_u} = 10^{[(C/N_0)_{a,u} - (C/N_0)]_{r,u}/10}$
n.a.	Downlink carriers		$\boldsymbol{n_d} = 10^{[(C/N_0)_{a,d} - (C/N_0)_{r,d}]/10}$
$n = 10^{[(C/N_0)_{ua} - (C/N_0)r]/10}$	Number of carriers		$n = \min(n_u, n_d)$
$EIRPc_u = EIRPt_u - 10 \log n$	EIRP per carrier, up	dBW	$\boldsymbol{EIRPc_u = EIRPt_u} - 10 \log \boldsymbol{n}$
$Pc_u = EIRPc_u - G_u + Lf_u$	Power per carrier, up	dBW	$\boldsymbol{Pc_u = EIRPc_u} - G_u + Lf_u$
$EIRPc_d = P_d - 10 \log n$	EIRP per carrier, down	dBW	$\boldsymbol{EIRPc_d = P_d} - 10 \log \boldsymbol{n}$

Lf_u = feed loss in earth station, dB
G_u = antenna gain at f_u, dBi

[a]In a transparent link, the number of carriers n is calculated from the total available (*including* M_{sys}) and required C/N_0 values.
[b]In a regenerative link, n_u is calculated from the total uplink EIRP controlled by $\mathbf{IBO}_u$ and the *selected* uplink BER; n_u is usually higher than n_d (except under E/S power limitations); and n_d is calculated from the total available and required C/N_0 values for the downlink. Each calculated n, n_u, and n_d is checked against its bandlimited value.

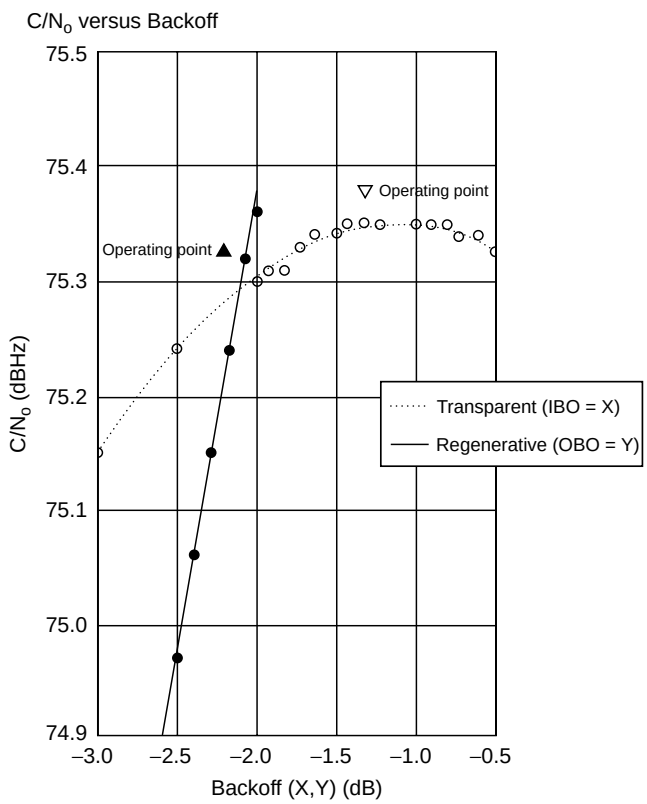


FIGURE 60.2 Selecting IBO for maximum C/N_0.

assumptions about the transponder, the Earth stations, and the service at 64-kb/s information rate. It is assumed that several large stations would act as hubs (gateways) and that multicarrier frequency-division multiple access (FDMA) is the access mode.

The importance of backoff is indicated with the variables X, $Y(X)$, and $Z(X)$; then the best IBO = X and **OBO** = $\boldsymbol{X_d}$ are selected, as already exemplified in Fig. 60.2. Table 60.7.4 has numbered entries with

TABLE 60.7.0 Numerical Examples, Hub to VSAT[a]

Transponder, INTELSAT VI	$B = 72.0$ MHz (78.6 dBHz)
	$\phi_{sat} = -77.6$ dBW/m^2, (high gain); $(G/T)_s = -9.2$ dB/K
	$EIRP_{sat} = 31.0$ dBW, (hemi beam)
Uplink, from several hubs	$D_u = 5.0$ m; $f_u = 6.130$ GHz; $APA_u = 0$ dB; $G_{1m2} = 37.2$ dBi/m$_2$
	$IBO_{cci} = -3$ dB; $(HPA_{IM0L}) = 21.0$ dBW/4 kHz; $Lf_u = 0.7$ dB;
Downlink, to many VSATs	$D_d = 1.8$ m; $f_d = 3.905$ GHz; $APA_d = 1$ dB; $(G/T)_e = 15.0$ dB/K
	$OBO_{cci} = -2$ dB
Service parameters	$R_b = 64$ kbit/s (48.1 dBHz); $BER = 1 \times 10^{-6}$
	Availability = 99.6%; see Table 60.3

[a] $X = IBO$, $Y(X) = OBO$, $Z(X) = C/IM$.

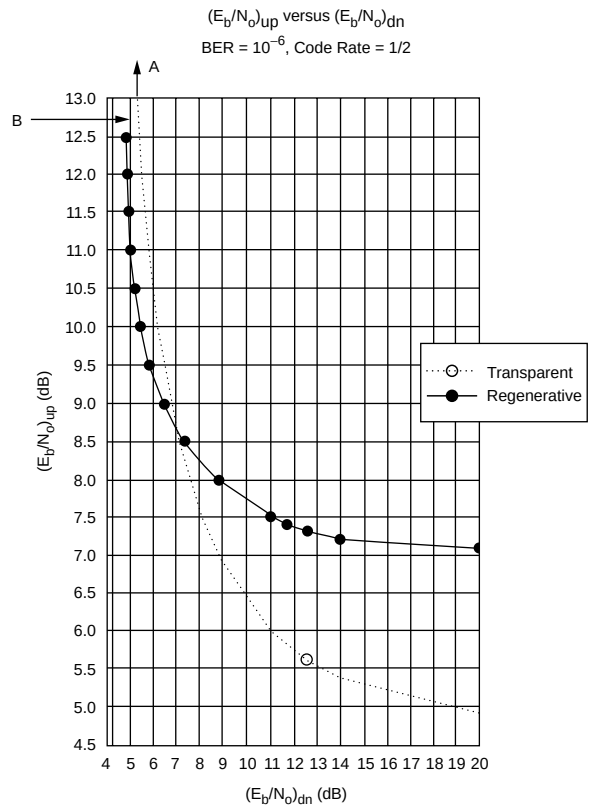


FIGURE 60.3 Iso-BER curves, comparing transparent and regenerative cases.

X, $Y(X)$, and $Z(X)$ and also numerical values when the best operating point is utilized. In Table 60.7.3, M_{sys} is determined, from which the required C/N_0 for the transparent case is calculated.

Table 60.7.6 shows that 69 carriers can be supported in the transparent link, whereas 75 carriers are possible with regeneration; the downlink e.i.r.p.s are similar; the small differences are due to the downlink-limited condition with the small (1.8-m) antennas. However, Table 60.7.6 also shows a design on the right in which the regenerative uplink can be backed off an additional 12 dB, to yield an uplink power of only 6.6 dBW per carrier, which is less than 5 W. In the transparent link, the hub(s) needs 79 W per carrier, and so a large high power amplifier (HPA) is required. There would be similar differences in the direction from VSAT to hub.

TABLE 60.7.1 Uplink and Downlink Thermal Equations, Clear Sky

Transparent	Name	Unit	**Regenerative**
$\phi = (-77.6) + X - 0.0$ $= -77.6 + X$	Operating flux density	dBw/m^2	$\boldsymbol{\phi} = (-77.6) + \boldsymbol{X_u} - 0.0$ $= -77.6 + \boldsymbol{X_u}$
$\left.\frac{C}{N_0}\right\|_u = \phi - 37.2 + (-9.2) - (-228.6)$ $= 104.6 + X$ #1*t*	Uplink carrier power-to-noise density	dBHz	$\left.\frac{\boldsymbol{C}}{\boldsymbol{N_0}}\right\|_u = \phi - 37.2 + (-9.2) - (-228.6)$ $= 104.6 + \boldsymbol{X_u}$ #1*r*
$P_d = 31.0 + Y(X)$	Downlink EIRP	dBW	$\boldsymbol{P_d} = 31.0 + \boldsymbol{Y}(\boldsymbol{X_d})$
$\left.\frac{C}{N_0}\right\|_d = P_d + 1.0 + 15.0 - (-228.6)$ $-196.4 - 0.3$ $= 78.9 + Y(X)$ #2*t*	Downlink carrier power-to-noise density	dBHz	$\left.\frac{\boldsymbol{C}}{\boldsymbol{N_0}}\right\|_d = \boldsymbol{P_d} + 1.0 + 15.0 - (-228.6)$ $- 196.4 - 0.3$ $= 78.9 + \boldsymbol{Y}(\boldsymbol{X_d})$ #2*r*

TABLE 60.7.2 Interference and Intermodulation Equations

Transparent	Name	Unit	**Regenerative**
$\left.\frac{C}{I_{0\,\mathrm{cci}}}\right\|_u = X - (-3.0) + 19.3 + 78.6$ $= 100.9 + X$ #3*t*	Carrier-to-CCI density up	dBHz	$\left.\frac{C}{\boldsymbol{I_{0cci}}}\right\|_u = \boldsymbol{X_u} - (-3.0) + 19.3 + 78.6$ $= 100.9 + \boldsymbol{X_u}$ #3*r*
$\left.\frac{C}{I_{0\,\mathrm{cci}}}\right\|_d = Y(X) - (-2.0) + 20.1 + 78.6$ $= 100.7 + Y(X)$ #4*t*	Carrier-to-CCI density down	dBHz	$\left.\frac{C}{\boldsymbol{I_{0cci}}}\right\|_d = \boldsymbol{Y}(\boldsymbol{X_d}) - (-2.0) + 20.1 + 78.6$ $= 100.7 + \boldsymbol{Y}(\boldsymbol{X_d})$ #4*r* #4*r*
$EIRPt_u = \phi - 37.2 + 200.4 + 0.3$ $= 85.9 + X$	Total EIRP up	dBW	$\boldsymbol{EIRPt_u} = \phi - 37.2 + 200.4 + 0.3$ $= 85.9 + \boldsymbol{X_u}$
$\left.\frac{C}{I_{0\,\mathrm{asi}}}\right\|_u = EIRPt_u - (-32.8)$ $= 118.7 + X$ #5*t*	Carrier-to-ASI density up	dBHz	$\left.\frac{C}{\boldsymbol{I_{0asi}}}\right\|_u = \boldsymbol{EIRPt_u} - (-32.8)$ $= 118.7 + \boldsymbol{X_u}$ #5*r*
$\left.\frac{C}{I_{0\,\mathrm{asi}}}\right\|_d = P_d + 1.0 + 12.4 - (-43.7)$ $= 88.1 + Y(X)$ #6*t*	Carrier-to-ASI density down	dBHz	$\left.\frac{C}{\boldsymbol{I_{0asi}}}\right\|_d = P_d + 1.0 + 12.4 - (-43.7)$ $= 88.1 + \boldsymbol{Y}(\boldsymbol{X_d})$ #6*r*
$C/IM_{0\,\mathrm{hpa}} = EIRPt_u - 21.0 + 36.0$ $= 100.9 + X$ #7*t*	Carrier-to-IM density E/S HPA	dBHz	$\boldsymbol{C/IM_{0\,hpa}} = \boldsymbol{EIRPt_u} - 21.0 + 36.0$ $= 100.9 + \boldsymbol{X_u}$ #7*t*
$C/IM_0 = Z(X) + 78.6$ $= 78.6 + Z(X)$ #8*t*	Carrier-to-IM density (satellite)	dBHz	$\boldsymbol{C/IM_0} = \boldsymbol{Z}(\boldsymbol{X_d}) + 78.6$ $= 78.6 + \boldsymbol{Z}(\boldsymbol{X_d})$ #8*r*

60.9 Other Issues

The link designs and examples are for star VSAT networks, which work with a few large hubs. Even more interesting would be a mesh VSAT designs where there is a need to conserve uplink and downlink power, and interference from similar systems is more severe due to lower isolations. Also, these link designs have excluded bandwidth improvement factors for intermodulation, which take into account sparse use of spectrum; likewise, bandwidth allocation factors and overheads have been excluded. In particular for small to medium size antennas, outer coding can be usefully applied. These designs have no demand

TABLE 60.7.3 Allowances, Losses and Margins

Transparent		Name	Unit	**Regenerative**	
$AOL = 0.5 + 0.9 = 1.4$		*TERN* and *OL*	dB	AOL_u/AOL_d	1.0/1.3
		Propagation:			
$L_{p,u} = 1.0$ & $1.2 = 1.6$		Prop losses, up	dB	$L_{p,u} = L_{r,u}$ & $L_{s,u}$	1.6/—
$L_{p,u} = 1.0$ & $1.2 = 1.6$		Prop losses, down	dB	$L_{p,d} = L_{D,d}$ & $L_{s,d}$	—/1.6
$M_m = 1.7$		Coded modem margin	dB	$M_{m,u}/M_{m,d}$	1.5/1.7
$(C/N_0)_{\text{other};\Delta} = [74.6 \ominus 76.9] + 1.6 = 80.1$	opt	Other than thermal	dBHz	n.a.	
$(C/N_0)_{d;D} = 76.9 + 1.6 - 1.6 = 76.9$	opt	Degraded	dBHz	n.a.	
$M_{d;D} = 74.6 + 1.6 - (80.1 \oplus 76.9) = 1.0$	opt	Available margin, down	dB	n.a.	
$M_{\text{sys}} = 1.6$		Up, downlink margins	dB	$M_u = 1.6/M_d = 1.6$	

TABLE 60.7.4 Sums of Link Equations

Transparent			Name	Unit	**Regenerative**		
$(C/N_0)_{su} = [104.6 + X]$	$\{103.3$	#1t	Sum uplink	dBHz	$(\boldsymbol{C/N_0})_{su} = [104.6 + \boldsymbol{X_u}$	$\{103.3$	#1r
$\oplus [100.9 + X]$	99.6	#3t	C-to-N_0		$\oplus [100.9 + \boldsymbol{X_u}$	99.6	#3r
$\oplus [118.7 + X]$	117.4	#5t	X = −1.3		$\oplus [118.7 + \boldsymbol{X_u}$	117.4	#5r
$\oplus [100.9 + X]$	99.6 $\}$	#7t	(Figure 60.2)		$\oplus [100.9 + \boldsymbol{X_u}] - 1.6$	99.6 $\}$ − 1.6	#7r
$= (97.0 + X)$	= 95.7	opt	$Y = -2$		$= 97.6 + \boldsymbol{X_u}] - 1.6$	= 94.1	
			(Figure 60.1)				
n.a.			Available, up	dBHz	$(\boldsymbol{C/N_0})_{a,u} = 94.0 - 1.0$		
			C-to-N_0		93.0 *including* M_u		
$(C/N_0)_{sd} = [78.9 + Y(X)]$	$\{76.9$	#2t	Sum downlink	dBHz	$\boldsymbol{C/N_0})_{sd} = [78.9 + Y(X_d) - 1.6]$	$\{75.2$	#2r
$\oplus [100.7 + Y(X)]$	98.7	#4t	C-to-N_0		$\oplus [100.7 + Y(X_d)]$	98.6	#4r
$\oplus [88.1 + Y(X)]$	86.1	#6t	$Z = 10.3$		$\oplus [88.1 + Y(X_d)]$	86.0	#6r
$\oplus [78.6 + Z(X)]$	88.9 $\}$	#8t			$\oplus [78.6 + Z(X_d)]$	89.0 $\}$	#8r
	= 76.1	opt				= 74.7	opt
$(C/N_0)_{ta} = \{95.7 \otimes 76.1\}$			(Total)	dBHz			
$- 1.4$			Available		$(\boldsymbol{C/N_0})_{a,d} = 74.7 - 1.3$		
$= 74.6$ (*excluding* M_{sys})		opt	C-to-N_0		$= 73.4$ *including* M_d		opt

Transparent: $X_{\text{opt}} = -1.3$, $Y_{\text{opt}} = -2.0$, $Z = 10.3$, Regenerative: $X_u = -14.3$, $Y_{\text{opt}} = -2.1$, $Z = 10.4$.

TABLE 60.7.5 Designed BER and Required C/N_0

Transparent	Name	Unit	**Regenerative**
$BER = 1.0 \times 10^{-6}$	Designed bit error ratio		$\boldsymbol{BER} = 1.0 \times 10^{-12} + 1.0 \times 10^{-6}$
Designed BER determines E_b/N_0	Energy-per-bit to noise-density	dB	$(E_b/N_0)_u = 12.8$ $(E_b/N_0)_d = 4.8$
n.a.	Required C/N_0 up	dBHz	$\left.\boldsymbol{\frac{C}{N_0}}\right\|_{r,u} = 12.8 + 48.1 + 1.5$ $= 62.4$ *excluding* M_u
$\left.\frac{C}{N_0}\right\|_r = 4.8 + 48.1 + 1.7 + 1.6$ $= 54.6$ (*including* M_{sys})	Required C/N_0 (total), down	dBHz	$\left.\frac{C}{N_0}\right\|_{r,d} = 4.8 + 48.1 + 1.7$ $= 54.6$ *excluding* M_d

TABLE 60.7.6 Number of Carriers, EIRP and Power per Carrier

Transparent		Name	Unit	**Regenerative**		Design
n.a.		Uplink carriers		$n_u = 10^{[93.1 \ominus 62.4]/10}$ $= 1174$		$n_u = 10^{(-81.1-62.4)/10}$ $= 75$
n.a.		Downlink carriers		$n_d = 10^{[73.4-54.6]/10}$ $= 75$		$n_d = 75$
$n = 10^{[74.6-56.2]/10}$ $= 69$	opt	Number of carriers		$\boldsymbol{n} = \min(n_u, n_d)$ $= 75$	opt	
$EIRPc_u = 85.9 + (-1.3)$ $-18.4 = 66.2$	opt	EIRP per carrier, up	dBW	$\boldsymbol{EIRPc_u} = 85.9 + (\mathbf{-1.3})$ $-18.8 = 65.8$		$X_u = -12.0$ $EIRPc_u = 53.8$
$Pc_u = 66.2 - 47.9 + 0.7$ $= 19.0$ {79 W}	opt	Power per carrier, up	dBW	$\boldsymbol{Pc_u} = 65.8 - 47.9 + 0.7$ $= 18.6$ {72 W}		$Pc_u = 6.6$ (5 W)
$EIRPc_d = 31.0 + (-2.1)$ $-18.4 = 10.6$	opt	EIRP per carrier, down	dBW	$\boldsymbol{EIRPc_d} = 31.0 + (\mathbf{-2.1})$ $-18.8 = 10.1$	opt	

assignment multiple access (DAMA) signalling, which would need to be robustly designed. As to other issues, data-only links can take advantage of automatic repeat request (ARQ) so that throughput and basic BER can be selected accordingly. Links with voice/video-only can be operated with relatively poor BER, and effective error concealement techniques can be applied.

60.10 Summary

This chapter has presented a unified but simplified set of link designs for transparent and regenerative links. Several numerical examples illustrate the methods and the differences in link designs.

All link designs must properly take into account thermal, interference, and intermodulation contributions, as well as allowances, losses, and margins. Transparent links, which are most prevalent now, suffer noise power addition of up to ten entries. Regenerative links, which are being developed in many forms and can be expected in next-generation satellites, separate the uplink and downlink so that each has only up to four entries. Regenerative links require an uplink modem margin (for the payload demodulator), but other allowances, losses, and margins are generally lower than for transparent links.

Appendix A: Some Calculations

Boltzmann's constant:

$$k = -228.6 \text{ dBJ/K}$$

Propagation velocity in free space:

$$c = 2.998 \times 10^8 \text{ m/s}$$

Wavelength (Lambda):

$$\lambda = \frac{c}{f} \text{ m}$$

where f is the frequency of operation [up f_u, down f_d], in Hertz.

The maximum antenna gain relative to isotropic for a parabolic antenna, up and down:

$$G = 10 \log_{10}\left(\eta\left(\frac{\pi D}{\lambda}\right)^2\right) \text{ dBi}$$

where η is the antenna efficiency (nominal 0.6) and D is the aperture diameter in meters.

Gain of ideal 1 m^2:

$$G_{1m2} = 21.455 + 20 \log F \text{ dBi/m}^2$$

where F is the (approximate) frequency in gigaHertz.

Typical atmospheric loss, up and down:

$$\begin{aligned} L_a = {} & 0.25 \text{ dB} \quad 2 < F < 5 \quad \text{GHz} \\ & 0.33 \quad\quad\quad 5 < F < 10 \text{ GHz} \\ & 0.53 \quad\quad 10 < F < 13 \text{ GHz} \\ & 0.73 \quad\quad 13 < F \quad\quad \text{GHz} \end{aligned}$$

Free space spreading loss up and down:

$$L_{sp} = 32.45 + 20 \log_{10}(\text{FR}) \quad \text{dB}$$

where R is the (slant-range) distance to the satellite in meters. [Typically, at 10° elevation angle, R = 40,587,000 m to geostationary Earth orbit (GEO).]

Appendix B: Calculation of Antenna Discrimination[4]

Gain mask:

$$\begin{aligned} G_m &= G - 0.0025[\theta D/\lambda]^2 && \text{mainlobe} \quad \text{dBi} \\ G_1 &= 2 + 15 \log_{10}(D/\lambda) && \text{first sidelobe} \\ G_s &= 32 - 25 \log_{10}\theta && \text{other sidelobes} \end{aligned}$$

where G is the maximum gain (as previously calculated) and θ is the satellite spacing in degrees (nominal 3°).

$$\begin{aligned} \delta_m &= (20\lambda/D)\sqrt{(G - G_1)} \quad \text{degree} \\ \delta_m &= \max[1°, 100\lambda/D°] \end{aligned}$$

Antenna discrimination for receive Earth station:

$$\begin{aligned} &\text{IF } \theta < \delta_m && \Gamma_{\text{asi},d} = G - G_m && \text{of mainlobe} && \text{dB} \\ &\text{IF } \delta_m \le \theta < \delta_1 && \Gamma_{\text{asi},d} = G - G_1 && \text{of first sidelobe} && \text{dB} \\ &\text{IF } \delta_1 \le \theta < 48 && \Gamma_{\text{asi},d} = G - G_s && \text{of other sidelobes} && \text{dB} \\ &\text{IF } 48 \le \theta && \Gamma_{\text{asi},d} = G - (-10) && && \text{dB} \end{aligned}$$

[4] Based on International Telecommunications Union (ITU) Rec. 465–4 and Report 391–6.

Defining Terms

Allowance: An allowance in decibels is subtracted from a sum of carrier-to-noise densities which would have been available in a link; usually, such an allowance accounts for factors external to a given system (and beyond the system designer's control).

Availability: Usually expressed as a percentage of a year, availability is the total time duration over which the specified BER is met or exceeded.

Clear-sky condition: Only atmospheric losses are added to spreading losses; this excludes weather effects.

Code rate: The code rate (actually a *ratio*) for forward error correcting codes is the ratio of input to output bits (or symbols), such as rate 1/2, rate 3/4 for binary codes, and rate 239/255 for a typical Reed–Solomon code.

Degraded-sky condition: In the uplink there is additional attenuation due to rain; in the downlink there is rain attenuation and also an increased effective receiver system temperature due to the increased sky noise.

Discrimination: Antenna discrimination is the differential gain compared to maximum for an antenna in the specified direction; usually, masks are provided for the main lobe, the first sidelobe, and other sidelobes.

Interference: The generic term that includes all *unwanted* signals that affect a telecommunications link, excluding thermal noise and intermodulation products; interference can be man-made or natural; it can be internal or external to a system.

Intermodulation: A generic term; *intermodulation products* are produced in nonlinear devices (amplifiers) and they are dependent on the wanted signal levels, modulation formats, and frequency plan.

Iso-BER curve: A graph of constant (iso) bit error ratio with uplink and downlink parameters on the y- and x-axis, respectively; C/N_0 or E_b/N_0 are usually used.

Isolation: An antenna beam's minimum discrimination to a co-channel or cross-polarization channel.

Margin: In decibel, a margin is added to required values, for example to E_b/N_0 required for a target BER, and reflects uncertainties in equipment performance; this is sometimes called *implementation margin*.

Propagation loss or downlink degradation: In decibel, propagation loss is subtracted from all carrier-to-noise (equivalent) densities where C (signal power) is affected but I_o is not; rain attenuation, increased sky noise, depolarization, and tropospheric scintillation are typical examples of such losses.

Regenerative: A link that, in comparison to transparent, has additional functions in the satellite payload, such as demultiplexing and demodulation of the communications signals in the uplink and modulation for the downlink; this is sometimes called *onboard regeneration* (*OBR*). OBR and additional functions such as possibly FEC decoding, re-encoding, baseband switching or IF routing, and TDM constitute what are usually called *onboard processing* (*OBP*).

Transparent: A transparent link uses a transponder that has an output signal format that is exactly the same as the input format.

Transponder: Consists of a wideband receiver, frequency-translating equipment, and a power amplifier; its associated input and output multiplex filters determine its center frequency and bandwidth; it would be more correctly called a *repeater*.

Acknowledgment

The authors wish to acknowledge William Salazar and Luis Gonzales for their dedicated work in preparing the link budget programs.

References

Brown, M.P., Duesing, R.W., Nguyen, L.N., Sandrin, W.A., and Tehrani, J.A. 1990. INTELSAT VI transmission design and computer system models for FDMA services. *COMSAT Tech. Rev.,* 20(2):373–399.

INTELSAT, 1992. INTELSAT Intersystem Coordination Manual, IICM-314, Rev. 1A, Dec.

Mathews, N. 1988. Performance evaluation of regenerative digital satellite links with FEC codecs. *Satellite Integrated Communications Networks,* Proc. Tirrenia Workshop, North-Holland, U.K.

Nuspl, P., Peters, R., Abdel-Nabi, T., and Mathews, N., 1987. On-board processing for communications satellites: Systems and benefits. *Int. J. Satellite Commun.,* 5(2):65–76.

Further Information

For a general treatment of propagation topics, see the book by J.E. Allnutt, 1989, *Satellite-to-Ground Radiowave Propagation, Theory, Practice and System Impact at Frequencies Above 1 GHz*, Peter Peregrinus Ltd., UK.

For another treatment of (transparent) link budgets, see Chapter 6 of *Satellite Communications*, by D. Roddy, 1989, Prentice Hall, Englewood Cliffs, NJ.

There was a special issue on on-board processing, edited by P. Nuspl, *International Journal of Satellite Communications*, vol. 5, no. 2, Wiley-Interscience, London, UK, April–June, 1987. See B. Pontano, P. de Santis, L. White, and F. Faris, "On-Board Processing Architectures for International Communications Satellite Applications," Proceedings of the 13th America Institute of Aeronautics and Astronautics (AIAA) International Communication Satellite Systems Conference and Exhibit, Los Angeles, CA, Paper 90-0853, March 1990. Also, see several sessions in Proceedings of International Conference on Digital Satellite Communications ICDSC-8, Guadeloupe, April 1989; ICDSC-9, Copenhagen, May 1992; 14th AIAA Conference, Washington, D.C., March 1992; and 15th AIAA Conference, San Diego, March 1994.

Also, trends for future payloads are described in "Advanced Communications Technologies for Future INTELSAT Systems," P.P. Nuspl, 1992, Proceedings, Advanced Methods for Satellite and Deep Space Communications, 1992 DLR Seminar, Bonn, Germany, September (DLR is the German Aerospace Research Center near Munich, Germany).

61

The Calculation of System Temperature for a Microwave Receiver

Wilbur Pritchard
W. L. Pritchard & Co., Inc.

61.1 Noise, Antenna, and System Temperatures

Thermal Noise Density

The ultimate capacity of any communications channel is determined by the ratio C/N_o of carrier power to thermal **noise density**, and the purpose of RF link analysis is to calculate this ratio. N_o is the composite noise spectral density and is given by $N_o = kT_s$, where T_s is the composite system temperature. It is the temperature to which a resistor at the system input, with the system having only its gain and generating no **noise**, would have to be heated to generate the same amount of random noise as is observed at the real system receiver output. This temperature characterizes the entire system noise performance and is intended to comprise the effects of **thermal noise**, receiver noise, and noise available at the antenna terminals. The latter noise itself comprises two components, the thermal noise of the surroundings with which the antenna interchanges energy and external sources of noise in the sky.

All noise considerations are based on the irreducible thermal noise due to the random motion of electrons in conductors. The full quantum expression for the spectral density of this noise is

$$N_o = \frac{hf}{e^{\frac{hf}{kT}} - 1} - hf \tag{61.1}$$

In the range of frequencies of usual interest in satellite communication $kT \gg hf$ and we have Nyquist's familiar formula for thermal noise density

$$N_o = kT \tag{61.2}$$

In a bandwidth B, the resulting voltage across a resistance R is given by Nyquist's formula $\sqrt{4kTBR}$. If we consider a source of impedance R generating this voltage, then the power available into a matched load R is simply kTB. This concept of available power is used almost exclusively in considering the noise performance of composite systems.

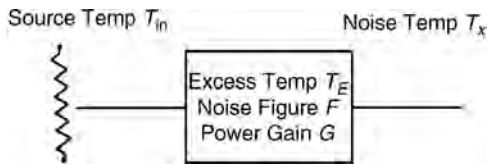


FIGURE 61.1 Noise parameters: basic network.

Noise Temperature

Figure 61.1 is an arbitrary two-port network characterized by a power gain G and excess noise characterized by parameters to be defined.

We define the **noise temperature** at a point as that temperature to which a matched resistor must be heated to generate the same available noise as is measured at that point. Thus, the noise at any point x in the chain is given by $N_x = kT_xB$. If the point in question is the input, then the noise temperature is often written as T_i, and, more specifically, if an antenna is at the input, then the temperature is written as Ta.

A different and altogether basic characterization is found in the **system temperature** T_s, which is the temperature to which a resistor at the input must be heated to generate a noise equal to the observed total noise output. Other sources of noise are neglected. The output noise can now be written as $N_x = GkT_sB$, and $GT_s = T_x$. The system temperature will clearly be higher than the input source temperature by virtue of the extra noise generated by the network. We use this idea to define an **excess temperature** T_E, characteristic of the network itself, and to write the noise temperature at any point T_x, by the equations

$$\begin{aligned} T_s &= T_{\text{in}} + T_E \\ T_x &= G(T_E + T_{\text{in}}) \end{aligned} \tag{61.3}$$

Performance figures and noise temperatures quoted by receiver manufacturers are usually this excess temperature T_E, naturally enough, since it is a measure of the receiver quality, but it is a frequent error to use that temperature in the calculation of system performance. That computation must use the system temperature, which requires the addition of the **antenna temperature**, and consideration of other sources of noise.

Noise Figure

An older, and still much used, figure of merit for network noise performance is **noise figure** F. It is based on the idea that the reduction in available **carrier-to-noise ratio** (CNR) from the input to the output of a network is a measure of the noise added to the system by the network and, thus, a measure of its noise quality. The difficulty with that concept in its naive form is that the same network will seem better if the input CNR is poorer and vice versa. To avoid this anomaly, the definition is based on the assumption that the available noise power at the input is always taken at room temperature, arbitrarily taken in radio engineering work to be 290 K.

$$F = \frac{C_{\text{in}}/N_{\text{in}}}{C_{\text{out}}/N_{\text{out}}} \tag{61.4}$$

This can be rewritten several ways. Note that the input and output carrier levels are related by the available **power gain** and, in accordance with our definition, the input noise level is simply kT_oB. One form, occasionally taken as a definition, is that the output noise is given by

$$N_{\text{out}} = FGkT_oB \tag{61.5}$$

Note well that, although the definition of noise figure assumes the input at room temperature, it can be measured and used at any input temperature if the calculations are made correctly. We can understand this better by rewriting the preceding equation as

$$N_{\text{out}} = GkT_oB + (F-1)GkT_oB \tag{61.6}$$

$$N_{\text{out}} = GkT_{\text{in}}B + (F-1)GkT_oB \tag{61.7}$$

In the form of Eq. (61.6) we can identify the first term as attributable to the amplified input noise and the second term as the excess noise generated by the network. In fact, we can change the term T_o to T_{in} where T_{in} is an arbitrary input source temperature, for instance, that of a receiver antenna. By comparing the second term of Eq. (61.7) to that of Eq. (61.3), we see that they are both terms for the excess noise and thus write

$$T_E = (F-1)T_o \tag{61.8}$$

This is a particularly useful relation inasmuch as the need to go back and forth between the two characterizations of excess noise is frequent.

The Hot Pad

A very important special case is the noise characterization of a plain attenuator, at either room temperature 290 K or some arbitrary temperature T. Transmission line loss between an antenna and the first stage of RF amplification, typically a **low-noise amplifier** (LNA), the loss in the atmosphere due to atmospheric gases such as oxygen and water vapor, and the loss due to rain can all be considered **hot pads**. It can be shown [Mumford and Scheibe, 1968] that the noise temperature of such an attenuator is

$$T_N = \frac{T_{\text{in}} + (L-1)T}{L} \tag{61.9}$$

Networks in Tandem

To be able to apply our set of definitions and relations to a real microwave receiver, we need a relation between the characteristics of two individual networks and their composite characteristics when the networks are in tandem or cascade (see Fig. 61.2).

The calculation of the noise performance of two stages in tandem is done by equating the noise output as calculated from the combined parameters, either F or T_E, with the noise output as calculated by taking the noise output of the first stage as the noise input for the second stage. From Eq. (61.3), we have, for the output noise in terms of the noise at x,

$$\begin{aligned} T_x &= G_1(T_{\text{in}} + T_{e1}) \\ T_{\text{out}} &= G_2(T_x + T_{e2}) \\ T_{\text{out}} &= G_1G_2(T_{\text{in}} + T_{e1}) + G_2T_{e2} \end{aligned} \tag{61.10}$$

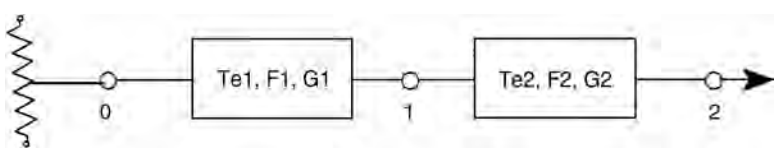


FIGURE 61.2 Two networks in tandem.

We can also write the output noise as $G_1G_2(T_{in} + T_e)$, again by using Eq. (61.3). By equating the two expressions for output noise, we arrive at the fundamental relation for the composite

$$T_e = T_{e1} + \frac{T_{e2}}{G_1} \tag{61.11}$$

The identical approach can be used with excess noise being characterized by noise figures, or Eq. (61.11) can be converted using the relation between excess noise temperature and the noise figure given by Eq. (61.7). In either case, the result is

$$F = F_1 + \frac{(F_2 - 1)}{G_1} \tag{61.12}$$

Antenna Temperature

We have been using the term "antenna temperature" casually with the understanding that the input to a microwave receiver is usually connected to an antenna and that in some way the antenna can be considered as generating an equivalent input temperature. This is indeed the case, and we can make the idea more precise by defining the antenna temperature as the weighted average temperature of the surroundings with which the antenna exchanges energy. For example, if two narrow beam antennas communicated with each other with no loss to external radiation, and if one of these antennas were connected to a matched resistor at a temperature T_A, then T_A would be the antenna temperature of the other antenna.

An Earth station antenna typically exchanges energy with the sky in the vicinity of the satellite at which it is pointed, the ground via its **side lobes**, and the sun, usually in the antenna side lobe structure but occasionally in the main beam. A satellite antenna exchanges energy, principally with the ground, but also with the sky and occasionally with the moon and sun. Obviously, the satellite orbit and antenna coverage determine the situation. Antenna temperature is much affected by atmospheric losses and especially by rain attenuation. By way of example, a narrow beam antenna pointed at the clear sky might have a temperature of 20 K, but this will increase to more than 180 K with 4.0 dB of rain loss, as is easily calculated from the hot pad formula (61.9). With a **receiver excess temperature** of 50 K, the system temperature will increase from 70 to 230 K or 5.0 dB. Note that this is a greater loss than that due to the signal attenuation by the rain. This is a frequently overlooked effect of rain attenuation. The implication for system design, at those frequencies where rain loss is anything other than negligible, is significant. The margin achieved using high-performance, low-noise receivers, is something of an illusion. The margin disappears quickly just when one needs it the most. The performance improvement achieved with a larger antenna is better than that achieved with a low-noise receiver because the deterioration of the former because of rain attenuation is only that due to the signal loss, whereas the latter suffers the signal loss and the increase in antenna temperature.

The antenna temperature can, in principle, be calculated from the definition by evaluating the integral of the temperature, weighted by the **antenna gain**, both as a function of solid angle. Due allowance must be made for reflections. The exact expression is

$$T = \frac{1}{4\pi} \int G_1 T_{sky}\, d\Omega + \frac{1}{4\pi} \int G_2[(1-\rho^2)T_g + \rho^2 T_{sky}]\, d\Omega \tag{61.13}$$

where:

G_1 = antenna gain in a sky direction
G_2 = antenna gain in a ground direction
Ω_1 = solid angle in sky directions
Ω_2 = solid angle in ground directions
T_g = ground temperature
T_{sky} = sky temperature
ρ = voltage reflection factor of the ground

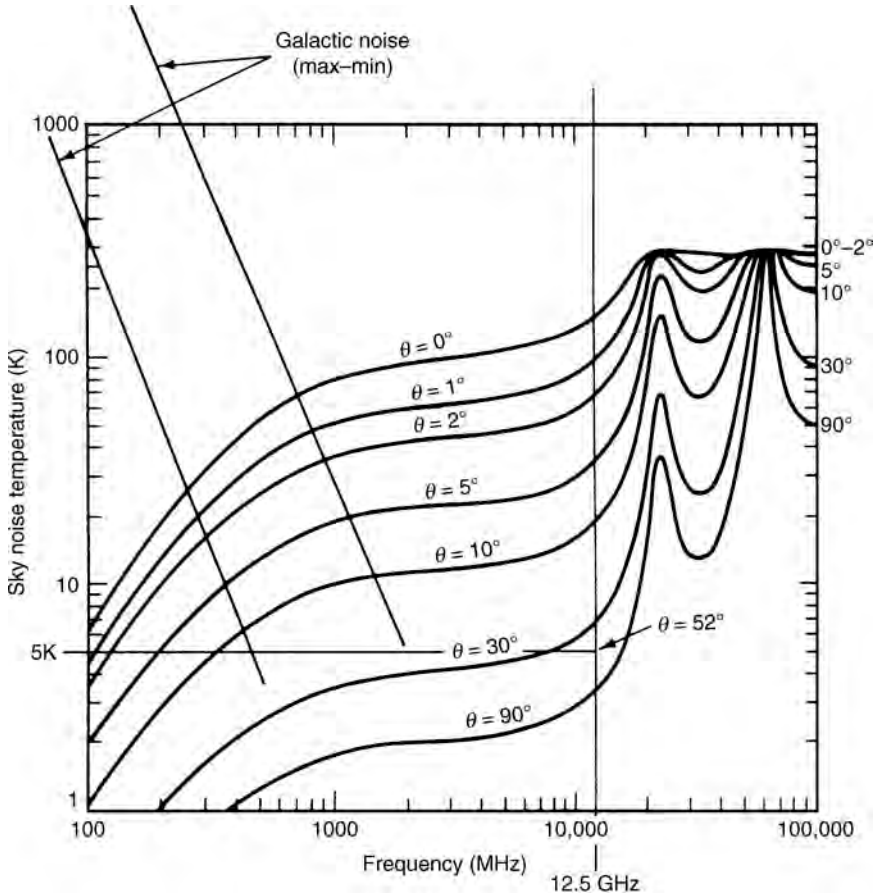


FIGURE 61.3 Sky noise temperature calculated for various elevation angles.

The **sky temperature**, the starting point for most calculations of antenna temperature, is found by first determining the clear sky temperature, normally very low. The spectral absorption of atmospheric gases raises this temperature in accordance with the hot pad effect and yields a sky temperature that is a function of frequency and angle of elevation. Figure 61.3 is a plot of this sky temperature vs. frequency with angle of elevation as a parameter. Note the presence of a microwave window in the frequency region high enough for the galactic noise to have fallen to a low value, nominally the 3.5-K microwave background, a remnant of the big bang, and low enough to avoid the increases due to atmospheric gas absorption. The frequency sensitivity largely mirrors the absorption spectra of these gases (shown in Fig. 61.4), and the dependence on angle of elevation seen in Fig. 61.5 results from the increased path length through the atmosphere at low angles. Note that there are regions in the sky where there is radio noise at a higher level than shown and generated by various galactic sources. Some of these sources of radio noise are used to calibrate the noise performance of Earth stations.

Ground and rain temperatures are typically taken to be T_o = 290 K. The integration requires a knowledge of the antenna pattern in spherical coordinates and the location of ground masses, reflecting surfaces, reflection factors, and sky temperatures. It can be carried out numerically but normally it is complicated and not worth the trouble. A result sufficiently accurate for most practical purposes can be found using

$$T_a = a_1 T_{\text{sky}} + a_2 T_g + a_3 T_{\text{sun}} \tag{61.14}$$

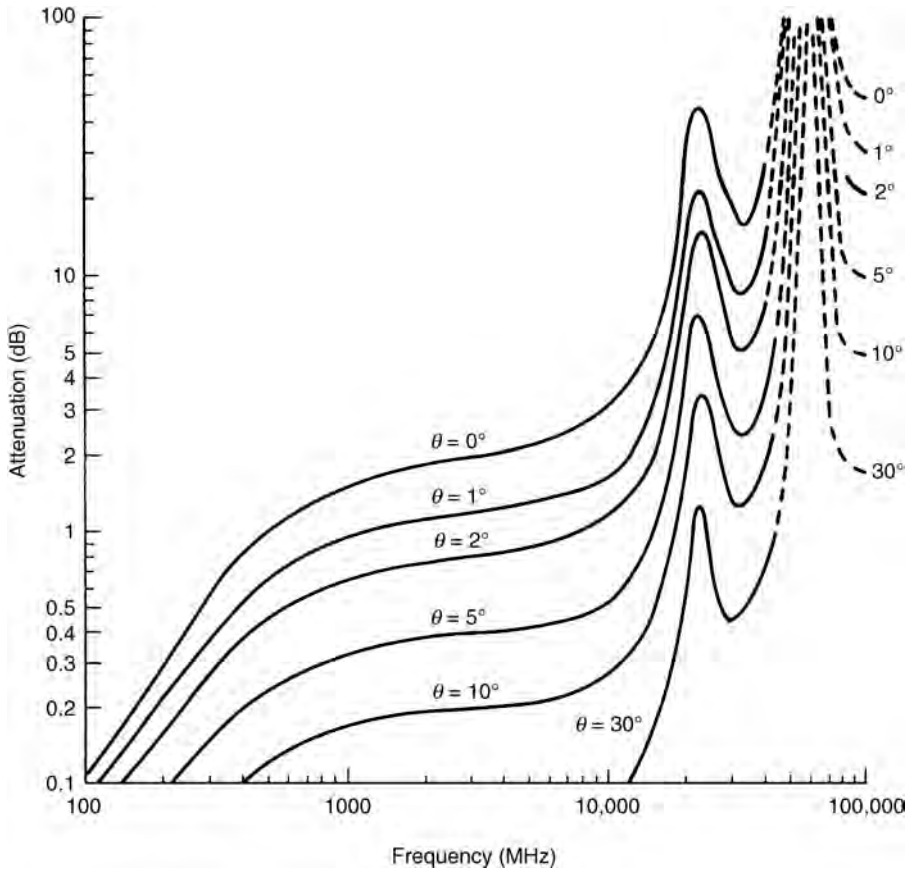


FIGURE 61.4 One way signal attenuation through troposphere at various elevation angles.

where the terms a_1, a_2, and a_3 are the relative solid angles covered by the main beam, side lobes looking at the ground, and side lobe looking at the sun, respectively. They can be approximated by

$$
\begin{aligned}
a_1 &= \frac{1}{4\pi}(G_{\text{sky}}\Omega_{\text{sky}} + \rho^2 G_g \Omega_g) \\
a_2 &= \frac{\Omega_g}{4\pi} G_g (1-\rho^2) \\
a_3 &= \rho \frac{\Omega_s}{4\pi} G_s / L_r
\end{aligned}
\tag{61.15}
$$

The idea is to obtain an average of the temperature weighted by the antenna gain. The W and G are the solid angles and gains, respectively, of the main beam, side lobes that see the ground, and side lobe that sees the sun. Here, ρ is the voltage reflection factor of the ground, and p is a factor to allow for the sun's radiation being randomly polarized while the antenna receives linear or circular polarization only; p is typically 1/2. The assumption is that the sun is not in the main beam but rather in a side-lobe where the gain is much lower, and this reduces the effect of the sun's high temperature. If the sun is in the main beam, and if the antenna beam is less than or equal to the sun's subtended angle (0.5∞), then the increase in temperature is formidable and produces a **sun outage**. The sun's apparent temperature can be found from Fig. 61.6.

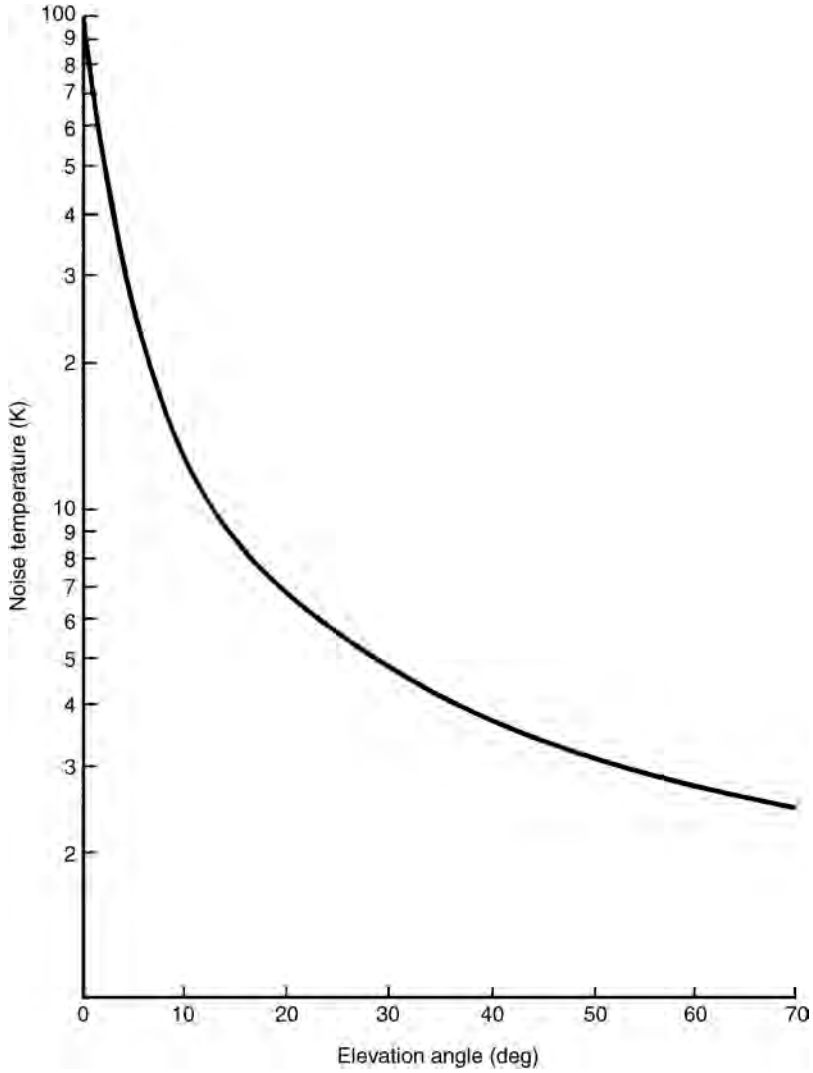


FIGURE 61.5 Sky noise temperature at 4 GHz vs. elevation angle.

Note that rain loss L_r has the good effect of shielding the receiver from the sun. On a beam reasonably wide compared to the sun, the margin normally available to protect against rain loss, say 6.0 dB or so at Ku band, is sufficient to ameliorate the **sun interference** to a bearable level even with the sun in the main beam. An antenna of 50 cm in diameter for the direct reception of satellite TV has a beamwidth of 3.5–4.0∞ and a solid angle 50 times that of the sun. In effect, this reduces the sun's temperature by 50 times.

The General Microwave Receiver

Figure 61.7 shows the functional diagram of a general microwave receiver. Almost all satellite and Earth station receivers fit a diagram of this sort.

We can use Eqs. (61.11) and (61.12) pairwise to combine the stages and yield an overall result. The antenna noise, discussed in the last section, is taken as the input, and we assume that there are no significant noise contributions beyond the mixer **down converter**. An important and useful point concerns the noise figure of a matched passive attenuator. Such a device has the result of reducing the available

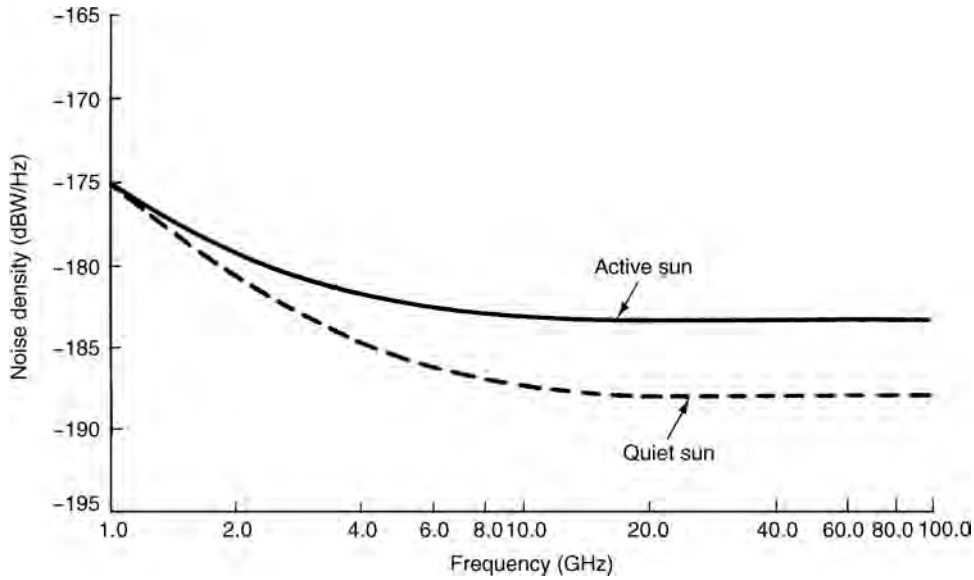


FIGURE 61.6 Noise spectral density of quiet and active sun. (From *NASA Propagation Effects Handbook for Satellite System Design,* ORI TR 1983.)

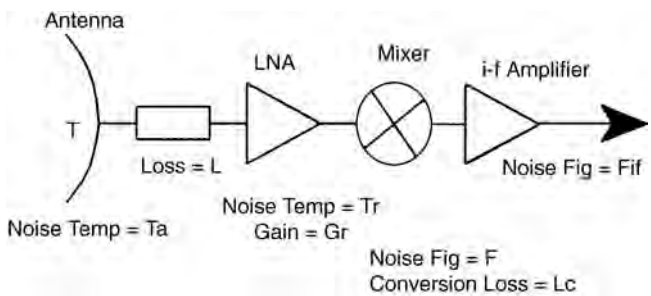


FIGURE 61.7 General microwave receiver.

carrier power by its loss in decibels, while not changing at all the available thermal noise. Thus, its noise figure, defined as the ratio of input to output carrier to noise ratios, is simply equal to its loss. The result, from combining the stages of the receiver two at a time, is

$$T_s = T_a + (L-1)T_o + LT_r + L(F-1)\frac{T_o}{G_r} \tag{61.16}$$

This equation is informative. Besides showing the more or less obvious advantages of low-noise receivers, it highlights the unexpectedly devastating effect of the loss L in the line between the antenna and the LNA. Even 1.0 dB here produces a temperature increase of 75 K. Such a loss is by no means unrealizable in a long line such as might be necessary between a prime focus feed and a convenient receiver location in a large antenna. The need to place the LNA directly at the horn feed, or to use a Cassegrainian type of antenna so that the LNA can be placed at the reflector vertex, is clear.

Defining Terms

Antenna gain: Directivity, ratio of power transmitted in a particular direction to the power that would be transmitted in that direction if the antenna were an isotropic radiator.

Antenna temperature: Noise temperature at the antenna terminals.

Carrier to noise ratio: Ratio of carrier to available noise power.

Down converter: Mixer to reduce RF carrier frequency to more convenient and constant value for amplification.

Excess noise temperature: Apparent temperature of extra noise above input noise added by any black box.

Hot pad: Input attenuator at some temperature higher than room temperature (rain, for example, at room temperature compared to sky noise at a few Kelvins).

Low-noise amplifier (LNA): First amplifier in receiver, just after antenna, with very good excess noise characteristics.

Noise: Random electrical signals that interface with signal detection.

Noise density: Noise in watts per unit bandwidth.

Noise figure: Ratio of input to output carrier to noise ratio.

Noise temperature: Apparent temperature of resistance producing the observed or calculated noise.

Power gain: Ratio of available power at output of a network to available power at the input.

Receiver temperature: Excess temperature of entire receiver including low-noise amplifiers and down converter.

Side lobe: Part of the antenna radiation pattern off the main beam but still radiating measurable energy.

Sky temperature: Apparent noise temperature of the sky.

Sun interference: Increase in antenna noise because antenna is pointed at the sun.

Sun outage: Sun interference so great that receiver is blinded completely.

System temperature: Apparent temperature of resistance producing the observed noise of the entire system if connected to the input.

Thermal noise: Noise due to random electron motions dependent on temperature.

References

Bousquet, M., Maral, G., and Pares, J. 1982. *Les Systémes de Télécommunications par Satellites,* Masson et Cie, Éditeurs, Paris.

Ippolito, L.J., 1986. *Radio Wave Propagation and Satellite Communication Systems,* Van Nostrand, New York.

Ippolito, L.J., Kaul, R.D., and Wallace, R.G. 1986. *Propagation Effects Handbook for Satellite Systems Design,* NASA Ref. Pub. 1082(03), 3rd ed., Washington, D.C., June.

Jasik, J. 1961. *Antenna Engineering Handbook,* McGraw-Hill, New York.

Kraus, J.D. 1950. *Antennas,* McGraw-Hill, New York.

Lin, S.H. 1976. Empirical rain attenuation model for Earth-satellite path. *IEEE Trans. Commun.* COM-27(5):812–817.

Mumford, W.W. and Scheibe E.H. 1968. *Noise Performance Factors in Communications Systems,* Horizon House-Microwave Inc., Dedham, Mass.

Pritchard, W.L., Suyderhoud, H.G., and Nelson, R.A. 1993. *Satellite Communication Systems Engineering,* 2nd ed., PTR Prentice-Hall, Englewood Cliffs, NJ.

62

Onboard Switching and Processing

A. K. Elhakeem
Concordia University

62.1 Introduction

The **onboard (OB)** combination of **satellite switching (SS)** and spot beam utilization is intended to enable a large number of small and inexpensive mobile as well as fixed site terminals to communicate and integrate with commercial broadcasting and integrated and various voice, data, and video applications. This interconnection should not only serve as a standby for terrestrial and optical networks, but also as a competing alternative. With the launching of low Earth orbit satellites (LEOs) [Louie, Rouffet, and Gilhausen, 1992] approaching, the roundtrip satellite propagation delays will become an order of magnitude less than those of geostationary Earth orbits (GEOs). Also, ground terminals will be cheaper, thus providing the edge for competing with ground-based networks. The concept of a bent pipe satellite will be long gone and users now operating in few zones around the globe will be divided into many zones with one satellite antenna supporting local communication of each zone and the switch onboard the satellite routing the traffic from zone to zone within the cluster (Fig. 62.1). If a constellation of LEOs or medium Earth orbit satellites (MEOs) is to cover the globe, then each satellite will handle one cluster, but intrasatellite links have to be provided by a ground network or by a space-based network for users belonging to different clusters to intracommunicate. OBSS systems apply to GEO, LEO, and MEO constellations. Operating in this environment while accommodating various **packet switching (PS)** and **circuit switching (CS)** types of traffic affects the selection of the OB switch type and technology. Before looking at the various OB switch types, details, performance, and fault tolerance, we state their main advantages and applications.

Satellite Switching (SS) Advantages

1. Terminal cost is reduced by lowering transmission power (smaller user antennas used). Also, reducing the number of users per zone reduces user bit rate and, hence, cost.
2. Up- and downlinks are decoupled, thus allowing their separate optimization. For example, for stream (CS) and wideband Integrated Services Digital Network (WBISDN) users in a certain zone,

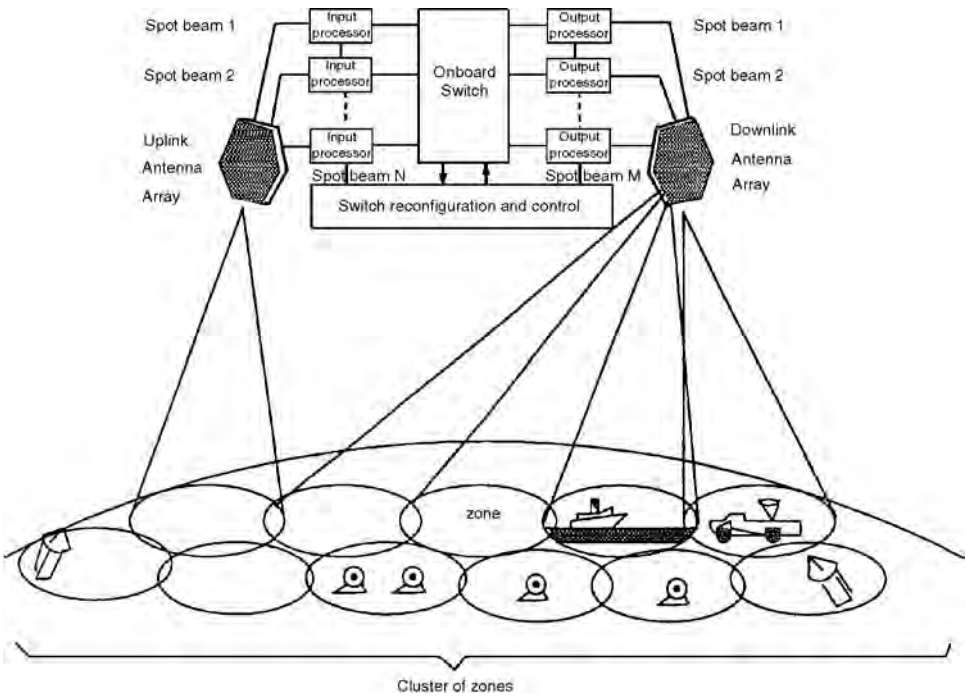


FIGURE 62.1 Onboard (OB) processing satellite using multiple-beam antennas and onboard switching.

time-division multiplexing (TDM) uplink and TDM downlink techniques are adequate. On the other hand, for a multiple of bursty Earth stations, in another zone, a **demand assignment multiple access–time-division multiple access (DAMA–TDMA)** uplink/TDM downlink technique allows the optimal sharing of the uplink capacity. A third zone in heavy rain or under shadowing fading conditions might prefer a switching satellite that has code-division multiple access (CDMA) in the uplink and/or downlink.

3. The number of possible virtual connections is reduced from an order of N^2 to the order of N where N is the number of users.
4. Flexibility and reconfigurability includes, among other things, reprogramming of onboard processing (OBP) control memories, reconfiguration of Earth stations (time/frequency plans for each connection setup), accommodation of both circuit and packet switched traffic, adopting different flow and congestion control measures for each zone depending on current traffic loads and types, and accommodation of multirate and/or dynamically changing bit rate and bit rate conversion on board.
5. SS has natural amenability to DAMA and ISDN services.
6. There is enhanced fault tolerance; a single satellite failure handling a cluster of zones can be easily replaced by another within hours or minutes, thus averting the disasters of losing customers for a few days or months. A single failing antenna onboard the satellite can be easily replaced by another hopping antenna.
7. SS increases the total capacity of the satellite system in terms of number of circuits and services provided within the same bandwidth.

These advantages come with a price, that is, the cost and complexity of the satellite itself.

Applications

- Internetworking, low-cost and low-power user terminals, fixed site, and mobile communications
- Video conferencing

- Interconnection of business oriented private networks, local area networks (LANs), metropolitan area networks (MANs), wide area networks (WANs)
- Narrowband and wideband ISDN
- Interconnection of all of the preceding to and through the public telephone network (PTN)
- Supercomputers networking
- Multiprotocol interconnection
- Universal personal communication (UPC)

62.2 Onboard Switching Types

- Reconfigurable static IF switching, also called **microwave switch matrix (MSM)**. This has fixed (nonhopping) beams. The access techniques could be **frequency-division multiple access (FDMA)**, TDMA, or multifrequency-TDMA (MF-TDMA), giving rise to three kinds of switched satellites, that is, SS-FDMA, SS-TDMA, SS-MF-TDMA. The first, SS-FDMA, is the most common.
- Reconfigurable static intermediate frequency (IF) switching with regeneration and forward error correction but no baseband processing
- Baseband switching with fixed or hopping (scanning) spot beams
- Fast packet switching
- Photonic baseband switches
- Asynchronous transfer mode (ATM) oriented switches

Reconfigurable Static IF Switching

Both the uplinks and downlinks access techniques could be TDMA, FDMA, or MF-TDMA. However, the most commonly used are FDMA and TDMA on the uplinks and downlinks, respectively. This choice results in a reduced terminal ground station cost. FDMA is selected on the uplinks to eliminate the need for high-power amplifiers (HPA) at the user terminal [an FDMA user transmits all of the time but with a fraction of the bandwidth (BW)]. On the other hand, to utilize the full power of the OB (HPA), it should operate near saturation, thus increasing the downlinks signal and allowing very small aperture terminals (VSAT) operation. Simply a high-power high bit rate TDM operation in the downlinks is efficient. No demodulation or error correction takes place onboard. Fixed beams are used, and the number of such beams in the uplinks is typically less than that of the downlinks. Stations use frequency synthesizers to transmit the information using an RF carrier corresponding to destination. This form of FDMA is called single channel per carrier (SCPC) and is appropriate for low rate users (64 kb/s). This carrier is demand assigned to the user call via a signalling channel (which could be ground based or an out of band SCPC channel). Onboard, the various RF carriers pass through the low noise amplifier then become down converted to IF (Fig. 62.2). Demultiplexing separates the various user signals that form the inputs to the static MSM. These are routed through the MSM to their destinations. The connection pattern of the switch (which input connected to which output) is configured by the master ground station or ground network control center (NCC) control channel. However, some of the connections may be permanent to allow broadcasting. Taking the example of the advanced communications technology satellite (ACTS) system of NASA [Cashman, 1992], the uplinks access technique is TDMA not FDMA, and the MSM is controlled by a digital controller and a dual storage bank of memories. The MSM switch is configured according to the 5-kb/s, telemetry, tracking, and command signal (TT&C) issued from the master ground station at the beginning of each superframe of 1-min duration. The TDMA frame is 1 or 32 ms and the switching time is approximately 100 ns. One bank of memory will be in active control of the switch while the other is waiting for TT&C new configurations. In the case of SCPC uplinks, the switch is configured on a call basis (slower basis). Three active fixed or hopping beams exist on the MSM input and output. However, the MSM size is 4×4 to allow for redundancy. The output signals of the MSM switch going

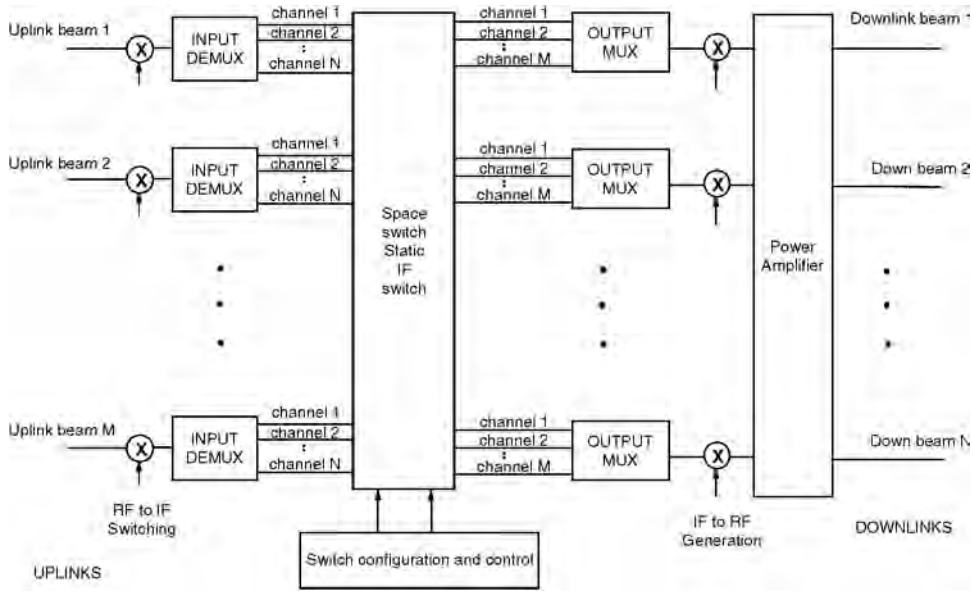


FIGURE 62.2 Nonregenerative onboard IF switching.

to same zone are grouped (multiplexed). The resultant is subsequently translated to RF frequency, amplified by solid-state power amplifiers (SSPA), and transmitted via output beams. An in-band or out-of-band signalling scheme is used by the user terminals to establish the call. The master control station finally assigns the right resources to the call, for example, one out of a pool on the FDMA uplinks and one time slot of a certain TDMA frame on the downlinks for that call. This reservation goes to the OBS, which becomes correspondingly configured for the call duration. We notice here the isolation between the up- and downlinks provided by OBS, for example, in the preceding case the uplink user transmits continuously using the assigned FDMA SCPC, i.e., disjointed from the TDM process taking place on the downlinks. Amplifying these FDMA signals OB creates intermodulation products (IMP). SCPC also suffers from its inability to meet variable bit rate (VBR) services. Multichannel per carrier (MCPC)/TDMA or MF/TDMA is typically employed to reduce these IMP. Here, a number of users (terminals) are allocated channels (time slots) on a TDM frame, which is transmitted on a single carrier.

Other disadvantages of FDMA-oriented MSM are the difficulty of dynamic networks reconfiguration without data loss, the difficulty of operation with hopping beams, and broadcasting of multiple carrier transmissions to many destinations.

Regenerative IF Switching

This is similar to the previous static IF switching in the following respects: switching of IF signals, employing TDMA, FDMA, or MF-TDMA access techniques with the FDMA uplinks/MF-TDMA downlinks as the preferred technique, no baseband processing being involved, a station simultaneously transmitting/receiving from many stations having to use a different carrier for each source (destination), programming reconfiguration via fixed connections and/or under the TT&C commands, use of low noise amplifiers (LNAs) and SSPAs, and RF to IF and IF to RF translations.

However, it differs in the following respects. The IF signals (following RF to IF translation and demultiplexing) are demodulated, forward error corrected, then forward error encoded [using error detecting and correcting coding (EDAC)] and then modulated [continuous phase shift keying (CPSK) or minimum shift keying (MSK)] again before feeding them to the MSM switch. If both uplinks and downlinks are of the MF-TDMA kind, then the various uplinks user TDM channels (slots) of each beam form the IF inputs to the switch. At the switch output, output IF slots are multiplexed again,

and the resultant modulates the RF carrier of the downlinks beam. The regeneration involved in the demodulation–modulation processes allows the separation of uplinks and downlinks noises, fading, interference effects, and so forth. Also, the decoding–encoding could be done at different rates, thus providing for more protection against rain in the downlinks, for example, by decreasing the forward error correction (FEC) encoding rate. These regenerations lead to better links performances and better amenability to MF-TDMA generation that needs higher transmission rates. Nevertheless, regenerative IF switching suffers the same aforementioned disadvantages of static IF switching.

Implementation of Regenerative and Static IF Switches via MSM Technology

Basically, this is a solid-state programmable crossbar switch. Signal switching at the matrix crosspoints is performed using dual-gate field-effect transistor (FET) active switches (employing GaAs technology for switches with IF bandwidths equal to or exceeding 2.5 GHz). In recent implementations, such as the NASA ACTS OBS [Cashman, 1992], input signal distribution and output signal combining are performed using passive recursive couplers (Fig. 62.3).

Among many alternatives, such as single pole, rearrangeable switching, and so forth, the coupler crossbar implementation was selected by both ACTS and other NASA systems for its simple control, low insertion loss, and amenability to broadcasting. In this architecture, a switching element [Fig. 62.3(b)] is situated between two directional couplers. If the data is to flow from input i to output j, the switch

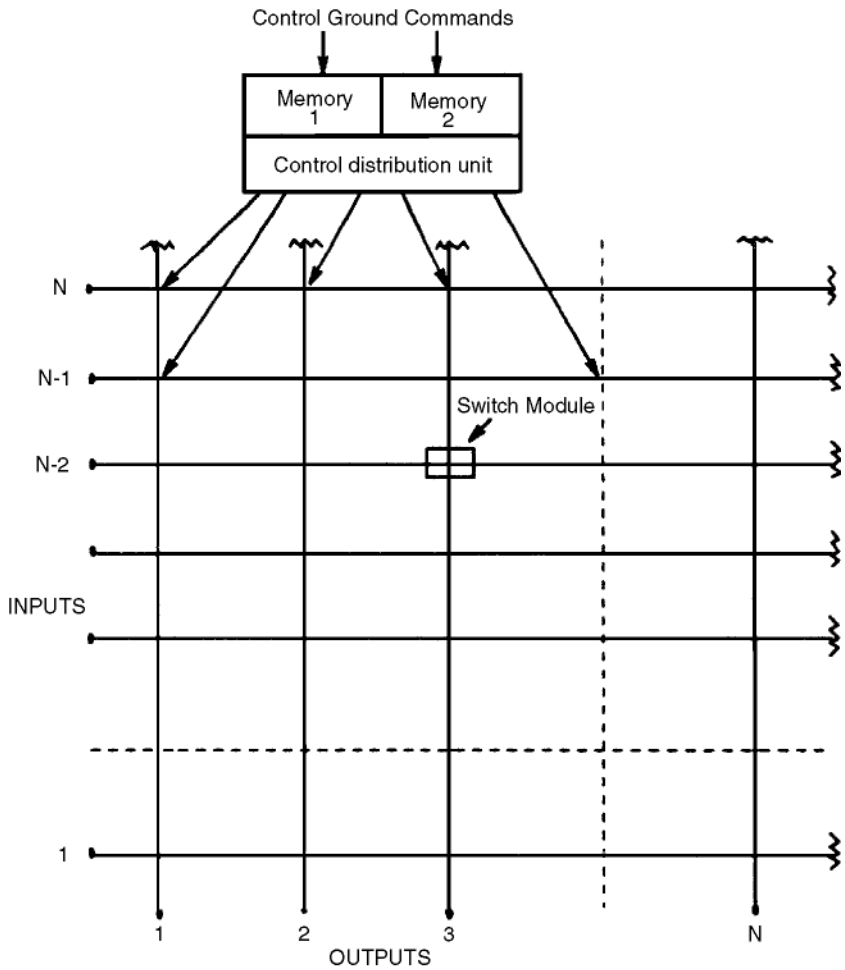


FIGURE 62.3(a) $N \times N$ IF switch matrix using crossbar configuration.

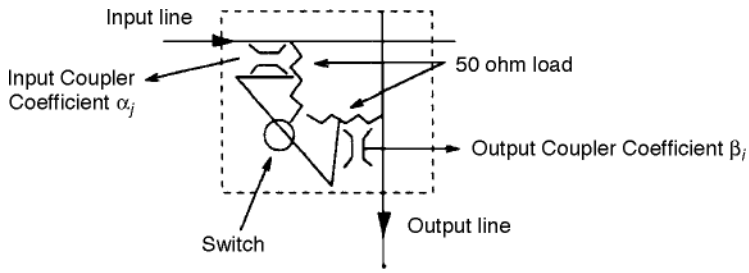


FIGURE 62.3(b) Switch module units: α_j = couplering coefficient of *j*th output cross point on the same input line; β_i = couplering coefficient of *i*th output cross point on the same output line.

will be ON and the input and output couplers pass the data; otherwise, the switch is OFF and the data is reflected and absorbed by the coupler loads. To equalize the output power at all N input couplers connected to the N output lines, the coupler coefficients increase as the output number increases, that is, $\alpha_j = 1/(N - j + 2)$. The same logic α_j applies to the output coupler coefficient β_i as in Fig. 62.3(b). Switches at the various cross points are independently controlled to provide broadcasting as well as point-to-point connections. Moreover, the switch design could be modular both electrically and mechanically (as in NASA designs); for example, different switch points or parts of the switch can be repaired and tested independently, thus improving cost and reliability, or larger switches can be built from smaller modules.

The switch elements could be built using ferrite switches, p-i-n diodes, or single gate FETs. The first option is too slow, the second needs larger amounts of DC power, the third yields no gain during ON states; dual gate FETs are typically preferred. These use Schottky barrier diodes with very low junction capacitance, thus allowing the switching from gain to high insertion loss in about 100 ps. Two FETs are used, with the RF input signal applied to the first and the switch control applied to the second. This provides for excellent isolation between the DC control (which ranges from 30-dB loss to 10-dB gain) and the RF input. One could envision an IF switch with constant coupling coefficients (α_j, β_i) and variable switch amplifier gain. Typically, however, the earlier variable coupling gains implementation is preferred.

The switch states (configuration) and latency of each cross point are decided upon by ground control signals. This control unit is typically called a control distribution unit (CDU). One bank of the dual control memory will be in active control of the switch states while the other bank receives the fresh ground control commands. These ground control signals also cause the exchange of the two memory bank identities to take account, for example, termination and/or new call establishments.

The switch sequence (logical signals) that controls the switching state of all cross points is periodically dumped from the control memory until fresh control commands are loaded to the other memory bank from the ground control station, in which case the switch control signals are dumped from the other memory bank. The reconfiguration rate is a few microseconds and the switching time is in the 2–10 ns range.

The reliability of the IF switch matrix is highly dependent on the failure rate of the GaAs FET, which is roughly equal to 23×10^{-9} (including the associated passive elements). This reliability can be improved by adding extra rows and columns so as to provide alternate routes from inputs to outputs that do not pass by the failed cross points [the dashes of Fig. 62.3(a)]. Additional reliability could be achieved by using multiple ($N \times N$) switching planes in parallel at the expense of cost and weight. For the switch control unit, using the memory decoder yields a reliability of 0.8331 for 10 years operation. This can be improved further by using simple forward error correction techniques to correct for possible memory read or write errors.

For an $N \times N$ cross point IF switch, and if the time frame has M data units, the control memory size is given by $M(N \log_2 N)$, since N cross points are controlled by $\log_2 N$ bits of memory (if decoding is used).

Implementation of Regenerative and Static IF Switches Using Optical Waveguides

Future low orbit satellite clusters will probably communicate through intersatellite optical links at data rates of the order of few gigabits per second. This will make the instant and immediate switching of these signals in the optical domain a great asset compared to the classic tedious and slow approach of optical to electric conversion, switching of IF or baseband signals, then electrical to optical conversion at the switch outputs. In both IF and baseband implementations of optical switches, a directional coupler is the basic switching element.

The directional coupler (Fig. 62.4) consists of two waveguides situated close enough to allow the propagating modes (carrying the information signals to be switched) to interact (couple) and transfer power from one waveguide to another. Electrical control is still applied to allow the ON or OFF switching of the input signal to one or both outputs. This controls the index of refraction and, therefore, the coupling coefficients.

In the OFF state, a 3-dB optical isolation is achievable between outputs. If the switching voltage is 25 V, coupling lengths lie in the range of 1 to 2 cm, the loss associated with single mode fiber termination is of the order of a few decibels, and switching times lie in the 1 to 100 ns range. It is possible, however, to lower these to 50 ps by using traveling wave electrodes or by using free space optical switching techniques.

To reduce the N^2 complexity of MSM cross point switches using optical waveguides, different configurations are used. The one shown in Fig. 62.4(b) uses only $N(N - 1)/2$, where $N = 6$ couplers, thus saving, for example, $(36 - 15) = 21$ couplers for a (6×6) switch.

Optical waveguides are not without disadvantages; for example, switch size is limited by the area of available substrate, high insertion loss (due to fan in losses of waveguides), high level of crosstalk, internal blocking in Benes, Banyan, or general multistage types, which eventually lead to difficulties in operation in broadcasting modes.

To enhance the reliability of optical wave guide IF switching architectures, replication in the case of $N \times N$ cross point structure or addition of a few input and output lines, and dilation (addition of more parallel lines between stages) in the case of Benes and multistage structures are generally recommended.

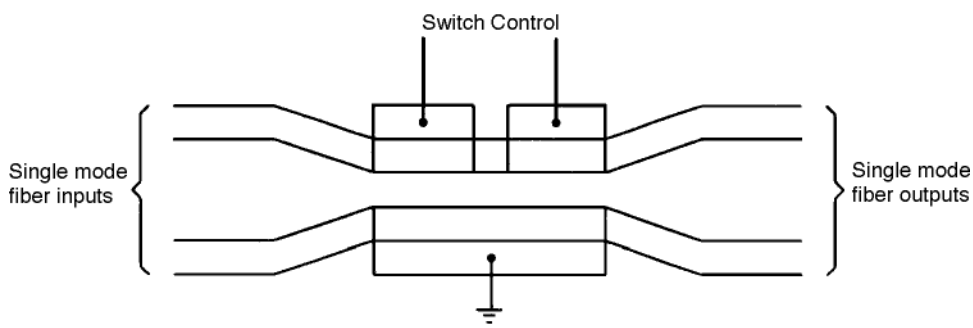


FIGURE 62.4(a) A 2×2 all optical switch module using direcitonal coupler.

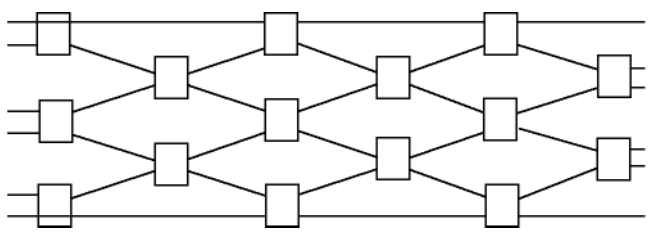


FIGURE 62.4(b) Multistage implementation of an optical $N \times N$ switch using $N(N - 1)/2$.

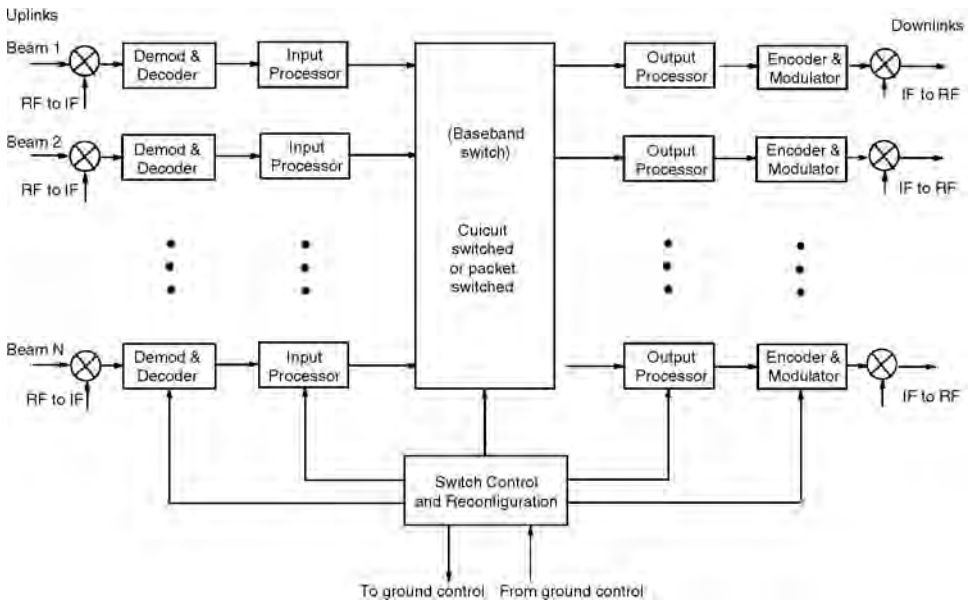


FIGURE 62.5 Building blocks of onboard baseband switching.

Onboard Baseband Switching

The uplink signals are downconverted, demodulated (possibly FEC decoded), switched, and then modulated, FEC encoded, upconverted, and transmitted downlink (Fig. 62.5). Switching takes place at baseband level, thus allowing for a much needed flexibility. For example, a station with only one carrier at hand can transmit to many destinations via different TDMA slots. These transmissions could reach the destinations on the different carriers, the different bit and FEC rates, the different modulation formats, and so forth, that best suit the fading environment and other conditions in the downlink. Multicasting and accommodation of a mixture of high and low bit rates without needing multicarriers as in IF switching are some of the added benefits.

Baseband switching can be oriented toward circuit or packet oriented traffic. When large portions of the network traffic and/or most of the users are of the circuit switched type, and/or switch reconfiguration is infrequent and onboard, baseband circuit switching is more efficient. For a satellite network carrying mostly packet switched traffic, and/or circuit switched traffic with frequent configuration, and/or a combination thereof, a packet switch is more efficient. Hybrid packet/circuit baseband switching is also a viable alternative that has gained attention recently.

Baseband Circuit Switching Categories

Time–space–time (T–S–T), common memory (T-stage), distributed memory, and high-speed fiber optic ring are the main types of this category.

(T–S–T) Baseband Switching. This is used for multigigabit per second operation and has been adopted in the NASA ACTS system and for European Space Agency (ESA) prototypes. It has also been in use for some time in terrestrial networks (AT&T no. 4ESS switch). The basic idea is to reduce the space switch complexity by premultiplexing n of input data units (e.g., bytes), thus reducing the number of inputs and outputs of the space switch. However, this reduced switch has to switch data much faster, as will follow.

Following RF to IF conversion, filtering, and data demodulation and forward error decoding in each of the (N/n) uplink banks [Fig. 62.6(a)], the n baseband data units of each multiplexer (with a combined rate of nR_b b/s) are further multiplexed via the time slot interchange (TSI) box [Fig. 62.6(b)] to yield a bit rate at the TSI equal to kR_b b/s, where $k \geq 2n$ to allow for a strictly internally nonblocking switch.

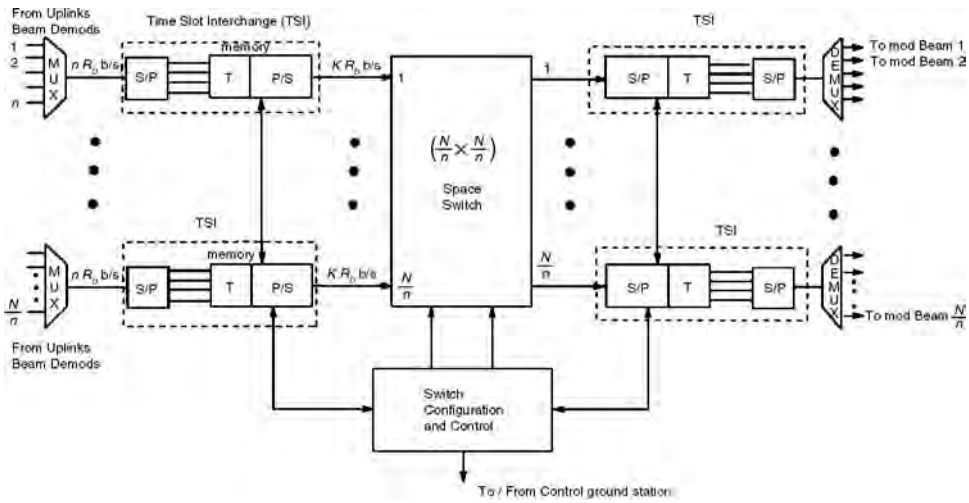


FIGURE 62.6(a) T–S–T baseband switching.

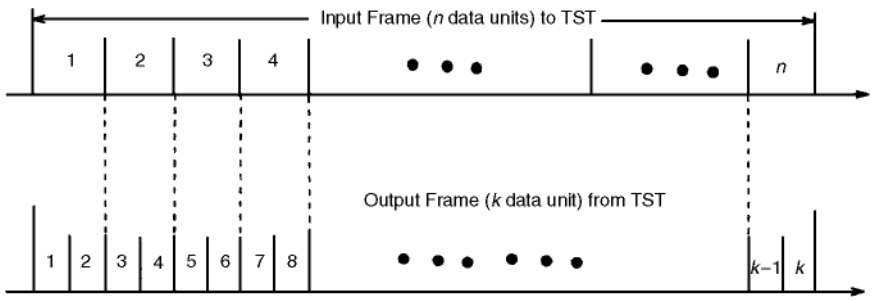


FIGURE 62.6(b) Time slot interchange (TSI) reading from T-stage is at least twice as fast as writing into this stage ($k \geq 2n - 1$).

Internal nonblocking means that there are enough virtual paths inside the switch to connect any of the (N/n) inputs to any of the (N/n) outputs. The TSI is basically a serial to parallel converter followed by a T-stage (memory), which stores the data units of the uplink frames, followed by a parallel to serial converter. At the space switch output, another TSI executes a reversed sequence of events that finally feeds all data units of the downlink frame to the data forward error encoder and modulator, after which IF to RF conversion and subsequent transmission in the downlink takes place. To see the reduced complexity of this approach compared to a conventional cross point switch, we take $N = 1024$, multiplexing of $n = 128$ slots (each can serve one unit of data), $k = 2n = 256$. The size of the space switch is ($N/n \times N/n$), that is, (8×8) a great simplification compared to the corresponding conventional cross point switch of (1024×1024) complexity. The relation $k \geq 2n$ was derived in the 1950s for a class of three-stage Clos space switch, which resembles the T–S–T switch, and it effectively says that the speed of the space stage inputs (output of TSI unit) should be at least twice as fast as that of the TSI input [Fig. 62.6(b)]. Memory read and write from the T stages limit the speed of the external lines (from data demodulator and FEC decoders). If we take the frame time $T = 4$ ms, 1 byte as the data unit, input data rate of $R_b = 4$ Mb/s per input beam, the T-stage total memory size should be $2n \cdot T_F \cdot R_b \cong 4$ Mb. The total memory requirements for all input and output T-stages become $4(8 + 8) = 64$ Mb/s. The total switch throughput $= NR_b = 1024 \cdot 4$ Mb/s ^{a} 4 Gb/s, which can be increased by increasing (N/n), that is, the number of pins of the space switch, and the memory access time per word of the input or output T-stages (assuming a 16-b-wide memory bus) is given by $1/(KR_b/16) = 16$ ns.

If current technology does not permit, one has to reduce n and/or the input data rate R_b. Dual memories (T-stages) of size $2n \cdot T_F \cdot R_b$ bytes are used on both input and output TSIs such that simultaneous read and write to the same memory can be achieved. Also, pipelined operation is assumed such that while bytes of the new frame are shifted at the input speed and stored in T-stages, bytes of the previous frames are shifted out to the output stage at the output speed.

To accommodate a multiplexity of user terminals of high and low speeds, additional multiplexing of lower rate users [not shown in Fig. 62.6(a)] is implemented at the input lines (prior to input TSI stage). This makes for a square $N/n \times N/n$ switch structure.

A sophisticated control processor (consisting mainly of memories and TSI timers) reconfigures the switch in a conflict free manner during the call establishment period for CS traffic and each frame for PS traffic. This consumes time and hardware since all three stages (input TSI, space switch, output TSI) have to be correspondingly adjusted. For example, if a user input beam 3 has a call in slot 5 of the input frame to a user in output beam 6 who listens in slot 9 of the output frame, then the space switch control part makes sure that input beam (3) is connected to output beam (6) in each frame for one slot time. The TSI control at the switch inputs and outputs makes sure that the input byte is written in the fifth location in input memory and the output byte is read from the ninth location.

To improve T–S–T reliability, (N/n) active T-stages and r redundant ones are used in both input and output sides of the space switch, which is by itself replicated. Also, error detection and correction can be used to protect the memories from permanent and transient errors. For example, 16 out of 20 (N/n out of $N/n + r$) redundancy and 1 out of 2 redundancy for space switch yields 99% reliability for 10 years.

It is also noted that the T–S–T architecture could, in principle, be used for routing packet switched traffic as well. However, switch control will be derived in this case from the individual input packet header, a process that slows the switch compared to the case of CS traffic. It is also generally assumed that all corresponding time slots of all inputs to TSIs are bit synchronized. This requirement will add costly hardware, especially in the case of hybrid PS/CS traffic; for example, some preambles will be stripped at the switch input, while some central headers will be routed through the switch. Furthermore, T–S–T switching has the advantage of expandability to higher speeds, and the disadvantage of encountering two switching TDM frame delays (corresponding to the input and output TSI stages).

Baseband Memory Switch (MS). This is the simplest and least costly architecture [Evans et al., 1986]. During a certain TDM frame, demodulated data bytes from all uplink beams are stored in a single double (dual) data memory (Fig. 62.7) in a sequential manner under the control of a set of counters or control memory.

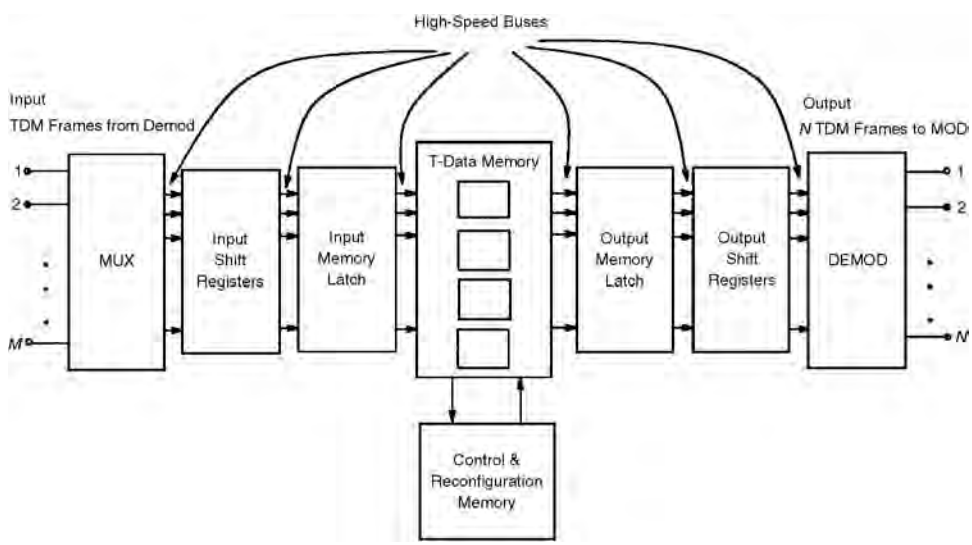


FIGURE 62.7 A baseband memory switch.

Parallel in time, the stored bytes from the previous switching TDM frame are served to the appropriate switch output under the supervision of the control memory. For one double dual memory to replace the many T-stages of the previous T–S–T architecture, this memory has to be large enough and fast enough to store the data bytes of the synchronized M input TDM frames in one frame time. A high-speed bus and a multiplexer (MUX) are thus used, and a demultiplexer (DEMUX) distributes the data memory contents into the various output channels under the control of simple counters and the configuration memory. Memory switches could be asymmetric (number of inputs π number of outputs) and are internally nonblocking. The data memory size is $(2NS)$, where S is the number of TDM slots per frame and N is the number of output beams. This is half the memory requirements of the T-stages of T–S–T switches and is typically organized into 4 or 8 smaller memory blocks.

The contents of control memory serve as the addresses to the data memory and together with a simple counter associated with each output they read the right location of data memory in the right TDM slot of each output frame. The control memory is reconfigured with each new call. However, this is less intensive in time and in control hardware compared to the T–S–T case.

This architecture could possibly route packet switched traffic, in which case the switch control is derived from the destination in the packet header, which is more involved and time critical.

Bus speed and memory access time upperbounds the capacity of this architecture compared to the T–S–T that grows independently of these. Also, the memory is a reliability hazard avoided by memory replication, and immunity to soft (intermittent) errors is provided by error detection and correction between data memory banks and read and write latches.

Power and weight are important considerations, and in recent designs (for example, the PSDE-SAT2 of the ESA), memory switches were approximately four times lighter than T–S–T switches and they took only 1/10th of the chips.

Distributed Memory Switch (DMS). To improve the reliability of the **memory switch (MS)** architecture, the memory could be split into N T-stages at the output side (Fig. 62.8). This way, failure of one or more of the T-stages means that only a few (not all) beams are out of service.

Multiplexing of the uplink baseband TDM frames and sequential writing to the TDM bus takes place as in MS. All T-stages hear the broadcast TDM data; however, under distributed control and timers, only the appropriate data is sequentially written at the right time slot into the applicable T-stage. Reading data from T-stage to output beams is done via random access in different time slots of the output TDM frame. The memory size is still the same as in MS, and the distributed memory switch (DMS) is a strictly nonblocking switch. Redundancy is achieved by duplicating each T-stage; the MUX is also duplicated. Control units are also distributed and duplicated for fault tolerance. Reconfiguration of the T-stages is simple compared to T–S–T switches.

However, T–S–T architecture has better expansion capabilities than MS or DMS structures. Though bus speed is not a problem in the DMS, the bus is overloaded by the presence of the many T-stages, thus

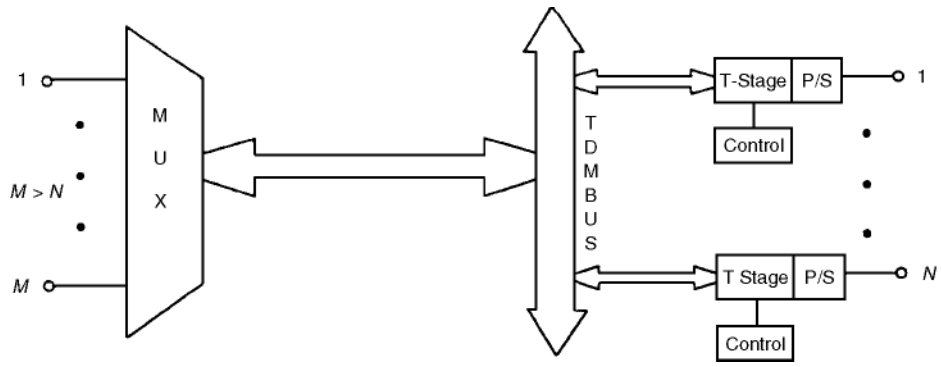


FIGURE 62.8 Distribution memory switch (DMS).

increasing the bus driver power consumption. Error detecting and correcting codes are also applied to the individual T-stages to protect against reading and writing errors.

Baseband Fast Packet Switches

Circuit switches are more or less amenable to handling packet switched traffic [Inukai and Pontano, 1993]. For example, the MS architecture becomes self-routing (and thus able to handle packet switched traffic) by stripping the headers of the packets that contain the destination and storing them in the control memory, which then selects the appropriate data unit from the data memory to the appropriate output register. In the case of the DMS architecture, these stripped packet headers are routed (using an address bus) to the distributed control units at the various output beams. A fiber optical token access scheme ring such as that of the fiber distributed data interface (FDDI) standard is directly adaptable to both CS and PS traffic.

MS, DMS, and T–S–T switch throughputs for PS traffic are reduced by the described header stripping, bus contention, memory size, and so forth. A self-routing multistage switch is generally preferred for handling PS traffic if throughputs higher than a few gigabits per second are required. A control memory is unnecessary in this case since part of the packet header is used to route the switch through each stage. Error detection and correction are used to protect against misrouting the packet due to header errors. However, complex data and control memories are not used; rather a set of cascaded (2×2) module switches is used. A typical multistage switch of the Banyan switch (BS) type is shown in Fig. 62.9(a) [the sorted **Banyan network (BN)** or Batcher Banyan network] [Hui and Arthurs, 1987].

The heart of this is the Banyan network (also called delta network), which is detailed in Fig. 62.9(b) for a 16×16 BN. The packet finds its path automatically by self-routing within the various stages in the $N \times N$ rectangular Banyan network, which is composed of $(\log_2 N = 4)$ stages. Each stage consists of $(N/2)$ smaller switching elements. Each element is a small (2×2) binary switch. A control bit of 0 or 1 moves the input packet to the upper or lower element output, respectively.

A BN has the property that there is exactly one path from any input to any output but has the disadvantages of exhibiting internal blocking, that is, two packets from two inputs colliding at the same intermediate switching element output, even if they were destined to different final outputs. All of the previous CS described earlier have the advantage of being strictly nonblocking in this sense. However, both types (internally blocking and nonblocking switches) may have final output blocking, which is due to two packets from different inputs destined to same output. To solve the two problems of internal and output blocking, the sorting network of Fig. 62.9(a) should be used or one has to replace the whole figure by an alternate buffered Banyan network (Fig. 62.10). In the latter, a limited buffering capacity exists at each (2×2) module input. If two input packets are destined to the same output, one packet is routed to the output and the other is temporarily stored for subsequent routing in the following time slot.

Among recent buffered Banyan implementations is the Caso–Banyan baseband switch of Toshiba [Shobtake and Kodama, 1990]. In that implementation, the FIFO property of the 2×2 switch element was violated such that when two head of line data units of the two input buffers contended for the same output, only one is forwarded to this output and another packet (not the head of line) of the other input is forwarded to the other output. It has been found that the unit data (packet, cell, byte) blocking rate is

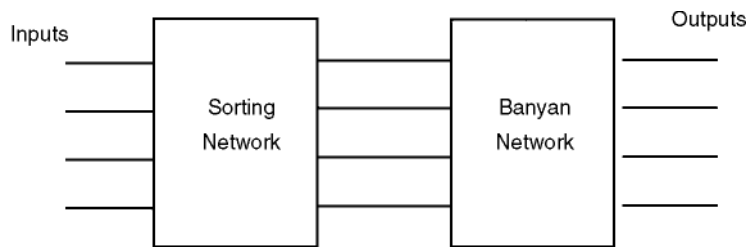


FIGURE 62.9(a) A sorter-Banyan baseband switch.

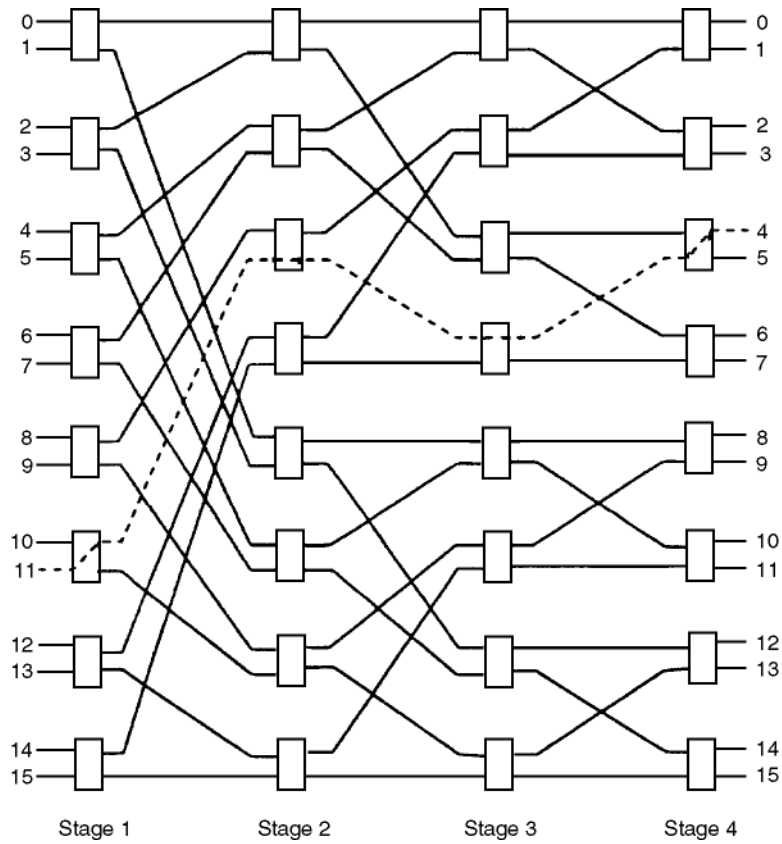


FIGURE 62.9(b) Banyan network with four stages and 16 input terminals. The dotted path corresponds to a packet header of 0100.

1% of that given by conventional buffered Banyan networks at normalized input utilization factor $\rho = 0.85$. Now the sorting network approach of Fig. 62.9(a) [Hui and Arthurs, 1987] is used to reorder packets at the N input requests so they appear ordered according to their packet destination at the sorter N outputs (on the first phase). Also, in this phase, only short connection requests are made by various inputs. This packet rearrangement guarantees no internal blocking when feeding these to the subsequent BN.

Output blocking is avoided by purging contending packets and acknowledging only the input requests that have priority (in the second phase) and forwarding only one of the sorter output packets to the BN (in the third phase).

Unacknowledged inputs will try retransmission of their output blocked packets in subsequent time slots. This three-phase algorithm slows the switching speed and lowers the achieved throughputs. Other implementations, such as the Starlite switch (Fig. 62.11), still have BN and use buffers at the output to store contending packets that did not win the output contention for subsequent routing in the following time slots, but do not have a separate request (phase 1) of the three-phase algorithm underlying Fig. 62.9(a). The Starlite approach [Huang and Knauer, 1984], however, has more hardware (buffers and control) and may produce packets out of sequence at the switch outputs.

Having realized that increasing the speed of Banyan type networks (reducing the blocking) was achieved only at the expense of increasing the hardware, a straightforward solution of *replicated-Banyan switches* provided modularity as well. Here, the number of internal paths (space diversity) is increased by distributing input traffic among many Banyan switches in different planes, thus reducing internal blockings. Output blocking may still be decreased by incorporating buffers at the various outputs of each switching plane. *Internal speed up* could be used to speed the data passage through the switch elements in all of

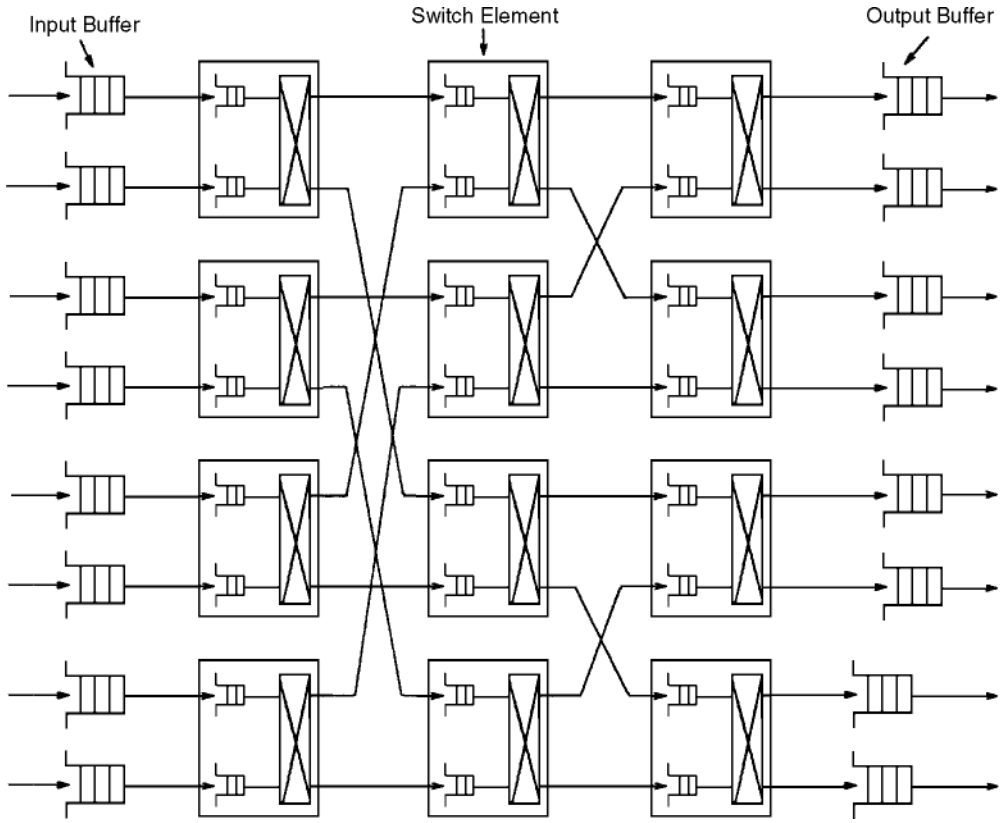


FIGURE 62.10 A buffered Banyan network.

the previous architectures (sorted Banyan, replicated Banyan, buffered Banyan, etc.) in a manner similar to that used in baseband CS (T–S–T, MS, DM). This leads to an almost strictly internal nonblocking switch, as was previously outlined, and a *space* switch component with less inputs and outputs.

Similar to baseband CS, the Banyan switches suffer from both *stuck-at* and intermittent faults. To increase their reliability, replication of switching planes, *dilation* (addition of parallel links and more switching stages within the same plane), and using EDAC codes to protect the packet headers before and after the baseband switches are the state-of-the-art techniques of fault tolerance in the baseband PS case [Kumar, 1989]. Also, self-testing hardware and testing sequences have been designed and added onboard to detect the presence of stuck-at faults.

Gamma switching networks [sometimes called inverse augmented data manipulator (lADM)] have a multistage structure, which uses a 3×3 switching element and allows broadcasting, a property not readily available in the Banyan switching networks presented.

Omega and *close* switching networks also belong to the multistage group but could be made strictly nonblocking under certain conditions. Lack of amenability to broadcasting is the common thread in most of the Banyan-based and multistage networks. To have this capability, replicated or dilated Banyan networks are essential and/or *copy* networks could be used.

Photonic Baseband Switches

These are ultra fast and efficient switches that can handle a mix of CS and PS traffic and have an inherent broadcasting capability. TDM-based optical rings, optical token rings, **wave division multiplexing (WDM)** optical rings, and knockout switches are the basic types in this category. A single fiber-based ring can support a multitude of switched traffic to a total sum of a few gigabits per second.

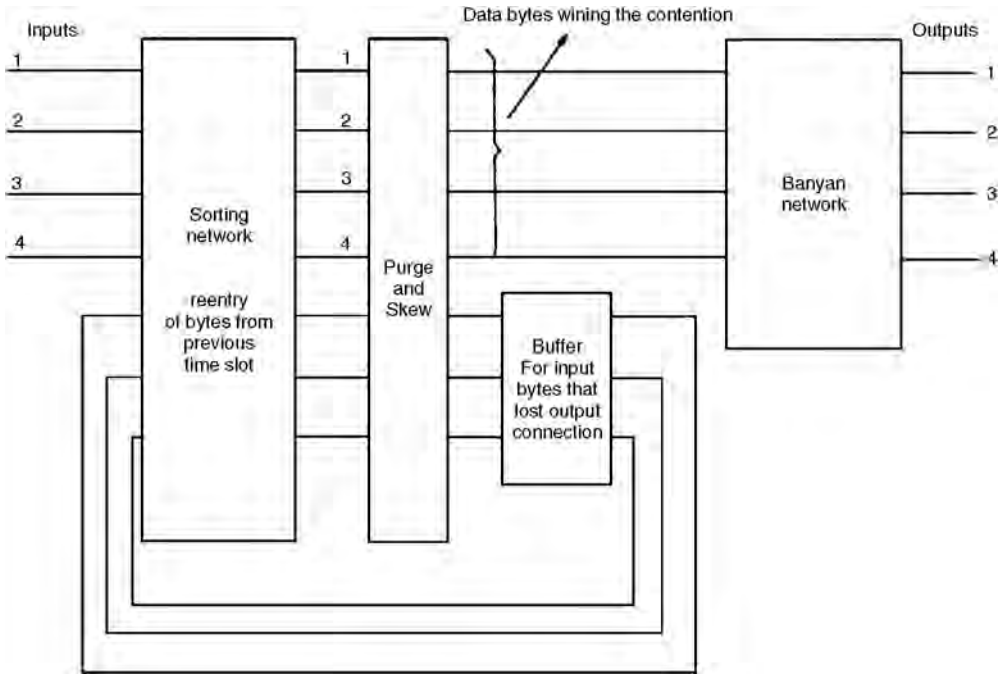


FIGURE 62.11 A Starlite network.

Optical technology (fibers, waveguides, free space optical transmission) provides a switching media whose bandwidths lie in the terahertz region and, hence, are bounded only by the processing speed of input and output interfaces and stringent bit synchronization requirements around the ring, etc. Optical switch throughput exceeding 80 Gb/s can be accommodated with present day technology.

A TDM Optical Ring Switch

Each beam will have one input and one output processor on the optical ring, see Fig. 62.12, which mainly shows the baseband stages. The input processor includes not only the optical transmitter, but also EDAC decoding, packet assembly, descrambling, deinterleaving, burst and/or frame synchronization, input traffic demultiplexing, and possibly bit stuffing for CS traffic. The time frame on the optical ring is divided into slots (that serve, for example, one byte for CS traffic or one packet for PS traffic) and each input processor of an uplink beam is allocated a certain part of this ring switching frame (say, 8 slots) for its traffic to all output processors of the downlink. The optical ring traffic controller, which has a dedicated control packet at the ring TDM frame start, keeps the ring in sync, issues the frame start and ring management commands, and reconfigures (allocates slots) the TDM ring frame among the users (input processors).

For PS traffic, the output processor of each beam receives all packets of the ring frame but processes only the ones destined to this beam by a filtering mechanism to recognize the destination address in each packet. For CS traffic, the output processor selects the appropriate slots of the frame based on the controller commands that are updated by the arrival of each CS call. Alternately, the input processor may insert dummy destination bits (addresses) in each CS packet only for routing purposes within the switch (TDM ring). These addresses will be filtered by the intended output processor. Note that dummy bits could be used even in the first approach to have the same packet size for both CS and PS traffic flowing through the ring so as to simplify synchronization and control hardware. Broadcasting is easily achieved by having all output processors identifying certain broadcast bits in the packet header and processing the packet accordingly.

The output processor of each beam will have the optical receiver, modulator, scrambler, interleaver, EDAC encoding circuits, and packet reassembly and multiplexing of the various packets on the TDM

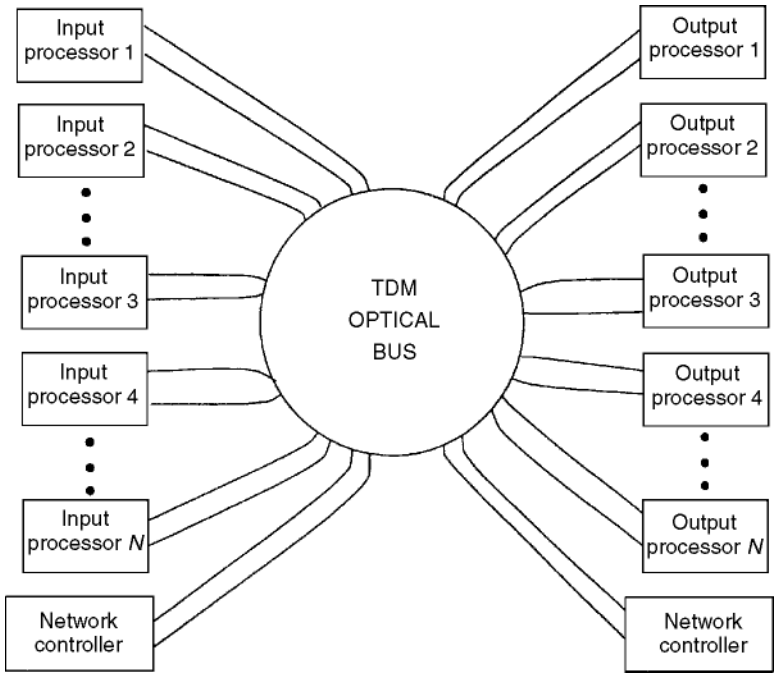


FIGURE 62.12 Onboard TDM optical ring switch.

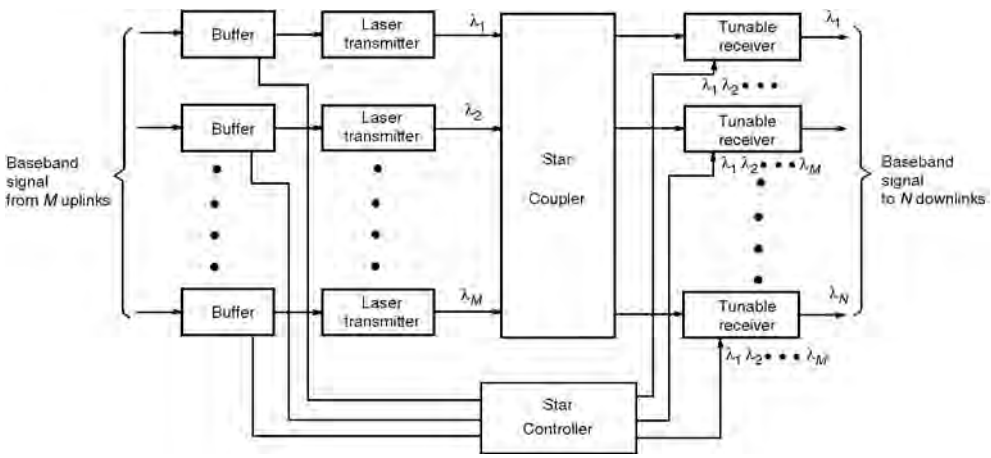


FIGURE 62.13 A WDM-based onboard baseband switch.

frame to the downlink. The reliability of the ring can be increased by using dual rings as in the FDDI standard, where failure of some parts or a whole ring does not affect the ring operation. Also, a bypass switch allows the isolation of faulty input and/or output processors. It is to be noted, however, that the access technique in FDDI is token ring and not TDM.

Wave Division Multiplexed (WDM) Baseband Switches

Figure 62.13 shows a WDM switch that uses a star coupler [Hinton, 1990]. This is a device with M optical inputs and N optical outputs that combines the M optical input signals, then divides them among the N optical outputs. Packetized baseband information is buffered at the various inputs. The input laser

has a fixed wavelength. The output receiver should be tuned to the same wavelength as that of the intended input. Once both the transmitter and receiver lasers are triggered by the star controller, the input buffer serially shifts the data bits into the laser transmitter. This controller also resolves the contention for switch outputs. This way, only one of the input packets is shifted, while the other contending input packet is left in the input buffer for the next time slot (packet time).

This represents the PS mode of operation, and the amenability to broadcasting is clear: if a certain input packet wavelength is to be broadcast to a set of outputs, all involved output receivers are tuned to the same input wavelength. For CS operation, the control commands are received from the ground station via out-of-band signalling channels. These reconfigure the star controller of the WDM switch by the arrival or termination of each new call. These commands may also be easily derived from a signalling TDM subframe slots (in-band signalling).

A mix of CS and PS traffic is easily accommodated, however, dummy bits would be inserted into CS traffic or the packet header stripped from PS traffic before entering the switch so as to have the same packet size passing through the WDM switch. Also, this switch becomes very efficient for directly switching CS intersatellite links in the optical domain.

For a cluster of low orbit satellites (LEO) (64 in the Global Star proposal [Louie, Rouffet, and Gilhausen, 1992]), optical intersatellite communications may render ground gateways (hubs) unnecessary and will, hence, reduce the end-to-end user delay. However, with an efficient all optical switch such as WDM, we save the expensive and delay prone optical/electronic and electronic/optical conversions. Reconfiguration control of CS traffic is still done electronically by the arrival of each new call and does not have to be ultrafast. This could be achieved by an out-of-band signal emanating from the ground control station. This signal carries the time slot plan to each involved LEO to enable it to communicate with its neighboring satellites. WDM technology has reached its maturity in terrestrial applications. Among WDM experiments is the HYPASS, where receiver wavelengths are fixed and the input laser transmitters are correspondingly tuned to the destination address. There is nothing in the weight, size, or cost that will forbid the use of these WDM switches in intersatellite switching applications so as to provide a total switch throughput over 50 Gb/s.

Asynchronous Transfer Mode (ATM)-Oriented Switches

Onboard switches handling PS and CS traffic could become more involved with the passing of wideband ISDN and multimedia multirate services through the multibeam satellite. With the advent of ATM-based WBISDN, signals of different kinds (video, voice, data, facsimile, etc.) are all physically transported on a standard size packet called a *cell*.

Processing a standard size cell through the OB self-routing fast packet switch (multistage or Banyan types) makes the dummy bits addition and header stripping circuits mentioned earlier unnecessary. All services (PS, CS, variable bit rate, etc.) will have the same cell format. The cell has a header part and a data part. The cell size is 53 bytes, and the header mainly consists of two parts (Fig. 62.14). The first is the 12-b virtual path identifier (VPI) and the second is a 16-b virtual circuit identifier (VCI), which is

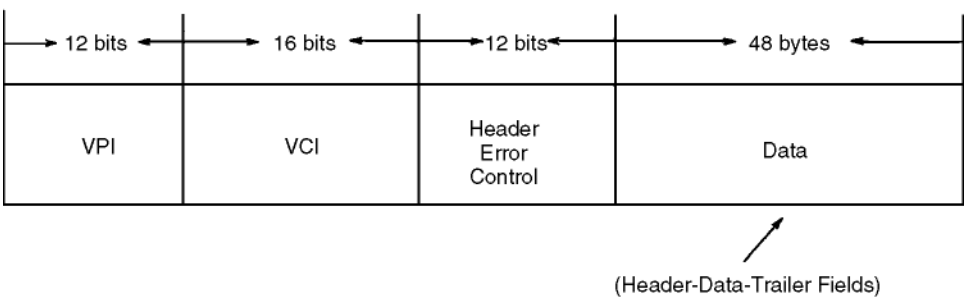


FIGURE 62.14 ATM cell user network interface (UNI).

followed by 12 b of error detection. The VPI is used to route the cell from satellite to next satellite or to ground, etc., enroute to the destination across the internetwork. The second cell field is a VCI between the two communicating entities. Cells find their way onboard the satellite fast packet switch by a route identifier. This route identifier is obtained by feeding the VPI to a translation lookup table.

However, this translation process may become unnecessary if the VPI is configured such that its first few bits are the routing tag through the satellite packet type switch. This is one of many possible scenarios. If the VPIs are not permanently assigned to ground user stations, the satellite should also have OB translation tables (to obtain switching routing tag) for both the VPI and the VCI of each incoming cell, otherwise different cells from different source–destination pairs may have the same value in their VPI and VCI fields. This implies that the satellite will have the baseband routing switch (ABN type, for example) as well as an ATM switch. This complicates the OB processing, slows down the operation, and is not recommended. Thus far, we have assumed an ATM type of user station. For stations launching packets, not cells, to the satellite, the satellite should still have an ATM switch onboard and a packet processing facility to divide the packet into cells, to add VPIs and VCIs to cells, and so forth. For ground user stations launching CS traffic, the ATM switch onboard will assign a VPI and a VCI to the incoming traffic and then route the traffic through the onboard baseband switch (BN, photonic, etc.). It is easily seen that unless the ATM users of the world have the ATM interface, the traffic arriving to the satellite during the early migration years of ATM technology may have cell, packet, or CS format, thus complicating onboard processing. However, this might be inevitable if satellite vendors are to provide the same service to ground-based optical network users on standby or even on a competitive basis. Once ATM supersedes other switching modes, OB processing will be less cumbersome except for the VPI and VCI processors referred to earlier.

62.3 Summary

Numerous switching techniques are in use or being developed for ground-based networks. We have tried to present the most relevant to the satellite environment. Knockout switches, token ring switches, and back-to-back hybrid packet/circuit switching, such as that of the ACTS system, are some of the techniques that could find application OB.

To compare the different switching architectures, one has to look at their throughput blocking, queueing delay, and reliability performance (see the Further Information section).

Cost of hardware, weight, power consumption, and size of the OB switch are crucial considerations. Comparing the different architectures based on these could be a less quantifiable process. Comparing them on a basis such as amenability to migration to ATM environments, and to different uplinks and downlinks access techniques, could be both art and science.

Defining Terms

Banyan network (BN): A self-routing switching technique for baseband signals.

Circuit switching (CS): A channel sharing technique where users are assigned physical or virtual circuits throughout their communications.

Demand assignment multiple access (DAMA): A multiaccess technique where each user is assigned a varying size time and/or frequency slot depending on his call needs.

Frequency-division multiple access (FDMA): A multiaccess technique where each user is assigned a certain frequency band for his call.

Memory switch (MS): A baseband switching technique for OB processing.

Microwave switch matrix (MSM): A switching technique for microwave signals.

Onboard (OB): Part of the satellite payload.

Packet switching (PS): A transmission and networking technique where the data units are called packets.

Satellite switching (SS): A satellite that redirects signals in the time and/or frequency and/or space domains.

Time-division multiple access (TDMA): A multiaccess technique where each user is assigned one or more slots in a certain time frame for his call.

Time–space–time (T–S–T): A baseband switching technique based on multiplexing in the time and space domains.

Wave-division multiplexing (WDM): A multiplexing technique for optical fiber networks also proposed for OB switching.

References

Cashman, W.F. 1992. ACTS multibeam communications package, description and performance characterization. In *Proc. AIAA 92 Conference,* AIAA, Washington, DC, March, 1151–1161.

Evans, B.G. et al. 1986. Baseband switches and transmultiplexers for use in an on-board processing mobile/business satellite system. *IEE Proc.,* Pt. F, 133(4):336–363.

Hinton, S.H. 1990. Photonic switching fabrics. *IEEE Commun. Mag.,* 28(4):71–89.

Huang, A. and Knauer. 1984. Starlite—a wideband digital switch. In *Proc. IEEE GlobeCom-84,* November, 5.3.1–5.3.5.

Hui, J.Y. and Arthurs, E. 1987. A broadband packet switch for integrated transport. *IEEE J. Selec. Areas Commun.,* SAC-5 (8):1264–1273.

Inukai, T., Faris, F., and Shyy, D.J. 1992. On-board processing satellite network architectures for broadband ISDN. In *Proceedings of the AIAA 92 Conference,* AIAA, Washington, DC, March, 1471–1483.

Inukai, T. and Pontano, A.B. 1993. Satellite unheard baseband switching architectures. *European Trans. Telecomm.,* 4(1):53–61.

Karol, M.J., Hluchyj, M.G., and Morgan, S.P. 1987. Input versus output queueing on a space-division packet switch. *IEEE Trans. Com. Comm.,* COM-35(Dec.):1347–1356.

Kumar, V.P. and Reibman, A.L. 1989. Failure dependent performance analysis of a fault-tolerant multistage interconnection network. *IEEE Trans. Comp.,* 38(12):1703–1713.

Kwan, R.K. et al. 1992. Technology requirements for mesh VSAT applications. In *Proc. AIAA 92 Conference,* AIAA, Washington, DC, March, 1304–1314.

Louie, M., Rouffet, D., and Gilhausen, K.S. 1992. Multiple access techniques and spectrum utilization of the GlobalStar mobile satellite system. In *Proc. AIAA 92 Conference,* AIAA, Washington, DC, March, 903–911.

Maronicchio, F. et al. 1988. Italsat—The first preoperational SS-TDMA system. In *Proc. IEEE GlobeCom 88 Conference,* March, 53.6.1–53.6.5.

Oie, Y., Suda, S., Murata, M., and Miyahara, H. 1990. Survey of the performance of nonblocking switches with FIFO input buffers. In *Proc. IEEE ICC 90 Conference,* April, 316.1.1–316.1.5.

Saadawi, T., Ammar, M., and Elhakeem, A.K. 1994. *Fundamentals of Telecommunications Networks,* Wiley Interscience, New York, chap. 12.

Shabtake, Y. and Kodama, T. 1990. A cell switching algorithm for the buffored Banyan network. In *Proc. IEEE ICC 90 Conference,* April, 316.4.1–316.4.7.

Further Information

Kwan et al. [1992] and Cashman [1992] provide a good treatment of the architectures and applications of microwave IF switches. Maronicchio et al. [1988], Evans et al. [1986], and Inukai and Pontano [1993] cover on-board baseband switches. The architectures and analysis of Banyan and other terrestrial baseband classes are investigated in Hui and Arthurs [1987], Huang and Knauer [1984], Carol, Hluchyj, and Morgan [1987], Oie et al. [1990], and Saadawi, Ammar, and Elhakeem [1994]. Photonic switching and LEO satellites are covered in Hinton [1990] and Louie, Rouffet, and Gilhausen [1992]. Fault tolerant switches are discussed in Kumar and Reibman [1989], which also contains an exhaustive list of references.

63

Path Diversity

Curt A. Levis
The Ohio State University

63.1 Concepts

Propagation impairments along the path between a satellite and Earth station are often spatially inhomogeneous. For example, cells of severe rain are generally a few kilometers in horizontal and vertical extent; the regions of destructive interference due to scintillations or to multipath generally have dimensions of a few meters, and so do the field variations due to roadside trees in a satellite communications system for land–mobile use. One way to ameliorate such impairments is to vary the path. When an additional Earth station is used for this purpose, it is called site diversity [Fig. 63.1(a)]; when an alternate satellite is used, it is termed orbital diversity [Fig. 63.1(b)].

Site and orbital diversity are not mature disciplines; there are very few data on the operational performance of complete systems. Research has focused on the propagation aspects and, to a lesser extent, on processor design and performance.

In principle, site and orbital diversity may be used on both uplinks and downlinks. In practice, the downlink problem is usually the more critical because of the limited power available aboard satellites, whereas the uplink problem is the more difficult when the processing is performed on the ground where the synchronization of signals at the satellite cannot be observed directly. For these reasons, site diversity processing has centered primarily on downlinks. Of course, attenuation data are applicable to both downlink and uplink design.

63.2 Site Diversity Processing

For site diversity reception, the signals from the various terminals must be linked to a common point, synchronized, and combined or selected according to some control algorithm. In principle, this can be done at intermediate frequency (IF), but baseband seems to be the preferred implementation. The links must have high reliability, e.g., coaxial lines or fiber optics. For wide-area service using very small aperture

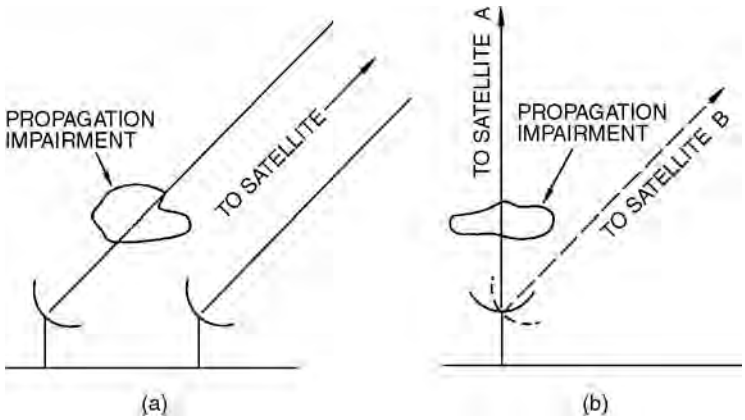


FIGURE 63.1 Path diversity concepts: (a) site diversity and (b) orbital diversity.

terminals (VSATs) accessed through a metropolitan area network (MAN), the MAN itself has been proposed as the link [Spracklen, Hodson, and Heron, 1993].

Sychronization generally involves both a fixed and a variable delay. For low data rates (e.g., 2 Mb/s), the fixed delay may be a digital first-in first-out (FIFO) memory; for higher rates (e.g., 34 Mb/s), a surface acoustic wave (SAW) delay line at IF has been proposed. The variable delay may be an elastic digital memory, which may be implemented as an addressable latch coupled to a multiplexer [Di Zenobio et al., 1988].

One method of combining signals is to switch to the best of the diversity branches, with carrier-to-noise ratio (C/N) or bit error rate (BER) as criterion. To avoid excessive switching when signal qualities are approximately equal, hysteresis in the switching algorithm is helpful. One algorithm causes switchover when

$$(E_0 \geq N) \wedge (E_s \leq N/H) = \text{TRUE} \tag{63.1}$$

where E_0 denotes the number of errors in a given time interval in the channel currently on line, E_s is the same for the current standby channel, and H is a hysteresis parameter appropriate for the current signal quality. The time interval is chosen to give N errors for a desired BER threshold [Russo, 1993].

A second approach is a weighted linear combination of the signals, with the weight for each branch determined on the basis of either C/N or BER. Considerable theory is available for designing optimized weight controls [Hara and Morinaga, 1990a, 1990b]. Such systems should outperform switching; however, their application has been considered only recently. Linear combining techniques developed for base-station site diversity in cellular systems may be adaptable to satellite systems.

63.3 Site Diversity for Rain-Fade Alleviation[1]

The concept of path diversity for rain-fade alleviation was first utilized on line-of-sight microwave paths; recognition of its applicability to satellites seems to date to the mid-1960s. It is most effective with systems having an appreciable rain margin, since cells of heavy rain generally do not extend far horizontally (Fig. 63.2). It is less effective at low elevation angles, for which substantial attenuation may result from stratiform rain.

[1]This section includes material from Lin, K.T. and Levis, C.A. 1993. Site diversity for satellite Earth terminals and measurements at 28 GHz. *Proc. IEEE,* 81(6):897–904. With permission.

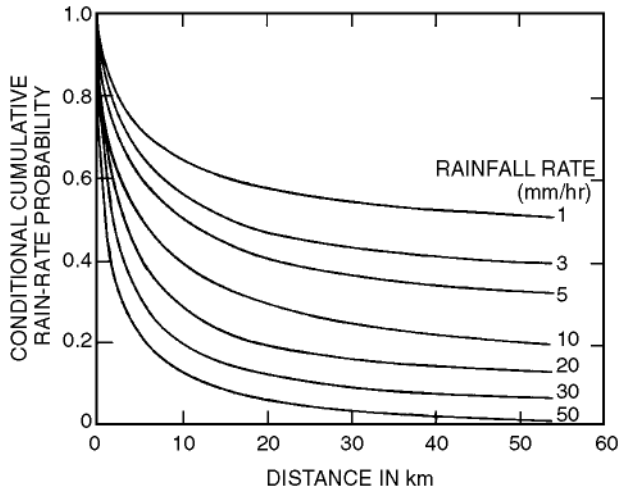


FIGURE 63.2 Probability that a rain-rate will be exceeded simultaneously at two stations, divided by the probability that it will be exceeded at a specific one of the stations. (*Source:* Adapted from Yokoi, H., Yamada, M., and Ogawa, A. 1974. Measurement of precipitation attenuation for satellite communications at low elevation angles. *J. Rech. Atmosph.* (France), 8:329–338. With permission.)

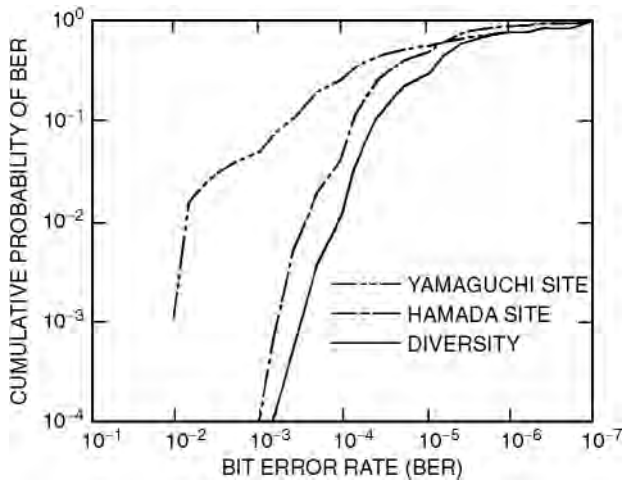


FIGURE 63.3 Bit error rate probability for an operational experiment in Japan. (*Source:* Watanabe et al. 1983. Site diversity and up-path power control experiments for TDMA satellite link in 4/11 GHz bands. In *Sixth International Conference on Digital Satellite Communications,* IEEE, New York, IX-21–IX-28. With permission.)

Measurements

An experiment with up- and downlink site diversity using switching based on BER determined from the forward error correcting (FEC) code was conducted in Japan with Intelsat *V* at approximately 6° elevation at 14 GHz (up) and 11 GHz (down) with a site separation of 97 km. Up- and downlinks were switched simultaneously. The data rate was 120 Mb/s. Results during a rain event appear in Fig. 63.3.

Diversity rain-fade measurements have been obtained by four different methods. The most direct measures the signal from a satellite beacon at two or more Earth stations; unfortunately, opportunities for such measurements have been quite limited. A second approach utilizes the thermodynamic law that absorbing materials at a finite temperature emit noise. By pointing radiometers in the proposed satellite

TABLE 63.1 Some Site-Diversity Propagation Experiments

Ref.	Location	Frequency (GHz)	Baseline (km)
Bostian et al. [1990] (B, R)[a]	Blacksburg, VA	11.4	7.3
Färber et al. [1991] (B)[a]	Netherlands	12.5, 30	10
Fionda, Falls, and Westwater [1991] (S)[a]	Denver, CO	20.6, 31.6	50
Goldhirsh [1984] (R)[a]	Wallops Island, VA	19.0, 28.6	£35
Goldhirsh [1982] (R)[a]	Wallops Island, VA	28.6	£35
Goddard and Cherry [1984] (R)[a]	Chilbolton, UK	11.6	£40
Lin, Bergmann, and Pursley [1980] (S)[a]	NJ, GA, IL, CO	13–18	11–33[b]
Lin and Levis [1993] (S)[a]	Columbus, OH	28.6	9.0
Pratt, Bostian, and Stutzman [1989] (R, B)[a]	Blacksburg, VA	11.5	7.3
Rogers and Allnutt [1990] (S)[a]	Atlanta, GA	12.0	37.5
Witternigg et al. [1993] (S)[a]	Graz, Austria	12.0	9–26[b]

[a](B) = satellite beacon, (S) = sky noise, (R) = radar.
[b]Includes multiple (3- or 4-site) diversity.

direction and measuring the excess sky noise, signal attenuation along the path can be inferred. It is difficult to measure attenuations in excess of about 15 dB by this method, and there are some uncertainties in converting the noise to equivalent attenuation; nevertheless, much useful data has been obtained from sky noise measurements. An efficient third means of data acquisition is to measure the radar reflectivity along the proposed paths. Through careful calibration and data processing, useful data may be obtained, although some assumptions must be made about the raindrop size distributions in converting reflectivity data to attenuation. An advantage of this method is that several diversity site configurations can be investigated in one experiment, provided the attenuations on all paths occur within the scanned radar volume. Finally, fast-response measurements with rain-gauge networks were the earliest means to infer propagation data, and rain-gauge measurements are still often correlated with propagation measurements.

A tabulation of site-diversity experiments prior to 1983 appears in the literature [Ippolito, 1989]. Table 63.1 lists additional experiments, with no pretension to completeness.

Data Presentation

The natural form of the data is a time series of attenuation samples for each path over one or more years. Statistical means are used to reduce the data to more convenient forms. A common graphical form of data presentation is in the form of fade-level statistics, where the ordinate represents the percentage of the time during which the attenuation specified by the abscissa is exceeded. Figure 63.4 shows the fade-level statistics for the two individual sites of a specific site diversity experiment and the corresponding statistics for ideal switching and maximal-ratio combining, both based on C/N. The effectiveness of the use of diversity is shown by the displacement, downward and to the left, of the combined-signal curves relative to those of the individual sites.

Propagation experiments and calculations have been interpreted traditionally in terms of switching to the best signal, with hysteresis neglected. Unless otherwise indicated, all diversity data presented here and in the literature should be interpreted accordingly.

For the system designer, statistics of the duration of fades and of the time between fades are also of interest. Figure 63.5 shows the fade-duration statistics for single sites and the corresponding diversity statistics; the **interfade interval** statistics for the same diversity experiment are shown in Fig. 63.6. The benefit of diversity is evident from the general reduction of fade durations and increase in the time between fades.

Fade-level statistics for a 3-site diversity configuration are shown in Fig. 63.7. It has been shown that average fade durations may be shortened by adding additional diversity sites, even if the fade-level statistics do not show a corresponding improvement [Witternig et al., 1993].

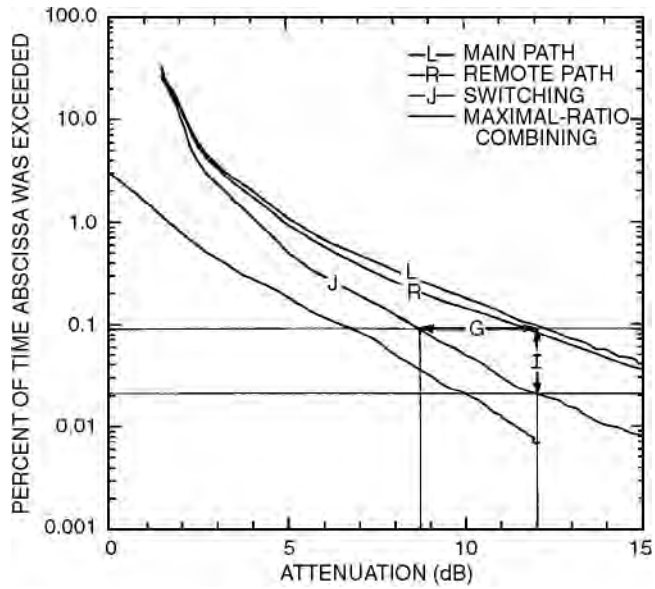


FIGURE 63.4 Fade-level statistics for single sites and two methods of processing at 28.6 GHz in Columbus, OH, with 9-km site separation and 25.6∞ elevation. (*Source:* Adapted from Lin, K.-T. and Levis, C.A. 1993. Site diversity for satellite Earth terminals and measurements at 28 GHz. *Proc. IEEE,* 81(6):897–904. With permission.)

Diversity Improvement Factor and Diversity Gain

Two criteria of diversity effectiveness have been defined in terms of fade-level distributions. One is the **diversity improvement factor**, defined [CCIR, 1990] as the ratio p_1/p_2, where p_1 denotes the percent of the time a given attenuation will be exceeded for the single-site distribution, and p_2 is the corresponding percentage for the combined-signal distribution. The diversity improvement factor (many authors use the term diversity advantage) may be defined with respect to any of the individual sites. If the system is balanced, that is, if the statistics for all of the participating sites are essentially the same, it is common practice to average the single-site attenuations (in decibel) at each probability level for use as the single-site statistics. The diversity improvement factor ***I*** is a measure of the vertical distance between the single-site and diversity curves, and it is proportional to that distance if the time-percentage scale is logarithmic. In Fig. 63.4, the diversity improvement factor at the 12-dB single-site attenuation level for switching-combining is shown to be about (0.09/0.021) = 4.3 for that experiment. The diversity improvement factor specifies the improvement in system reliability due to use of diversity. For the example of Fig. 63.4, if the system has a 12-dB rain margin, the probability of rain outage would be reduced by a factor of 4.3.

The most commonly used criterion for diversity effectiveness is the **diversity gain**, defined as the attenuation difference (in decibel) between the single-site attenuation and the diversity system attenuation at a given exceedence level. It is a measure of the horizontal distance between the single-site and diversity fade statistics curves, and it is proportional to that distance if the attenuation scale is linear in decibel. It represents the decrease, due to diversity, in the required rain fade margin. In Fig. 63.4, the distance G shows the diversity gain for that experiment for switching-combining at the 0.09% exceedence level as (12.0 - 8.7) = 3.3 dB. For an allowed rain failure rate of 0.09%, the fade margin could be reduced from 12 dB for single-site operation to 8.7 dB for switching-diversity operation.

Although diversity gain is defined as the attenuation difference between single-site and diversity systems at a given exceedence level p, it should be stated and graphed as a function of the corresponding single-site attenuation level $A_s(p)$; for example, in Fig. 63.4, G(12 dB) = 3.3 dB. This point is trivial when dealing with a single set of fade-level statistics, but it is significant when different sets of statistics are compared.

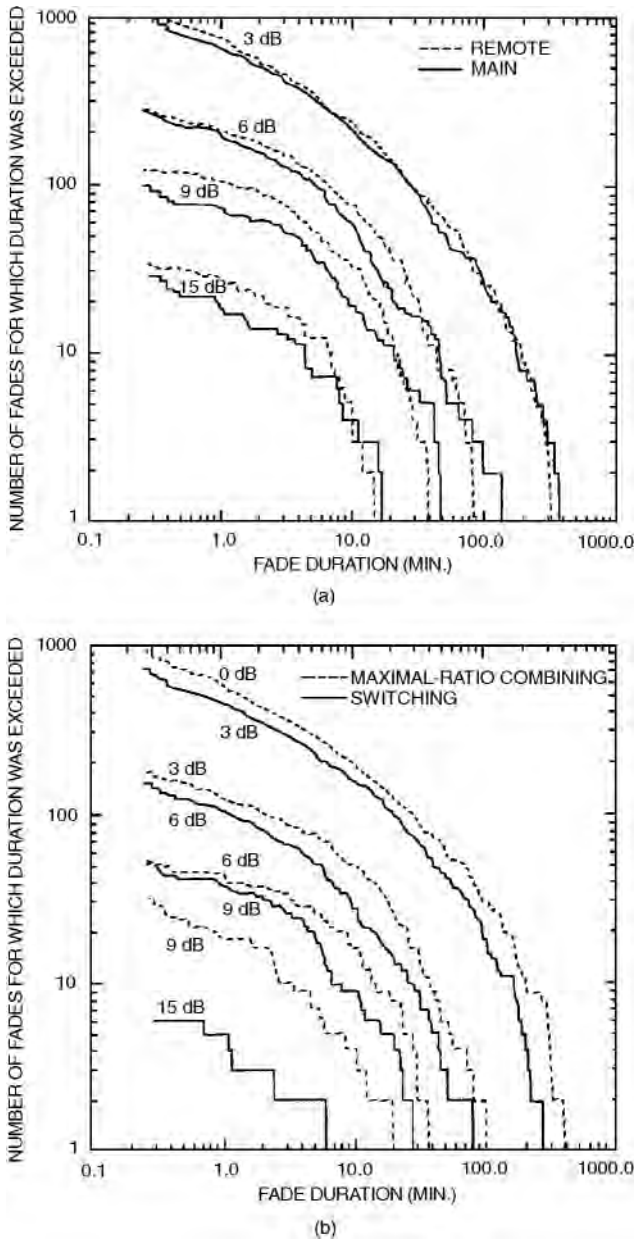


FIGURE 63.5 Fade-duration statistics for one year at 28.6 GHz in Columbus, OH, with 9-km site separation and 25.6∞ elevation: (a) single sites and (b) diversity for two methods of processing. (*Source:* Adapted from Lin, K.-T. and Levis, C.A. 1993. Site diversity for satellite Earth terminals and measurements at 28 GHz. *Proc. IEEE*, 81(6):897–904. With permission.)

The dependence of G on A_S has been found to be relatively independent of single-site statistics; this is not true of the dependence of G on p. Conversely, the model adopted by the CCIR for diversity improvement factor implies that the diversity improvement factor I is relatively independent of single-site statistics when it is specified as a function of p, but not when specified as a function of A_s.

Diversity gain and improvement are clearly not independent. Given the single-site fade-level statistics and the diversity improvement factor at all attenuation levels, the diversity fade-level statistics curve can

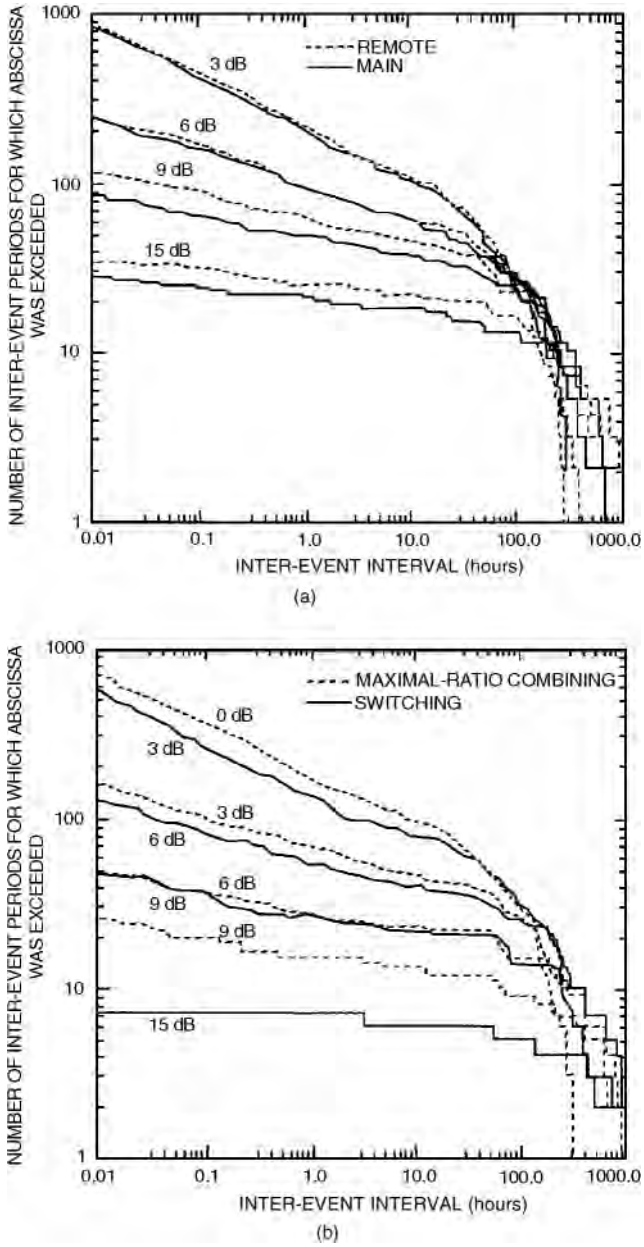


FIGURE 63.6 Interfade interval statistics for one year at 28.6 GHz in Columbus, OH, with 9-km site separation and 25.6° elevation: (a) single sites and (b) diversity for two methods of processing. (*Source:* Adapted from Lin, K.-T. and Levis, C.A. 1993. Site diversity for satellite Earth terminals and measurements at 28 GHz. *Proc. IEEE.*, 81(6): 897–904. With permission.)

be determined and, from it, diversity gain. Similarly, the diversity improvement factor can be determined if the single-site fade-level statistics and the diversity gain are given. An exception occurs in the practical case where reception is limited (e.g., by noise considerations) to a certain attenuation level. In this case, diversity data may extend to lower time percentages than the single-site data, and diversity gain will be undefined for such time percentages (e.g., in Fig. 63.4 for $p < 0.04\%$). Diversity gain and improvement factor can be defined for all methods of diversity signal processing (e.g., both switching and maximal-ratio

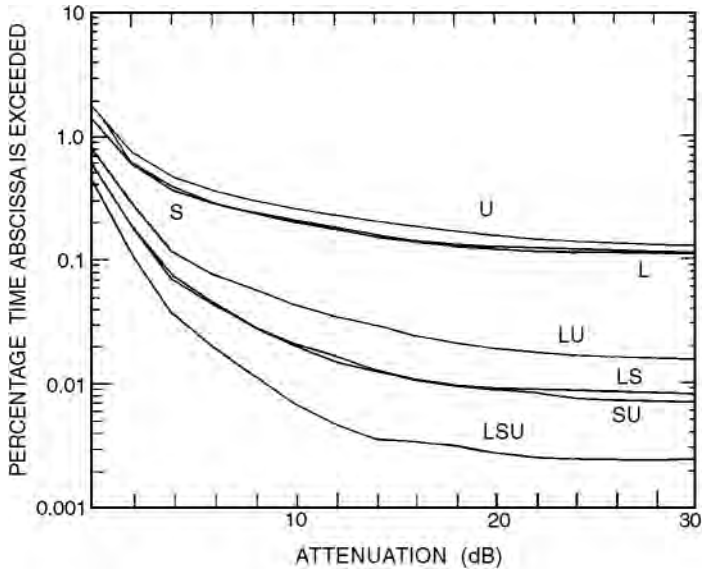


FIGURE 63.7 Single-site, two-site diversity, and triple diversity fade-level statistics at 19 GHz for three sites near Tampa, FL, with separations of 11, 16, and 20 km and 55∞ elevation. (*Source:* Tang, D.D., Davidson, D., and Bloch, S.C. 1982. Diversity reception of COMSTAR satellite 19/29-GHz beacons with the Tampa Triad, 1978–1981. *Radio Sci.,* 17(6):1477–1488. With permission.)

combining in Fig. 63.4), but current literature references ordinarily imply switching to the best available signal.

A **relative diversity gain** can be defined [Goldhirsh, 1982] as the ratio of the diversity gain for a given site separation to the diversity gain for a very large site separation, say, 35 km, with other system parameters held constant. The relative diversity gain is determined by the rain-rate profile along the Earth–satellite path and should be essentially independent of frequency and of rain drop-size distribution.

Isolation Diversity Gain

For systems utilizing polarization diversity to allow frequency reuse, isolation-level statistics, analogous to fade-level statistics, may be generated (Fig. 63.8). **Isolation diversity gain** may then be defined as

$$G_{\mathrm{ID}}(p) = I_j(p) - I_s(p) \tag{63.2}$$

where I_j and I_s denote, respectively, the diversity and single-site isolation, both at a selected time percentage p for which these isolations were not exceeded, for example, the horizontal displacement of the diversity curve with respect to a single-site curve in Fig. 63.8.

Prediction Models

Empirical models have been generated, generally by proposing a mathematical relationship on theoretical or observational grounds and adjusting the coefficients by regression analysis. Analytical models are based on theoretical models of the structure of rain or of the attenuation statistics of signals in the presence of rain. In practice, empirical models are much easier to apply, but they give much less information, being limited to a specific parameter, such as diversity gain.

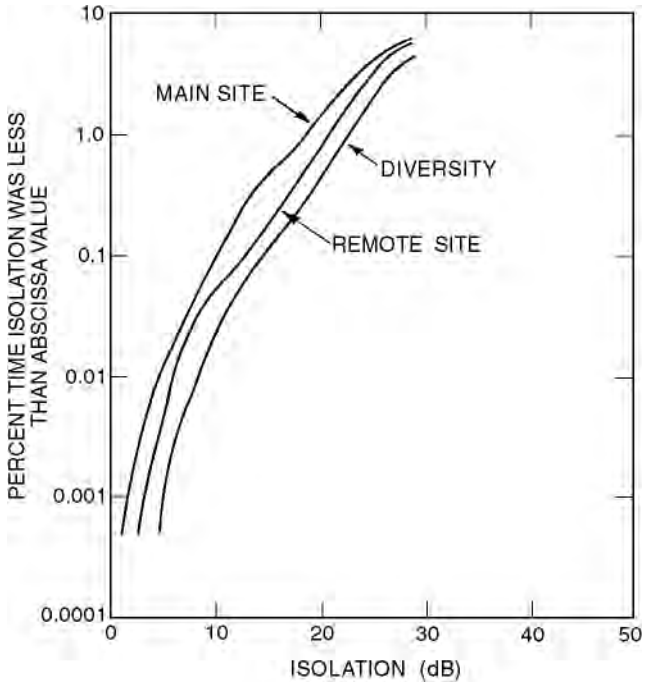


FIGURE 63.8 Polarization isolation statistics at 11.6 GHz in Blacksburg, VA, for 7.3 km site separation and 10.7° elevation. (*Source:* Towner et al. 1982. Initial results from the VPI & SU SIRIO diversity experiment. *Radio Sci.,* 17(6):1489–1494. With permission.)

Empirical Models

An empirical model for diversity gain in decibels [Hodge, 1982] has been adopted by the CCIR [1990],

$$G_D = G_d G_f G_q G_f \tag{63.3}$$

where

$$G_d = a\left(1 - e^{-bd}\right) \tag{63.4}$$

$$a = 0.78A - 1.94\left(1 - e^{-0.11A}\right) \tag{63.5}$$

$$b = 0.59\left(1 - e^{-0.1A}\right) \tag{63.6}$$

$$G_f = e^{-0.025f} \tag{63.7}$$

$$G_q = 1 + 0.006\theta \tag{63.8}$$

$$G_f = 1 + 0.002\phi \tag{63.9}$$

where d is the site separation in kilometers, A is the single-site attenuation in decibel, θ is the path elevation angle in degrees, f is the frequency in gigahertz, and $\phi \leq 90°$ is the baseline orientation angle

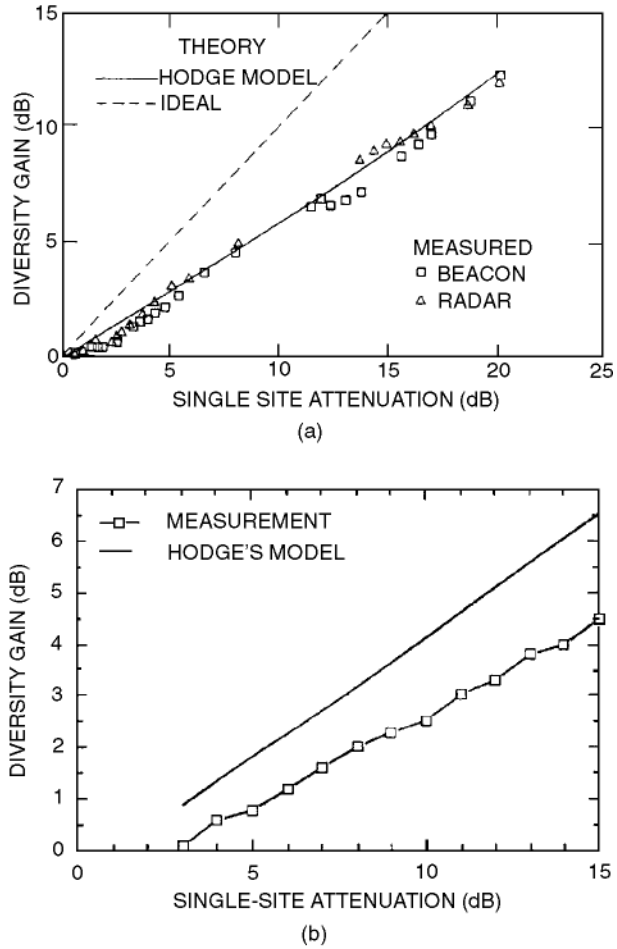


FIGURE 63.9 Comparison of Hodge's model with experiment: (a) 11.4-GHz, Blacksburg, VA, 7.3-km baseline, 18∞ elevation. (*Source:* Bostian et al. 1990. Satellite path diversity at 11.4 GHz: Direct measurements, radar observations, and model predictions. *IEEE Trans. Ant. Prop.*, AP-38(7):1035–1038. With permission.) (b) 28.6-GHz, Columbus, OH, 25.6∞ elevation, 9-km baseline. (*Source:* Adapted from Lin, K.T. and Levis, C.A. 1993. Site diversity for satellite Earth terminals and measurements at 28 GHz. *Proc. IEEE,* 81(6):897–904. With permission.)

with respect to the azimuth direction of the propagation path. When tested against the CCIR data set, the arithmetic mean and standard deviation were found to be 0.14 dB and 0.96 dB, respectively, with an rms error of 0.97 dB. Many comparisons of this model with experiments are found in the literature, mostly with good agreement. Two examples are given in Fig. 63.9.

Another empirical model [Rogers and Allnutt, 1984] is based on the typical behavior of diversity gain as a function of single-site attenuation (see Fig. 63.10). In this model, the offset between the ideal and observed diversity gain is attributed to correlated impairment on both paths due to stratified rain. For modeling purposes, a time percentage of 0.3% is assumed for the correlation, and the offset is computed as the single-site path attenuation for the rain-rate, which is exceeded 0.3% of the time. A straight line parallel to the ideal diversity-gain curve ($G_D = A_s$) is thus determined. The *knee* is determined by the rain rate at which the two paths become uncorrelated, that is, the onset of the convective regime, which is taken as 25 mm/h. Thus, the knee attenuation is calculated as the single-path attenuation for a rain-rate of 25 mm/h. The model is completed by a straight line from the knee to the origin.

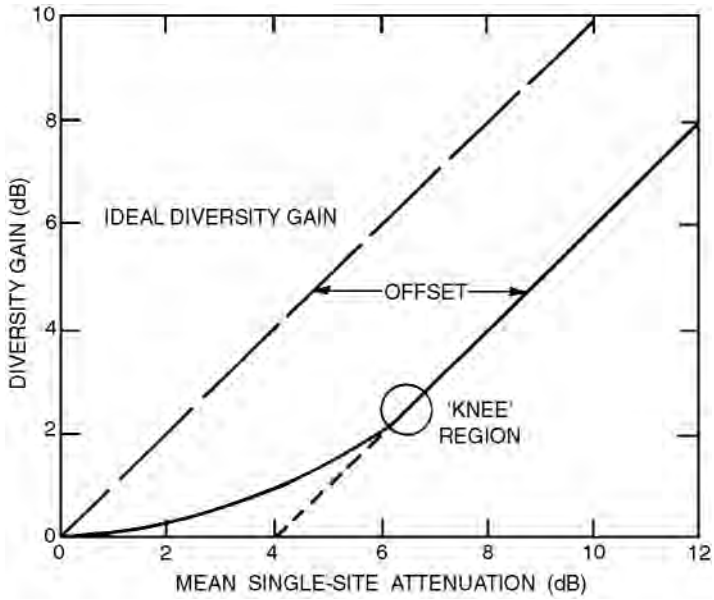


FIGURE 63.10 Site diversity model. (*Source:* Rogers, D.V. and Allnutt, J.E. 1984. Evaluation of a site diversity model for satellite communications systems. *IEE Proc. F.,* 131(5):501–506. With permission.)

An empirical expression for the relative diversity gain is

$$G_{D,\text{rel}} = 1 - 1.206e^{-0.531\sqrt{d}} \quad 1 \le d \le 30 \tag{63.10}$$

where d is the site spacing in kilometer [Goldhirsh, 1982].

An empirical model for the diversity improvement factor, which has been adopted by the CCIR [1990] but is not as widely supported in the literature as the Hodge diversity gain model, is

$$I = \frac{1}{1+\beta^2} \times \left(1 + \frac{100\beta^2}{P_1}\right) \tag{63.11}$$

where

$$b^2 = 10^{-4} d^{1.33} \tag{63.12}$$

the site separation in kilometer again is d, and P_1 represents the single-site time percentage of exceedence of a given attenuation, that is, the ordinate of a fade-level statistics diagram such as Fig. 63.4. For the usual case of small β^2, this reduces to

$$I = 1 + \frac{100\beta^2}{P_1} \tag{63.13}$$

In Fig. 63.11, experimental data is compared with this model.

Analytical Models

Analytical models are usually more cumbersome and computationally intensive than empirical models, but they are potentially applicable to a wider class of problems, for example, the determination of diversity gain for three or more sites.

TABLE 63.2 Mathematically Based Models

Ref.	Probability Distribution	Spatial Correlation
Matricciani [1983]	lognormal	$e^{-\beta d}$
Kanellopoulos and Koukoulas [1987]	lognormal	$G/(G^2+d^2)^{1/2}$
Kanellopoulos and Koukoulas[a] [1990]	lognormal	$G/(G^2+d^2)^{1/2}$
Kanellopoulos and Ventouras [1990a]	lognormal	$e^{-\alpha\sqrt{d}}$
Kanellopoulos and Ventouras [1990b]	lognormal	$G/(G^2+d^2)^{1/2}$
Kanellopoulos and Ventouras[a] [1992]	gamma	$G/(G^2+d^2)^{1/2}$
Koukoulas and Kanellopoulos [1990]	gamma, lognormal	$e^{-\alpha\sqrt{d}}$

[a]Three-site diversity.

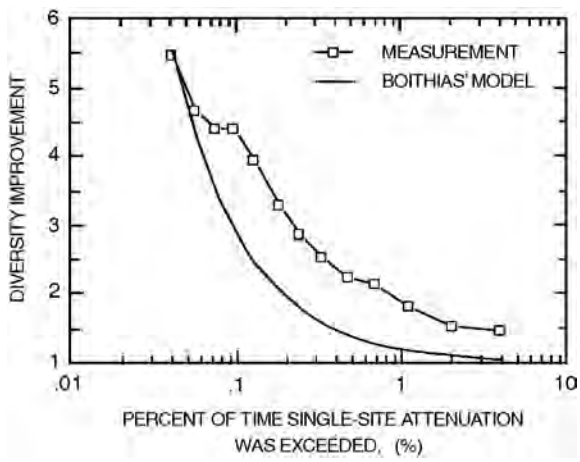


FIGURE 63.11 Comparison of CCIR diversity improvement factor model, due to Boithias, with experiment at 28.6 GHz in Columbus, OH, for 9-km site separation, 25.6° elevation. (*Source:* Adapted from Lin, K.-T. and Levis, C.A. 1993. Site diversity for satellite Earth terminals and measurements at 28 GHz. *Proc. IEEE,* 81(6):897–904. With permission.)

Predictions based on a rain model consisting of storm cells and debris which are mutually uncorrelated, with distributions and parameters based on quite detailed meteorological considerations, show rms deviations of 59% in attenuation and 222% in probability when compared to 48 diversity pairs in the CCIR data base [Crane and Shieh, 1989].

A model based on circularly cylindrical rain cells of constant rain rate and uniform horizontal distribution was used in Mass [1987]. The cell diameters are lognormally distributed, with the median dependent on rain rate, and the cylinder heights are given by four equally probable values, which are also dependent on rain rate. Calculated attenuation differences are compared graphically with experimental values for some 20 paths, principally at yearly exceedence levels of 0.1, 0.01, and 0.001%. A bias of 0.2 dB and rms deviation of 0.9 dB can be calculated from these data.

When a suitable n-dimensional joint rain-rate probability distribution is assumed for n diversity paths, and a suitable spatial correlation coefficient is assumed for the rain rate, more detailed meteorological considerations can be avoided. Models using this approach are listed in Table 63.2. In most cases, the model results are compared with a limited set of experimental data.

A comparison of seven models for diversity gain with numerous measurements [Bosisio, Capsoni, and Matricciani, 1993] should be used with caution because of the weight it accords differences at low attenuations, which generally have little influence on system performance.

63.4 Optical Satellite Site Diversity

For optical Earth–space paths, cloud cover is the most serious source of obscuration, with correlation distances extending, in some cases, to hundreds of miles. The link availability using a number of NASA sites singly and in three- and four-site diversity configurations has been calculated from meteorological data; the best single-site availability was 75%, although several four-site configurations were predicted to have 98% availability [Chapman and Fitzmaurice, 1991].

63.5 Site Diversity for Land–Mobile Satellite Communications

In the UHF and 1.5-GHz frequency bands considered for the land-mobile service, the principal impairment mechanisms are shadowing (for example, by the woody parts of trees) and multipath due to terrain. Site separations of 1 to 10 m have been found useful for country roads, and an expression

$$I(A,d) = 1 + [0.2\log_e(d) + 0.23]A \tag{63.14}$$

for the diversity improvement factor has been proposed, where d denotes distance in meters and A the single-site fade depth in decibel [Vogel, Goldhirsh, and Hase, 1992].

Because fade margins are likely to be low in these applications, linear signal combining is attractive. An optimization of weights has been proposed, based on modeling the signal as having a coherent (direct) and an incoherent (scattered) component in the presence of fast Rician noise. Land, sea, and aircraft applications of this model differ in the representation of the incoherent component [Hara and Morinaga, 1990a, 1990b]. A Bayes-test algorithm for weight optimization has also been proposed [Sandström, 1991].

63.6 Site Diversity for Arctic Satellite Communications

Propagation measurements at 6 GHz in the Canadian arctic at 80°N latitude have shown clear-air fading and enhancements in the summer under conditions of sharply defined temperature inversions at the low elevation angles required for geostationary satellites [Strickland, 1981]. These variations were equally well correlated for horizontal site separations of 480 m and vertical separations of 20 m, but were qualitatively uncorrelated at vertical separations of 180 m. These results appear to be predicted well by modeling the interface as sinusoidally corrugated, resulting in vertical coverage gaps of size

$$g = 1.039 \times 10^{-3}\left[\frac{\lambda^2}{A}\left(\frac{\Delta N}{\phi}\right)^2 h^2\right]^{1/3} \tag{63.15}$$

where λ is the corrugation wavelength, A its amplitude, h its height above the observer, and ϕ the elevation angle of the path, with g, λ, h, and A in meters, ΔN in N-units, and ϕ in degrees.

In a systems experiment, a diversity system at this location communicated with a single southern station via the Anik B satellite in the 6/4-GHz frequency bands. Both C/N and BER were used, separately, as switching criteria, with data rates up to 1.544 Mb/s. On the S–N path, a single frequency was transmitted; on the N–S path, each diversity site transmitted on its own carrier to allow signal selection at the southern station. Figure 63.12 shows the experiment configuration and results.

63.7 Microscale Diversity for VSATs

The large separations required for site-diversity protection from rain fades may be too expensive for low-cost systems utilizing very small aperture terminals, but smaller separations can protect against fading due to scintillations [Cardoso, Safaai-Jazi, and Stutzman, 1993]. The improvement may be small in cases

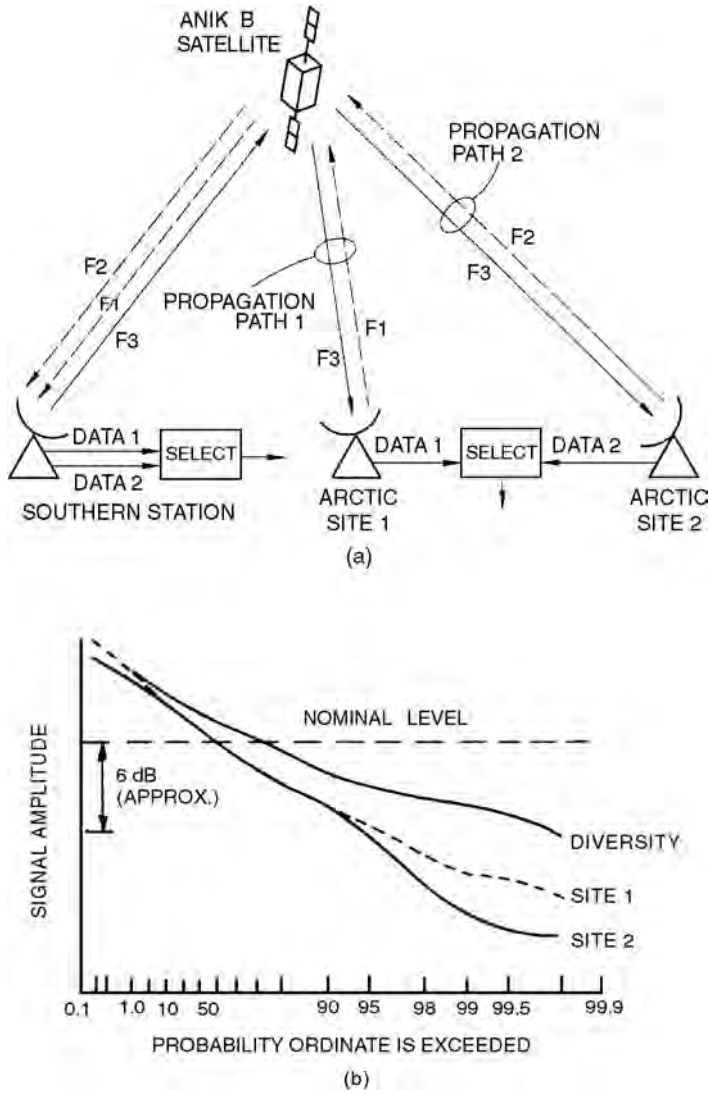


FIGURE 63.12 Signal-level statistics for a diversity experiment at 4/6 GHz at Eureka, Canada (80°N) with 150-m vertical antenna separation: (a) experiment configuration and (b) signal statistics. (*Source:* Adapted from Mimis, V. and Smalley, A. 1982. Low elevation angle site diversity satellite communications for the Canadian arctic. *IEEE International Conference on Communications,* 4A.4.1–4.A.4.5. With permission.)

where rain fades predominate. Fade-level statistics for a separation of 48 m are shown in Fig. 63.13. The correlation of the scintillations is examined experimentally and theoretically in Haidara [1993].

63.8 Orbital Diversity

Orbital diversity is likely to be less effective than site diversity because the propagation impairment always is near the Earth terminals, so that propagation effects tend to be more correlated for orbital diversity. The calculated tradeoff between the required satellite separation angle, as viewed from the Earth station, for orbital diversity and the separation of sites for site diversity, which will give identical diversity gain, at 19 GHz at Fucino, Italy, is shown in Fig. 63.14. Economic factors favor putting the

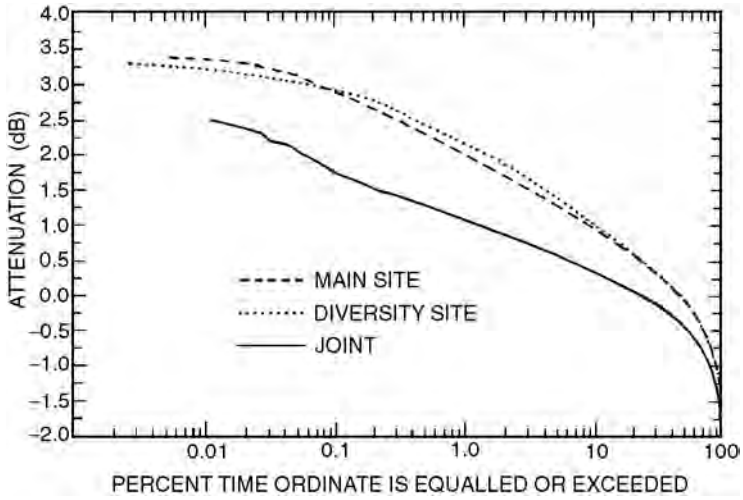


FIGURE 63.13 Microscale attenuation statistics at 20 GHz in Blacksburg, VA, with a 48-m baseline, 14° elevation, from a 1-h sample. (*Source:* Cardoso, J.C., Safaai-Jazi, A., and Stutzman, W.L. 1993. Microscale diversity in satellite communications. *IEEE Trans. Ant. Prop.*, 41(6):801–805. With permission.)

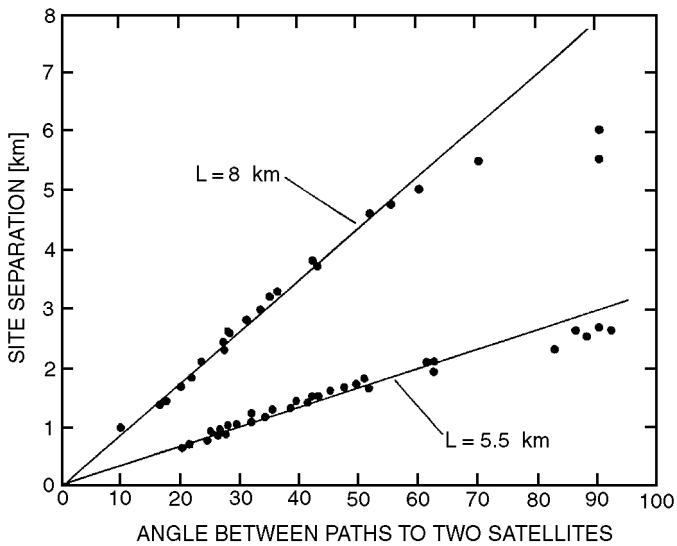


FIGURE 63.14 Relationship between site separation (for site diversity) and angular separation (for orbital diversity) for the same gain, calculated from radar data for the Fucino plain, Italy, at 19 GHz. L is the slant-path length through the rain of the site-diversity systems. (*Source:* Matricciani, E. 1983. An orbital diversity model for earth to space links under rain and comparisons with site diversity. *Radio Sci.,* 18(4):583–588. With permission.)

redundancy on the ground unless the system is usable by a great many Earth stations. A system utilizing spare channels on the Italsat and Olympus satellites and 100 Earth stations has been analyzed ignoring the economic aspects [Matricciani, 1987]. Although significant power margin reductions are attainable, it was concluded that many problems remain to be solved before orbital diversity can be considered practical.

Defining Terms

Diversity gain: The difference in the attenuation, measured in decibel, which is exceeded a given percentage of the time at a single reference site, and the attenuation which is exceeded the same percentage of the time for the combined diversity signal.

Diversity improvement factor: The ratio of the percent of the time that a given attenuation will be exceeded at a single reference site to the percent of the time that the same attenuation will be exceeded for the combined diversity signal.

Interfade interval: A time interval during which the attenuation remains less than a given threshold, while it is exceeded immediately before and after the interval.

Isolation diversity gain: The difference between a polarization isolation value in decibel of a diversity system, such that the observed isolation is less than this value a given percentage of time, and the corresponding value for a single reference site.

Relative diversity gain: The ratio of the diversity gain for a given site separation to that for a very large site separation, typically 35 km, with other system parameters held constant.

References

Bosisio, A.V., Capsoni, C., and Matricciani, E. 1993. Comparison among prediction methods of site diversity system performances. In *8th International Conference on Antennas and Propagation (ICAP 93)*, Vol. 1, 60–63 (IEE Conf. Publ. No. 370).

Bostian, C.W., Pratt, T., Stutzman, W.L., and Porter, R.E. 1990. Satellite path diversity reception at 11.4 GHz: direct measurements, radar observations, and model predictions. *IEEE Trans. Ant. Prop.,* AP-38(7): 1035–1038.

CCIR 1990. Propagation data and prediction methods required for Earth-space telecommunication systems. In *Reports of the CCIR, 1990,* Rept. 564-4, Annex to Vol. V, Propagation in nonionized media, International Telecommunication Union, Geneva, Switzerland, 458–463, 486–488.

Cardoso, J.C., Safaai-Jazi, A., and Stutzman, W.L. 1993. Microscale diversity in satellite communications. *IEEE Trans. Ant. Prop.,* 41(6):801–805.

Chapman, W. and Fitzmaurice, M. 1991. Optical space-to-ground link availability assessment and diversity requirements. In *Free-Space Laser Communications Technology III* (*Proc. SPIE* 1417), 63–74.

Crane, R.K. and Shieh, H.-C. 1989. A two-component rain model for the prediction of site diversity performance. *Radio Sci.,* 24(6):641–665.

Di Zenobio, D., Lombardi, P., Migliorni, P., and Russo, E. 1988. A switching circuit scheme for a satellite site diversity system. In *Proceedings of 1988 IEEE International Symposium on Circuits and Systems,* Vol. 1, 119–122 (IEEE Cat. 88CH2458-8).

Färber, K., Mawira, A., Quist, J., and Verhoef, G.J.C. 1991. 12 & 30 GHz Olympus propagation beacon experiment of PTT-Netherlands: space diversity and frequency dependence of the co-polar phase. In *7th International Conference on Antennas and Propagation (ICAP 91),* Part 1, 476–479 (IEE Conf. Publ. 333).

Fionda, E., Falls, M.J., and Westwater, E.R. 1991. Attenuation statistics at 20.6, 31.65, and 52.86 GHz derived from emission measurements by ground-based microwave radiometers. *IEE Proc., H* 138(1): 46–50.

Goddard, J.W.F. and Cherry, S.M. 1984. Site diversity advantage as a function of spacing and satellite elevation angle, derived from dual-polarization radar data. *Radio Sci.,* 19(1):231–237.

Goldhirsh, J. 1982. Space diversity performance prediction for Earth-satellite paths using radar modeling techniques. *Radio Sci.,* 17(6):1400–1410.

Goldhirsh, J. 1984. Slant path rain attenuation and path diversity statistics obtained through radar modeling of rain structure. *IEEE Trans. Ant. Prop.,* AP-32(1):54–60.

Haidara, F.M. 1993. Characterization of tropospheric scintillations on Earth-space paths in the Ku and Ka frequency bands using the results from the Virginia Tech OLYMPUS Experiment. Ph.D. Dissertation, Bradley Dept. of Elect. Eng., Virginia Polytechnic Institute and State University, Blacksburg, VA.

Hara, S. and Morinaga, N. 1990a. Optimum post-detection diversity of binary DPSK system in fast Rician fading channel. *Trans. Inst. Electron. Inf. Commun. Eng.* (Japan), E-73(2):220–228.

Hara, S. and Morinaga, N. 1990b. Post-detection combining diversity improvement of four-phase DPSK system in mobile satellite communications. *Electron. Commun. Japan 1,* 73(7):68–75. [Translated from *Denshi Joho Tsushin Gakkai Ronbunshi,* 1989. 72-BII(7):304–309.]

Hodge, D.B. 1982. An improved model for diversity gain on earth-space propagation paths. *Radio Sci.,* 17(6):1393–1399.

Ippolito, L.J. 1989. *Propagation Effects Handbook for Satellite Systems Design,* 4th ed., NASA Ref. Publ. 1082(04):7-78–7-107.

Kanellopoulos, J.D. and Koukoulas, S.G. 1987. Analysis of the rain outage performance of route diversity systems. *Radio Sci.,* 22(4):549–565.

Kanellopoulos, J.D. and Koukoulas, S.G. 1990. Prediction of triple-site diversity performance in Earth-space communication. *J. Electromagn. Waves Appl.* (Netherlands), 4(4):341–358.

Kanellopoulos, J. and Ventouras, S. 1990a. A modification of the predictive analysis for the multiple site diversity performance taking into account the stratified rain. *Eur. Trans. Telecommun. Relat. Tech.,* 1(1):49–57.

Kanellopoulos, J. and Ventouras, S. 1990b. A unified analysis for the multiple-site diversity outage performance of single/dual-polarized communication systems. *Eur. Trans. Telecommun. Relat. Tech.,* 1(6):625–632.

Kanellopoulos, J.D. and Ventouras, S. 1992. A model for the prediction of the triple-site diversity performance based on the gamma distribution. *IEICE Trans. Commun.* (Japan), E75-B(4):291–297.

Koukoulas, S.G. and Kanellopoulos, J.D. 1990. A model for the prediction of the site diversity performance based on the two-dimensional gamma distribution. *Trans. Inst. Elec. Inf. Commun. Eng.* (Japan), E-73(2):229–236.

Lin, K.-T. and Levis, C.A. 1993. Site diversity for satellite Earth terminals and measurements at 28 GHz. *Proc. IEEE,* 81(6):897–904.

Lin, S.H., Bergmann, H.J., and Pursley, M.V. 1980. Rain attenuation on Earth-satellite paths—summary of 10-year experiments and studies. *Bell Syst. Tech. J.,* 59(2):183–228.

Mass, J. 1987. A simulation study of rain attenuation and diversity effects on satellite links. *COMSAT Tech. Rev.,* 17(1):159–188.

Matricciani, E. 1983. An orbital diversity model for earth to space links under rain and comparisons with site diversity. *Radio Sci.,* 18(4):583–588.

Matricciani, E. 1987. Orbital diversity in resource-shared satellite communication systems above 10 GHz. *IEEE J. Sel. Areas Commun.,* SAC-5(4):714–723.

Mimis, V. and Smalley, A. 1982. Low elevation angle site diversity satellite communications for the Canadian arctic. In *IEEE International Conference on Communications,* 4A.4.1–4A.4.5. (IEEE Cat. 0536-1486/82).

Pratt, T., Bostian, C.W., and Stutzman, W.L. 1989. Diversity gain and rain height statistics for slant paths from radar measurements. In *6th International Conference an Antennas and Propagation (ICAP 89),* 340–344. (IEE Conf. Publ. 301.)

Rogers, D.V. and Allnutt, J.E. 1984. Evaluation of a site diversity model for satellite communications systems. *IEE Proc. F,* 131(5):501–506.

Rogers, D.V. and Allnutt, J.E. 1990. Results of a 12-GHz radiometric site diversity experiment at Atlanta, Georgia. *COMSAT Tech. Rev.,* 20(1):97–103.

Russo, E. 1993. Implementation of a space diversity system for K_a-band satellite communications. In *Proceedings of IEEE International Conference on Communications '93,* 1468–1474. (IEEE Cat. 0-7803-0950-2/93).

Sandström, H., 1991. Optimum processing of QPSK signals for site diversity. *Int. J. Satel. Commun.,* 9:93–97.

Spracklen, C.T., Hodson, K., and Heron, R. 1993. The application of wide area diversity techniques to Ka band VSATs. In *Electron. Div. Colloq. Future of Ka Band for Satell. Commun.,* 5/1–5/8. (IEE Colloq. Dig. 1993/215).

Strickland, J.I., 1981. Site diversity measurements of low-angle fading and comparison with a theoretical model. *Ann. Télécommun.* (France), 36(7–8):457–463.

Tang, D.D., Davidson, D., and Bloch, S.C. 1982. Diversity reception of COMSTAR satellite 19/29-GHz beacons with the Tampa Triad, 1978–1981. *Radio Sci.,* 17(6):1477–1488.

Towner, G.C., Marshall, R.E., Stutzman, W.L., Bostian, C.W., Pratt, T., Manus, E.A., and Wiley, P.H. 1982. Initial results from the VPI&SU SIRIO diversity experiment. *Radio Sci.,* 17(6):1489–1494.

Vogel, W.J., Goldhirsh, J., and Hase, Y. 1992. Land-mobile-satellite fade measurements in Australia. *J. Spacecr. Rockets,* 29(1):123–128.

Watanabe, T., Satoh, G., Sakurai, K., Mizuike, T., and Shinonaga, H. 1983. Site diversity and up-path power control experiments for TDMA satellite link in 14/11 GHz bands. In *Sixth International Conference on Digital Satellite Communications,* IEEE, New York, IX-21–IX-28.

Witternig, N., Kubista, E., Randeu, W.L., Riedler, W., Arbesser-Rastburg, B., and Allnutt, J.E. 1993. Quadruple-site diversity experiment in Austria using 12 GHz radiometers. *IEE Proc. H,* 140(5):354–360.

Yokoi, H., Yamada, M., and Ogawa, A. 1974. Measurement of precipitation attenuation for satellite communications at low elevation angles. *J. Rech. Atmosph.* (France), 8:329–338.

Further Information

Louis J. Ippolito, 1989, gives a very comprehensive treatment of propagation impairments and their amelioration, including path diversity systems. The fine treatment of the general propagation problem in *Satellite-to-Ground Radiowave Propagation,* 1989, by J.E. Allnutt, includes material on path diversity in Chapters 4 and 10.

The excellent text *Radiowave Propagation in Satellite Communications,* by Louis J. Ippolito, 1986, includes a good overview of space diversity in Chapter 10, including the application of Hodge's model to the performance prediction of switching diversity systems.

Although Lin et al., 1980, is limited to the research of one particular organization, it is so broad as to furnish much insight into the general problem of impairments and of space diversity as a means of alleviation.

A good review of early work with copious references appears in J.E. Allnutt, 1978, "Nature of space diversity in microwave communications via geostationary satellites: a review," *Proc. IEE,* 125(5):369–376.

A discussion of rain storms useful for path diversity applications is found in R.R. Rogers, 1976, "Statistical rainstorms models; their theoretical and physical foundation," *IEEE Trans. Antennas Propag.,* AP-24(4):547–566.

64

Mobile Satellite Systems

John Lodge
Communications Research Centre

Michael Moher
Communications Research Centre

64.1 Introduction

Mobile satellite systems are capable of delivering a range of services to a wide variety of terminal types. Examples of mobile satellite terminal platforms include land vehicles, aircraft, marine vessels, and remote data collection and control sites. Services can also be provided to portable terminals, which can range in size from that of a briefcase to that of a handheld telephone. Many services include position determination, using a combination of the inherent system signal processing and inputs from a Global Positioning System (GPS) receiver, and reporting the position estimates over the mobile satellite communications system. The types of the communication channels available to the user can be subdivided into four categories: store-and-forward packet data channels, interactive packet data channels, broadcast channels, and circuit-switched channels. Store-and-forward packet channels, which are the easiest to implement, allow for the transmission of small quantities of user data with delivery times that can be several minutes or more. This channel type is usually acceptable for services such as vehicle position reports (for example, when tracking truck trailers, rail cars, or special cargoes), paging, vehicle routing messaging, telemetry, telexes, and for some emergency and distress signaling. However, for many emergency and distress applications, as well as for interactive messaging (such as inquiry-based services), a delay of up to several minutes is unacceptable. For these applications, the second channel type, interactive packet data channels, is required. Services such as television and audio programming are delivered through broadcast channels, which are point-to-multipoint in nature. The quantities of user data per transmission are typically quite large, and delivery is relatively delay-insensitive. Finally, for applications involving real-time voice communications or the transmission of large amounts of data (for example, facsimile and file transfers), circuit-switched channels are needed. In many cases, voice services provide a mobile telephone capability similar to those offered by cellular telephone, but offering much broader geographical coverage. In these cases, a gateway station, which may provide an interface between the satellite network and the public switched telephone network (PSTN), communicates with the mobile via the satellite. In other cases, mobile radio voice services can be offered to a closed user group (such as a government or large company), with the satellite communications being provided between the mobiles and a base station.

Mobile satellite systems share many aspects with other satellite systems as well as with terrestrial mobile radio systems. With respect to other satellite systems, the main differences are due to the small size and mobility of the user's terminal. The largest antennas used by mobile satellite terminals are 1 m parabolic dish antennas for large ship-board installations. At the other extreme are antennas small enough to be part of a handheld terminal. Also, the transmit power level of the terminal is limited due to small antenna size and often due to restrictions on available battery power. The above factors have the consequence that mobile satellite systems operate at a significantly lower power margin than most mobile radio systems. Therefore, mobile satellite services are very limited unless there is line-of-sight transmission to the satellite.

64.2 The Radio Frequency Environment and Its Implications

When studying any radio system, it is generally instructive to begin by considering the basic link equation. This is particularly true for mobile satellite systems because the combination of limited electrical power on the satellite and at the mobile terminals and relatively small antennas on the mobile terminals results in a situation where transmitted power is a critical resource that should be conserved subject to the requirement of providing a sufficiently high signal-to-noise ratio to support the desired grade of service. Here, we only consider a simplified link equation with the objective of providing some context for subsequent issues.

Consider the following simplified link equation:

$$C/N_o = P_t + G_t - L_p + G_r - T_r - k - L \quad \text{dB-Hz} \tag{64.1}$$

where:

C/N_o: the ratio of the signal power to the noise power spectral density after being received and amplified (dB-Hz)
P_t: RF power delivered to the transmitting antenna (dBW)
G_t: gain of the transmitting antenna relative to isotropic radiation (dBi)
L_p: free-space path loss (dB)
G_r: gain of the receiving antenna (dBi)
T_r: composite noise temperature of the receiving system (dBK)
k: Boltzmann's constant expressed in decibels (-228.6 dBW/K-Hz)
L: composite margin accommodating losses due to a variety of transmission impairments (dB)

A link equation of this type can be applied to either the **forward direction** or the **return direction** as well as to the corresponding links between the satellite and the gateway or base station. Many satellites simply amplify and frequency translate the received signal from the uplink prior to transmission over the downlink. Therefore, the end-to-end signal-to-noise ratio must incorporate both uplink and downlink noise. Further examination of some of the constituents of the link equation can provide some interesting insights. For instance, the transmitting antenna gain is a measure of an antenna's ability to convert the electrical radio signal into an electromagnetic waveform and then to focus that energy toward the receiving antenna. Similarly, the receiving antenna gain is a measure of an antennas ability to collect the transmitted energy and then to convert the energy into an electrical signal that can be amplified and processed by the receiver. In general, the greater the gain of an antenna, the narrower its beam. As an example, consider a circular parabolic antenna. Its antenna gain is given by the following expression:

$$G = 10\log_{10}(\Omega\, \pi^2 D^2/\lambda^2) \text{ dBi} \tag{64.2}$$

where:

W: efficiency of the antenna
D: diameter of the antenna (m)
λ: wavelength of the RF signal (m)

While the circular parabolic antenna is only one example of many possible antennas, this equation provides an order-of-magnitude indication of the size of antenna needed to provide a desired amount of gain.

The free space path loss is given by

$$L_p = 10\log_{10}[(4\pi r)^2/\lambda^2] \quad \text{dB} \tag{64.3}$$

where r is the distance in meters between the terminal and the satellite.

Note that the path loss and both the transmitting and receiving antenna gains increase with the square of the radio frequency. From Eq. (64.1), this would imply that, given constraints on the dimensions of the antennas, there is a net benefit in using relatively high radio frequencies. While there is some truth to this, the benefits are not as great as one might think because the composite propagation losses also tend to increase with the radio frequency. Frequencies used for delivering mobile satellite services range from about 130 MHz up to 30 GHz. Use of the lowest frequency bands is dominated by low data rate store-and-forward services provided with low earth-orbit (LEO) satellites. Most mobile satellite telephone services, with typical transmission rates ranging from several kbit/s up to 64 kbit/s, are provided in frequency bands located between 1500 MHz and 2500 MHz. Interactive packet data and circuit-switched services, at transmission rates up to several Mbit/s, are planned for the higher frequency bands. This correlation between transmission rates and carrier frequencies is a result of both the above link considerations as well as the fact that larger spectrum allocations are available in the higher frequency bands than in the lower ones.

Efficient use of the satellite's two critical resources, spectrum and power, can be achieved if the satellite's antenna system covers the appropriate area of the earth's surface with multiple beams instead of one large beam. The total allocated system bandwidth is divided into a number of distinct subbands, which need not be of equal size, and each beam is then assigned a subband in such a way that some desired minimum isolation between beams with the same subband is maintained. This situation is illustrated in Fig. 64.1.

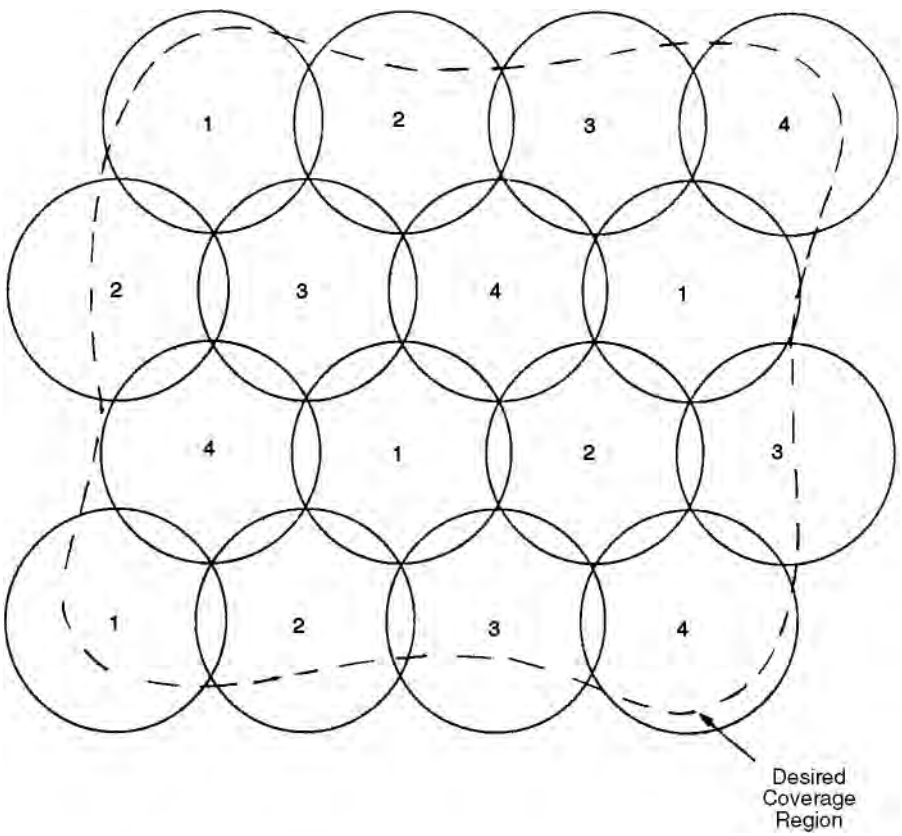


FIGURE 64.1 An example illustrating frequency re-use with a multiple-beam satellite. Here, the total allocated system frequency band is divided into four subbands. The assignment of the subbands to the beams is denoted by the number shown in each beam.

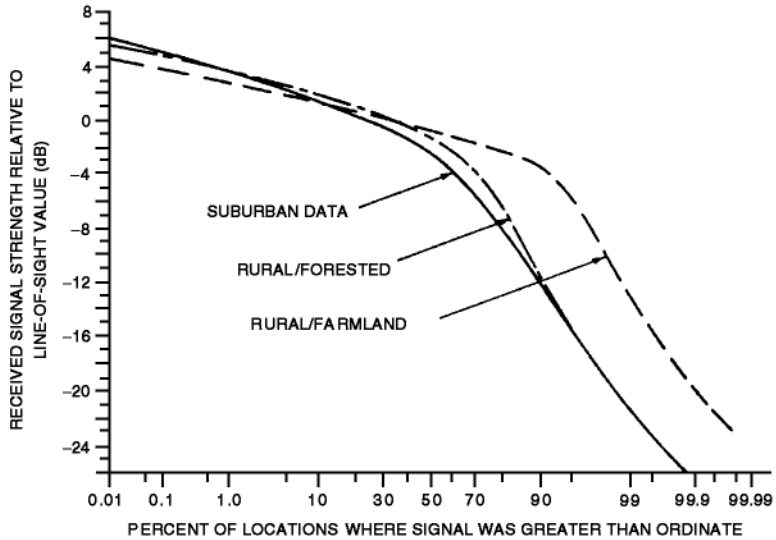


FIGURE 64.2 Distribution functions for data recorded near Ottawa, Canada in June 1983. The measurements were made with an omnidirectional antenna at a frequency of 1542 MHz. The elevation angle (the angle between a plane parallel to the earth's surface and the line-of-sight path to the satellite) was approximately 20° [Butterworth, 1984].

While not necessarily the case in an operational system, this example illustrates the beams as being circular and equal in area. Here, there are four distinct subbands and sixteen beams. Note that in this example each frequency subband is used four times. In general, the frequency reuse factor is defined to be the ratio of the number of beams to the number of distinct frequency subbands. Improved power efficiency is a direct result of the higher satellite antenna gain implicit in the smaller beam area. Ideally, the satellite should be able to apportion dynamically its total transmit power among the individual beams depending upon the loading in each beam.

The satellite mobile channel is subject to multipath fading and shadowing just as the terrestrial mobile channel. The physical mechanisms for these propagation phenomena are the same. Shadowing is due to foliage, buildings, or terrain blocking the line-of-sight signal from the satellite. Many propagation campaigns have been performed to investigate the statistics of shadowing in mobile satellite environments. For example, Fig. 64.2 shows some results of data gathered near Ottawa, Canada in June, 1983 [Butterworth, 1984]. Note that these curves exhibit a "knee" between 50 and 90%. For the portion of the curves to the left of the knee, the characteristics are dominated by multipath fading, whereas shadowing is the dominating factor to the right of the knee. With adequate margins, it is reasonable to expect mobile satellite systems to maintain communications links in environments characterized by fading and light shadowing, but it is not practical to overcome heavy shadowing for most services. For this reason, availability targets between 90 and 95%, corresponding to margins of less than 10 dB, are common for land-mobile satellite systems delivering services to users in rural, highway, and suburban environments. Availability figures for urban environments tend to be much lower. Exceptions include very low-capacity services, such as paging, which may require margins as high as 30 dB to allow for building penetration. Of course, the margin that is required to meet a given level of availability in a particular set of user environments is dependent upon many factors including the elevation angle to the satellite, the type of terrain, the density and height of the vegetation, and the number of man-made obstacles.

The shadowing is often modeled as a log-normal process [Butterworth, 1984; Loo, 1984] where the direct path power has a distribution given by

$$P[S < x_{\mathrm{dB}}] = 0.5 + 0.5\,\mathrm{erf}\left(\frac{x_{\mathrm{dB}} - m_s}{\sqrt{2}\,\sigma_s}\right) \tag{64.4}$$

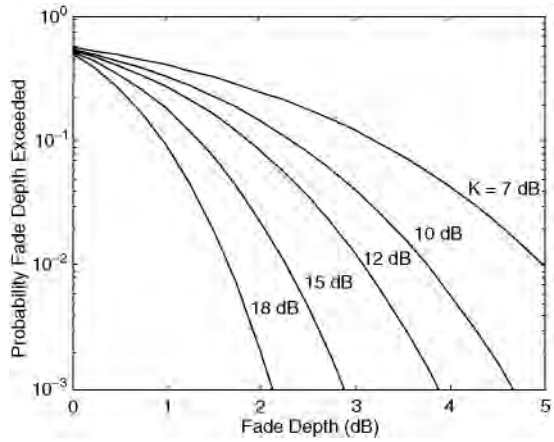


FIGURE 64.3 Cumulative fade distribution for the Rician channel as a function of K-factor.

where m_s is the received signal strength in dB, S is the average dB shadowing loss, and σ_s is the dB variance of the shadowing loss. The parameters of this distribution depend not only on environment but on the elevation angle to the satellite and on frequency. For a variety of situations, estimates of these parameters can be derived from the empirical results of these propagation campaigns provided in ITU-R Recommendation PN.681 and PN.681-1 and Stutzman [1988].

Multipath fading results from interference caused by reflections of the radio signal from the immediate environment, with the time-varying nature being caused by motion of the terminal through the environment. In satellite applications, the reflections are generally too weak to be useful by themselves but not so weak as to be ignored. Consequently, the satellite mobile channel is usually modeled as a Rician fading channel consisting of an unshadowed direct path and a diffuse component which has a complex Gaussian distribution (Rayleigh distributed amplitude and uniformly distributed phase). The amplitude distribution of the resulting channel is the non-central Chi-squared distribution with two degrees of freedom,

$$P[R < r] = 1 - e^{-(a^2+b^2)/2} \sum_{k=0}^{\infty} \left(\frac{a}{b}\right)^k I_k(ab) \tag{64.5}$$

where $a = \sqrt{K}$, $b = r/\sigma$, and $I_k(.)$ is the kth order modified Bessel function. The quantity K is the ratio of the direct path power to the average power of the multipath, $2\sigma^2$. Rician channels are characterized by their K-factor. The probability of a fade of a given depth is shown in Fig. 64.3 for various K-factors. Note that there is a greater than 50% probability that the diffuse component will increase the signal strength. K-factors with vehicle[1]-mounted antennas are typically 10 dB and higher, with the worst case value of 10 dB occurring at low elevation angles. For handheld terminals, where the antenna has little directivity, lower K-factors can be expected.

The time-varying nature of the fading, in particular, the duration of the fades, is characterized by its spectrum. For land-mobile applications, the Clarke model [Clarke, 1968], which assumes uniformly distributed scatterers in a plane about the moving terminal, is often used. The maximum Doppler frequency observed is related to the transmission frequency by the expression $f_d = v/\lambda$, where v is the velocity of the terminal, and λ is the transmission wavelength. The Clarke model for the diffuse spectrum is given by

$$S(f) = \frac{1}{\sqrt{1-(f/f_d)^2}} \tag{64.6}$$

[1]Vehicle refers to automobile, aircraft, or ship.

and the corresponding time correlation of the diffuse component is given by

$$R(\tau) = J_o(2\pi f_d \tau) \tag{64.7}$$

where $J_o(.)$ is the zeroth order Bessel function.

For the land-mobile channel, some effort has been devoted to providing an appropriate model combining the effects of shadowing and multipath fading which is appropriate for satellite applications. An example of these is the Lutz-model [Lutz et al., 1991], which consists of two states, a good state and a bad state. In the good state, the channel is Rician. In the bad state, the line-of-sight signal is essentially blocked due to significant shadowing. The parameters of the model are the distribution of the dwell times in each state, the K-factor, and the distribution of the shadowing, which are selected to model the environment of interest.

In aeronautical-mobile, the fading spectrum is generally modeled as Gaussian. However, measurements in this case indicate that the fading bandwidth is not strongly dependent on the aircraft velocity. The fading tends to be slow and is due to reflections from a limited portion of the earth's surface and from the aircraft's body. At L-band frequencies (1.6 GHz), the fading bandwidths for aeronautical-mobile channels range from 20 to 100 Hz with 20 Hz being more prevalent [Neul et al., 1987].

The physics involved in determining an appropriate model for the maritime channel are quite similar to that for the aeronautical case [Moreland, 1987]. Consequently, the fading spectrum is generally modeled as Gaussian. However, the fading bandwidth is typically less than 1 Hz [Schweikert and Hagenauer, 1983].

In general, multipath in satellite-mobile channels is not time-dispersive. Measurements indicate that for practical bandwidths (<2 MHz), this is true for wideband channels up to 2 MHz [Jahn et al., 1995]. A possible exception to this is the high-altitude aeronautical channel where one may occasionally receive distant reflections from the sea surface or ice. For this case, the strength of the reflected path is strong enough to be problematic only at elevation angles below 20° .

64.3 Satellite Orbits

Initial mobile satellite systems used geostationary orbits and were generally aimed at mobile terminals with significant antenna gain (>8 dB). As satellite technology evolved, applications have expanded to include terminals with much lower antenna gain (0 dBi). These include a number of regional mobile satellite services as well as the Inmarsat international service. The geostationary orbits provide good coverage in the low- to mid-latitude regions. Service availability to mobile units in the mid- to higher-latitude regions can be compromised due to blockage at the lower elevation angles. To achieve a higher service availability to a greater portion of the earth's surface necessitates the use of non-geostationary orbits.

With non-geostationary orbits, coverage from one satellite is time-varying, although the orbits are often chosen to be a submultiple of either the solar day (24 h) or the sidereal day (23 h 56 m 4 s). With the latter, the satellite ground tracks are periodic, while with the former, the ground tracks drift approximately 1° longitudinally each day. As a result of this time-varying behavior, a constellation of satellites must be employed to provide continuous coverage.

A variety of non-geostationary orbits and constellations have been proposed. These range from the low earth orbit (LEO) systems at altitudes from 500 to 2000 km, through the medium earth orbit (MEO) systems at altitudes from 9000 to 14,000 km, to the highly elliptical orbit (HEO) systems. One constraint on orbits is the Van Allen radiation belts which encircle the planet [Hess, 1968]. The primary belt from 2000 to 9000 km and the secondary belt from 14,000 to 19,000 km can restrict operation at these altitudes because of the greater radiation hardness requirements on the satellite components.

Constellations are generally configured as a number of orbital planes with a number of satellites in each plane. The number of satellites in the constellation generally decreases with the altitude. Most examples fall into the category of inclined circular orbits [Ballard, 1980], where the inclination refers to the angle

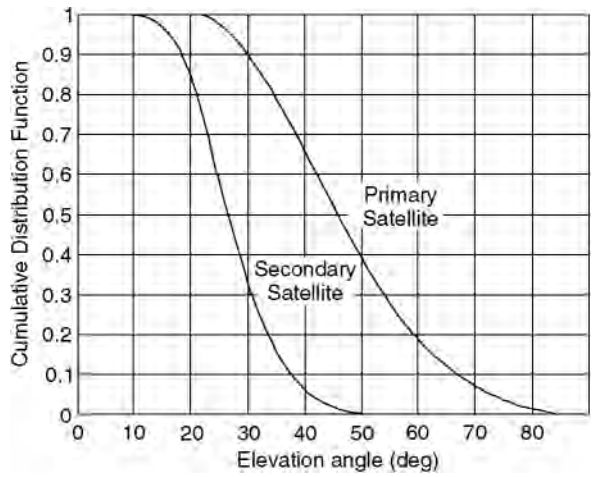


FIGURE 64.4 Cumulative distribution of elevation angle to primary satellite with a constellation of four inclined planes of three satellites at an altitude of 10,355 km.

the orbital plane makes with the equator. The extreme case of these are the polar-orbiting constellations [Adams and Rider, 1987]. The coverage provided by all these constellations is global in nature, but the minimum elevation angle and number of satellites visible simultaneously at any location depends on the constellation and its altitude. In general, simple constellations can provide good coverage of the earth. As an example, the cumulative distribution of the elevation angles (averaged over time and location) for a twelve satellite constellation is shown in Fig. 64.4. This figure indicates that the elevation to the primary visible satellite is always greater than 23° ; and 90% of the time/locations it is always greater than 30° . Figure 64.4 also indicates that a second satellite is always visible with an elevation of at least 10° .

An alternative to the inclined circular orbits are the elliptical orbits of which the Molniya [Ruddy, 1981] and Loopus [Dondl, 1984] are well-known examples. A constellation of satellites in highly elliptical orbits can approximate the effect of an overhead geo-synchronous satellite for high latitude regions. However, the useful coverage with elliptical orbits is restricted (generally they are only active near the orbit apogee), and they are intended for regional services.

While non-geostationary orbits ameliorate a number of the propagation phenomena affecting communications with a mobile terminal, they also introduce a number of challenges. In particular, with non-geostationary satellites, a number of quantities such as path delay and Doppler shifts become not only position-dependent but also time-varying, depending on the satellite orbital speed. The period of a satellite in an elliptic or circular orbit around the earth is

$$T = \frac{2\pi a^{3/2}}{\sqrt{GM}} \tag{64.8}$$

where a is the semimajor axis of the orbit (the radius in the case of a circular orbit) and GM is the earth's gravitational parameter ($GM = 3.99 \times 10^{14}$ m^3/s^2).

As with a geostationary satellite, the propagation delay and path loss vary with user position (range), and the general expression for satellite range as a function of elevation angle θ is given by

$$r(\theta) = R_e\left(\sqrt{(1 + h/R_e)^2 - \cos^2\theta} - \sin\theta\right) \tag{64.9}$$

where R_e is the earth's radius (6378 km) and h is the satellite altitude. The corresponding delay is $\tau = r/c$. Unlike geostationary satellite systems, there can be significant Doppler and propagation delay variations

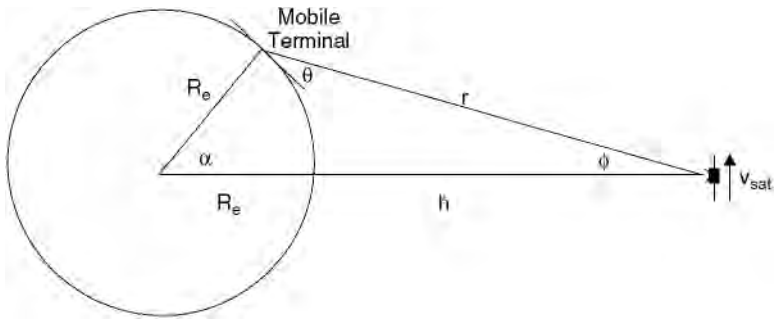


FIGURE 64.5 Illustration of satellite geometry for calculating range and range rates.

due to the relative velocity of the user and the satellite with non-geostationary mobile satellite systems. Both of these quantities are proportional to the range rate, with the delay variation given by

$$\frac{d\tau}{dt} = \frac{1}{c}\frac{dr}{dt} \tag{64.10}$$

where τ is the propagation delay, r is the range, and the Doppler shift is given by

$$f_s = \frac{1}{\lambda}\frac{dr}{dt} \tag{64.11}$$

Note that this Doppler frequency is distinct from the maximum Doppler rate used to characterize multipath fading. The latter corresponds to the mobile terminal velocity relative to its immediate environment. There are a number of contributors to the range rate. They include the satellite motion, the earth's rotation, and the velocity of the mobile terminal relative to the earth. In a number of simple cases involving circular orbits, the contribution of these various components can be calculated explicitly. When the user lies in the orbital plane, such that the geometry is as shown in Fig. 64.5, then the range rate due to satellite motion alone is given by

$$\left(\frac{dr}{dt}\right)_{\text{sat}} = -\frac{2\pi R_e}{T}\cos\theta \tag{64.12}$$

In addition, the velocity of the mobile terminal due to the earth's rotation is given by

$$v_r = \frac{2\pi R_e}{T_e}\cos\gamma \tag{64.13}$$

in a direction of constant latitude, where γ is the latitude of the mobile terminal, and T_e is the period of the earth's rotation (23 h 56 m 4s). In the general case, the range rate is the vector sum of the velocities of these different contributors projected along the range vector. This depends on the mobile terminal position as well as velocity and does not have a simple explicit formula [Vilar and Austin, 1991]. However, the general circular-orbit case can be upperbounded by Eqs. (64.12) and (64.13), projected in the direction of the satellite, to give

$$\left|\frac{dr}{dt}\right| \le 2\pi R_e \cos\theta\left(\frac{1}{T} + \frac{\cos\gamma}{T_e}\right) \tag{64.14}$$

This shows that both the Doppler and delay rate are zero at the subsatellite point and increases toward the satellite horizon, although the Doppler increase depends upon the relative direction. The Doppler can

range from zero for a geostationary orbit to greater than 42 kHz for a LEO satellite (1000 km) at 2.0 GHz. The corresponding delay rate can vary from zero to 21 microseconds per second.

Another parameter of interest is the rate of change of Doppler. There is no simple expression for the Doppler rate but, in a similar manner, it can be upper bounded by

$$\left|\frac{df_o}{dt}\right| \leq \frac{(2\pi)^2 R_e \sin\theta}{\lambda T}\left(\frac{1}{T} + \frac{\cos\gamma}{T_e}\right) \tag{64.15}$$

Equations (64.14) and (64.15) do not include the effect of mobile terminal motion relative to the earth. The Doppler rate can range from zero for a geostationary satellite to 100 Hz/sec for a LEO satellite at 1000 km at 2.0 GHz depending upon the positions of the user and the satellite.

In addition to these potential influences of non-geostationary satellites on modem design, there can also be a significant impact on the multiple access strategy due to differential delay and differential Doppler across a spotbeam. Assuming the differential effects of the earth's rotation across a spotbeam can be neglected, the differential delay can be upper bounded by

$$\Delta\tau \leq \frac{r(\theta_{\min}) - r(\theta_{\max})}{c} \tag{64.16}$$

where $\theta_{\min}$ and $\theta_{\max}$ are the minimum and maximum elevations for the spotbeam of interest. The greatest differential delay occurs at the edge of coverage. For many scenarios, this differential delay can be of the order of milliseconds, which can have an impact on multiple access strategies which require the time-synchronization of users.

The differential Doppler across a spotbeam depends upon the location of the spotbeam relative to the orbital plane of the satellite. If one neglects the differential effects of the earth's rotation across a spotbeam, one can bound the differential Doppler in spotbeams along the orbital plane by

$$\Delta f_s \leq \frac{2\pi R_e}{\lambda T}(\cos\theta_{\min} - \cos\theta_{\max}) \tag{64.17}$$

This bound assumes that the whole spotbeam lies completely in either the fore or aft coverage of the satellite, but indicates that the maximum differential Doppler occurs across a spotbeam covering the subsatellite point. For many scenarios, the differential Doppler across a spotbeam can be on the order of kilohertz and, thus, can also influence the design of a multiple access strategy.

It should be noted that lower altitude satellites must provide the same RF power flux density as higher altitude satellites. Typically, however, higher altitude satellites are used to cover larger areas and, to provide an equivalent return link service, must have larger antennas and more spotbeams. Consequently, while fewer of the higher altitude satellites are required to provide global coverage, they are generally larger and more complex.

64.4 Multiple Access

Satellites intended for mobile services are frequently designed with antennas large enough to provide multiple spotbeams. This not only reduces the power requirements but also allows frequency re-use between the beams. The result is a system which is similar to terrestrial cellular systems with the following differences:

1. Isolation between co-channel users is controlled by the spacecraft antenna characteristics and not by propagation losses.

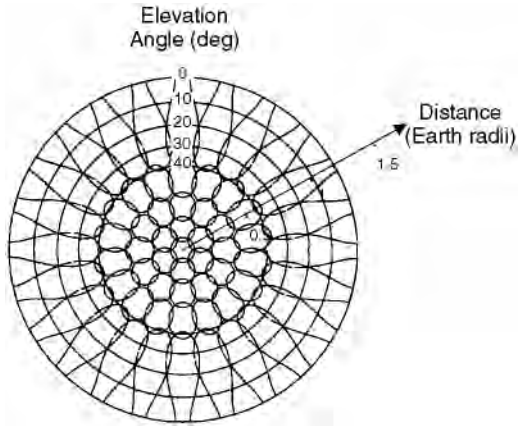


FIGURE 64.6 Spotbeam pattern projected onto the earth's surface: 4-db contours with edge of coverage at 10∞ elevation.

2. The near–far problem is less critical so power control is less of a factor. Due to delay constraints in satellite systems, power control is generally limited to counteracting the effects of variation in antenna gain and average propagation losses. This, however, means that co-beam users in the forward direction will generally be received at different power levels.
3. Due to the distortion caused by the curvature of the earth's surface, the spotbeams do not cover the same surface area. In particular, the outer beams can be significantly extended in the radial direction, as illustrated in Fig. 64.6; this figure is a two-dimensional projection of the coverage pattern on the earth that preserves radial proportions but results in some angular distortion. The consequence of this distortion is that the outer beams may be required to serve significantly more users. Furthermore, the outer beams are subject to more adjacent beam interference for similar reasons. A complicating factor is that the outer beams are also the most likely to be in the field of view of other satellites in the constellation. The satellite antenna can be designed to provide equal size spotbeams on the earth, but this policy penalizes the communications performance of a large number of users for the benefit of none.
4. With many constellations there can be significant overlapping coverage from multiple satellites, and this allows the possibility of diversity transmissions. A whole range of diversity techniques are possible for all multiple access strategies depending on the power, bandwidth, and complexity constraints. The simplest approach is selection diversity during the call setup. More complex approaches include dynamic techniques such as **switched diversity** in the forward direction based on return-link signal strength and **maximal ratio combining** in the return direction.
5. This brings up a final difference with respect to terrestrial systems in that, with a fixed antenna cluster, the satellite cells and "base station" are moving with respect to the mobile rather than vice versa. This raises issues regarding handover procedures and, in light of multiple satellite coverage, the orthogonality of user accesses.

In general, multiple access techniques for mobile satellite applications can be divided into two groups: wideband techniques, such as CDMA [Gilhousen et al., 1990], and narrowband techniques, such as FDMA and **narrowband TDMA**. Each technique has its strengths and weaknesses which depend on the application and implementation, but there are a number of common aspects. The power efficiency of all techniques can be improved by adding forward error correction coding but with a bandwidth penalty. The exception to this is asynchronous CDMA, which does not suffer a bandwidth penalty but generally must tolerate greater co-user interference. There are synchronous versions of all these techniques (synchronization at the chip level for CDMA and at the bit and frequency level for FDMA and TDMA) that can be used to improve spectral efficiency on the forward link where one can easily synchronize transmissions.

For applications in the 1 GHz and 2 GHz range, the propagation characteristics of wideband techniques (<2 MHz) and narrowband techniques are similar.

Interference tolerance, in particular multiple access interference, is one aspect which differs between wideband and narrowband systems. With CDMA, interference due to co-users is averaged over all co-users and affects all users. With narrowband techniques, interference generally results in greater degradation, but only of individual carriers. It is this interference tolerance which determines the frequency re-use capability and the bandwidth-limited capacity. When comparing the bandwidth-limited capacities of various approaches, care must be taken to be equitable. Interference has a probability distribution due to the random location (and isolation) of co-channel users relative to the desired user. In both wideband and narrowband techniques, it is the tails of this distribution which determine the effect of interference on the service availability [Moher et al., 1996].

An additional issue is the design of the ancillary channels for signaling and control for a multi-beam satellite system and their integration with the traffic services with a minimum of power and bandwidth requirements.

64.5 Modulation and Coding

Along with the choice of multiple access strategy, modulation and coding schemes must be selected to meet the stringent bandwidth and power efficiency requirements of the mobile satellite communications system. With respect to the choice of modulation scheme, binary phase-shift keying (BPSK) has been popular for low data rate store-and-forward systems due to its simplicity and robustness. However, BPSK is very spectrally inefficient when transmitted from a terminal with a nonlinear power amplifier (e.g., a power-efficient class-C amplifier). Aeronautical satellite communications systems have improved upon the spectral efficiency, without sacrificing robustness or compatibility with nonlinear amplifiers, by using **π/2-BPSK** coupled with 40% square-root raised-cosine pulse shaping.

The most popular choice of modulation technique for higher-rate services has been quadrature phase-shift keying (QPSK). QPSK can be considered as two BPSK waveforms orthogonally modulated: one by a sine wave at the carrier frequency, and the other by the corresponding cosine wave. Consequently, for a given choice of pulse shaping filter and for a linear channel, the spectral efficiency (in terms of the ratio of the number of bit/s transmitted to the occupied bandwidth in Hz) for QPSK is twice that for BPSK. With standard QPSK, the symbol timing is identical in the sine and cosine channels. A fairly linear transmitting power amplifier is required in the mobile terminal if good bandwidth efficiency is to be maintained with standard QPSK. As is the case for BPSK, variations of QPSK can be used that are reasonably tolerant to amplifier nonlinearity. Figure 64.7 [Patenaude and Lodge, 1992] illustrates the transmitted power spectrum of π/4-QPSK (60% square-root raised-cosine pulse shaping) and offset-QPSK (60% and 100% square-root raised-cosine pulse shaping) after being amplified by a hard-limiting amplifier (which is the extreme example of a nonlinear power amplifier). Note that even in this severe case, offset-QPSK with 60% pulse shaping is quite spectrally efficient. In fact, this is the choice of modulation scheme for a number of mobile satellite communications systems including Inmarsat's M, B, and Aeronautical systems.

The viability of mobile satellite systems is dependent upon excellent power efficiency. Consequently, detection is often based on phase-coherent approaches rather than the less power-efficient options of differentially-coherent and noncoherent detection that are commonly used for terrestrial systems. Increasingly, traditional carrier tracking schemes, such as phase-locked loop based techniques, are being replaced by digital schemes such as the one described in [Viterbi and Viterbi, 1983]. The fading and shadowing nature of the channel can result in frequent cycle slips of the carrier recovery algorithm, resulting in a phase ambiguity (which is inherent in the PSK signal) that must be resolved. Differential phase encoding of the transmitted symbols is a fairly traditional approach for eliminating the detrimental effects of the phase ambiguity. However, differential encoding tends to increase the average bit error rate for a given symbol error rate and is consequently avoided in many mobile satellite system signal designs. A more common approach is to periodically insert a known bit or symbol (referred to as a *reference symbol*) or

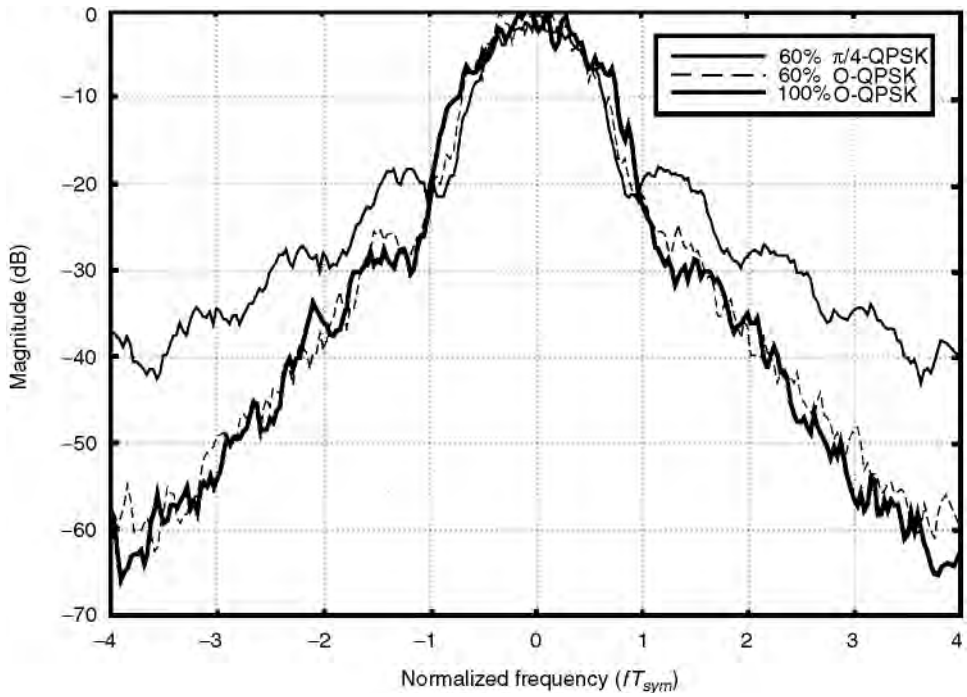


FIGURE 64.7 The transmitted power spectrum of p/4-QPSK (60% square-root raised-cosine pulse shaping) and offset-QPSK (60% and 100% square-root raised-cosine pulse shaping) following hard-limiting amplification. [Patenaude and Lodge, 1992].

group of bits (referred to as a *unique word*) into the transmitted signal, which can be used to detect the occurrence of cycle slips and then to resolve the resulting carrier phase ambiguity. In the extreme case, for which the period of insertion is sufficiently greater than the required tracking bandwidths, carrier tracking can be based solely on the reference symbols [Moher and Lodge, 1989].

Forward error control (FEC) coding is an important tool for achieving good power and bandwidth efficiency. The choice of FEC coding technique is dependent upon the number of bits that are to be coded as an identifiable group. First, consider the case for which only several tens of bits are to be coded into a codeword. This situation naturally arises in a number of circumstances including short data packets containing signaling information or user data, and for digital voice transmission for which additional delays (due to coding) of greater than a couple of voice frames is unacceptable and only the most sensitive bits in the frame are protected with FEC coding. In these cases, block codes such as the Golay code and Bose–Chaudhuri–Hocquenghem (BCH) codes are common choices. Another possibility is the use of convolutional coding with puncturing [Yasuda et al., 1984] used to achieve the desired code rate. Convolutional codes with a constraint-length of 7 have been popular choices for early systems, with constraint-length 9 codes being proposed for later ones. Usually prior to convolutional encoding, a field of "flush bits" is appended to the end of the field of data bits so that the last data bits to be decoded have a similar level of integrity as the rest of the data bits. The field of flush bits can represent a significant (and undesirable) overhead for the transmission of very short data blocks. Fortunately, the requirement for such a field can be obviated if tail biting [Ma and Wolf, 1986] is used. Usually, convolutional coding is preferable to block coding because it is more amenable to the use of soft decisions.

The second scenario to consider is the transmission of a continuous data stream (e.g., digital voice) or data in large fields. In this case, convolutional coding is a reasonable choice for services for which high average bit error rates (i.e., higher than 10^{-4}) are acceptable. For applications requiring higher quality data, reasonable choices include concatenated coding (with a convolutional code as the inner code and

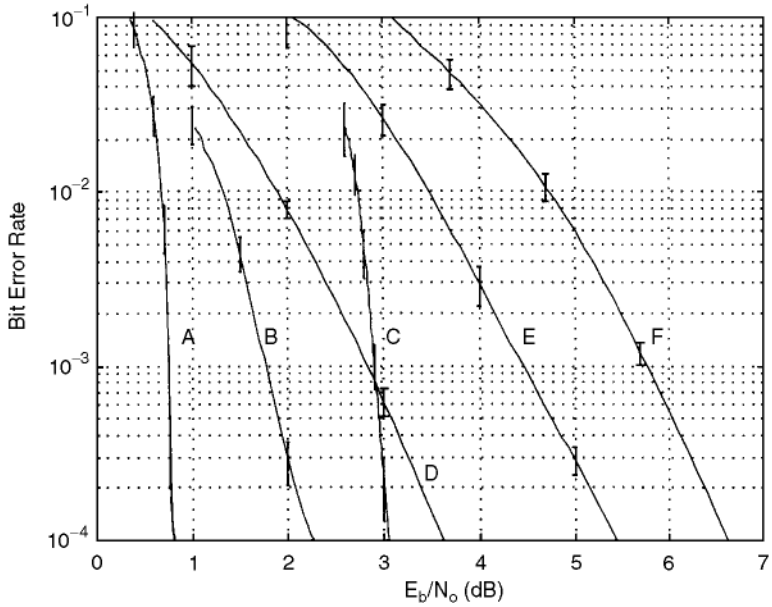


FIGURE 64.8 The performance of various coding schemes in an additive white Gaussian noise environment. A: rate-1/2 turbo code with a block size of 10,000 b; B: rate-1/2 turbo code with a block size of 300 b; C: rate-1/2 concatenated convolution/Reed–Solomon code; D: rate-1/2 constraint length 7 convolution code; E: rate-3/4 punctured constraint length 7 convolutional code; F: rate-1 pragmatic constraint length 7 convolutional code. (Provided by P. Guinand, Communications Research Centre, Ottawa, Canada.)

a Reed-Solomon code as the outer code) and turbo coding [Berrou et al., 1993]. As an example of the relative power efficiency, Fig. 64.8 shows the performance of several of these alternatives where the only channel impairment is additive white Gaussian noise. Curves A, B, C, and D are all for rate-1/2 coding. Curve D is a constraint length-7 convolutional code. Curve C is the same convolutional code, punctured to rate-3/5, concatenated with a (255, 213) Reed–Solomon code. Curves A and B represent the performance of turbo coding for block lengths of 10,000 and 300 data bits, respectively. Curves E and F provide an indication of the trade-off between bandwidth and power efficiency. These latter two curves use the convolutional codec of curve D, so the level of code complexity is fixed. Curve E is punctured to a rate of 3/4, while curve F uses pragmatic coding [Viterbi et al., 1989] to achieve a rate of 1 bit per dimension. For the range of spectral efficiencies considered here and for a fixed level of coding complexity, the rule of thumb is that a 1 dB improvement in bandwidth efficiency results in approximately a 1 dB degradation in power efficiency. Note that trading bandwidth against power efficiency, using puncturing or pragmatic coding, is equally applicable for concatenated and turbo coding schemes.

Fading and shadowing can cause reductions in the level of the received signal to a point where reliable transmission is no longer possible. If these signal outages are of short duration, they can successfully be compensated for by interleaving the order of the coded data prior to modulation in the transmitter and then deinterleaving the data prior to decoding [Clark and Cain, 1981]. This effectively shuffles the received bits having low received signal power with received bits having higher signal power. Of course, for the combination coding and interleaving to eliminate errors in the decoded data bits, the interleaving must be performed over a signal duration that is several times greater than the duration of the outage. Interleaving adds to the total end-to-end delay, and, consequently, the duration of outages that can be dealt with is dependent upon the delay constraint of the given mobile satellite service.

Inevitably, some of the outages will have a duration too great to be handled by interleaving and coding. In such cases, many services will simply be unavailable. However, some source codecs (e.g., certain vocoders) have the capability to gracefully handle several sequential missing or erroneous frames. Also, applications

that can tolerate long delays can overcome fairly long outages by employing automatic repeat request (ARQ) schemes.

64.6 Trends in Mobile Satellite Systems

An increasing variety of services are being offered to travelers via mobile satellite systems, with the data rates for the high-end services increasing steadily. In parallel with this, there is continuing pressure to reduce the size of the mobile terminals. The result is that the size of the system's space segment is growing through some combination of an increased number of satellites and larger, more complex satellites. For a given orbit altitude, the size of the satellite antenna is growing dramatically. A corresponding increase can be seen in system complexity, with the apportioning of this complexity between the space segment and the ground infrastructure varying from one system to another. System designs that place much of this complexity in the space segment are characterized by on-board switching and processing and, in some cases, intersatellite links. System designs that place much of the complexity in the ground infrastructure are characterized by a large number of coordinated and very sophisticated gateway stations.

The number of customers for mobile satellite services has begun to grow significantly, and explosive growth in the future is forecast. Many organization are designing new systems to meet this anticipated demand, with new frequency bands being allocated to accommodate them. Even with these new bands, increasing frequency congestion will be experienced if the forecasted traffic volumes materialize. The first generation mobile satellite systems were predominantly power limited. However, subsequent generations will be increasingly bandwidth limited.

The traffic volumes for many early generation mobile satellite systems are fairly symmetric (in the sense that the volume of traffic in the forward direction is approximately the same as that for the return direction) because voice telephone applications are dominant. However, it is anticipated that an increasing portion of the traffic will support services such as database access, message retrieval (e.g., e-mail and voice mail), and broadcast information. In general, these services result in much greater volumes of data being transmitted in the forward direction than in the return direction. Fortunately, this asymmetric traffic requirement is a good match with the technological situation in a couple of respects. Firstly, in the forward direction, improved bandwidth efficiency can be achieved by taking advantage of the fact that the system can be designed to accurately coordinate the carrier frequency, carrier phase, and symbol timing of each transmission within a beam. Consequently, techniques such as synchronous CDMA, orthogonal frequency-division multiplexing, and wavelet-packet multiple access can be employed. Secondly, unlike the return direction for which the mobile terminal's power amplifier places fairly severe constraints on the dynamic range of the transmitted waveform, the forward link channel is incrementally linear. *Incrementally linear* means that each carrier appears to be transmitted through a linear channel, with intermodulation due to the nonlinearity of the satellite transponder's power amplifier resulting in an additive noise-like process. Given that both the traffic requirements and the technological situation are asymmetric, the transmission schemes selected for these mobile satellite services will differ significantly between the forward and the return direction.

Defining Terms

Forward direction: Direction of transmission that is from the satellite to the mobile.

Maximal ratio combining: A combining technique for which signals from two or more communication paths are aligned in phase, weighted proportional to their voltage-to-noise power ratios, and then summed [Proakis, 1989].

Narrowband TDMA: A multiple access strategy where a small number of users, typically less than ten, share a single frequency channel in time. Larger numbers of users are accommodated by using a number of narrowband TDMA carriers of different frequencies.

Offset-QPSK: A variation of QPSK for which the symbol timing in the sine channel is offset by half of a symbol period relative to that in the cosine channel.

Return direction: Direction of transmission that is from the mobile to the satellite.

Switched diversity: The process of switching to the best of two or more communication paths [Jakes, 1974].

π/2-BPSK: A variation of BPSK for which consecutive symbols experience a relative phase shift of $\pi/2$ radians.

π/4-QPSK: A variation of QPSK for which consecutive symbols experience a relative phase shift of $\pi/4$ radians.

References

Adams, W.S. and Rider, L., Circular polar constellations providing continuous single or multiple coverage above a specified latitude, *J. Astron. Sci.,* 35, 155, 1987.

Ballard, A.H., Rosette constellations of earth satellites, *IEEE Trans. Aero. Elec. Sys.,* AES-16(5), 656, 1980.

Berrou, C., Glavieux, A., and Thitimajshima, P., Near Shannon limit error-correcting coding and decoding: turbo-codes, *Proc. IEEE Int. Conf. Comm.,* ICC'93, 1064, 1993.

Butterworth, J.S., Propagation measurements for land-mobile satellite systems at 1542 MHz, Communications Research Centre Tech. Note No.: 723, Department of Communications, Canada, August 1984.

Clark, G.C., Jr. and Cain, J.B., Interleaver structures for coded systems, *Error-Correction Coding for Digital Communications,* Plenum Press, New York, 1981, Section 8.3.

Clarke, R.H., A statistical theory of mobile radio reception, *Bell Syst. Tech. J.,* 47, 957, 1968.

Dondl, P., Loopus opens a new dimension in satellite communications, *Int. J. Sat. Comm.,* 2, 241, 1984.

Gilhousen, K.S., Jacobs, I.M., Padovani, R., and Weaver, L.A., Jr., Increased capacity using CDMA for mobile satellite communications, *IEEE J. Sel. Areas Comm.,* 8, 503, 1990.

Hess, W.N., *The Radiation Belt and Magnetosphere,* Blaisdell, Waltham, MA, 1968.

ITU, Propagation data required for the design of earth-space land mobile telecommunication systems, Recommendation ITU-R PN.681 (1992) and ITU-R PN.681-1 (1995).

Jahn A., Lutz, E., Sforza, M., and Buonomo, S., A wideband channel model for land mobile satellite systems, *Proc. Int. Mob. Sat. Conf.,* Ottawa, June 1995, 122.

Jakes, W.C., Ed., *Microwave Mobile Communications,* IEEE Press, New York, 1993.

Loo, C., A statistical model for land-mobile satellite link, *Proc. Int. Conf. Comm.,* 588, 1984.

Lutz, E., Cygan, D., Dippold, M., Dolainsky, M., and Papke, W., The land mobile satellite communication channel—recording, statistics and channel model, *IEEE Trans. Veh. Tech.,* 40(2), 375, 1991.

Ma, H.H. and Wolf, J.K., On tail biting convolutional codes, *IEEE Trans. Commun.,* COM-34(2), 104, 1986.

Moher, M.L. and Lodge, J.H., TCMP—a modulation and coding strategy for Rician fading channels, *IEEE J. Sel. Areas Commun.,* 7(9), 1347, 1989.

Moher, M.L., Erup, L., and Lyons, R.G., Interference statistics for multi-beam satellites, *Proc. 2nd Eur. Workshop Mobile/Personal Satcom,* 366, 1996.

Moreland, K.W., Ocean scatter propagation models for aeronautical and maritime satellite communication applications, M.Sc. Thesis, Carleton University, 1987.

Neul, A., Hagenauer, J., Papke, W., Dolainsky, F., and Edbauer, F., Propagation measurements for the aeronautical satellite channel, *Proc. IEEE. Veh. Tech. Conf.,* Tampa, FL, June 1987, 90.

Patenaude, F. and Lodge, J., Non-coherent detection of hardlimited O-QPSK, *Proc. 1992 Int. Conf. Sel. Topics Wireless Comm.,* Vancouver, Canada, June 1992, 93.

Proakis, J.G., *Digital Communications,* 2nd ed., McGraw-Hill, New York, 1989.

Ruddy, J.M., A novel non-geostationary satellite communications system, *Proc. Int. Conf. Comm.,* 3, Denver, June 1981.

Schweikert, R. and Hagenauer, J., Channel modelling and multipath compensation with forward error correction for small satellite ship earth stations, *Proc. 6th Int. Conf. Dig. Sat. Comm.,* Phoenix, AZ, 1983, XII-32.

Stutzman, W.L., Ed., Mobile satellite propagation measurements and modeling: a review of results for systems engineers, *Proc. Int. Mob. Sat. Conf.,* Pasadena, CA, June 1988, 107.

Vilar, E. and Austin, J., Analysis and correction techniques of Doppler shift for non-geosynchronous communication satellites, *Int. J. Sat. Comm.,* 9, 123, 1991.

Viterbi, A.J. and Viterbi, A.M., Nonlinear estimation of PSK-modulated carrier phase with applications to burst digital transmission, *IEEE Trans. Info. Theory,* IT-29(4), 543, 1983.

Viterbi, A.J., Wolf, J.K., Zehavi, E., and Padovani, R., A pragmatic approach to trellis-coded modulation, *IEEE Commun. Mag.,* 11, 1989.

Yasuda, Y., Kashiki, K., and Hirata, Y., High rate punctured convolutional codes for soft decision Viterbi decoding, *IEEE Trans. Commun.,* COM-32, 315, 1984.

Further Information

A number of tutorial paper on mobile satellite systems have appeared in recent years. These include:

Kato, S., Personal communication systems and low earth orbit satellites, *Proceedings of the Space and Radio Science Symposium*, U.R.S.I., Brussels, Belgium, April 1995, 30.

Lodge, J., Mobile satellite communications systems: toward global personal communications, *IEEE Commun. Mag.*, 29(11), 24, 1991.

Wu, W.W., Miller, E.F., Pritchard, W.L., and Pickholtz, R.L., Mobile satellite communications, *Proc. IEEE*, 82(9), 1431, 1994.

Special issues of journals, dedicated to mobile satellite communications, are:

Kucar, A.D., Kato, S., Hirata, Y., and Lundberg, O., Eds., Special Issue on Satellite Systems and Services for Travelers, *IEEE J. Sel. Areas Commun.,* 10(8), 1992.

Maral, E., Ed., Special Issue on Personal Communication via Satellite, *Int. J. Sat. Commun.,* 12(1), 1994.

A comprehensive book that contains a number of sections that are highly relevant to the discussion of multiple access techniques is:

Abramson, N., Ed., *Multiple Access Communications: Foundations for Emerging Technologies,* IEEE Press, New York, 1993.

65

Satellite Antennas

Yeongming Hwang
Pinnacle EM Wave

Youn Ho Choung
TRW Space and Electronics Group

65.1 Introduction

Satellite antennas provide not only for telecommunications, but also communications for controlling and commanding the satellite in a prescribed **orbit**: an invisible tower in space [Hwang, 1992]. Their capabilities have evolved from being only a radio relay station to being an onboard processing base station. Most communications satellites are boosted into geosynchronous Earth orbit (GEO). At a particular altitude, with an angular speed exactly matching the Earth's rotational rate, the satellite is motionless as viewed from the Earth. Three such satellites in synchronous orbit can cover over 90% of all inhabited regions of the Earth. Signals sent up from the ground and transmitted down from the satellite travel 19,300 nmi and suffer large propagation loss. A high-gain antenna is required to compensate for the path loss. The antenna can be a simple one for generating a circular pencil **beam** for Earth coverage or a much more complicated antenna generating multiple shaped beams and/or reconfigurable beams for **frequency reuse** with spatial and polarization diversities. As mobile communications services evolve from a regional to a global scale, satellites play an important role in providing the low-power-transmitters needed to broadcast from relatively unobtrusive base stations distributed throughout space.

In order to provide low-power, personal handheld telephone services, the coverage area would normally be divided into cells, each cell covered by a separate antenna beam. For a GEO satellite, this beam would require a very large antenna, which could not be implemented even with deployable antenna technology. In addition, the time delay of the half-hop propagation for a GEO satellite varies between 119 and 139 μs, becoming a problem in voice communication. This time delay can be reduced by the use of low Earth orbit (LEO) or medium Earth orbit (MEO) satellites. LEO and MEO orbits are particularly suitable for cellular mobile communication systems since the cell size remains constant throughout the orbit. Several LEO and MEO systems have been proposed, among them Iridium, Globalstar, and Odyssey. The range of possible communication distance, propagation loss, propagation delay, and satellite view time vs. orbital attitude are key parameters in the system tradeoff and design.

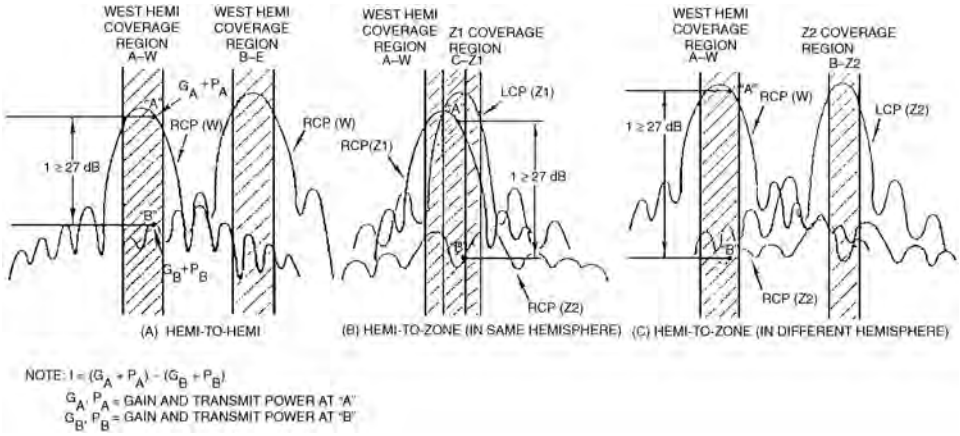


FIGURE 65.1 Transmit isolation requirements.

To meet these requirements, the antenna radiation pattern varies from omnidirectional to highly directional. It can be fixed or reconfigured to meet the needs for each system configuration. A shaped beam is required in order to constrain the energy to conform to the coverage area to improve equivalent isotropically radiated power (EIRP) and reduce interference between beams. It can be achieved either by a multiple beam antenna system or a shaped reflector system. Beam reconfiguration is required to accommodate the change of radiation pattern for a time-domain multiple access (TDMA) system.

Conventional **antenna parameters** such as half-power beamwidth (HPBW), antenna radiation efficiency, and aperture efficiency are insufficient to describe the performance of the system. In addition to the gain ripple and gain slope over the coverage and frequency band, isolation or interference among these beams is important in the system. They are assessed in terms of spatial and polarization isolation, and both must be considered. Figure 65.1 depicts a 27-dB isolation between two transmit beams. Similarly, the receive isolation is defined as the isolation of two received signals of equal incident flux density. Beam shaping efficiency is used to evaluate the gain over a highly shaped coverage. An ideally shaped beam is one that has maximum antenna gain without sidelobes and achieves constant antenna radiation intensity over the area. Therefore, the theoretically achieved maximum gain is

$$G = 41253/A_{\text{cov}}$$

where A_{cov} is the coverage area in degrees square. The performance is evaluated in terms of beam shaping efficiency η defined as

$$\eta = G_{\min} A_{\text{cov}}/41253$$

where $G_{\min}$ is the minimum gain over A_{cov}. Practically, the gain-area product is between 10,000 and 18,000, depending on the size of antenna aperture and the shape of coverage.

The tracking, telemetry, and command (TT&C) antenna provides ranging, telemetry, and command operation throughout all mission phases after launch vehicle separation. An autotracking antenna is required for two reasons: to improve the pointing accuracy of the highly shaped beam where the performance degradation is not accepted in the system due to satellite pointing error and to communicate with a moving target or user.

Successful observation of weather and climate from satellites has led to other new applications such as the measurement of atmospheric parameters by seeing through cloud canopies. A microwave radiometer is a good candidate for this application. However, at least a 4.4-m antenna would be required to provide 30-km nadir spatial resolution. This type of antenna is still under development. The most common types of antennas being used to meet these requirements are summarized in Table 65.1.

TABLE 65.1 Satellite Antennas

Area Coverage	Function	Antenna Types	Antenna Gain	Design Challenge	Name of Satellites
Earth coverage	±9∞FOV	Corrugated horn Multiflare horn	Medium	Single horn with diplexer to cover both up- and down-links, with Good axial ratio	Most of all communication satellites
Omni coverage Spot coverage	TT&C with 2 operation modes: 1. transfer orbit 2. on-orbit	Conspiral antenna Biconical antenna Horn antenna	Low Low Medium/ low	Nearly omni-directional Spacecraft body blockage Reliability	Most of all communication satellites
Pencil beam coverage	Limited portion of FOV	Reflector antennas Single feed Multiple feeds	High	Gimbal Steering Beam forming network (BFN)	INTELSAT-VIIs NASA ACTS MILSTAR
Shaped beam coverage	Multiple beams for frequency reuse	Multibeams antennas Shaped reflector antennas	High	Isolation between beams Switching BFN	INTELSAT-V INTELSAT-VI INTELSAT-VII
Nulling coverage	Maximize signal-to-interference power ratio	Reflector antennas Multiple feeds Adaptive BFN Phased array antennas Adaptive BFN	High	Adaptive nulling algorithm Adaptive BFN Calibration and characterization of BFN	MILSTAR
Finite coverage	Passive microwave radiometer	Reflector antennas	High	Size of reflector Surface tolerence	

Satellite antennas differ from other antennas in several respects. A satellite must be designed to withstand the dynamic mechanical and thermal stress during the launch and in the space environment. The materials used become a key factor. Other design constraints are size and weight. The antenna has to fit the shroud of the launching vehicle. As a result, the reflector used for the high-gain antenna must be stowed using either a solid reflector or a meshed reflector. Weight cannot exceed that weight budget allocated in the system. The challenge of satellite antenna design for the future will be to make them lighter, smaller, and cheaper.

65.2 Horn Antennas and Reflector Feeds

Currently, the reflector antenna is the most widely used satellite antenna, but future satellites may feature phased array antennas capable of rapid and low-loss beam scan, achievable with the design maturity of solid-state microwave amplifiers and receivers. The reflector antenna may have either a single feed or multiple feeds, depending on the coverage requirement; the phased array antenna, in contrast, has many radiating elements. The best candidates for feeds and radiating elements are horns and cup dipoles. Above 6 GHz, horns are usually used since the waveguide is low loss and small in size. Cup dipoles are typically used below 4 GHz since coaxial cables are smaller and lighter than waveguides.

Horn Antennas

There are five types of horn antennas: open-ended waveguide, conical horn, dual-mode horn, corrugated horn, and multiflare horn. Table 65.2 compares pros and cons among these five horns. The open-ended waveguide or the conical horn are most frequently used for multiple beam antennas (MBA) with phased

TABLE 65.2 Comparison Among Various Horns

	Open-End Waveguide	Conical Horn	Dual-Mode Horn	Corrugated Horn	Multiflare Horn
Bandwidth	Wide	Wide	Narrow	Wide	Moderate
Pattern Symmetry	No	No	Yes	Yes	Yes
Phase Center Location	On the aperture	Inside aperture	Inside aperture	Inside aperture	On/Inside the aperture
VSWR for Large Aperture	Good	Excellent	Frequency sensitive	Poor without matching section	Excellent
Fabrication	Easy	Easy	Easy	Difficult	Easy
Applications	Array element	Array element	High-efficiency feed for reflector	High-efficiency feed for reflector	High-efficiency feed for reflector

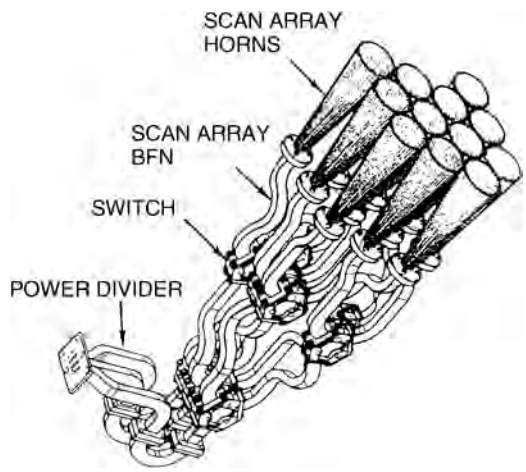


FIGURE 65.2 Feed array for NASA ACTS MBA.

array elements since their apertures can be small. Figure 65.2 shows a 30-GHz feed array for the NASA Advanced Communication Technology Satellite (ACTS). The advantage of this type of horn is its simplicity and ease of fabrication. The disadvantage is that the radiation pattern is not circularly symmetrical, and so antenna efficiency is low and a high axial ratio is obtained for the circular polarization.

The dual-mode horn antenna, also called the Potter horn, generates a circularly symmetric pattern by adding transverse electric (TE_{11}) and transverse magnetic (TM_{11}) modes in proper amplitude and phase [Potter, 1963]. The TM_{11} mode is excited by a simple waveguide step. However, due to the phase velocity difference, proper phasing of the two modes occurs only in a limited-frequency band with less than 5% bandwidth. Therefore, the dual-mode horn is only used in narrow bandwidth applications.

The corrugated horn carries a hybrid mixture of TE_{11} and TM_{11} modes through the grooved horn walls; this fundamental hybrid mode is called the hybrid electric (HE_{11}) mode. The TE and TM components in the HE_{11} mode have the same cutoff and phase velocity; therefore, the beamwidth limitation of the dual-mode (or multiflare) horn is removed. A detailed theoretical analysis has been given by Clarricoats and Saha [1971]. One design problem of the corrugated horn is that without a matching section at the horn throat region it has a poor voltage standing wave ratio (VSWR), especially over the lower portion of the operating frequency band. Two kinds of matching sections can be used to improve VSWR: varying groove depths and ring-loaded grooves [James, 1982]. The corrugated horn can easily be designed up to 60% bandwidth using a varying grooves matching section and, using a ring-loaded matching section, can even be designed to a 2:1 bandwidth. One drawback of the corrugated horn is that

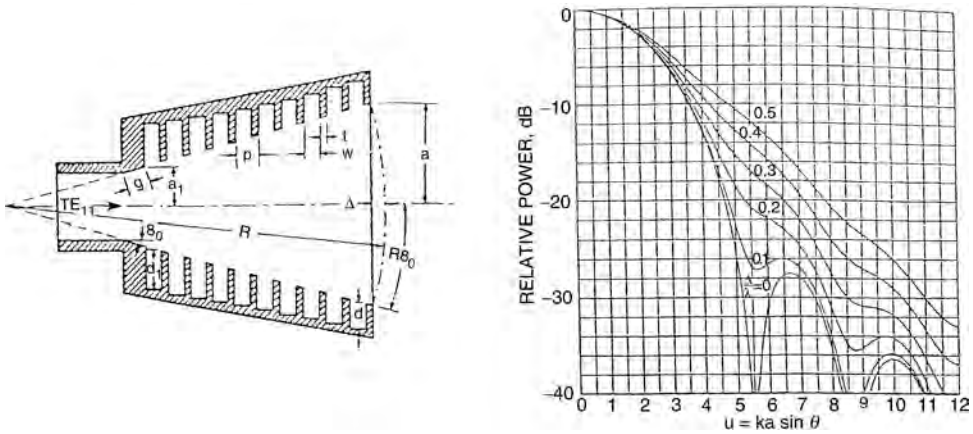


FIGURE 65.3 Narrow-flare corrugated horn.

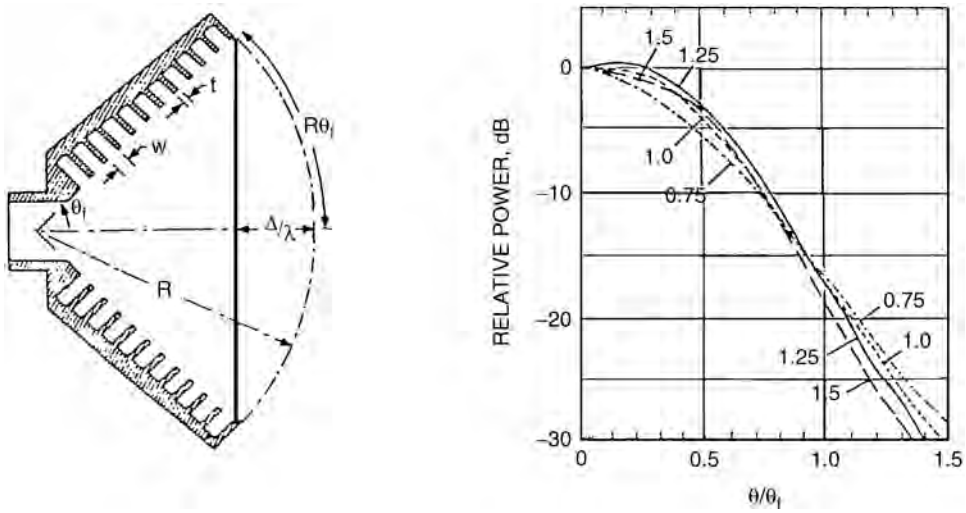


FIGURE 65.4 Wide-flare corrugated horn.

the fabrication of the grooved walls is difficult, resulting in high manufacturing cost. Yet the corrugated horn is the most used reflector feed and is the best candidate for the Earth coverage antenna. A corrugated horn is classified either as a narrow- or a wide-flare horn. Figure 65.3 shows the antenna geometry and the normalized radiation patterns for the narrow-flare corrugated horn. Figure 65.4 shows those for the wide-flare corrugated horn.

A multiflare horn is formed of several flared waveguide sections. This mechanically simple structure leads to very low fabrication cost. The smooth wall, as well as the flare structure, provides good RF match to the waveguide. A low VSWR is always achieved. A conical multiflare horn generates higher order TE_{1n} and TM_{1n} modes at the flare change. The circularly symmetric pattern is obtained by adding these higher modes (mainly TM_{11}) to the fundamental TE_{11} in proper ratio. The analysis is based on mode matching at the flare change to compensate for phase front curvature change. The differential phase shift among modes in a multiflare horn is less than that in a dual-mode horn. Thus, the multiflare horn is capable of achieving about 15% bandwidth. A three-section multiflare horn is comparable (or even smaller) in size to a corrugated horn with the same beamwidth [Chen, Choung, and Palma, 1986]. The multiflare horn

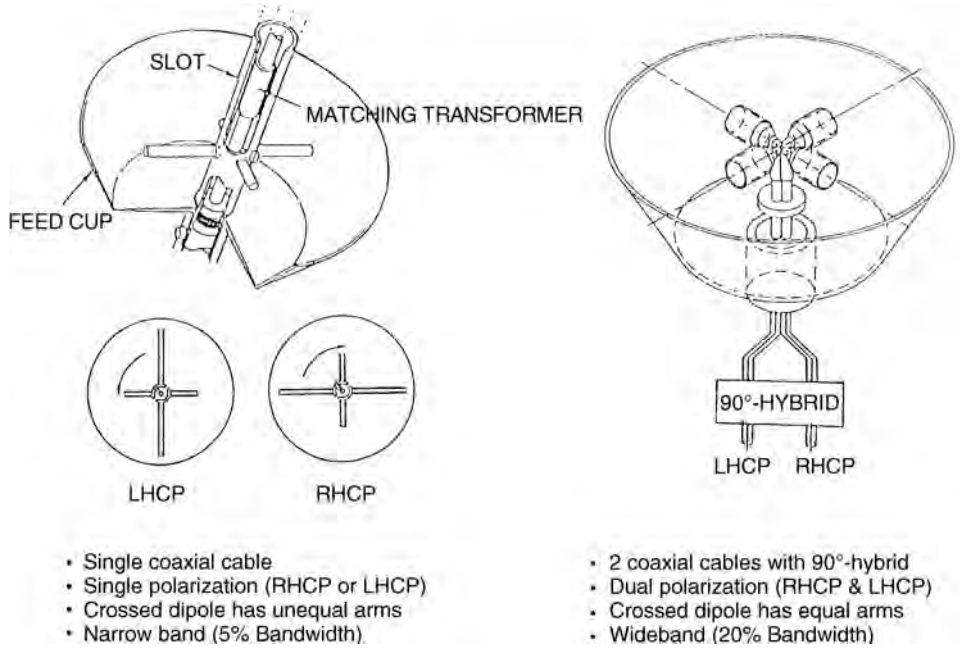


FIGURE 65.5 Two types of CP cup dipole.

has advantages over the dual-mode horn in terms of bandwidth and over the corrugated horn in terms of manufacturing simplicity. Multiflare horns are currently utilized as spot beam feeds for NASA ACTS.

Cup-Dipole Antennas

A cup-dipole antenna has crossed dipoles in a cylinder and has good axial ratio for circular polarization. The antenna is compact and easy to manufacture. There are two types, as shown in Fig. 65.5: unequal arms with single coaxial cable and equal arms using two coaxial cables and a 90∞ hybrid. The cup dipole with unequal arms is mechanically simple but can only generate single circular polarization with less than 5% frequency bandwidth. The cup dipole with equal arms has dual circular polarization with about 20% bandwidth, but it is mechanically complex. Several design variations of the cup dipole antenna such as the short backfire and cavity-back helix are described in Chapter 4 of Kumar and Hristrov [1989]. The dipole is about 0.45λ and the cup diameter is from 0.7λ to 3l depending on the beamwidth and/or the array spacing.

Frequency Reuse Feed and Polarization Diversity

A given frequency band can be used more than once at the satellite to increase the communications capability of a given link. Since different communications channels will exist at the same carrier frequencies, each reuse channel must be isolated.

There are two techniques to achieve this required isolation. One way is to use two orthogonal polarizations. One channel carries a vertical polarization [or right-hand circular polarization (RHCP)], and the other channel carries a horizontal polarization [or left-hand circular polarization (LHCP)] with more than 30-dB isolation. This can be done using a polarizer (for the CP case) and an orthomode transducer placed right after the feed horn. Another method is to use two spot beams pointing at different locations with high (about 30 dB) spatial isolation. This can be achieved using two feeds (or feed clusters) of the reflector to generate low sidelobes. Low sidelobes can be achieved with minimum spillover energy by the use of a high taper feed on the reflector.

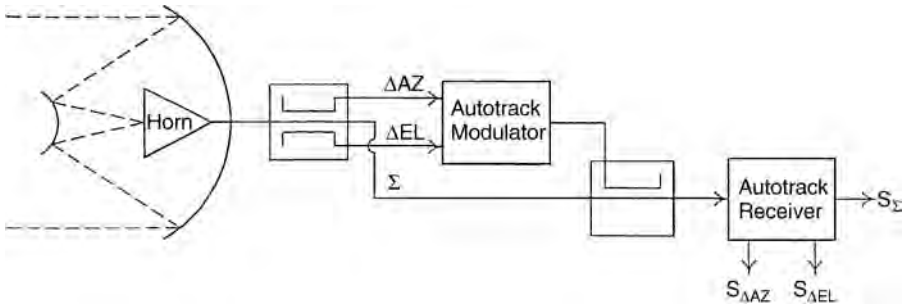


FIGURE 65.6 Block diagram of monopulse tracking.

Monopulse Feed

Point-to-point satellite communications require a narrow beamwidth to obtain high antenna gain. But the satellite jitters, causing a pointing error. In order to maintain reliable communications, a satellite antenna must be accurately pointed toward a signal source. To achieve accurate pointing, the satellite employs an autotracking system, providing tracking signals related to pointing errors in elevation and azimuth. The tracking signals control a feedback servoloop of the satellite to orient the satellite as required, positioning the antenna accurately with respect to the signal source. A block diagram of the monopulse tracking is shown in Fig. 65.6. It is achieved using a single corrugated horn and a tracking coupler. The coupler uses a longitudinal traveling wave mechanism and utilizes the circular TE_{11} and TE_{21} modes that provide the sum and the difference channels, respectively, to achieve wideband monopulse tracking. The main beam is obtained on the boresight axis of the antenna when only the dominant mode (TE_{11}) is excited at the aperture of the feed. When the beam is off the boresight axis, higher order modes (TE_{21}) are excited in the feed. The amplitude of the TE_{21} mode is proportional to the angle of boresight, and a null in the higher order mode is obtained at the boresight.

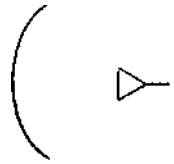
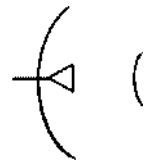
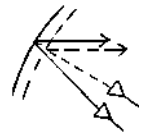
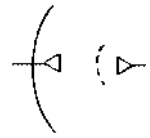
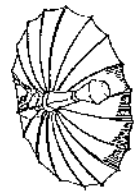
65.3 Reflector Antennas

The reflector antenna, a high-gain antenna used to support high data rates with low transmitter power, is the most desirable type for communication antennas because of its light weight, structural simplicity, and design maturity. Its disadvantage is that the reflector has to be offset to avoid the aperture blockage. The offset destroys the rational symmetry of the optical aperture and limits the achievable performance as compared to that of a lens antenna.

A reflector antenna system consists of one or more **reflector surfaces**. The surface can be solid, gridded, printed substrate, or meshed. It may be paraboloid, hyperboloid, spheroid, ellipsoid, or general in shape. Seven antenna types that have been most commonly used for satellite communication systems are summarized in Table 65.3. They include the parabolic reflector and the Cassegrain reflector, both either center fed or offset fed, the overlapped dual-gridded reflector, a multireflector system with a frequency selective surface (FSS) subreflector, and the deployable reflector antenna.

The center-fed parabolic and Cassegrain reflectors are used for a high-gain antenna where the radiation pattern is of circular shape. When the radiation pattern must be shaped to form a contour beam, a multiple beam reflector antenna system or shaped reflector antenna system can be utilized. One of the most complicated multiple beam antenna systems in use was designed for the INTELSAT V-VII series. The INTELSAT VII satellite used a 2.44-m **graphite fiber reinforced plastic (GFRP)** reflector for the 4-GHz frequency band over the Atlantic, Pacific, and Indian Oceans. The antenna provides shaped hemispheric and zone coverage, as shown in Fig. 65.7. The coverage was determined by the traffic demand in each region. The isolation between zone and hemispheric beams is accomplished by polarization diversification,

TABLE 65.3 Reflector Antenna Configurations Used in Communication Antenna Systems

Type	Features and Application	Shape
1. Center-fed parabolic reflector	Simple, symmetrical light weight structure Aperture blockage causing sidelobe level and reducing peak gain High feed line loss Pencil beam	
2. Cassegrain reflector system	Short, low feed line loss Shaped reflector may be used to increase efficiency Aperture blockage causing sidelobe level and reducing efficiency Pencil beam	
3. Offset parabolic reflector	No aperture blockage Convenience for spacecraft mounting with feed embedded below the platform Shaped beam by using multihorns or shaping reflector surface	
4. Offset Cassegrain (or Gregorian) reflector	No or minimum aperture blockage Worse scan loss Fixed or scanning beam	
5. Overlapping dual-gridded parabolic reflector	Offset configuration Polarization reuse with 33-dB isolation Shaped beam Thermal distortion of the gridded surface	
6. Multiband dual FSS reflector	FSS subreflector or overlapping Subreflectors in which one is an FSS Same main reflector aperture for multifrequency band operation Compact Thermal distortion of FSS	
7. Deployable reflector	Meshed reflector that can be folded Allow to use larger reflector aperture size Tradeoff between weight and surface accuracy Passive intermodulation may be generated Use mostly for lower frequency band system	

whereas the isolation between the east and west hemispheric beams is accomplished by spatial diversification. There are 110 feeds used in the feed system. Each feed is a 1.13 wavelength GFRP circular horn element arranged in triangular grid, as depicted in Fig. 65.8. The beam forming network consists of air supported striplines and power dividers sandwiched between ground planes. The gain of the hemispheric beam is about 15 dBi, and the gain of the zone beam is 22 dBi (decibels over an isotropic radiator). The isolation achieved for both polarization and spatial diversification is about 27 dB.

Another complex reflector antenna in flight is NASA ACTS [Choung and Wong, 1968] operating at Ka-band (30 GHz for receive and 20 GHz for transmit) (see Fig. 65.9). The MBA is designed to cover the continental U.S. (CONUS). The coverage and polarization plan are shown in Fig. 65.10. The MBA provides three fixed spot beams (Cleveland, Atlanta, and Tampa) for high-burst-rate (HBR) operations and two scanning beams for low-burst-rate (LBR) operations. The scanning beams provide coverage for

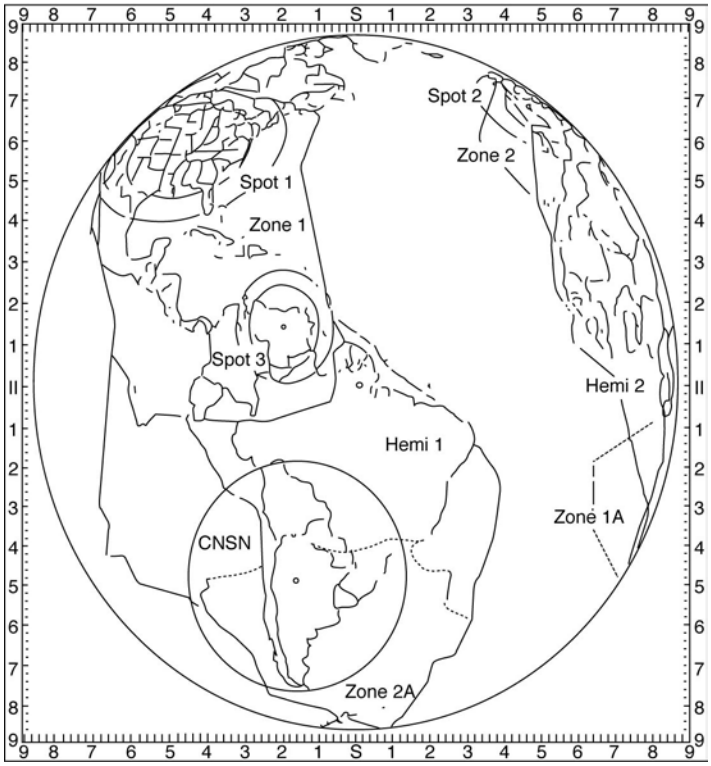


FIGURE 65.7 INTELSAT VII hemi/zone coverage.

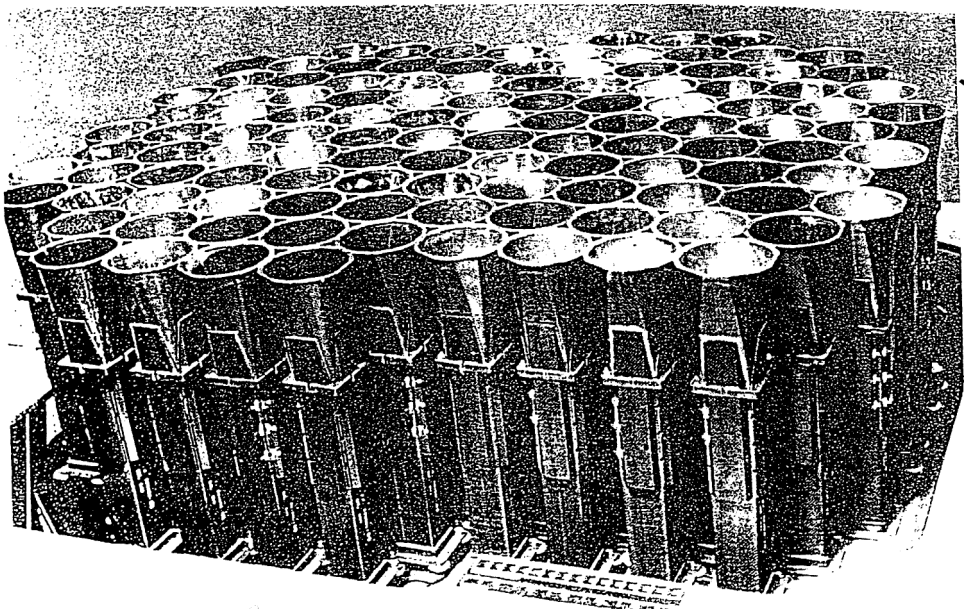
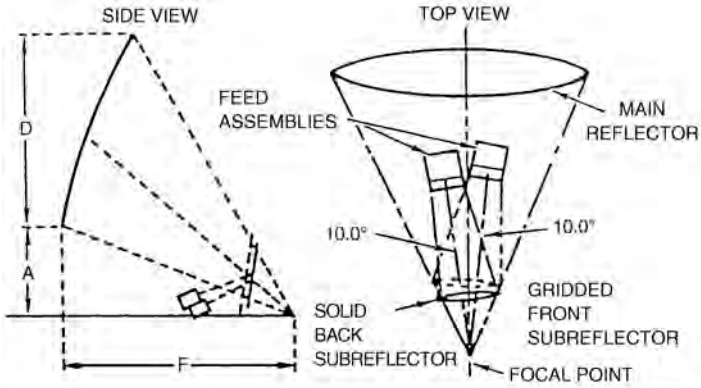


FIGURE 65.8 INTELSAT VII transmit feed array.



PARAMETER	20 GHz	30 GHz
D (MAIN REFLECTOR DIAMETER) (METERS)	3.3	2.2
A (OFFSET DISTANCE) (INCHES)	50.0	44.0
F (MAIN REFLECTOR FOCAL LENGTH) (INCHES)	132.0	88.0
GRIDDED FRONT SUBREFLECTOR		
FOCAL LENGTH (INCHES)	69.0	44.5
MAGNIFICATION FACTOR	2.0	2.0
SOLID BACK SUBREFLECTOR		
FOCAL LENGTH (INCHES)	52.4	32.4
MAGNIFICATION LENGTH	2.0	2.0

FIGURE 65.9 NASA ACTS MBA geometry.

FIGURE 65.10 NASA ACTS coverage and polarization plan.

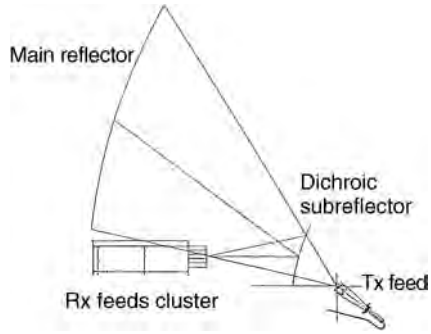


FIGURE 65.11 A typical FSS reflector antenna.

two scan sectors (east and west) and 13 scan spots (Seattle, San Francisco, Los Angeles, Phoenix, Denver, White Sands, Dallas, Kansas City, New Orleans, Memphis, Huntsville, Houston, and Miami). ACTS has two electrically identical MBAs (one for transmit and one for receive) with half-power beamwidth (HPBW) on the order of 0.35° . Each MBA shown in Fig. 65.10 has one main reflector, a dual polarized subreflector, and two orthogonal feed assemblies. Both the low sidelobe fixed beams and the provision for overlapping fixed and scanning beams can be achieved by using an offset Cassegrain antenna with a dual-gridded subreflector consisting of two offset hyperbolas arranged in a piggyback configuration. The front subreflector is gridded to pass one sense of polarization and reflect the orthogonal polarization. The back subreflector is solid and reflects the polarization transmitted by the front subreflector. The focal axes of the two subreflectors are symmetrically displaced from the symmetry plane of the main reflector by an angle of 10° so that two orthogonally polarized feed assemblies can be placed in their respective focal regions without mechanical interference.

A multireflector antenna system with an FSS subreflector is used for multiband operation. The FSS is designed to be transparent for one frequency band and to be reflective for another band. The feeds are placed at two foci of the reflector system. The typical FSS antenna has a main reflector, an FSS subreflector using etched Jerusalem cross elements, a cluster of receive feeds, and a transmit feed, as shown in Fig. 65.11. The receive feed cluster may allow adaptive nulling, and the main reflector may be gimballed to allow beam scan within the Earth coverage. This design concept is applicable to a MILSTAR MBA.

Successful deployment and application of the 30-ft-diam reflector on the ATS-6 satellite demonstrated the feasibility of large deployable reflectors. The deployable reflector antenna became a space proven technology that has been used for satellite applications. The use of a deployable reflector antenna provides an allowable larger aperture that can be stowed for launch, then deployed on orbit. The reflector utilizes GFRP ribs to shape and support the reflective mesh surface, as shown in Fig. 65.12. The number of ribs is based on a tradeoff considering surface tolerance and weight. As the number of ribs increases, the surface error decreases, while weight increases. This type of antenna has also been used for NASA's Tracking and Data Relay Satellite system (TDRSS) single access (SA) antenna, as shown in Fig. 65.13. Each TDRSS spacecraft has two SA antennas that are used for communication with user satellites in low Earth orbit.

65.4 Phased Array Antennas

The complexity of a phased array is highly dependent on the grating lobe requirements. A phased array antenna for a communication satellite, especially for geosynchronous orbit, has a limited field of view (FOV) and, thus, does not require grating lobe free antenna design. It only requires that the grating lobe be out of the FOV. Thus, the antenna design can be simpler, using high-gain elements with larger spacing.

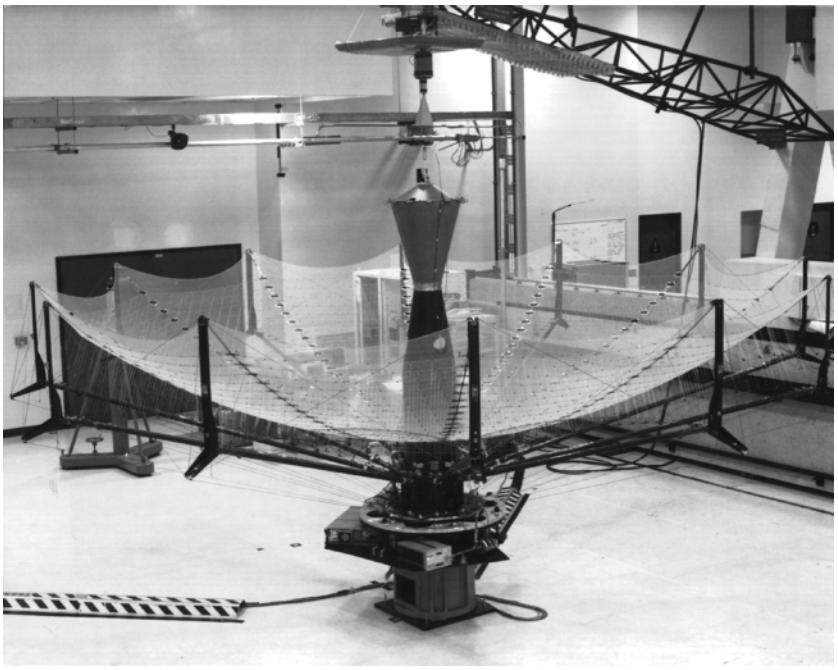


FIGURE 65.12 TDRSS mesh antenna.

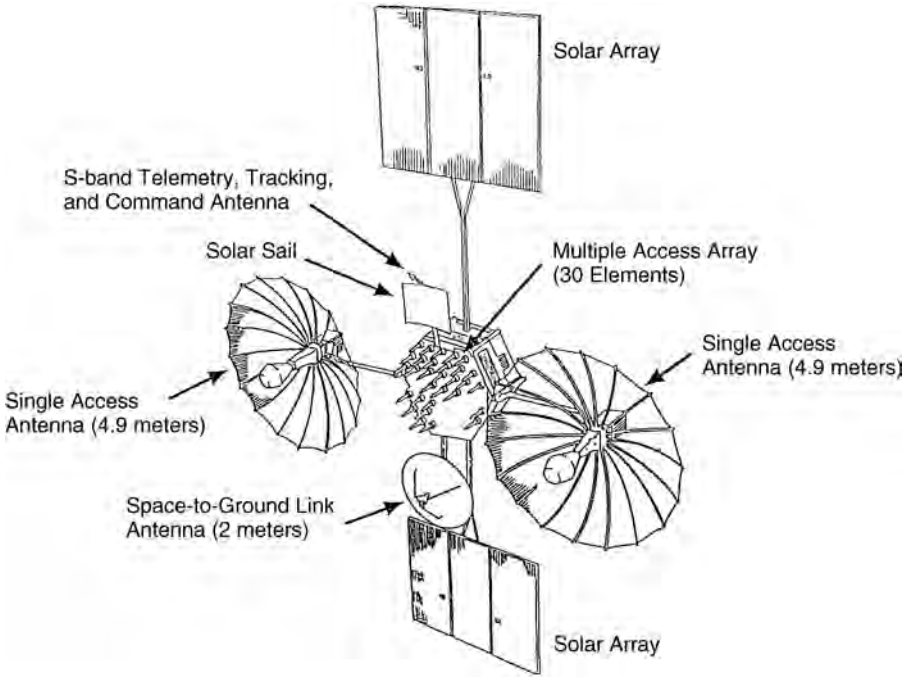


FIGURE 65.13 Tracking and data relay satellite.

TABLE 65.4 Types of Satellite Phased Array Antenna System

Type	Features and Applications
Horn array antenna	Good electrical performance Used as a feed array in a multibeam antenna system
Dipole array antenna	Low profile and simple
Helix array antenna	High gain is required Mobile antenna system
Waveguide slot array antenna	Low profile High power Meanderline polarizer is required for circular polarization
Active MMIC phased array	High gain as the BFN loss is compensated For by active device

The array gain is approximately expressed as

$$G = 10 \operatorname{Log}_{10} N + G_e$$

where N is the number of elements and G_e is the element gain in decibels. A larger element has higher gain, and, thus, the number of elements is minimized for the given gain. Also, mutual coupling between elements is insignificant if the element spacing is larger than two wavelengths. If the gain requirement is stringent, a large number of elements is required and the mechanical complexity increases. In this case, the trade should be performed with comparison to a multibeam antenna.

The most widely used radiating elements in a satellite phased array antenna system are the horn, the dipole, the helix, and the microstrip patch, as summarized in Table 65.4. The horn is most frequently used. The helix is used for a high-gain antenna system used in mobile communication or in a multiple access (MA) system like TDRSS. The waveguide slot array antenna is also used for a high-gain and high-power communications system. Most of waveguide slot array antennas are linearly polarized. If circular polarization is required, it can be obtained by putting a layer of meanderline polarizer in front of the array to transform a linearly polarized wave to a circularly polarized one. The microstrip patch is used where a low profile and compact size are required.

The TDRSS MA antenna [Imbriale and Wong, 1977] is an S-band, 30-element phased array antenna for both transmit and receive. Between transmit and receive, 12 elements are diplexed, and the remaining 18 elements are receive only. The MA antenna is spacecraft body mounted with a fixed ground plane and points to the spacecraft nadir, as shown in Fig. 65.13. The radiating elements are helices arranged in a triangular grid. The circularly polarized helical element is lightweight and provides 13 dB of gain over ±13.5° FOV. Figure 65.14 shows a block diagram of the transmit array. The transmitter splits the input signal (2.1064 GHz) with a 12-way power divider. Each element is individually phase shifted and amplified to form a transmit beam. Figure 65.15 shows a block diagram of the receive array. The receive signals (2.2875 GHz) from each of the MA users are received and preamplified at each of the 30 receive elements. The preamplified signals are then multiplexed. The beam forming is performed at IF level. Recent progress made in the areas of GaAs monolithic microwave integrated circuit (MMIC) design and processing in conjunction with advanced microwave packaging techniques has made MMIC phased array antennas attractive for the satellite that provides one or more pencil beams capable of being steered at electronic speed. The beam forming network's high loss is compensated for by the active devices. COMSAT Laboratories has built a 64-element prototype Ku-band active phased array [Zaghloul et al., 1990] for application to INTELSAT coverage of the Atlantic Ocean region with four dual-polarized reconfigurable shaped beams. A 3.4λ square horn arranged in a triangular lattice is used. Figure 65.16 shows a single active element and the 64-element active phased array system.

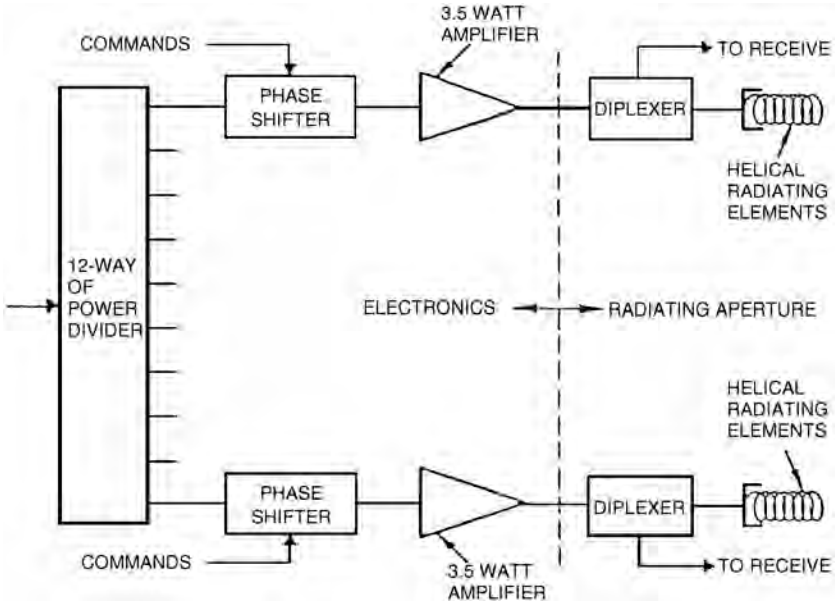


FIGURE 65.14 TDRSS MA transmit phased array antenna.

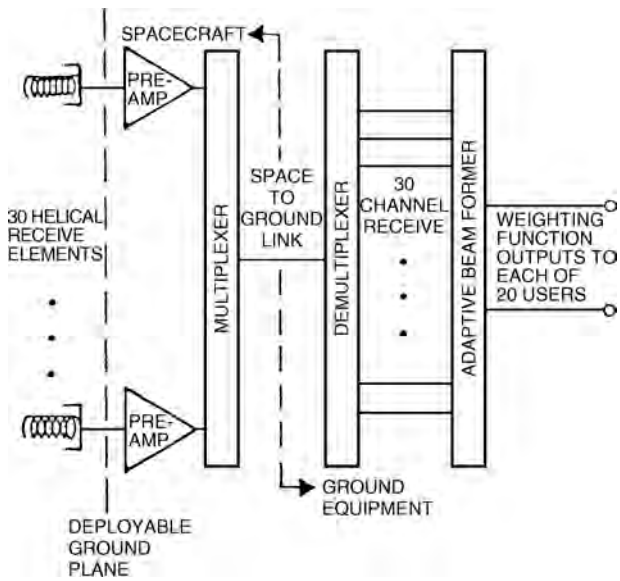


FIGURE 65.15 TDRSS MA receive phased array antenna.

65.5 Tracking, Telemetry, and Command Antennas

The TT&C antenna must provide broad coverage throughout the spacecraft mission, including initial acquisition, transfer orbit, and normal on-orbit operations. It must also provide nearly omnidirectional coverage for the unlikely event of altitude failure. If the satellite starts to tumble, satellite control commands from the ground must be received for any arbitrary orientation of the satellite. During on-orbit operation, the antenna provides Earth coverage for the lifetime of the satellite.

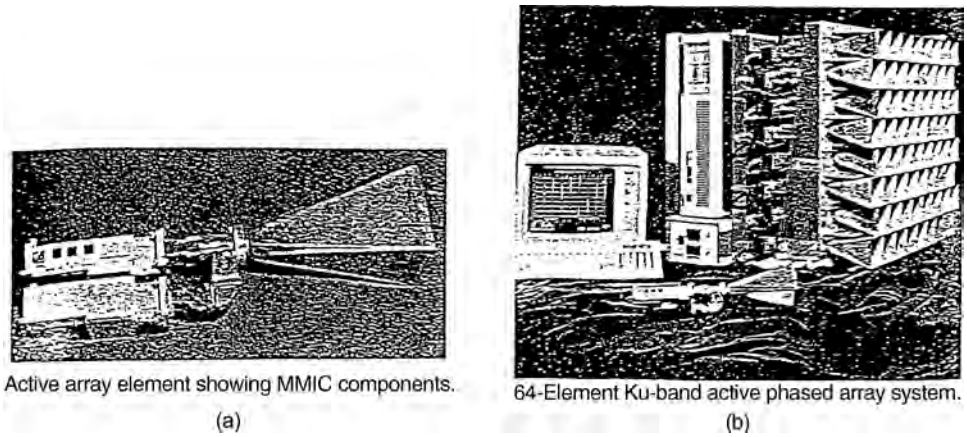


FIGURE 65.16 COMSAT Ku-band MMIC phased array antenna: (a) element and (b) array.

To achieve omnidirectional coverage, two hemispheric antennas can be mounted on opposite sides of the satellite and the appropriate antenna selected. For on-orbit Earth coverage, either one of the hemispheric antennas or the separate horn antenna can be used. In this case, high system margin is achieved, but switching complexity between antennas remains an issue. The design challenge for the TT&C antenna is to use only a single antenna having a broad beam cardioid radiation pattern with lower system margin.

Most frequently used TT&C antennas are conical spiral antennas [Yeh and Mei, 1967] or biconical antennas [Barrow, Chu, and Jansen, 1939]. A loosely wound four-arm conical spiral antenna mounted at the end of a deployable boom is ideal for the single antenna approach. Figure 65.17 illustrates various TT&C operation modes using a single antenna. This approach is used for the geostationary operational environmental satellite (GOES) weather satellite. This antenna is simple and lightweight, provides nearly identical patterns, and has good polarization purity. During initial acquisition, the antenna is deployed, but no solar panel deployment has taken place. The coverage region of interest is a 45∞ conical region about the z-axis. In transfer orbit operation, a single solar panel is deployed, and the coverage region of interest is a 45° conical region about the x-axis. For on-orbit operation with full deployment of the solar panels and the solar sail, the coverage region is only a 9° cone about the z-axis.

The conical spiral antenna can provide satisfactory coverage under free-space conditions, but the interference and shadowing caused by the spacecraft could cause serious pattern degradations that result in inadequate coverage in some regions of space. To calculate the antenna pattern in the presence of the spacecraft, the geometrical theory of diffraction (GTD) [Lee, Schenuum, and Hwang, 1991] is used as the best method of analysis. The GTD is a high-frequency solution. Using this approach, a ray optics technique can be applied to determine the fields incident on the various scatterers such as edges, corners, and curved surfaces. The fields diffracted are found by using the GTD solutions in terms of rays summed with the geometrical optics components and edge-diffracted components in the far field. The basic approach is to model the spacecraft with simple structural forms such as flat plates and cylinders; this has been used very successfully in the past.

65.6 Current and Planned LEO/MEO Mobile Satellite Antennas

Most of the LEO/MEO satellite antennas are designed to provide frequency reuse for cellular systems. As a result, the coverage areas are divided into cells. Each cell is covered by an antenna beam. The LEO/MEO satellites orbit in the Van Allen radiation belts, which lie between the magnetosphere and the ionosphere. The environment of these two orbital satellites is somewhat different from that of GEO. The Van Allen belts consist of two regions containing energetic protons and electrons traveling around the Earth's

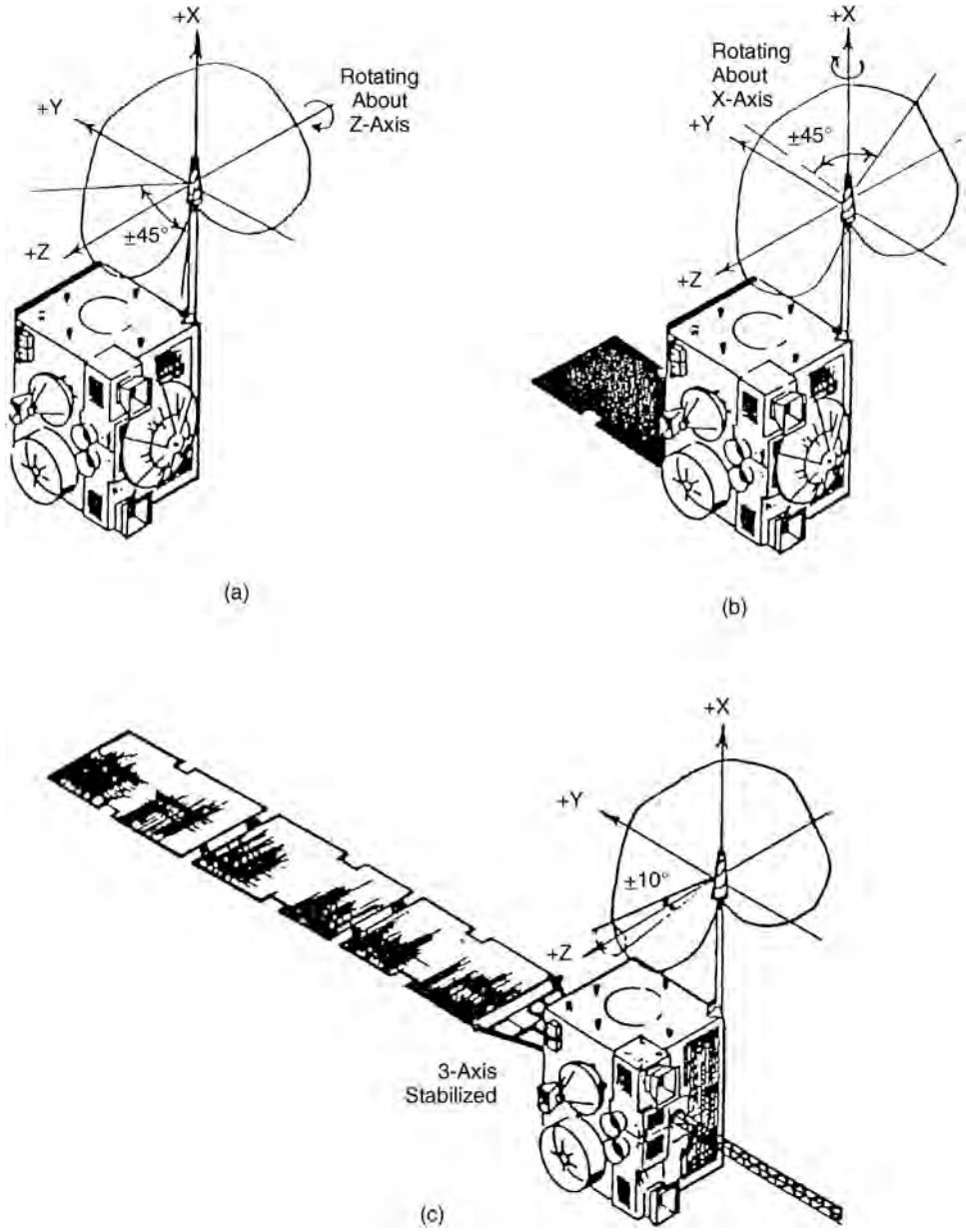


FIGURE 65.17 TT&C antenna and its operation modes: (a) initial acquisition, (b) transfer orbit, and (c) on-orbit operation.

magnetic field lines. LEO satellites are orbiting under the most intense and damaging portions of the lower Van Allen belts. The energetic particles can damage the exposed antenna components. The eclipse portions and the eclipse durations of the LEO/MEO satellites are also different from those of GEO. They may spend 35% or more of their time in the Earth's shadow.

The logical choice for the LEO/MEO satellite antennas is the multiple beam antenna system. The types of antennas selected can be phased arrays or reflector multiple beam antenna systems. Variable beamwidth, sidelobe control, and/or power level in each beam can also be incorporated in the system. The phased array can be an active MMIC phased array with distributed output power amplifiers.

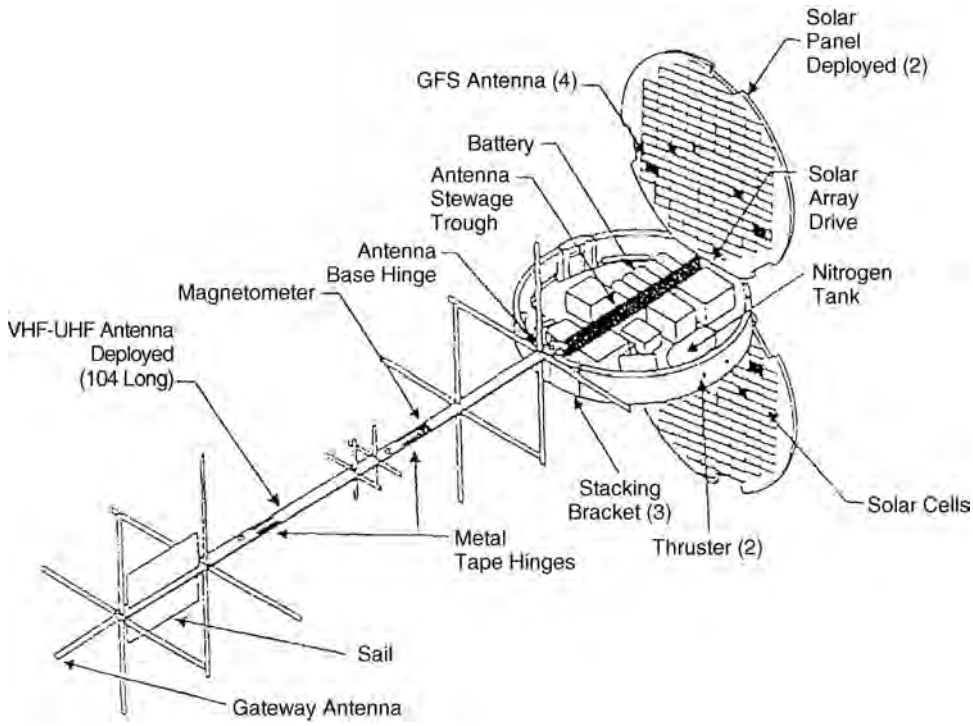


FIGURE 65.18 Orbcomm's satellite configuration.

Even though many mobile communication satellite systems have been proposed, at this writing, communication antennas are still in the design and development phase. The antennas can range from the simplest system, such as the VHF/UHF cross-dipole array antenna in Orbcomm's Microstar satellites, to the very complicated Ka-band MBA system in Teledesic satellites. Not all of the proposed antenna systems are described in the open literature. Figure 65.18 depicts the Orbcomm's spacecraft [Schoen and Locke, 1995] in which the antenna is designed for uplink frequencies from 148 to 150.05 MHz and downlink frequencies from 137 to 138 MHz.

The currently proposed antennas in the Globalstar mobile satellite system [Schindall, 1995] are L- and S-band active MMIC phased array antennas. The antenna design may be changed in the final development stage. Each antenna will generate 16 beams that cover the 5760-km-diam circle on Earth. The antenna beam is shaped to yield higher gain at the edge of coverage to compensate for the higher range losses, thus providing nearly constant power throughout the operating region. Figure 65.19 depicts the 91-element transmit array. The receive antenna is separated from the transmit antenna and is almost identical to the transmit antenna except that high-power amplifiers are replaced by low-noise amplifiers. The element consists of microstrip patch, filters, interfacing substrates, and hybrids used for generating circular polarization. Heat sink is used for heat dissipation. The beam forming network is implemented in stripline and comprises 32 layer Teflon dielectric substrates with required printed power dividers or combiners. The transmit antenna operates at 2.4835 to 2.5 GHz with minimum 50-dB gain, whereas the receive antenna operates at 1.61 to 1.6265 GHz with minimum 40-dB gain.

The counterpart in the Odyssey communication satellite system [Speltz, 1996] is a phased array antenna system. The antenna design may be changed in the final development stage. Two direct radiating arrays have been developed: an L-band downlink and an S-band uplink. Each antenna generates 61 spot beams. Hybrid networks driven by amplifiers provide real-time power reconfiguration to meet varying demands in each region, and the satellite antennas are designed to provide coverage to a portion of the total area visible to the satellite.

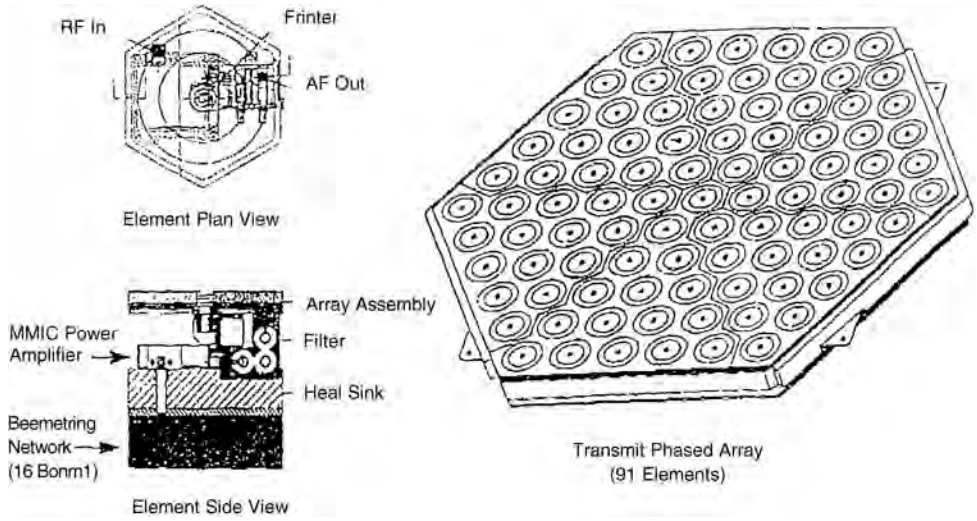


FIGURE 65.19 Globalstar active MMIC phased array antenna.

65.7 Radiometer Antennas

A satellite passive radiometer has unique properties that are not found in other measurement systems: it provides consistent viewing angles to a given location. The capability to observe the same location for a long time prolongs the sensing dwell time to achieve the required signal-to-noise ratio of the measurement within the spectral resolution limits of the instrument. Such an ability to provide a temperature profile at frequent intervals has proven significantly beneficial to meteorological analysis and forecasting based on modeling simulation.

The radiometer is a multichannel microwave (90–230 GHz) system. Temperature profiles are made near the center and on the wings of the 118-GHz oxygen absorption region. Moisture profiles can be made using the 183-GHz water vapor absorption band. The determination of precipitation coverage and rate can be done at several frequencies at the atmospheric windows at 92, 104, 150, and 230 GHz. All channels can operate simultaneously from the same resolution area using coincidental beams. However, beamwidths of different channels need not be the same.

Key design parameters are beam efficiency, antenna losses, and antenna physical temperature, in addition to conventional requirements such as directional gain, low sidelobes, mismatch, and polarization purity. The beam efficiency is the fraction of total power contained within the main beam of the correct polarization. It should be greater than 85%.

An antenna of at least 4.4-m aperture is required to provide 30-km nadir spatial resolution. The 4.4 m is also the largest antenna that fits the space shuttle cargo bay without folding. However, it is difficult to build a lightweight antenna with dimensional stability and surface accuracy of 2 mils in a space thermal environment. The scanning of a 4.4-m antenna generates angular momentum, which must be compensated for by the spacecraft attitude control subsystem.

65.8 Space Qualification

Qualification tests of all flight hardware are required to show suitability of the design for the space environment. Although the space or operating environment cannot be exactly duplicated, the tests are conducted in a controlled manner so that the hardware that passes the tests will function successfully.

The tests can be grouped into two main categories: external environment during and after launch and operating environment. The external environment tests include zero gravity, high vacuum, solar radiation, g-radiation, neutron radiation, electromagnetic pulse, and mechanical launch environment.

Four main components of the launch environment are steady-state acceleration, low-frequency vibration (5–100 Hz), vibro/acoustic (20–2000 Hz), and pyrotechnic shock. Space is a vacuum with diurnal and seasonal temperature fluctuation; the thermal vacuum test is for survivability in such environment. If the component must survive nuclear weapons effects or laser attack, it may also undergo g-ray testing and electromagnetic compatibility testing.

The operating environment refers to the circumstances in which the components or communication system must operate. Two factors are considered most in the operating environment: **multipactor** and **passive intermodulation (PIM)**. Multipactor discharge can cause noise in a sensitive system, heating of the surface, and gas discharge into the local atmosphere. The temperature can rise to the point that it discolors or burns the components. PIM takes place in high-power applications, especially where the antenna is designed for both transmit and receive functions. PIM may interfere with the receive signals. Careful attention should be paid to the design and fabrication of any junction between microwave components. Materials play an important role in the design's ability to meet these stringent requirements. Graphite epoxy is widely used for the reflector and other composite materials.

Space qualification is a lengthy and expensive process. Any new hardware component or system has to go through a dedicated qualification phase. Components of lesser importance can be qualified by similarity, that is, the component and the environment are identical to previously qualified hardware.

65.9 Future Trends and Further Study

Communications satellite antenna systems have undergone drastic changes since the advent of satellites. The changes can be divided into five major phases. In the first phase, antenna systems produced only simple circular beams used in the first generation satellites in the early 1960s. The second phase started in the 1970s. In this phase, a shaped beam directed the energy to the coverage area to improve effective isotropic radiation power, reducing the output of the power amplifiers. A major change occurred in the third phase (1980s), when the demand for increased communication satellite capacity directed frequency reuse. Typical technology during this period used a multibeam antenna system to provide highly shaped patterns and several simultaneous beams, such as that used for INTELSAT V. Limited beam reconfiguration was required to accommodate the change of radiation pattern from one satellite orbital location to another. The fourth phase (1990s) was characterized by full exploitation of the beam reconfiguration. Beam reconfiguration was also extended from space to channel. Typical examples are the direct broadcast satellite and NASA ACTS, which can provide coverage of either half CONUS or full CONUS. The fifth phase (2000s and beyond) will represent rapid antenna pattern control capability, required to provide adaptive nulling and coverage for mobile communications satellites. Table 65.5 summarizes technology development and risks at each of the five phases.

The design of a satellite antenna system is driven by the system requirements, technology advancement, and maturity of developed components and composite materials. Recent research and development in MMIC technology, photonic array antennas, and superconductive materials in the communications system will be incorporated into future satellite antenna systems. Present satellite antenna systems are usually designed separately, then integrated into other subsystems, that is, filters, multiplexers, and power and receive amplifiers. Future communications systems will be designed and developed as integrated systems.

Communications antenna systems design involves a tradeoff between the space segment and the ground segment. Higher EIRP in the space segment enables a lower cost and a smaller ground antenna system. A lower EIRP satellite antenna system reduces the complexity and cost of the space segment; however, a larger ground antenna system is required to compensate for the space loss. This incurs higher cost for the ground segment.

TABLE 65.5 Evolution of Satellite Antenna System Technology

Year	Antenna Radiation Patterns	Key Components	Required Technology Development
1960s	Simple beam	Antenna connectors	Lightweight Space qualified materials
1970s	Simple shaped beam	Reflector Deployable mesh reflector Lens	Lightweight Space qualified materials
1998s	Reconfigurable over Space	Reflector Lens Gridded reflector Shaped reflector FSS reflector Variable, power divider (VPD)/ Variable phase shifter (VPS) Beam forming network (BFN) components,	Wider bandwidth Computer-aided design (CAD) tools Switches
1990s	Reconfigurable over space and channel	Ditto MMIC LNAs MMIC VPD/VPS BFN controller	CAD tools Higher frequency BFN Switches MMIC solid-state power amplifier (SSPA)
2000s and beyond	Reconfigurable over space, channel, and time	Large optics Photonic phased array Active MMIC phased array	MMIC phased array FSS Multiband BFN components

Defining Terms

Antenna performance parameters: Include beam shaping efficiency that is used to indicate how well a beam is shaped as compared to an ideally shaped beam; circular polarizations, right-hand circular polarization (RHCP), and left-hand circular polarization; copolarization and cross-polarization isolations/interferences; equivalent isotropically radiated power (EIRP); field of view (FOV); half-power beamwidth (HPBW); and voltage standing wave ratio (VSWR).

Beam: A ray of electromagnetic field, which may be omnidirectional, pencil shaped, and/or reconfigurable; the beam that will change as a function of time and space.

Frequency reuse: A frequency band can be used more than once via space and/or polarization diversifications.

Graphite fiber reinforced plastic surface (GFRP): A perfectly conducting material that is light and has a very low-thermal-expansion coefficient.

Multipactor and passive intermodulation (PIM): Interference or noises due the heating of the surface, gas discharge, or higher order modes generated in the system.

Orbits: Geosynchronous Earth orbit (GEO), an orbit that is 19,300 nmi away from the Earth; low Earth orbit (LEO); medium Earth orbit (MEO).

Reflector surfaces: Include solid, gridded for polarization purification, frequency selective surface (FSS) that will pass one frequency band and reject another band(s), and meshed.

Space qualification: Qualification test of flight hardware.

Types of antennas: Include cup-dipole antennas, deployable reflector antennas that are stowed for launch and deployed on orbit, horn antennas, monolithic microwave integrated circuit (MMIC) phased array antennas, radiometer antennas, shaped beam antennas whose radiation pattern is shaped to conform the contour of the coverage area, and tracking, telemetry, and command (TT&C) antennas that provide communications during initial acquisition, transfer orbit, and on-orbit operations.

References

Barrow, W., Chu, L., and Jansen, J. 1939. Biconical electromagnetic horns. *Proc. IRE,* (Dec.):769–779.

Chen, C., Choung, Y., and Palma, D. 1986. A compact multiflare horn design for spacecraft reflector antennas. *1986 IEEE AP-S Symp. Digest,* (June):907–910.

Choung, Y. and Wong, W. 1986. Multibeam antenna design and development for NASA advanced communications technology satellite. IEEE/GLOBECOM Conference, Houston, TX, Dec., 16.2.1–16.2.6.

Clarricoats, P. and Saha, P. 1971. Propagation and radiation behavior of corrugated feeds. *Proc. IEEE,* 118(Sept.):1176–1186.

Hwang, Y. 1992. Satellite antennas. *IEEE Proc.,* 80(1):183–193.

Imbriale, W. and Wong, G. 1977. An S-band phased array for multiple access communications. NTC'77, 19.3.1–19.3.7.

James, G. 1982. TE_{11} to HE_{11} mode converters for small angle corrugated horns. *IEEE Trans. AP,* AP-30(6):1057–1062.

Kumar, A. and Hristov, H. 1989. *Microwave Cavity Antennas,* Artech House, Norwood, MA.

Lee, E., Schenuum, G., and Hwang, Y. 1991. Scattering measurements for I-VII TC&R antennas on a spacecraft mockup. 1991 IEEE AP-S Symposium, London, Canada, (June):1690–1693.

Potter, P. 1963. A new horn antenna with suppressed sidelobes and equal beamwidths. *Microwave J.,* VI(June):71–78.

Schindall, J. 1995. Concept and implementation of the globalstar mobile satellite system. 1995 International Mobile Satellite Conference, Ottawa, Canada, A-11–16.

Schoen, D. and Locke, P. 1995. Orbcomm—initial operations. 1995 International Mobile Satellite Conference, Ottawa, Canada, 397–400.

Speltz, L. 1996. Personal communications satellite systems. 18th Annual Pacific Telecommunications Conference, Honolulu, HI, Jan. 14–18.

Yeh, Y. and Mei, K. 1967. Theory of conical equiangular-spiral antennas. *IEEE Trans. AP-S.,* Pt. 1, (Sept.): 634–639.

Yeh, Y. and Mei, K. 1968. Theory of conical equiangular-spiral antennas. *IEEE Trans. AP-S.,* Pt. 2, (Jan.): 14–21.

Zaghloul, A., Hwang, Y., Sobello, R., and Assal, F. 1990. Advances in multibeam communications satellite antennas. *IEEE Proc.,* 78(7):1214–1232.

66

Tracking and Data Relay Satellite System

Erwin C. Hudson
WildBlue Communications, Inc.

66.1 Introduction

The tracking and data relay satellite system (TDRSS) was developed in the late 1970s and early 1980s by NASA Goddard Space Flight Center (GSFC) with an industry team comprising Western Union Space Communication Company, TRW Space & Electronics Group, and Harris Government Communications Division. Second generation tracking and data relay (TDR) satellites were developed by Boeing Satellite Systems, Inc. (formerly Hughes Space and Communications) in the late 1990s for launch in 2000, 2001, and 2002. TDRSS is now operated for NASA under the Consolidated Space Operations Contract (CSOC) by contractors under Lockheed Martin Corporation.

The TDRSS mission is to provide communications and tracking services to low-Earth orbiting user satellites and manned space missions. TDRSS is the primary communication and tracking system for the Shuttle Space Transportation System (STS) and the international space station (ISS), providing near-continuous data, voice, and video links for the manned space program. User satellites have included the Landsat 4 and 5 multispectral imaging satellites, the Compton Gamma Ray Observatory (GRO) and the Cosmic Background Explorer (COBE). Current user satellites include the Hubble Space Telescope (HST), the Tropical Rain Measurement System (TRMS), and the Terra Multispectral Earth Imaging Mission.

The TDRS system consists of three segments: the space segment encompassing the tracking and data relay satellite (TDRS) vehicles in **geosynchronous orbit**; the ground segment, which includes the White Sands complex (WSC) in New Mexico and the Guam Rremote Ground Terminal (GRGT) on the island of Guam; and the user segment, which includes the TDRSS transponders on user satellites orbiting the Earth. As shown in Fig. 66.1, forward links to user satellites are established through the WSC up to a TDRS and then down to the user satellites in low-Earth orbits. Return links from the user satellites are relayed up through a TDRS and then down to the WSC directly or indirectly via the GRGT.

Prior to TDRSS, NASA's worldwide network of tracking stations provided communications in 5 to 15 min contacts as user satellites occasionally passed over a station's coverage area. The TDRS relay satellites, from their vantage point 35,786 km above the Earth, allow near-continuous long arc coverage of low-Earth orbiting satellites. TDRSS eliminates the necessity of operating a network of ground-based tracking stations, while providing lower cost and higher performance communications for NASA.

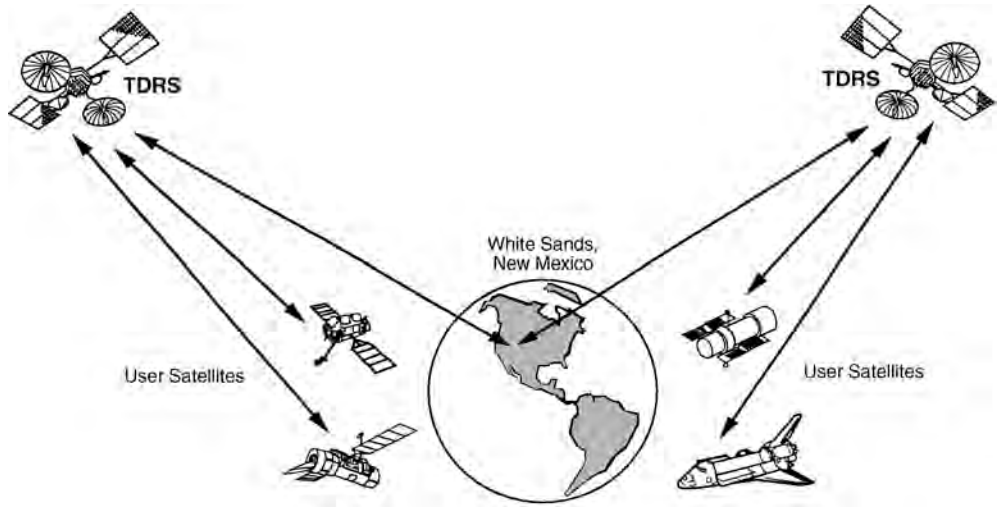


FIGURE 66.1 Tracking and data relay satellite system concept.

TDRS-A, the first TDRS satellite, was carried into orbit on shuttle mission STS-6 in April 1983. TDRS-B was lost in the space shuttle Challenger accident in January 1986. With shuttle mission STS-26 in September 1988, a second relay satellite, TDRS-C, was deployed, and the TDRS system became fully operational. The first second generation relay satellite, TDRS-H, was launched in June 2000. As of mid-2001, TDRSS is a mature system with a constellation of six relay satellites on orbit.

66.2 TDRS System Overview

TDRSS provides command, telemetry, range tracking, Doppler tracking, and wideband data relay services to user satellites operating in low-Earth orbits. User satellite payload operations control centers (POCCs) interface with TDRSS through the NASA communications (NASCOM) network. The user satellite operator sends a schedule request to the network control center (NCC), which determines whether TDRSS has the space and ground resources to honor the request. If the schedule request can be honored, the NCC sends a schedule order message through NASCOM to initiate service. At the scheduled time, the requested communication services are established between the user POCC and the user satellite, all automatically and without manual intervention.

TDRSS provides four user services: command, tracking, telemetry, and mission data communications. Command links are forward services, telemetry and mission data links are return services, and range tracking is both a forward and return (turn-around) service. TDRSS can also perform Doppler tracking using only a return service.

Command service. User satellite operators send instructions and other data to their space vehicles through command channels. Command links may be used, for example, to operate sensors and fire thrusters or to reconfigure onboard equipment to recover from failures.

Telemetry service. Satellites are designed to telemeter state-of-health and operational status data to their controllers. Satellite telemetry typically includes vehicle attitude, temperatures of sensitive equipment, power system voltages and currents, and other critical parameters.

Tracking service. Space missions such as Earth sensing require accurate knowledge of the orbital position of the sensor. In addition to Doppler frequency measurements, which provide an estimate of range rate, TDRSS performs user satellite ranging by sending a **pseudonoise (PN) coded** ranging signal on a round-trip path from the TDRSS ground station, to the user satellite, and back to the TDRSS ground station. The round-trip time delay is determined by correlating the transmitted

ranging code with the returned version of the code and adjusting time delays until the two copies match. With proper time delay calibration, the round-trip distance can be determined to within a few meters. TDRSS uses ground-based transponders at known locations to perform rapid and highly accurate orbit determination for each TDRS. The NASA Flight Dynamics Facility precisely determines the orbital elements of each user satellite using the TDRS orbital data and the ranging and Doppler measurements to the user transponders.

Mission data relay service. Mission data is the experiment data, scientific measurements, sensor data, or other digital information generated by a satellite in orbit. While many scientific missions have data rate requirements of 1 Mb/s or less, Earth sensing and imaging missions may require tens to hundreds of megabits per second. A typical TDRSS contact involves a forward link command service to a user satellite and a return link data relay service back through the WSC to the user POCC. Based on the schedule order parameters, the required TDRS and WSC equipment is assigned and configured prior to the contact. At the specified time, the TDRS points the assigned antenna toward the user satellite and activates the forward signal. The forward signal frequency is initially offset and then continuously adjusted to compensate for Doppler shift so that the user receives the signal at a constant frequency. The user satellite acquires the forward signal and establishes simultaneous two-way communications by transmitting a return link signal back to the TDRS. At the end of the contact, the TDRS and WSC equipment assignments are released, and the equipment is made available to support other users.

66.3 TDRSS Communications Design

TDRSS provides a complete set of communication capabilities for existing, planned, and future user satellites. The system supports a variety of frequencies, data rates, modulation types, and signal formats [Poza, 1979]. The relay satellites themselves are relatively simple analog repeaters; the communications processing is done in the ground facilities where the equipment can be maintained, upgraded, and modified as required.

The primary TDRSS constellation consists of two satellites over the Atlantic Ocean, at 41° West and 46° West longitude, and two satellites over the Pacific Ocean, at 171° West and 174° West longitude. These orbital locations were selected to maximize worldwide coverage while providing line-of-sight visibility to White Sands. The dry desert climate in southern New Mexico results in low rain losses and excellent link availability, even at low elevation angles; this is one of the principal reasons NASA selected White Sands for the TDRSS ground site.

The primary TDRSS constellation has an oval-shaped zone-of-exclusion (ZOE) centered over the Indian Ocean where coverage is not available below 1200 km due to blockage by the Earth. Since the early 1990s, NASA has stationed a TDRS over the Indian Ocean at 85° East longitude to provide TDRSS services in the ZOE. While the original ZOE implementation provided only limited S-band services to GRO using a temporary ground facility in Australia, TDRSS ZOE coverage is now fully operational for all TDRSS services using the GRGT located on Guam.

Each TDRS communicates with its ground control site over **Ku-band** space-to-ground links. User satellites communicate over space-to-space links with the TDRS relay satellites at **S-band**, Ku-band, and/or **Ka-band**. The TDRSS frequency plan is shown in Fig. 66.2.

TDRSS services are referred to as either single access or multiple access. Single access service is established using one of two large reflector antennas on the TDRS. Multiple access service is provided with an S-band array antenna on each TDRS. Multiple access service supports up to 20 simultaneous **code-division multiple access (CDMA)** users in the return direction while providing a single beam in the forward direction to one user satellite at a time.

Table 66.1 describes the four forward link services available through TDRSS. Multiple access (MA) forward service uses a 6-MHz bandwidth channel centered at a fixed frequency of 2.1064 GHz. S-band single access (SSA) forward channels are 20 MHz wide and are tunable over the 2.0304 to 2.1133 GHz band. Ku-band single access (KSA or KuSA) forward channels are centered at a fixed 13.7750 GHz and

TABLE 66.1 TDRSS Forward Service Capabilities

Forward Service	Frequency F_{fwd} (GHz)	Bandwidth (MHz)	Forward Link Data Rate	Ranging Rate ([a]3 Mchips/sec)
MA	2.1064	6	100 b/s–10 kb/s	$F_{fwd} \times 31/221/96$
SSA	2.0304–2.1133	20	100 b/s–300 kb/s	$F_{fwd} \times 31/221/96$
KSA	13.7750	50	1 kb/s–25 Mb/s	$F_{fwd} \times 31/1469/96$
KaSA	22.5500–23.5500	50	Up to 50 Mb/s	Ka ranging not available

TABLE 66.2 TDRSS Return Service Capabilities

Return Service	Frequency F_{rtn} (GHz)	Bandwidth (MHz)	Return Link Data Rate	Ranging Rate ([a]3 Mchips/sec)[a]
MA	2.2875	5	100 b/s–50 kb/s	$F_{rtn} \times 31/240/96$
SSA	2.2050–2.2950	10	100 b/s–6 Mb/s	$F_{rtn} \times 31/240/96$
KSA	15.0034	225	1 kb/s–300 Mb/s	$F_{rtn} \times 31/1600/96$
KaSA	25.2500–27.5000	225 or 550	Up to 800 Mb/s	Ka ranging not available

[a]Return link ranging is not available when the SSA data rate exceeds 3.15 Mb/s or when the KSA data rate exceeds 150.15 Mb/s.

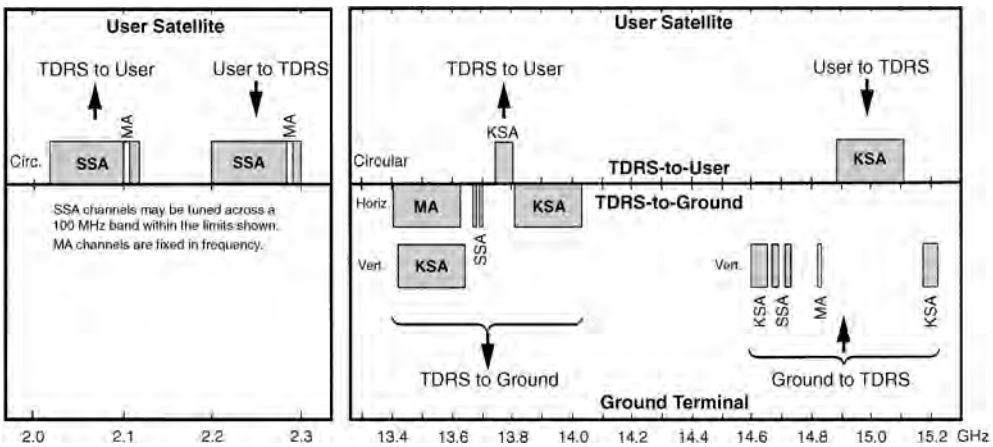


FIGURE 66.2 TDRSS misssion communication frequency plan.

have 50-MHz bandwidths. Ka-band single access forward channels, available only on the second-generation TDR satellites, have a bandwidth of 50 MHz and can be tuned over the 22.550–23.550 GHz band. The forward link ranging code rates, nominally 3 Mchips/sec, are coherently related to the forward link carrier frequency by the ratios in Table 66.1. All TDRSS frequency conversions are performed coherently to maintain this relationship between the code rate and the carrier frequency [Gagliardi, 1979].

Table 66.2 describes the four TDRSS return link services. MA return service provides users with a 5-MHz bandwidth channel centered at a fixed 2.2875 GHz. SSA return channels are 10 MHz wide and tunable over the 2.2050 to 2.2950 GHz band. KSA return channels, centered at a fixed 15.0034 GHz, are 225 MHz wide and support data rates up to 300 Mb/s. The MA and SSA return carrier frequencies are generated by multiplying the forward link frequencies by the ratio 240/221. The KSA return carrier is derived in the same manner using the turn-around ratio 1600/1469. Ka-band single access return channels, available only on the second-generation TDR satellites, have a bandwidth of either 225 MHz or 550 MHz and can be tuned over the 25.250 to 27.500 GHz band. The return link ranging code rates are the same as the forward link rates and are tied to the carrier frequency by the ratios given in Table 66.2.

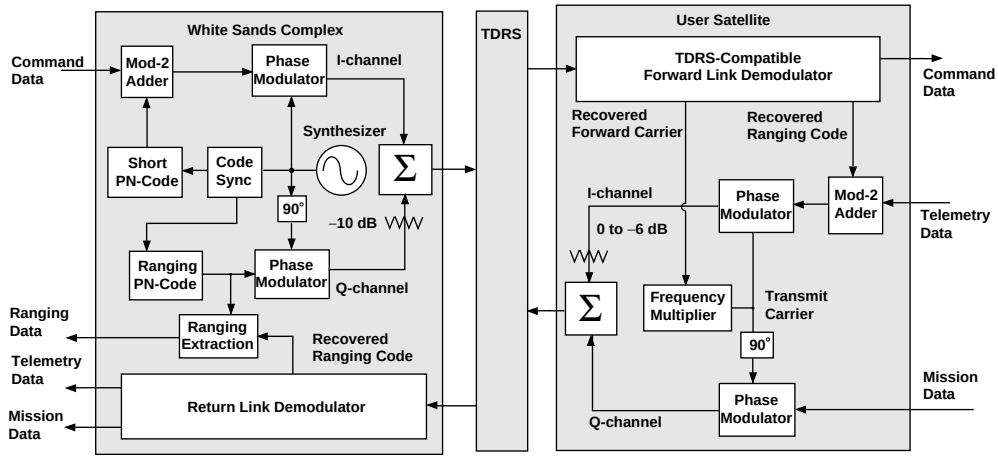


FIGURE 66.3 TDRSS forward and return link signal design.

Figure 66.3 shows a typical TDRSS forward signal transmitted up from the WSC to a user satellite and the return signal from the user satellite back down to the WSC. Unlike ground-based satellite tracking systems, the TDRS transmits high-power signals down from above with the Earth in the background. For data rates below 300 kb/s, TDRSS must use **spread-spectrum** signals rather than narrowband signals to avoid violating international limitations on microwave flux density [NASA, 1995, App. G].

A typical TDRSS forward signal is **quadrature phase-shift keying (QPSK)** modulated with the quadrature (Q) channel attenuated 10 dB below the in-phase (I) channel. A short 1023-bit PN-code running at a chip rate of approximately 3 MHz (per Table 66.1) is modulo-2 added to the command data and modulated onto the I-channel. The Q-channel is modulated with a long 1023×256 bit PN-code running at the same chip rate as the short code. The starting points or epochs of the short and long codes are synchronized to simplify user acquisition. The short code is used by the user satellite transponder for initial acquisition and the long code is used for turn-around ranging.

A typical TDRSS return signal is QPSK modulated with telemetry data on the I-channel and mission data on the Q-channel. The I-channel may be attenuated up to 6 dB relative to the Q-channel. The return link carrier is generated by frequency multiplying the recovered forward link carrier by the appropriate turn-around ratio. The user satellite telemetry data on the I-channel is modulo-2 added to the ranging code recovered from the forward signal. When the mission data rate on the Q-channel is less than 300 kb/s, a PN-code may be modulo-2 added to the data to improve interference rejection.

Although PN-coded QPSK is the most common TDRSS waveform, additional signal types, waveform designs, and communication modes are available. The *Space Network (SN) Users Guide* [NASA, 1995] is a complete reference for the entire set of frequency bands, waveforms, and data rates available through TDRSS.

66.4 First Generation TDRS Relay Satellite Design

Each first generation TDRS relay satellite consists of a communications payload module containing the antennas and electronics to perform the TDRSS mission and a spacecraft bus module that provides a three-axis stabilized platform and support functions for the payload [Harper and Woodson, 1981]. A sketch of the first generation TDRS vehicle is shown in Fig. 66.4.

The TDRS frame is constructed primarily of aluminum and consists of a six-sided outer structure built around a cylindrical inner section. The cylindrical section supports the antenna platform on the top, attaches to the launch vehicle adapter on the bottom, and contains two pressurized fuel tanks in its interior. The six-sided enclosure houses the communications payload and spacecraft bus electronics on

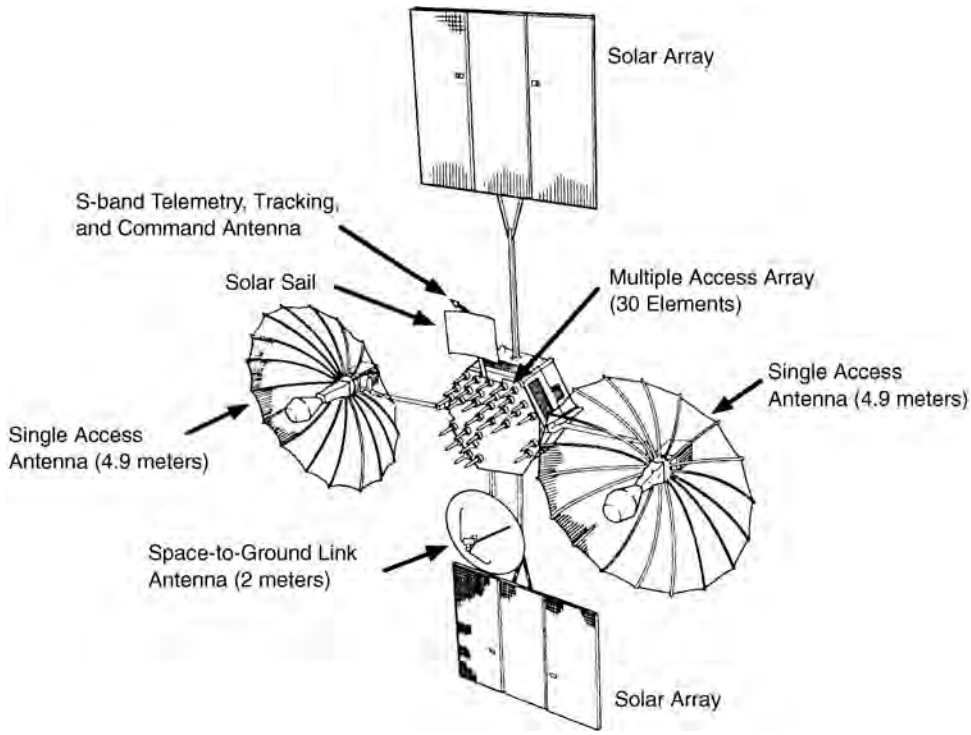


FIGURE 66.4 Tracking and data relay satellite (TDRS-G shown).

the interior facing sides of 12 rectangular panels. The weight of the fully assembled and fueled satellite, including the launch vehicle adapter, is 2230 kg.

Two rectangular solar arrays, each approximately 3.8 × 3.8 m, are attached to the north and south sides of the spacecraft. The two arrays contain 28,080 silicon solar cells and produce greater than 1700 W at end-of-life. A solar array drive motor slowly rotates each array once per 24 h to track the sun as the TDRS orbits the Earth.

The TDRS antenna suite consists of two 4.9-m single access antennas, a 30-element multiple access array, and a 2-m space-to-ground link antenna. The single access antennas are mechanically gimbaled and steerable over ±22.5° East–West and ±31.0° North–South. The multiple access array is electronically steerable over a ±13.0° conical field-of-view.

Attached directly to the base of each single access antenna is an electronics compartment, which houses the high-power transmitters and low noise receivers. Placing this equipment at the antenna minimizes signal losses between the antenna and the electronics, improving receiver sensitivity and radiated power.

The block diagram in Fig. 66.5 shows the signal flow through the TDRS communication payload. Backup equipment, frequency generators, TDRS command and telemetry processors, and other support equipment are not shown.

The TDRS interface with the WSC is at Ku-band through the space-to-ground link antenna. The frequency-division multiplexed forward signals from the WSC are downconverted and then demultiplexed in the forward processor. The forward processor performs automatic level control on the MA, SSA, and KSA forward signals and tunes the SSA signals to the commanded center frequencies. The forward processor then sends the SSA and KSA signals to the single access antennas and the MA signal to the MA transmitter.

The KSA forward signals are upconverted to 13.7750 GHz in the antenna compartment and transmitted through a 1.5-W power amplifier. The SSA forward signals are sent to the antenna centered at the commanded frequency between 2.0304 and 2.1133 GHz, and are transmitted through a 26-W power

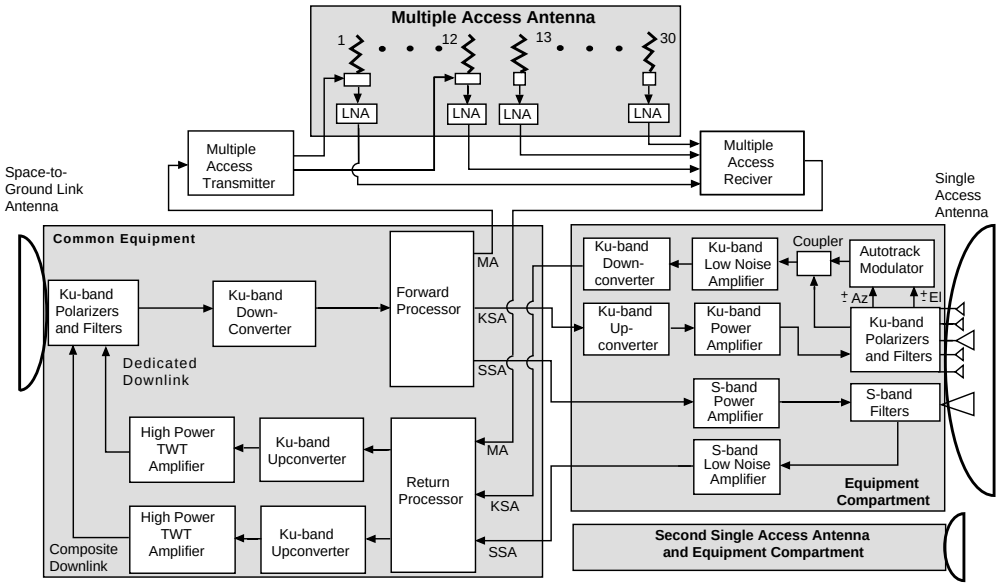


FIGURE 66.5 Tracking and data relay satellite payload block diagram.

amplifier. Both KSA and SSA transmitters may be used simultaneously through one single access antenna for dual-band user service.

The MA phased array has 30 elements, 12 of which can simultaneously transmit and receive; the remaining 18 elements are receive only. The MA transmitter splits the 2.1064-GHz MA forward signal using a 12-way divider and adjusts each of these components in phase to steer the forward beam. Eight signals are used for normal operations, and the remaining four are spares. Each transmit signal is routed through a 3.6-W power amplifier in the MA transmitter and a composite beam is formed in the direction of the user satellite.

All 30 MA array antenna elements have low noise amplifiers, providing 30-MA return signals centered at 2.2875 GHz to the MA receiver. The MA receiver frequency-division multiplexes these 30 signals on 7.5-MHz centers into a 225-MHz wide signal. The MA receiver sends the multiplexed MA return signal to the return processor for downlink transmission. MA return array processing is performed on the ground rather than onboard to minimize the complexity of the satellite.

The single access antennas have both S-band and Ku-band low noise amplifiers for SSA and KSA return services. The KSA return signal at 15.0034 GHz is downconverted in the antenna equipment compartment, and both the SSA and KSA return signals are sent to the return processor for downlink transmission. The return processor tunes the center frequencies of the SSA return channels over the 2.2050–2.2950 GHz receive band.

The single access antennas have narrow 0.27° half-power beamwidths at Ku-band, and auto-tracking is available to ensure accurate beam pointing. TDRS performs KSA autotracking using four error-sensing horns around the main Ku-band receive feed. The autotrack modulator scans these error signals and amplitude modulates azimuth and elevation errors onto the KSA return signal. The autotrack modulator has a low modulation index so that the autotrack modulation has essentially no effect on the phase modulated KSA user data. Autotrack error signals are recovered at the TDRS ground site and used to generate antenna pointing commands that are sent up to the TDRS. Autotrack improves pointing from a 99% circular error probability of 0.22° using open-loop pointing alone to better than 0.06° with closed-loop autotrack, reducing potential KSA pointing loss from 9.6 dB to 0.6 dB.

The return processor frequency-division multiplexes the 30-MA return element signals, the SSA return signals, and one of the two KSA return signals into a composite downlink. The composite downlink is amplified through a 30-W traveling-wave tube amplifier (TWTA) and transmitted on horizontal polarization.

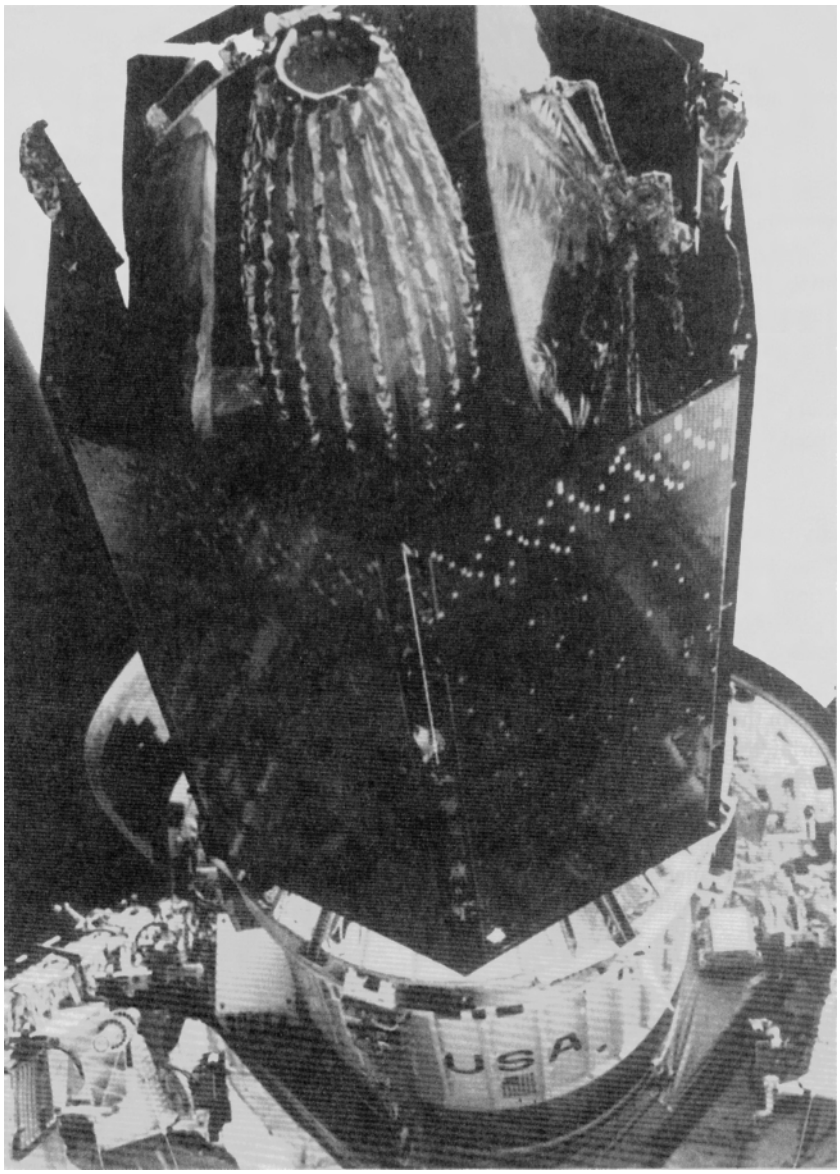


FIGURE 66.6 TDRS-A launched on shuttle mission STS-6, April 1983.

The return processor performs automatic level control on the return signals to keep the TWTA in its highly linear region, at an output power of approximately 3 W, to avoid intermodulation products between the many composite carriers. The remaining KSA return signal is transmitted on a dedicated downlink. The dedicated downlink uses an identical 30-W TWTA in its quasilinear region, at approximately 12-W output power, to prevent suppression of the autotrack modulation on the KSA return carrier. The dedicated downlink band overlaps the MA portion of the composite downlink band, and the dedicated downlink is transmitted down on vertical polarization to avoid interference.

The first generation TDRSS satellites were launched on the space shuttle and used Boeing's inertial upper stage (IUS) to fly to geosynchronous orbit. The TDRS/IUS vehicle, in the configuration shown in Fig. 66.6, is released from the cargo bay; the IUS performs an initial burn to place the TDRS into an elliptical transfer orbit; the first stage of the IUS is discarded. Several hours later at the peak or apogee of

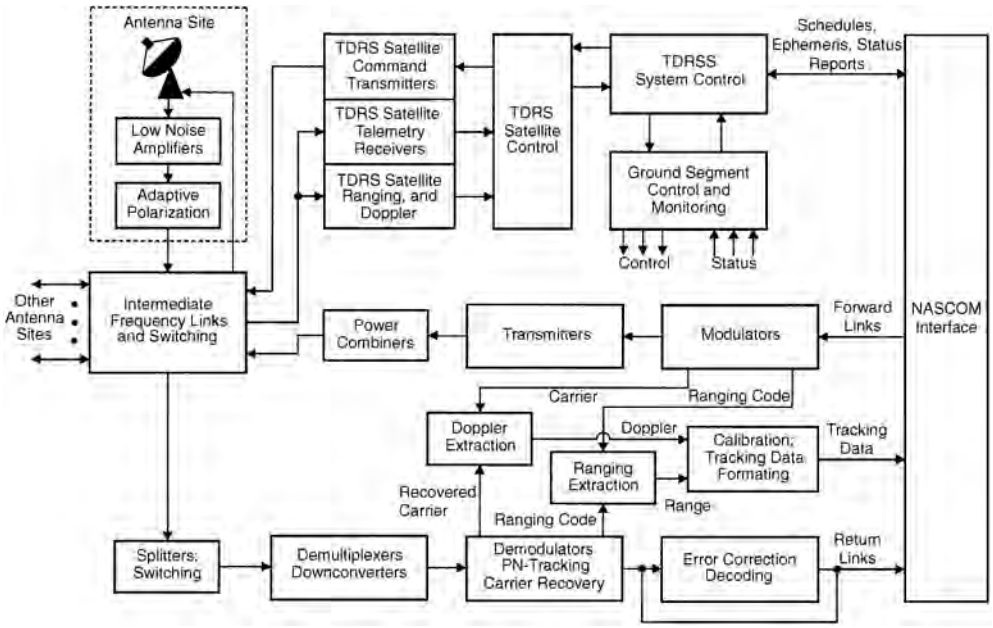


FIGURE 66.7 White Sands ground equipment block diagram.

the transfer orbit, the IUS second stage does an injection burn to insert the TDRS into geosynchronous orbit. The TDRS begins deployment of its solar arrays and antennas while still attached to the IUS to take advantage of the combined TDRS/IUS moment of inertia. After the IUS separates, the TDRS completes its deployments and initiates normal spacecraft operations. After approximately a month of on-orbit calibration and evaluation, the new TDRS is ready to begin providing relay services to user satellites.

White Sands Complex

The White Sands complex consists of two facilities, approximately 5 km apart, located on the western edge of the White Sands Missile Base in southern New Mexico. To the south is the original White Sands ground terminal (WSGT), which became operational in 1983. A new facility, called the second TDRSS ground terminal (STGT), was completed in 1995. As of 2001, WSGT has been upgraded to the STGT hardware and software configuration.

A diagram of the original White Sands ground equipment is shown in Fig. 66.7. Three 18-m Ku-band antennas provide the microwave interface with the TDRS relay satellites. The Ku-band low noise receivers are collocated with the antennas to minimize losses. Each White Sands ground terminal consists of satellite control equipment to operate the TDRS relay satellites and communications/tracking equipment to provide user services. Racks of command transmission, telemetry reception, and ranging/Doppler equipment perform the control and monitoring functions required to operate the relay satellites. Forward link user communications equipment consist of encoders, modulators, equalizers, transmitters, and multiplexers. Return link user communications equipment consists of demultiplexers, demodulators, equalizers, and decoders.

User Satellite Transponder

User satellites carry a TDRSS-compatible transponder to interface with the TDRS system over the space-to-space links. User transponders may be S-band, Ku-band, or have a combination of S-band and Ku-band capabilities. Users with mission data rate requirements below 6 Mb/s may be S-band only; users above 6 Mb/s usually have both S-band and Ku-band capabilities.

Figure 66.8 has a diagram of a typical user transponder subsystem capable of both S-band and Ku-band operations. The subsystem consists of a two-axis gimbaled antenna module and a signal processing module.

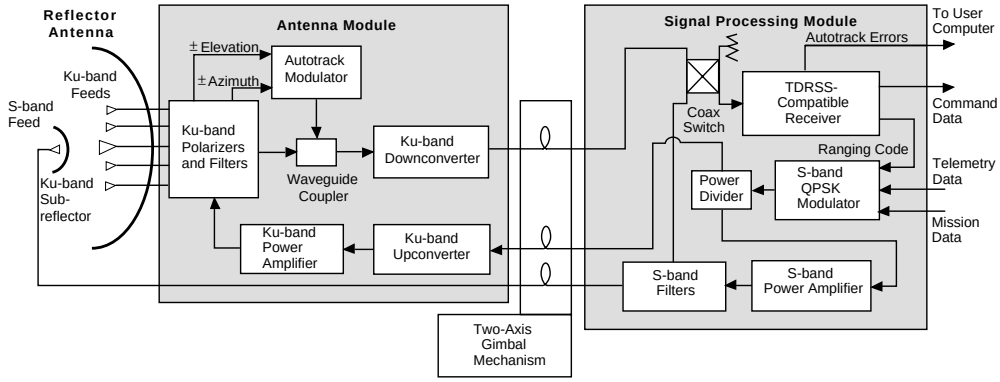


FIGURE 66.8 TDRSS user satellite transponder block diagram.

The antenna module may be mounted on a boom away from the body of the user satellite to allow an unobstructed line of sight to TDRS; the signal processing module is mounted directly to the user vehicle [Landon and Raymond, 1982].

The antenna module provides the mounting structure for a dual-frequency S/Ku-band reflector antenna, typically 1 to 2 m in diameter. The antenna compartment houses a Ku-band low-noise preamplifier/downconverter for receiving along with an upconverter and a 10 to 25 W Ku-band power amplifier for transmitting. Cable losses at S-band are low enough that amplification and frequency conversion at the antenna are not required. Most high data rate users also have Ku-band tracking feeds and autotrack electronics.

The antenna module and the signal processing modules are connected with a number of wires and coax cables. Service loops in the gimbal drive assembly allow the antenna module to be articulated for greater than hemispherical coverage.

The signal processing module contains a TDRSS-compatible spread-spectrum receiver that demodulates either the S-band forward signal or the downconverted Ku-band forward signal. The receiver recovers the command data, extracts the ranging code, and derives frequency references by phase tracking the forward link carrier. The spread-spectrum receiver also processes the autotrack error signals required for onboard autotracking.

A QPSK modulator in the signal processing module modulates the return signal with the mission data and turn-around ranging code. The return link signal goes directly to the S-band power amplifier for S-band return service or to the Ku-band upconverter and power amplifier in the antenna module for Ku-band return service.

66.5 TDRS Link Budget Examples

Appendix A of the *Space Network (SN) Users Guide* [NASA, 1995] describes the TDRS link budget methodology and contains example budgets for each user service.

On the forward links, TDRS guarantees a minimum **effective isotropic radiated power (EIRP)** for SSA, MA, and KSA service. The user satellite designer is responsible for having adequate receive sensitivity to close the link, allowing for maximum distance, polarization mismatch, and the bit error rate (BER) performance of the user satellite demodulator.

In the return direction, a minimum received power (P_{rec}) at each TDRS antenna aperture is specified in STDN 101.2 as a function of data rate to provide an end-to-end BER of 10^{-5} or better. The user satellite designer is responsible for transmitting enough power to close the link, accounting for maximum distance, polarization mismatch, and any signal distortions that exceed those allowed in the users' guide.

Table 66.3 summarizes the Gamma Ray Observatory link budget for SSA forward service at 2.1064 GHz. The waveform is 1:10 unbalanced QPSK with a PN ranging code on the I-channel and PN-spread command data on the Q-channel. A robust command channel is critical to ensuring the health and safety

TABLE 66.3 Gamma Ray Observatory SSA Forward Link Budget

Parameter	SSA Forward Service
TDRS SSA forward EIRP	43.5 dBW
Space loss from TDRS to user (44,000 km)	-191.8 dB
Polarization mismatch loss	-0.2 dB
GRO received power	-148.5 dBWi
GRO antenna gain (1.5 m, 48% efficiency)	27.6 dBi
Pointing loss (3σ error of 1.1∞)	-0.4 dB
Effective received power at GRO	-121.3 dBW
Effective thermal noise density (81,800 K)	-179.8 dBW-Hz
GRO power received-to-noise ratio (P_{rec}/N_o)	58.5 dB-Hz
Minimum P_{rec}/N_o for acquisition and ranging	39.5 dB-Hz
Acquisition and ranging link margin	19.0 dB
Mininium P_{rec}/N_o for 1-kb/s command data	43.7 dB-Hz
Command link margin (10^{-5} BER)	14.8 dB

TABLE 66.4 Gamma Ray Observatory SSA Return Link Budget

Parameter	SSA Return Service
GRO power amplifier (5.2 W)	7.2 dBW
Power losses from amplifier to antenna	-5.9 dB
Antenna gain (1.5 m, 52% efficiency)	27.9 dBi
GRO pointing loss (3s error of 1.1∞)	-0.4 dB
GRO transmitter EIRP	28.8 dBW
Polarization mismatch loss	-0.2 dB
Space loss from GRO to TDRS (44,000 km)	-192.5 dB
TDRS received power	-163.9 dBWi
I-channel power sharing (1/5)	-7.0 dB
I-channel distortion and interference losses	-0.7 dB
I-channel received power at TDRS	-171.6 dBWi
TDRS requirement (10^{-5} BER)	-167.8 dBWi
Link margin (Ranging, 32-Kb/s telemetry)	10.2 dB
Q-channel power sharing (4/5)	-1.0 dB
Q-channel distortion and interference losses	-1.7 dB
Q-channel received power at TDRS	-166.6 dBWi
TDRS requirement (10^{-5} BER)	-170.8 dBWi
Link margin (512 Kb/s mission data)	4.2 dB

of a satellite in orbit, and, as a result, GRO acquires the command signal with 19-dB margin and demodulates command data with almost 15-dB margin.

Table 66.4 shows a GRO link budget for one of several SSA return services that GRO is designed to use. The return center frequency, generated by multiplying the forward carrier frequency by 240/221, is 2.2875 GHz. The waveform is 1:4 unbalanced QPSK with ranging and 32 kb/s telemetry data modulated on the I-channel and with 512 kb/s mission data on the Q-channel. Like the command channel, telemetry is critical to spacecraft operations and is designed for greater than a 10 dB link margin. The mission data channel has a 4.2-dB margin, which ensures excellent performance while allowing for any uncertainties and unexpected variations over temperature and life.

66.6 Summary

TDRSS has established a complete space communications infrastructure, providing command, telemetry, tracking, and wideband data relay services to NASA satellites in low-Earth orbit. Using relatively simple relay satellites in geosynchronous orbits and a sophisticated ground complex at White Sands, TDRSS user satellites have access to a full complement of forward and return communication capabilities at S-band

and Ku-band. The ZOE of the original satellite constellation has been closed by deploying a TDRS at 85∞ East longitude, supported by a new facility (GRGT) on the island of Guam.

NASA has contracted with Boeing Satellite Systems, Inc. for three new TDRS satellites to allow replacement of the existing vehicles as they complete their useful lives on orbit. The new satellites, designated TDRS-H, -I, and -J, are backwards compatible with respect to the services provided by the original TDR satellites, have enhanced G/T and EIRP performance, and provide expanded user services in the Ka-band. Introduction of these next-generation relay satellites began with the successful Atlas/Centaur launch of TDRS-H in June 2000. TDRS-I and -J was launched in 2000 and 2001, respectively, allowing NASA to continue providing TDRSS communications services to the space shuttle and other user satellites well into the 21st century.

Defining Terms

Code-division multiple access: Use of families of near-orthogonal codes to allow several data streams to occupy the same frequency band and pass through a single channel.

Effective isotropic radiated power: Transmitted flux density expressed in terms of the equivalent input power to an ideal unity gain transmit antenna.

Geosynchronous orbit: A 24-h orbit, approximately 35,800 km above the Earth, in which a satellite appears, from the ground, to be almost stationary.

Ka-band: The frequency band between 18.0 and 40.0 GHz, in which several subbands are allocated for space communications and space research.

Ku-band: The frequency band between 12.5 and 18.0 GHz, in which several subbands are allocated for space communications and space research.

Pseudonoise code: Binary codes, generated using linear shift registers and sequential feedback, that appear random and have good correlation properties.

Quadrature phase-shift keying: A modulation technique in which pairs of data bits are encoded onto a carrier by rotating the phase to one of four biorthogonal positions.

S-band: The frequency band between 1.55 and 5.20 GHz, in which several subbands are allocated for space communications and space research.

Spread-spectrum: Techniques, such as modulo-2 adding a PN-code to binary communications data, that increase bandwidth and reduce the power density of the signal.

References

Agrawal, B.N. 1986. *Design of Geosynchronous Spacecraft,* Prentice–Hall, Englewood Cliffs, NJ.

Gagliardi, R.M. 1979. Phase coherency in tracking and data relay satellites. *IEEE Trans. Comm.,* COM-27(10):1527–1537.

Gagliardi, R.M. 1984. *Satellite Communications,* Lifetime Learning Pub., Belmont, CA.

Harper, W.R. and Woodson, W.L. 1981. The tracking and data relay satellite. *ISA Trans.,* 20(2):49–62.

Landon, R.B. and Raymond, H.G. 1982. Ku-band satellite communications via TDRSS, 9th AIAA Communication Satellite Systems Conference, 82–0457.

Martin, D.H. 2000. *Communication Satellites,* 4th ed., American Institute of Aeronautics.

NASA. 1995. *Space Network (SN) Users Guide,* Rev. 7, Goddard Space Flight Center, Greenbelt, MD. 530-SNUG (formerly STDN 101.2).

Poza, H.B. 1979. TDRSS telecommunications payload: an overview. *IEEE Trans. Aero. Elec. Syst.,* AES-15(3):414–429.

Further Information

TDRSS was a special topic at the 1977 National Telecommunications Conference (NTC'77), and the collection of TDRSS-related papers in the NTC'77 Conference Record (IEEE Publication 77CH12922 CSCB) provides an excellent overview of the TDRS system. TDRSS user services and the TDRS user

interfaces are fully specified in the *Space Network (SN) Users Guide*, Rev. 7 [NASA, 1995]. For a general overview of spacecraft design, including orbital, launch vehicle, and payload electronics considerations, see *Design of Geostationary Spacecraft* [Agrawal, 1986]. *Satellite Communications* [Gagliardi, 1984] is a good introductory text covering digital communications through satellite channels. *Communication Satellites,* 4th ed. [Martin 2000], contains a technical summary of a large number of military, scientific, and commercial satellites, including additional material on the new TDRS-H, -I, and -J satellites.

Acknowledgments

The author wishes to sincerely thank Paul J. Heffernan at NASA Goddard Space Flight Center for providing updated information on the tracking and data relay satellite system and for supporting the preparation of the second edition of this material.

VI

Wireless

67

Wireless Personal Communications: A Perspective

Donald C. Cox
Stanford University

67.1 Introduction

Wireless personal communications has captured the attention of the media and with it, the imagination of the public. Hardly a week goes by without one seeing an article on the subject appearing in a popular U.S. newspaper or magazine. Articles ranging from a short paragraph to many pages regularly appear in local newspapers, as well as in nationwide print media, e.g., *The Wall Street Journal, The New York Times, Business Week,* and *U.S. News and World Report*. Countless marketing surveys continue to project enormous demand, often projecting that at least half of the households, or half of the people, want wireless personal communications. Trade magazines, newsletters, conferences, and seminars on the subject by

many different names have become too numerous to keep track of, and technical journals, magazines, conferences, and symposia continue to proliferate and to have ever increasing attendance and numbers of papers presented. It is clear that wireless personal communications is, by any measure, the fastest growing segment of telecommunications. The explosive growth of wireless personal communications has continued unabated worldwide. Cellular and high-tier PCS pocketphones, pagers, and cordless telephones have become so common in many countries that few people even notice them anymore. These items have become an expected part of everyday life in most developed countries and in many developing countries around the world.

If you look carefully at the seemingly endless discussions of the topic, however, you cannot help but note that they are often describing different things, i.e., different versions of wireless personal communications [29,50]. Some discuss pagers, or messaging, or data systems, or access to the national information infrastructure, whereas others emphasize cellular radio, or cordless telephones, or dense systems of satellites. Many make reference to popular fiction entities such as Dick Tracy, Maxwell Smart, or *Star Trek*. In addition to the things noted above, the topic of wireless loops [24,30,32] has also become popular in the widespread discussions of wireless communications. As discussed in [30], this topic includes several fixed wireless applications as well as the low-tier PCS application that was discussed originally under the wireless loop designation [24,32]. The fixed wireless applications are aimed at reducing the cost of wireline loop-ends, i.e., the so-called "last mile" or "last km" of wireline telecommunications.

Thus, it appears that almost everyone wants wireless personal communications, but *what is it?* There are many different ways to segment the complex topic into different communications applications, modes, functions, extent of coverage, or mobility [29,30,50]. The complexity of the issues has resulted in considerable confusion in the industry, as evidenced by the many different wireless systems, technologies, and services being offered, planned, or proposed. Many different industry groups and regulatory entities are becoming involved. The confusion is a natural consequence of the massive dislocations that are occurring, and will continue to occur, as we progress along this large change in the paradigm of the way we communicate. Among the different changes that are occurring in our communications paradigm, perhaps the major constituent is the change from wired fixed place-to-place communications to wireless mobile person-to-person communications. Within this major change are also many other changes, e.g., an increase in the significance of data and message communications, a perception of possible changes in video applications, and changes in the regulatory and political climates. The fixed wireless loop applications noted earlier do not fit the new mobile communications paradigm. After many years of decline of fixed wireless communications applications, e.g., intercontinental HF radio and later satellites, point-to-point terrestrial microwave radio, and tropospheric scatter, it is interesting to see this rebirth of interest in fixed wireless applications. This rebirth is riding on the gigantic "wireless wave" resulting from the rapid public acceptance of mobile wireless communications. It will be interesting to observe this rebirth to see if communications history repeats; certainly mobility is wireless, but there is also considerable historical evidence that wireless is also mobility.

This chapter attempts to identify different issues and to put many of the activities in wireless into a framework that can provide perspective on what is driving them, and perhaps even to yield some indication of where they appear to be going in the future. Like any attempt to categorize many complex interrelated issues, however, there are some that do not quite fit into neat categories, and so there will remain some dangling loose ends. Like any major paradigm shift, there will continue to be considerable confusion as many entities attempt to interpret the different needs and expectations associated with the new paradigm.

67.2 Background and Issues

Mobility and Freedom from Tethers

Perhaps the clearest constituents in all of the wireless personal communications activity are the desire for mobility in communications and the companion desire to be free from tethers, i.e., from physical connections to communications networks. These desires are clear from the very rapid growth of mobile

technologies that provide primarily two-way voice services, even though economical wireline voice services are readily available. For example, cellular mobile radio has experienced rapid growth. Growth rates have been between 35 and 60% per year in the United States for a decade, with the total number of subscribers reaching 20 million by year-end 1994. The often neglected wireless companions to cellular radio, i.e., cordless telephones, have experienced even more rapid, but harder to quantify, growth with sales rates often exceeding 10 million sets a year in the United States, and with an estimated usage significantly exceeding 50 million in 1994. Telephones in airlines have also become commonplace. Similar or even greater growth in these wireless technologies has been experienced throughout the world. The explosive growth in cellular and its identical companion, high-tier PCS, has continued to about 55 million subscribers in the U.S. at year-end 1997 and a similar number worldwide. In Sweden the penetration of cellular subscribers by 1997 was over one-third of the total population, i.e., the total including every man, woman, and child! And the growth has continued since. Similar penetrations of mobile wireless services are seen in some other developed nations, e.g., Japan. The growth in users of cordless telephones also has continued to the point that they have become the dominant subscriber terminal on wireline telephone loops in the U.S. It would appear that, taking into account cordless telephones and cellular and high-tier PCS phones, half of all telephone calls in the U.S. terminate with at least one end on a wireless device.

Perhaps the most significant event in wireless personal communications since the writing of this original chapter was the widespread deployment and start of commercial service of personal handphone (PHS) in Japan in July of 1995 and its very rapid early acceptance by the consumer market [53]. By year-end 1996 there were 5 million PHS subscribers in Japan with the growth rate exceeding one-half million/month for some months. The PHS "phenomena" was one of the fastest adoptions of a new technology ever experienced. However, the PHS success story [41] peaked at a little over 7 million subscribers in 1997 and has declined slightly to a little under 7 million in mid-1998. This was the first mass deployment of a low-tier-like PCS technology (see later sections of this chapter), but PHS has some significant limitations. Perhaps the most significant limitation is the inability to successfully handoff at vehicular speeds. This handoff limitation is a result of the cumbersome radio link structure and control algorithms used to implement dynamic channel allocation (DCA) in PHS. DCA significantly increases channel occupancy (base station capacity) but incurs considerable complexity in implementing handoff. Another significant limitation of the PHS standard has been insufficient receiver sensitivity to permit "adequate" coverage from a "reasonably" dense deployment of base stations. These technology deficiencies coupled with heavy price cutting by the cellular service providers to compete with the rapid advancing of the PHS market were significant contributors to the leveling out of PHS growth. It is again evident, as with CT-2 phone point discussed in a later section, that low-tier PCS has very attractive features that can attract many subscribers, but it must also provide vehicle speed handoff and widespread coverage of highways as well as populated regions.

Others might point out the deployment and start of service of CDMA systems as a significant event since the first edition. However, the major significance of this CDMA activity is that it confirmed that CDMA performance was no better than other less-complex technologies and that those, including this author, who had been branded as "unbelieving skeptics" were correct in their assessments of the shortcomings of this technology. The overwhelming failure of CDMA technology to live up to the early claims for it can hardly be seen as a significant positive event in the evolution of wireless communication. It was, of course, a significant negative event. After years of struggling with the problems of this technology, service providers still have significantly fewer subscribers on CDMA worldwide than there are PHS subscribers in Japan alone! CDMA issues are discussed more in later sections dealing with technology issues.

Paging and associated messaging, although not providing two-way voice, do provide a form of tetherless mobile communications to many subscribers worldwide. These services have also experienced significant growth and have continued to grow since 1996. There is even a glimmer of a market in the many different specialized wireless data applications evident in the many wireless local area network (WLAN) products on the market, the several wide area data services being offered, and the specialized satellite-based message services being provided to trucks on highways. Wireless data technologies still

have many supporters, but they still have fallen far short of the rapid deployment and growth of the more voice oriented wireless technologies. However, hope appears to be eternal in the wireless data arena.

The topics discussed in the preceding two paragraphs indicate a dominant issue separating the different evolutions of wireless personal communications. That issue is the voice versus data communications issue that permeates all of communications today; this division also is very evident in fixed networks. The packet-oriented computer communications community and the circuit-oriented voice telecommunications (telephone) community hardly talk to each other and often speak different languages in addressing similar issues. Although they often converge to similar overall solutions at large scales (e.g., hierarchical routing with exceptions for embedded high-usage routes), the small-scale initial solutions are frequently quite different. Asynchronous transfer mode (ATM)–based networks are an attempt to integrate, at least partially, the needs of both the packet-data and circuit-oriented communities.

Superimposed on the voice-data issue is an issue of competing modes of communications that exist in both fixed and mobile forms. These different modes include the following.

Messaging is where the communication is not real time but is by way of message transmission, storage, and retrieval. This mode is represented by voice mail, electronic facsimile (fax), and electronic mail (e-mail), the latter of which appears to be a modern automated version of an evolution that includes telegraph and telex. Radio paging systems often provide limited one-way messaging, ranging from transmitting only the number of a calling party to longer alpha-numeric text messages.

Real-time two-way communications are represented by the telephone, cellular mobile radio telephone, and interactive text (and graphics) exchange over data networks. The two-way video phone always captures significant attention and fits into this mode; however, its benefit/cost ratio has yet to exceed a value that customers are willing to pay.

Paging, i.e., broadcast with no return channel, alerts a paged party that someone wants to communicate with him/her. Paging is like the ringer on a telephone without having the capability for completing the communications.

Agents are new high-level software applications or entities being incorporated into some computer networks. When launched into a data network, an agent is aimed at finding information by some title or characteristic and returning the information to the point from which the agent was launched. The rapid growth of the World Wide Web is based on this mode of communications.

There are still other ways in which wireless communications have been segmented in attempts to optimize a technology to satisfy the needs of some particular group. Examples include (1) user location, which can be differentiated by indoors or outdoors, or on an airplane or a train and (2) degree of mobility, which can be differentiated either by speed, e.g., vehicular, pedestrian, or stationary, or by size of area throughout which communications are provided. As noted earlier, wireless local loop with stationary terminals has become a major segment in the pursuit of wireless technology.

At this point one should again ask; wireless personal communications—*what is it*? The evidence suggests that what is being sought by users, and produced by providers, can be categorized according to the following two main characteristics.

Communications portability and mobility on many different scales:

- Within a house or building (cordless telephone, WLANs)
- Within a campus, a town, or a city (cellular radio, WLANs, wide area wireless data, radio paging, extended cordless telephone)
- Throughout a state or region (cellular radio, wide area wireless data, radio paging, satellite-based wireless)
- Throughout a large country or continent (cellular radio, paging, satellite-based wireless)
- Throughout the world?

Communications by many different modes for many different applications:

- Two-way voice
- Data

- Messaging
- Video?

Thus, it is clear why wireless personal communications today is not one technology, not one system, and not one service but encompasses many technologies, systems, and services optimized for different applications.

67.3 Evolution of Technologies, Systems, and Services

Technologies and systems [27,29,30,39,50,59,67,87], that are currently providing, or are proposed to provide, wireless communications services can be grouped into about seven relatively distinct groups, the seven previous groups are still evident in the technology but with the addition of the fixed point-to-multipoint wireless loops there are now eight, although there may be some disagreement on the group definitions, and in what group some particular technology or system belongs. All of the technologies and systems are evolving as technology advances and perceived needs change. Some trends are becoming evident in the evolutions. In this section, different groups and evolutionary trends are explored along with factors that influence the characteristics of members of the groups. The grouping is generally with respect to scale of mobility and communications applications or modes.

Cordless Telephones

Cordless telephones [29,39,50] generally can be categorized as providing low-mobility, low-power, two-way tetherless voice communications, with low mobility applying both to the range and the user's speed. Cordless telephones using analog radio technologies appeared in the late 1970s, and have experienced spectacular growth. They have evolved to digital radio technologies in the forms of second-generation cordless telephone (CT-2), and digital European cordless telephone (DECT) standards in Europe, and several different industrial scientific medical (ISM) band technologies in the United States.[1]

Personal handyphone (PHS) noted earlier and discussed in later sections and inserts can be considered either as a quite advanced digital cordless telephone similar to DECT or as a somewhat limited low-tier PCS technology. It has most of the attributes of similarity of the digital cordless telephones listed later in this section except that PHS uses $\pi/4$ QPSK modulation.

Cordless telephones were originally aimed at providing economical, tetherless voice communications inside residences, i.e., at using a short wireless link to replace the cord between a telephone base unit and its handset. The most significant considerations in design compromises made for these technologies are to minimize total cost, while maximizing the talk time away from the battery charger. For digital cordless phones intended to be carried away from home in a pocket, e.g., CT-2 or DECT, handset weight and size are also major factors. These considerations drive designs toward minimizing complexity and minimizing the power used for signal processing and for transmitting.

Cordless telephones compete with wireline telephones. Therefore, high circuit quality has become a requirement. Early cordless sets had marginal quality. They were purchased by the millions, and discarded by the millions, until manufacturers produced higher-quality sets. Cordless telephones sales then exploded. Their usage has become commonplace, approaching, and perhaps exceeding, usage of corded telephones.

The compromises accepted in cordless telephone design in order to meet the cost, weight, and talk-time objectives are the following.

- Few users per megahertz
- Few users per base unit (many link together a particular handset and base unit)

[1] These ISM technologies either use spread spectrum techniques (direct sequence or frequency hopping) or very low-transmitter power (<~1 mW) as required by the ISM band regulations.

- Large number of base units per unit area; one or more base units per wireline access line (in high-rise apartment buildings the density of base units is very large)
- Short transmission range

There is no added network complexity since a base unit looks to a telephone network like a wireline telephone. These issues are also discussed in [29,50].

Digital cordless telephones in Europe have been evolving for a few years to extend their domain of use beyond the limits of inside residences. Cordless telephone, second generation, (CT-2) has evolved to provide telepoint or phone-point services. Base units are located in places where people congregate. e.g., along city streets and in shopping malls, train stations, etc. Handsets registered with the phone-point provider can place calls when within range of a telepoint. CT-2 does not provide capability for transferring (handing off) active wireless calls from one phone point to another if a user moves out of range of the one to which the call was initiated. A CT-2+ technology, evolved from CT-2 and providing limited handoff capability, is being deployed in Canada. CT-2+ deployment was never completed. Phone-point service was introduced in the United Kingdom twice, but failed to attract enough customers to become a viable service. In Singapore and Hong Kong, however, CT-2 phone point has grown rapidly, reaching over 150,000 subscribers in Hong Kong [75] in mid-1994. The reasons for success in some places and failure in others are still being debated, but it is clear that the compactness of the Hong Kong and Singapore populations make the service more widely available, using fewer base stations than in more spreadout cities. Complaints of CT-2 phone-point users in trials have been that the radio coverage was not complete enough, and/or they could not tell whether there was coverage at a particular place, and the lack of handoff was inconvenient. In order to provide the alerting or ringing function for phone-point service, conventional radio pagers have been built into some CT-2 handsets. (The telephone network to which a CT-2 phone point is attached has no way of knowing from which base units to send a ringing message, even though the CT-2 handsets can be rung from a home base unit). CT-2 phone points in Hong Kong peaked at about 170,000 subscribers. There was then a precipitous decline as these subscribers abandoned the limited service CT-2 in favor of high-tier PCS and cellular services which had reduced their prices to compete with phone points. Phone points have now been removed from Hong Kong. CT-2 phone points were also deployed in Korea, again with initial success followed by decline. The message is clear from the CT-2 and PHS experiences. The attributes of cordless phone-like low-tier PCS are very attractive, but need widespread coverage and vehicle speed handoff in order to be long term attractive and viable in the market.

Another European evolution of cordless telephones is DECT, which was optimized for use inside buildings. Base units are attached through a controller to private branch exchanges (PBXs), key telephone systems, or phone company CENTREX telephone lines. DECT controllers can hand off active calls from one base unit to another as users move, and can page or ring handsets as a user walks through areas covered by different base units. DECT has increased in deployment in some countries, notably Italy, for several applications including cordless telephone and wireless loop. There are reported to be perhaps 7 million or more DECT users in Europe.

These cordless telephone evolutions to more widespread usage outside and inside with telepoints and to usage inside large buildings are illustrated in Fig. 67.1, along with the integration of paging into handsets to provide alerting for phone-point services. They represent the first attempts to increase the service area of mobility for low-power cordless telephones.

Some of the characteristics of the digital cordless telephone technologies, CT-2 and DECT, are listed in Table 67.1. Additional information can be found in [32,50]. Even though there are significant differences between these technologies, e.g., multiple access technology [frequency division multiple access (FDMA) or time division multiple access (TDMA)/FDMA], and channel bit rate, there are many similarities that are fundamental to the design objectives discussed earlier and to a user's perception of them. These similarities and their implications are as follows.

32-kb/s Adaptive Differential Pulse Code Modulation (ADPCM) Digital Speech Encoding: This is a low-complexity (low-signal processing power) speech encoding process that provides wireline speech quality and is an international standard.

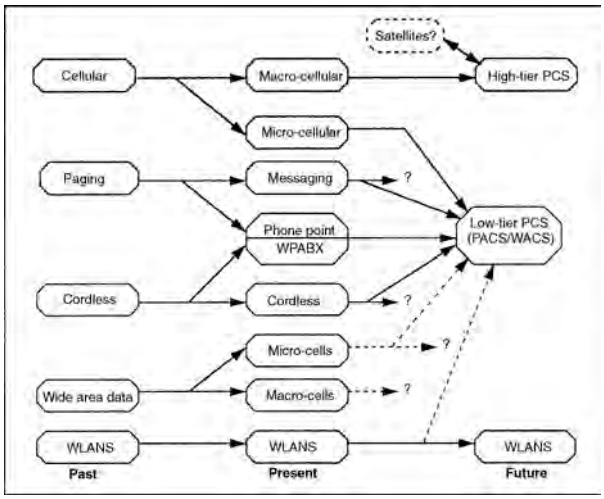


FIGURE 67.1

Average Transmitter Power ≤10 mW: This permits many hours of talk time with small, low-cost, lightweight batteries, but provides limited radio range.

Low-Complexity Radio Signal Processing: There is no forward error correction and no complex multipath mitigation (i.e., no equalization or spread spectrum).

Low Transmission Delay, e.g., <50 ms, and for CT-2 <10-ms Round Trip: This is a speech-quality and network-complexity issue. A maximum of 10 ms should be allowed, taking into account additional inevitable delay in long-distance networks. Echo cancellation is generally required for delays >10 ms.

Simple Frequency-Shift Modulation and Noncoherent Detection: Although still being low in complexity, the slightly more complex 4QPSK modulation with coherent detection provides significantly more spectrum efficiency, range, and interference immunity.

Dynamic Channel Allocation: Although this technique has potential for improved system capacity, the cordless-telephone implementations do not take full advantage of this feature for handoff and, thus, cannot reap the full benefit for moving users [15,19].

Time Division Duplex (TDD): This technique permits the use of a single contiguous frequency band and implementation of diversity from one end of a radio link. Unless all base station transmissions are synchronized in time, however, it can incur severe cochannel interference penalties in outside environments [15,16]. Of course, for cordless telephones used inside with base stations not having a propagation advantage, this is not a problem. Also, for small indoor PBX networks, synchronization of base station transmission is easier than is synchronization throughout a widespread outdoor network, which can have many adjacent base stations connected to different geographic locations for central control and switching.

Cellular Mobile Radio Systems

Cellular mobile radio systems are becoming known in the United States as high-tier personal communications service (PCS), particularly when implemented in the new 1.9-GHz PCS bands [20]. These systems generally can be categorized as providing high-mobility, wide-ranging, two-way tetherless voice communications. In these systems, high mobility refers to vehicular speeds, and also to widespread regional to nationwide coverage [27,29,50]. Mobile radio has been evolving for over 50 years. Cellular radio integrates wireless access with large-scale networks having sophisticated intelligence to manage mobility of users.

Cellular radio was designed to provide voice service to wide-ranging vehicles on streets and highways [29,39,50,82], and generally uses transmitter power on the order of 100 times that of cordless telephones

TABLE 67.1 Wireless PCS Technologies

	High-Power Systems				Low-Power Systems			
	Digital Cellular (High-Tier PCS)				Low-Tier PCS		Digital Cordless	
System	IS-54	IS-95 (DS)	GSM	DCS-1800	WACS/PACS	Handi-Phone	DECT	CT-2
Multiple Access	TDMA/FDMA	CDMA/FDMA	TDMA/FDMA	TDMA/FDMA	TDMA/FDMA	TDMA/FDMA	TDMA/FDMA	FDMA
Freq. band, MHz						1895–1907	1880–1990	864–868
Uplink, MHz	869–894	869–894	935–960	1710–1785	Emerg. Tech.*			
Downlink, MHz	824–849	824–849	890–915	1805–1880				
	(USA)	(USA)	(Eur.)	(UK)	(USA)	(Japan)	(Eur.)	(Eur. and Asia)
RF ch. spacing						300	1728	100
Downlink, KHz	30	1250	200	200	300			
Uplink, KHz	30	1250	200	200	300			
Modulation	$\pi/4$ DQPSK	BPSK/QPSK	GMSK	GMSK	$\pi/4$ QPSK	$\pi/4$ DQPSK	GFSK	GFSK
Portable txmit Power, max./avg.	600 mW/200 mW	600 mW	1 W/125 mW	1 W/125 mW	200 mW/25 mW	80 mW/10 mW	250 mW/10 mW	10 mW/5 mW
Speech coding	VSELP	QCELP	RPE-LTP	RPE-LTP	ADPCM	ADPCM	ADPCM	ADPCM
Speech rate, kb/s	7.95	8 (var.)	13	13	32/16/8	32	32	32
Speech ch./RF ch.	3	—	8	8	8/16/32	4	12	1
Ch. Bit rate, kb/s						384	1152	72
Uplink, kb/s	48.6		270.833	270.833	384			
Downlink, kb/s	48.6		270.833	270.833	384			
Ch. coding	1/2 rate conv.	1/2 rate fwd. 1/3 rate rev.	1/2 rate conv.	1/2 rate conv.	CRC	CRC	CRC (control)	None
Frame, ms	40	20	4.615	4.615	2.5	5	10	2

*Spectrum is 1.85–2.2 GHz allocated by the FCC for emerging technologies; DS is direct sequence.

(≈2 W for cellular). Thus, cellular systems can only provide reduced service to handheld sets that are disadvantaged by using somewhat lower transmitter power (<0.5 W) and less efficient antennas than vehicular sets. Handheld sets used inside buildings have the further disadvantage of attenuation through walls that is not taken into account in system design.

Cellular radio or high-tier PCS has experienced large growth as noted earlier. In spite of the limitations on usage of handheld sets already noted, handheld cellular sets have become very popular, with their sales becoming comparable to the sales of vehicular sets. Frequent complaints from handheld cellular users are that batteries are too large and heavy, and both talk time and standby time are inadequate. Cellular and high-tier PCS pocket handsets have continued to decrease in size and weight and more efficient lithium batteries have been incorporated. This has increased their attractiveness (more on this in the later section, Reality Check). For several years there have been many more pocket handsets sold than vehicular mounted sets every year. However, despite the improvements in these handsets and batteries, the complaints of weight and limited talk time still persist. The electronics have become essentially weightless compared to the batteries required for these high-tier PCS and cellular handsets.

Cellular radio at 800 MHz has evolved to digital radio technologies [29,39,50] in the forms of the deployed systems standards

- Global Standard for Mobile (GSM) in Europe
- Japanese or personal digital cellular (JDC or PDC) in Japan
- U.S. TDMA digital cellular known as USDC or IS-54

and in the form of the code division multiple access (CDMA) standard, IS-95, which is under development but not yet deployed. Since the first edition was published, CDMA systems have been deployed in the U.S., Korea, Hong Kong, and other countries after many months (years) of redesign, reprogramming, and adjustment. These CDMA issues are discussed later in the section "New Technology."

The most significant consideration in the design compromises made for the U.S. digital cellular or high-tier PCS systems was the high cost of cell sites (base stations). A figure often quoted is U.S. $1 million for a cell site. This consideration drove digital system designs to maximize users per megahertz and to maximize the users per cell site.

Because of the need to cover highways running through low-population-density regions between cities, the relatively high transmitter power requirement was retained to provide maximum range from high antenna locations.

Compromises that were accepted while maximizing the two just cited parameters are as follows.

- High transmitter power consumption.
- High user-set complexity, and thus high signal-processing power consumption.
- Low circuit quality.
- High network complexity, e.g., the new IS-95 technology will require complex new switching and control equipment in the network, as well as high-complexity wireless-access technology.

Cellular radio or high-tier PCS has also been evolving for a few years in a different direction, toward very small coverage areas or microcells. This evolution provides increased capacity in areas having high user density, as well as improved coverage of shadowed areas. Some microcell base stations are being installed inside, in conference center lobbies and similar places of high user concentrations. Of course, microcells also permit lower transmitter power that conserves battery power when power control is implemented, and base stations inside buildings circumvent the outside wall attenuation. Low-complexity microcell base stations also are considerably less expensive than conventional cell sites, perhaps two orders of magnitude less expensive. Thus, the use of microcell base stations provides large increases in overall system capacity, while also reducing the cost per available radio channel and the battery drain on portable subscriber equipment. This microcell evolution, illustrated in Fig. 67.1, moves handheld cellular sets in a direction similar to that of the expanded-coverage evolution of cordless telephones to phone points and wireless PBX.

Some of the characteristics of digital-cellular or high-tier PCS technologies are listed in Table 67.1 for IS-54, IS-95, and GSM at 900 MHz, and DCS-1800, which is GSM at 1800 MHz. The technology listed here as IS-54 has also become known as IS-136 having more sophisticated digital control channels. These technologies, IS-54/IS-136 are also sometimes known as DAMPS (i.e., Digital AMPS), as U.S. TDMA or North American TDMA, or sometimes just as "TDMA." Additional information can be found in [29,39,50]. The JDC or PDC technology, not listed, is similar to IS-54. As with the digital cordless technologies, there are significant differences among these cellular technologies, e.g., modulation type, multiple access technology, and channel bit rate. There are also many similarities, however, that are fundamental to the design objectives discussed earlier. These similarities and their implications are as follows.

Low Bit-Rate Speech Coding ≤13 kb/s with Some ≤8 kb/s: Low bit-rate speech coding obviously increases the number of users per megahertz and per cell site. However, it also significantly reduces speech quality [29], and does not permit speech encodings in tandem while traversing a network; see also the section on Other Issues later in this chapter.

Some Implementations Make Use of Speech Inactivity: This further increases the number of users per cell site, i.e., the cell-site capacity. It also further reduces speech quality [29], however, because of the difficulty of detecting the onset of speech. This problem is even worse in an acoustically noisy environment like an automobile.

High Transmission Delay; ≈200-ms Round Trip: This is another important circuit-quality issue. Such large delay is about the same as one-way transmission through a synchronous-orbit communications satellite. A voice circuit with digital cellular technology on both ends will experience the delay of a full satellite circuit. It should be recalled that one reason long-distance circuits have been removed from satellites and put onto fiber-optic cable is because customers find the delay to be objectionable. This delay in digital cellular technology results from both computation for speech bit-rate reduction and from complex signal processing, e.g., bit interleaving, error correction decoding, and multipath mitigation [equalization or spread spectrum code division multiple access (CDMA)].

High-Complexity Signal Processing, Both for Speech Encoding and for Demodulation: Signal processing has been allowed to grow without bound and is about a factor of 10 greater than that used in the low-complexity digital cordless telephones [29]. Since several watts are required from a battery to produce the high transmitter power in a cellular or high-tier PCS set, signal-processing power is not as significant as it is in the low-power cordless telephones; see also the section on Complexity/Coverage Area Comparisons later in this chapter.

Fixed Channel Allocation: The difficulties associated with implementing capacity-increasing dynamic channel allocation to work with handoff [15,19] have impeded its adoption in systems requiring reliable and frequent handoff.

Frequency Division Duplex (FDD): Cellular systems have already been allocated paired-frequency bands suitable for FDD. Thus, the network or system complexity required for providing synchronized transmissions [15,16] from all cell sites for TDD has not been embraced in these digital cellular systems. Note that TDD has not been employed in IS-95 even though such synchronization is required for other reasons.

Mobile/Portable Set Power Control: The benefits of increased capacity from lower overall cochannel interference and reduced battery drain have been sought by incorporating power control in the digital cellular technologies.

Wide-Area Wireless Data Systems

Existing wide-area data systems generally can be categorized as providing high mobility, wide-ranging, low-data-rate digital data communications to both vehicles and pedestrians [29,50]. These systems have not experienced the rapid growth that the two-way voice technologies have, even though they have been deployed in many cities for a few years and have established a base of customers in several countries. Examples of these packet data systems are shown in Table 67.2.

TABLE 67.2 Wide-Area Wireless Packet Data Systems

	CDPD[1]	RAM Mobile (Mobitex)	ARDIS[2] (KDT)	Metricom (MDN)[3]
Data rate, kb/s	19.2	8 (19.2)	4.8 (19.2)	76
Modulation	GMSK BT = 0.5	GMSK	GMSK	GMSK
Frequency, MHz	800	900	800	915
Chan. spacing, kHz	30	12.5	25	160
Status	1994 service	Full service	Full service	In service
Access means	Unused AMPS channels	Slotted Aloha CSMA		FH SS (ISM)
Transmit power, W			40	1

Note: Data in parentheses ( ) indicates proposed.
[1]Cellular Digital Packet Data
[2]Advanced Radio Data Information Service
[3]Microcellular Data Network

The earliest and best known of these systems in the United States are the ARDIS network developed and run by Motorola, and the RAM mobile data network based on Ericsson Mobitex Technology. These technologies were designed to make use of standard, two-way voice, land mobile-radio channels, with 12.5- or 25-kHz channel spacing. In the United States these are specialized mobile radio services (SMRS) allocations around 450 MHz and 900 MHz. Initially, the data rates were low: 4.8 kb/s for ARDIS and 8 kb/s for RAM. The systems use high transmitter power (several tens of watts) to cover large regions from a few base stations having high antennas. The relatively low data capacity of a relatively expensive base station has resulted in economics that have not favored rapid growth.

The wide-area mobile data systems also are evolving in several different directions in an attempt to improve base station capacity, economics, and the attractiveness of the service. The technologies used in both the ARDIS and RAM networks are evolving to higher channel bit rates of 19.2 kb/s.

The cellular carriers and several manufacturers in the United States are developing and deploying a new wide-area packet data network as an overlay to the cellular radio networks. This cellular digital packet data (CDPD) technology shares the 30-kHz spaced 800-MHz voice channels used by the analog FM advanced mobile phone service (AMPS) systems. Data rate is 19.2 kb/s. The CDPD base station equipment also shares cell sites with the voice cellular radio system. The aim is to reduce the cost of providing packet data service by sharing the costs of base stations with the better established and higher cell-site capacity cellular systems. This is a strategy similar to that used by nationwide fixed wireline packet data networks that could not provide an economically viable data service if they did not share costs by leasing a small amount of the capacity of the interexchange networks that are paid for largely by voice traffic. CDPD has been deployed in many U.S. cities for several years. However, it has not lived up to early expectations and has become "just another" wireless data service with some subscribers, but not with the large growth envisioned earlier.

Another evolutionary path in wide-area wireless packet data networks is toward smaller coverage areas or microcells. This evolutionary path also is indicated in Fig. 67.1. The microcell data networks are aimed at stationary or low-speed users. The design compromises are aimed at reducing service costs by making very small and inexpensive base stations that can be attached to utility poles, the sides of buildings and inside buildings and can be widely distributed throughout a region. Base-station-to-base-station wireless links are used to reduce the cost of the interconnecting data network. In one network this decreases the overall capacity to serve users, since it uses the same radio channels that are used to provide service. Capacity is expected to be made up by increasing the number of base stations that have connections to a fixed-distribution network as service demand increases. Another such network uses other dedicated radio channels to interconnect base stations. In the high-capacity limit, these networks will look more like a conventional cellular network architecture, with closely spaced, small, inexpensive base stations, i.e., microcells, connected to a fixed infrastructure. Specialized wireless data networks have been built to provide metering and control of electric power distributions, e.g., Celldata and Metricom in California.

A large microcell network of small inexpensive base stations has been installed in the lower San Francisco Bay Area by Metricom, and public packet-data service was offered during early 1994. Most of the small (shoe-box size) base stations are mounted on street light poles. Reliable data rates are about 75 kb/s. The technology is based on slow frequency-hopped spread spectrum in the 902–928 MHz U.S. ISM band. Transmitter power is 1 W maximum, and power control is used to minimize interference and maximize battery life time. The metricom network has been improved and significantly expanded in the San Francisco Bay Area and has been deployed in Washington, D.C., and a few other places in the U.S. However, like all wireless data services so far, it has failed to grow as rapidly or to attract as many subscribers as was originally expected. Wireless data overall has had only very limited success compared to that of the more voice-oriented technologies, systems, and services.

High-Speed Wireless Local-Area Networks (WLANs)

Wireless local-area data networks can be categorized as providing low-mobility high-data-rate data communications within a confined region, e.g., a campus or a large building. Coverage range from a wireless data terminal is short, tens to hundreds of feet, like cordless telephones. Coverage is limited to within a room or to several rooms in a building. WLANs have been evolving for a few years, but overall the situation is chaotic, with many different products being offered by many different vendors [29,59]. There is no stable definition of the needs or design objectives for WLANs, with data rates ranging from hundreds of kb/s to more than 10 Mb/s, and with several products providing one or two Mb/s wireless link rates. The best description of the WLAN evolutionary process is: having severe birth pains. An IEEE standards committee, 802.11, has been attempting to put some order into this topic, but their success has been somewhat limited. A partial list of some advertised products is given in Table 67.3. Users of WLANs are not nearly as numerous as the users of more voice-oriented wireless systems. Part of the difficulty stems from these systems being driven by the computer industry that views the wireless system as just another plug-in interface card, without giving sufficient consideration to the vagaries and needs of a reliable radio system. This section still describes the WLAN situation in spite of some attempts at standards in the U.S. and Europe, and continuing industry efforts. Some of the products in Table 67.3 have been discontinued because of lack of market and some new products have been offered, but the manufacturers still continue to struggle to find enough customers to support their efforts. Optimism remains high in the WLAN community that "eventually" they will find the "right" technology, service, or application to make WLANs "take off"—but the world still waits. Success is still quite limited.

There are two overall network architectures pursued by WLAN designers. One is a centrally coordinated and controlled network that resembles other wireless systems. There are base stations in these networks that exercise overall control over channel access [44].

The other type of network architecture is the self-organizing and distributed controlled network where every terminal has the same function as every other terminal, and networks are formed ad hoc by communications exchanges among terminals. Such ad hoc networks are more like citizen band (CB) radio networks, with similar expected limitations if they were ever to become very widespread.

Nearly all WLANs in the United States have attempted to use one of the ISM frequency bands for unlicensed operation under part 15 of the FCC rules. These bands are 902–928 MHz, 2400–2483.5 MHz, and 5725–5850 MHz, and they require users to accept interference from any interfering source that may also be using the frequency. The use of ISM bands has further handicapped WLAN development because of the requirement for use of either frequency hopping or direct sequence spread spectrum as an access technology, if transmitter power is to be adequate to cover more than a few feet. One exception to the ISM band implementations is the Motorola ALTAIR, which operates in a licensed band at 18 GHz. It appears that ALTAIR has been discontinued because of the limited market. The technical and economic challenges of operation at 18 GHz have hampered the adoption of this 10–15 Mb/s technology. The frequency-spectrum constraints have been improved in the United States with the recent FCC allocation of spectrum from 1910–1930 MHz for unlicensed data PCS applications. Use of this new spectrum requires implementation of an access etiquette incorporating listen before transmit in an attempt to

provide some coordination of an otherwise potentially chaotic, uncontrolled environment [68]. Also, since spread spectrum is not a requirement, access technologies and multipath mitigation techniques more compatible with the needs of packet-data transmission [59], e.g., multipath equalization or multicarrier transmission, can be incorporated into new WLAN designs. The FCC is allocating spectrum at 5 GHz for wideband wireless data for internet and next generation data network access, BUT it remains to be seen whether this initiative is any more successful than past wireless data attempts. Optimism is again high, BUT…

Three other widely different WLAN activities also need mentioning. One is a large European Telecommunications Standards Institute (ETSI) activity to produce a standard for high performance radio local area network (HIPERLAN), a 20-Mb/s WLAN technology to operate near 5 GHz. Other activities are large U.S. Advance Research Projects Agency (ARPA)–sponsored WLAN research projects at the Universities of California at Berkeley (UCB) and at Los Angeles (UCLA). The UCB Infopad project is based on a coordinated network architecture with fixed coordinating nodes and direct-sequence spread spectrum (CDMA), whereas, the UCLA project is aimed at peer-to-peer networks and uses frequency hopping. Both ARPA-sponsored projects are concentrated on the 900-MHz ISM band.

As computers shrink in size from desktop to laptop to palmtop, mobility in data network access is becoming more important to the user. This fact, coupled with the availability of more usable frequency spectrum, and perhaps some progress on standards, may speed the evolution and adoption of wireless mobile access to WLANs. From the large number of companies making products, it is obvious that many believe in the future of this market. It should be noted that the objective for 10 MB/s data service with widespread coverage from a sparse distribution of widely separated base stations equivalent to cellular is unrealistic and unrealizable. This can be readily seen by considering a simple example. Consider a cellular coverage area that requires full cellular power of 0.5 watt to cover from a handset. Consider the handset to use a typical digital cellular bit rate of about 10 kb/s (perhaps 8 kb/s speech coding + overhead). With all else in the system the same, e.g., antennas, antenna height, receiver noise figure, detection sensitivity, etc., the 10 MB/s data would require 10 MB/s $\prod$ 10 kb/s = 1000 times as much power as the 10 kb/s cellular. Thus, it would require $0.5 \times 1000 = 500$ watts for the wireless data transmitter. This is a totally unrealistic situation. If the data system operates at a higher frequency (e.g., 5 GHz) than the cellular system (e.g., 1 or 2 GHz) then there will be even more power required to overcome the additional loss at a higher frequency. The sometimes expressed desire by the wireless data community for a system to provide network access to users in and around buildings and to provide 10 MB/s over 10 miles with 10 milliwatts of transmitter power and costing $10.00 is totally impossible. It requires violation of the "laws of physics."

Paging/Messaging Systems

Radio paging began many years ago as a one-bit messaging system. The one bit was: someone wants to communicate with you. More generally, paging can be categorized as one-way messaging over wide areas. The one-way radio link is optimized to take advantage of the asymmetry. High transmitter power (hundreds of watts to kilowatts), and high antennas at the fixed base stations permit low-complexity, very low-power consumption, pocket paging receivers that provide long usage time from small batteries. This combination provides the large radio-link margins needed to penetrate walls of buildings without burdening the user set battery. Paging has experienced steady rapid growth for many years and serves about 15 million subscribers in the United States.

Paging also has evolved in several different directions. It has changed from analog tone coding for user identification to digitally encoded messages. It has evolved from the 1-b message, someone wants you, to multibit messages from, first, the calling party's telephone number to, now, short e-mail text messages. This evolution is noted in Fig. 67.1.

The region over which a page is transmitted has also increased from (1) local, around one transmitting antenna; to (2) regional, from multiple widely-dispersed antennas; to (3) nationwide, from large networks of interconnected paging transmitters. The integration of paging with CT-2 user sets for phone-point call alerting was noted previously.

TABLE 67.3 Partial List of WLAN Products

Product Company Location	Freq., MHz	Link Rate, Mb/s	User Rate	Protocol(s)	Access	No. of chan. or Spread Factor	Mod./ Coding	Power, mW	Network Topology
Altair Plus Motorola Arlington Hts, IL	18–19 GHz	15	5.7 Mb/s	Ethernet			4-level FSK	25 peak	Eight devices/radio; radio to base to ethernet
WaveLAN NCR/AT&T Dayton, OH	902–928	2	1.6 Mb/s	Ethernet-like	DS SS		DQPSK	250	Peer-to-peer
AirLan Solectek San Diego, CA	902–928		2 Mb/s	Ethernet	DS SS		DQPSK	250	PCMCIA w/ant.; radio to hub
Freeport Windata Inc. Northboro, MA	902–928	16	5.7 Mb/s	Ethernet	DS SS	32 chips/bit	16 PSK trellis coding	650	Hub
Intersect Persoft Inc. Madison, WI	902–928		2 Mb/s	Ethernet token ring	DS SS		DQPSK	250	Hub
LAWN O'Neill Comm. Horsham, PA	902–928		38.4 kb/s	AX.25	SS	20 users/chan.; max. 4 chan.		20	Peer-to-peer

WILAN Wi-LAN Inc. Calgary, Alberta	902–928	20	1.5 Mb/s/ chan.	Ethernet, token ring	CDMA/ TDMA	3 chan. 10–15 links each	unconven- tional	30	Peer-to-peer
RadioPort ALPS Electric USA	902–928		242 kb/s	Ethernet	SS	7/3 channels		100	Peer-to-peer
ArLAN 600 Telesys. SLW Don Mills, Ont.	902–928; 2.4 GHz		1.35 Mb/s	Ethernet	SS			1 W max	PCs with ant.; radio to hub
Radio Link Cal. Microwave Sunnyvale, CA	902–928; 2.4 GHz	250 kb/s	64 kb/s		FH SS	250 ms/hop 500 kHz space			Hub
Range LAN Proxim, Inc. Mountain View, CA	902–928		242 kb/s	Ethernet, token ring	DS SS	3 chan.		100	
RangeLAN 2 Proxim, Inc. Mountain View, CA	2.4 GHz	1.6	50 kb/s max.	Ethernet, token ring	FH SS	10 chan. at 5 kb/s; 15 sub-ch. each		100	Peer-to-peer bridge
Netwave Xircom Calabasas, CA	2.4 GHz	1/adaptor		Ethernet, token ring	FH SS	82 l-MHz chn. or “hops”			Hub
Freelink Cabletron Sys. Rochester, NH	2.4 and 5.8 GHz		5.7 Mb/s	Ethernet	DS SS	32 chips/bit	16 PSK trellis coding	100	Hub

Another evolutionary paging route sometimes proposed is two-way paging. This is an ambiguous and unrealizable concept, however, since the requirement for two-way communications destroys the asymmetrical link advantage so well exploited by paging. Two-way paging puts a transmitter in the user's set and brings along with it all of the design compromises that must be faced in such a two-way radio system. Thus, the word paging is not appropriate to describe a system that provides two-way communications. The two-way paging situation is as unrealistic as that noted earlier for wide-area, high-speed, low-power wireless data. This can be seen by looking at the asymmetry situation in paging. In order to achieve comparable coverage uplink and downlink, a 500-watt paging transmitter downlink advantage must be overcome in the uplink. Even considering the relatively high cellular handset transmit power levels on the order of 0.5 watt results in a factor of 1000 disadvantage, and 0.5 watt is completely incompatible with the low power paging receiver power requirements. If the same uplink and downlink coverage is required for an equivalent set of system parameters, then the only variable left to work with is bandwidth. If the paging link bit rate is taken to be 10 kb/sec (much higher than many paging systems), then the usable uplink rate is 10 kb/s/1000 = 10 B/s, an unusably low rate. Of course, some uplink benefit can be gained because of better base station receiver noise figure and by using forward error correction and perhaps ARQ. However, this is unlikely to raise the allowable rate to greater than 100 B/s which even though likely overoptimistic is still unrealistically low and we have assumed an unrealistically high transmit power in the two-way "pager!"

Satellite-Based Mobile Systems

Satellite-based mobile systems are the epitome of wide-area coverage, expensive base station systems. They generally can be categorized as providing two-way (or one-way) limited quality voice and/or very limited data or messaging to very wide-ranging vehicles (or fixed locations). These systems can provide very widespread, often global, coverage, e.g., to ships at sea by INMARSAT. There are a few messaging systems in operation, e.g., to trucks on highways in the United States by Qualcomm's Omnitracs system.

A few large-scale mobile satellite systems have been proposed and are being pursued: perhaps the best known is Motorola's Iridium; others include Odyssey, Globalstar, and Teledesic. The strength of satellite systems is their ability to provide large regional or global coverage to users outside buildings. However, it is very difficult to provide adequate link margin to cover inside buildings, or even to cover locations shadowed by buildings, trees, or mountains. A satellite system's weakness is also its large coverage area. It is very difficult to provide from Earth orbit the small coverage cells that are necessary for providing high overall systems capacity from frequency reuse. This fact, coupled with the high cost of the orbital base stations, results in low capacity along with the wide overall coverage but also in expensive service. Thus, satellite systems are not likely to compete favorably with terrestrial systems in populated areas or even along well-traveled highways. They can complement terrestrial cellular or PCS systems in low-population-density areas. It remains to be seen whether there will be enough users with enough money in low-population-density regions of the world to make satellite mobile systems economically viable. Some of the mobile satellite systems have been withdrawn, e.g., Odyssey. Some satellites in the Iridium and Globalstar systems have been launched. The industry will soon find out whether these systems are economically viable.

Proposed satellite systems range from (1) low-Earth orbit systems (LEOS) having tens to hundreds of satellites through (2) intermediate- or medium-height systems (MEOS) to (3) geostationary or geosynchronous orbit systems (GEOS) having fewer than ten satellites. LEOS require more, but less expensive, satellites to cover the Earth, but they can more easily produce smaller coverage areas and, thus, provide higher capacity within a given spectrum allocation. Also, their transmission delay is significantly less (perhaps two orders of magnitude!), providing higher quality voice links, as discussed previously. On the other hand, GEOS require only a few, somewhat more expensive, satellites (perhaps only three) and are likely to provide lower capacity within a given spectrum allocation and suffer severe transmission-delay impairment on the order of 0.5 s. Of course, MEOS fall in between these extremes. The possible evolution of satellite systems to complement high-tier PCS is indicated in Fig. 67.1.

Fixed Point-to-Multipoint Wireless Loops

Wideband point-to-multipoint wireless loop technologies sometimes have been referred to earlier as "wireless cable" when they were proposed as an approach for providing interactive video services to homes [30]. However, as the video application started to appear less attractive, the application emphasis shifted to providing wideband data access for the internet, the worldwide web, and future wideband data networks. Potentially lower costs are the motivation for this wireless application. As such, these technologies will have to compete with existing coaxial cable and fiber/coax distribution by CATV companies, with satellites, and with fiber and fiber/coax systems being installed or proposed by telephone companies and other entities [30]. Another competitor is asymmetric digital subscriber line technology, which uses advanced digital signal processing to provide high-bandwidth digital distribution over twisted copper wire pairs.

In the U.S. two widely different frequency bands are being pursued for fixed point-to-multipoint wireless loops. These bands are at 28 GHz for local multipoint distribution systems or services (LMDS) [52] and 2.5 to 2.7 GHz for microwave or metropolitan distribution systems (MMDS) [74]. The goal of low-cost fixed wireless loops is based on the low cost of point-to-multipoint line-of-sight wireless technology. However, significant challenges are presented by the inevitable blockage by trees, terrain, and houses, and by buildings in heavily built-up residential areas. Attenuation in rainstorms presents an additional problem at 28 GHz in some localities. Even at the 2.5-GHz MMDS frequencies, the large bandwidth required for distribution of many video channels presents a challenge to provide adequate radio-link margin over obstructed paths. From mobile satellite investigations it is known that trees can often produce over 15 dB additional path attenuation [38]. Studies of blockage by buildings in cities have shown that it is difficult to have line-of-sight access to more than 60% of the buildings from a single base station [55]. Measurements in a region in Brooklyn, NY [60], suggest that access from a single base station can range from 25% to 85% for subscriber antenna heights of 10 to 35 ft and a base station height of about 290 ft. While less blockage by houses could be expected in residential areas, such numbers would suggest that greater than 90% access to houses could be difficult, even from multiple elevated locations, when mixes of one- and two-story houses, trees, and hills are present. In regions where tree cover is heavy, e.g., the eastern half of the U.S., tree cover in many places will present a significant obstacle. Heavy rainfall is an additional problem at 28 GHz in some regions. In spite of these challenges, the lure of low-cost wireless loops is attracting many participants, both service providers and equipment manufacturers.

Reality Check

Before we go on to consider other applications and compromises, perhaps it would be helpful to see if there is any indication that the previous discussion is valid. For this check, we could look at cordless telephones for telepoint use (i.e., pocketphones) and at pocket cellular telephones that existed in the 1993 time frame.

Two products from one United States manufacturer are good for this comparison. One is a third-generation hand-portable analog FM cellular phone from this manufacturer that represents their second generation of pocketphones. The other is a first-generation digital cordless phone built to the United Kingdom CT-2 common air interface (CAI) standard. Both units are of flip phone type with the earpiece on the main handset body and the mouthpiece formed by or on the flip-down part. Both operate near 900 MHz and have 1/4 wavelength pull-out antennas. Both are fully functional within their class of operation (i.e., full number of U.S. cellular channels, full number of CT-2 channels, automatic channel setup, etc.). Table 67.4 compares characteristics of these two wireless access pocketphones from the same manufacturer.

The following are the most important items to note in the Table 67.4 comparison.

1. The talk time of the low-power pocketphone is four times that of the high-power pocket-phone.
2. The battery inside the low-power pocketphone is about one-half the weight and size of the battery attached to the high-power pocketphone.

TABLE 67.4 Comparison of CT-2 and Cellular Pocket Size Flip-Phones from the Same Manufacturer

Characteristics/Parameter	CT-2	Cellular
Weight, oz		
Flip phone only	5.2	4.2
Battery[1] only	1.9	3.6
Total unit	7.1	7.8
Size (max.dimensions), in		
Flip phone only	5.9 × 2.2 × 0.95 8.5 in^3	5.5 × 2.4 × 0.9 —
Battery[1] only	1.9 × 1.3 × 0.5 internal	4.7 × 2.3 × 0.4 external
Total unit	5.9 × 2.2 × 0.95 8.5 in^3	5.9 × 2.4 × 1.1 11.6 in^3
Talk-time, min (h)		
Rechargeable battery[2]	180 (3)	45
Nonrechargeable battery	600 (10)	N/A
Standby time, h		
Rechargeable battery	30	8
Nonrechargeable battery	100	N/A
Speech quality	32 kb/s telephone quality	30 kHz FM depends on channel quality
Transmit power avg., W	0.005	0.5

[1]Rechargeable battery.
[2]Ni-cad battery.

3. The battery-usage ratio, talk time/weight of battery, is eight times greater, almost an order of magnitude, for the low-power pocketphone compared to the high-power pocketphone!
4. Additionally, the lower power (5 mW) digital cordless pocketphone is slightly smaller and lighter than the high-power (500 mW) analog FM cellular mobile pocketphone.

Similar comparisons can be made between PHS advanced cordless/low-tier PCS phones and advanced cellular/high-tier PCS pocketphones. New lithium batteries have permitted increased talk time in pocketphones. Digital control/paging channels facilitate significantly extended standby time. Advances in solid-state circuits have reduced the size and weight of cellular pocketphone electronics so that they are almost insignificant compared to the battery required for the high power transmitter and complex digital signal processing. However, even with all these changes, there is still a very significant weight and talk time benefit in the low complexity PHS handsets compared to the most advanced cellular/high-tier PCS handsets. Picking typical minimum size and weight handsets for both technologies results in the following comparisons.

	PHS	Cellular
weight, oz total unit	3	4.5
size total unit	4.2 in^3	—
talk-time, h	8	3
standby time, h	600	48

From the table, the battery usage ratio has been reduced to a factor of about 4 from a factor of 8, but this is based on total weight, not battery weight alone as used for the earlier CT-2 and cellular comparison.

Thus, there is still a large talk time and weight benefit for low-power low-complexity low-tier PCS compared to higher power high-complexity, high tier PCS and cellular.

The following should also be noted.

1. The room for technology improvement of the CT-2 cordless phone is greater since it is first generation and the cellular phone is second/third generation.
2. A digital cellular phone built to the IS-54, GSM, or JDC standard, or in the proposed United States CDMA technology, would either have less talk time or be heavier and larger than the analog FM phone, because: a) the low-bit-rate digital speech coder is more complex and will consume more power than the analog speech processing circuits; b) the digital units have complex digital signal-processing circuits for forward error correction—either for delay dispersion equalizing or for spread-spectrum processing—that will consume significant amounts of power and that have no equivalents in the analog FM unit; and c) power amplifiers for the shaped-pulse nonconstant-envelope digital signals will be less efficient than the amplifiers for constant-envelope analog FM. Although it may be suggested that transmitter power control will reduce the weight and size of a CDMA handset and battery, if that handset is to be capable of operating at full power in fringe areas, it will have to have capabilities similar to other cellular sets. Similar power control applied to a CT-2-like low-maximum-power set would also reduce its power consumption and thus also its weight and size.

The major difference in size, weight, and talk time between the two pocketphones is directly attributable to the two orders of magnitude difference in average transmitter power. The generation of transmitter power dominates power consumption in the analog cellular phone. Power consumption in the digital CT-2 phone is more evenly divided between transmitter-power generation and digital signal processing. Therefore, power consumption in complex digital signal processing would have more impact on talk time in small low-power personal communicators than in cellular handsets where the transmitter-power generation is so large. Other than reducing power consumption for both functions, the only alternative for increasing talk time and reducing battery weight is to invent new battery technology having greater density; see section on Other Issues later in this chapter.

In contrast, lowering the transmitter power requirement, modestly applying digital signal processing, and shifting some of the radio coverage burden to a higher density of small, low-power, low-complexity, low-cost fixed radio ports has the effect of shifting some of the talk time, weight, and cost constraints from battery technology to solid state electronics technology, which continues to experience orders-of-magnitude improvements in the span of several years. Digital signal-processing complexity, however, cannot be permitted to overwhelm power consumption in low-power handsets; whereas small differences in complexity will not matter much, orders-of-magnitude differences in complexity will continue to be significant.

Thus, it can be seen from Table 67.4 that the size, weight, and quality arguments in the preceding sections generally hold for these examples. It also is evident from the preceding paragraphs that they will be even more notable when comparing digital cordless pocketphones with digital cellular pocketphones of the same development generations.

67.4 Evolution Toward the Future and to Low-Tier Personal Communications Services

After looking at the evolution of several wireless technologies and systems in the preceding sections it appears appropriate to ask again: wireless personal communications, *what is it*? All of the technologies in the preceding sections claim to provide wireless personal communications, and all do to some extent. All have significant limitations, however, and all are evolving in attempts to overcome the limitations. It seems appropriate to ask, what are the likely endpoints? Perhaps some hint of the endpoints can be found by exploring what users see as limitations of existing technologies and systems and by looking at the evolutionary trends.

In order to do so, we summarize some important clues from the preceding sections and project them, along with some U.S. standards activity, toward the future.

Digital Cordless Telephones

- Strengths: good circuit quality; long talk time; small lightweight battery; low-cost sets and service.
- Limitations: limited range; limited usage regions.
- Evolutionary trends: phone points in public places; wireless PBX in business.
- Remaining limitations and issues: limited usage regions and coverage holes; limited or no handoff; limited range.

Experience with PHS and CT-2 phone point have provided more emphasis on the need for vehicle speed handoff and continuous widespread coverage of populated areas and of highways in between.

Digital Cellular Pocket Handsets

- Strength: widespread service availability.
- Limitations: limited talk time; large heavy batteries; high-cost sets and service; marginal circuit quality; holes in coverage and poor in-building coverage; limited data capabilities; complex technologies.
- Evolutionary trends: microcells to increase capacity and in-building coverage and to reduce battery drain; satellite systems to extend coverage.
- Remaining limitations and issues: limited talk time and large battery; marginal circuit quality; complex technologies.

Wide Area Data

- Strength: digital messages.
- Limitations: no voice, limited data rate; high cost.
- Evolutionary trends: microcells to increase capacity and reduce cost; share facilities with voice systems to reduce cost.
- Remaining limitations and issues: no voice; limited capacity.

Wireless Local Area Networks (WLANs)

- Strength: high data rate.
- Limitations: insufficient capacity for voice, limited coverage; no standards; chaos.
- Evolutionary trends: hard to discern from all of the churning.

Paging/Messaging

- Strengths: widespread coverage; long battery life; small lightweight sets and batteries; economical.
- Limitations: one-way message only; limited capacity.
- Evolutionary desire: two-way messaging and/or voice; capacity.
- Limitations and issues: two-way link cannot exploit the advantages of one-way link asymmetry.

Fixed Wireless Loops

- Strength: high data rates.
- Limitations: no mobility.

There is a strong trajectory evident in these systems and technologies aimed at providing the following features.

High Quality Voice and Data

- To small, lightweight, pocket carried communicators.
- Having small lightweight batteries.

- Having long talk time and long standby battery life.
- Providing service over large coverage regions.
- For pedestrians in populated areas (but not requiring high population density).
- Including low to moderate speed mobility with handoff. It has become evident from the experience with PHS and CT-2 phone point that vehicle speed handoff is essential so that handsets can be used in vehicles also.

Economical Service

- Low subscriber-set cost.
- Low network-service cost.

Privacy and Security of Communications

- Encrypted radio links.

This trajectory is evident in all of the evolving technologies but can only be partially satisfied by any of the existing and evolving systems and technologies! Trajectories from all of the evolving technologies and systems are illustrated in Fig. 67.1 as being aimed at low-tier personal communications systems or services, i.e., low-tier PCS. Taking characteristics from cordless, cellular, wide-area data and, at least moderate-rate, WLANs, suggests the following attributes for this low-tier PCS.

1. 32 kb/s ADPCM speech encoding in the near future to take advantage of the low complexity and low power consumption, and to provide low-delay high-quality speech.
2. Flexible radio link architecture that will support multiple data rates from several kilobits per second. This is needed to permit evolution in the future to lower bit rate speech as technology improvements permit high quality without excessive power consumption or transmission delay and to provide multiple data rates for data transmission and messaging.
3. Low transmitter power (≤25 mW average) with adaptive power control to maximize talk time and data transmission time. This incurs short radio range that requires many base stations to cover a large region. Thus, base stations must be small and inexpensive, like cordless telephone phone points or the Metricom wireless data base stations. The lower power will require somewhat closer spacing of base stations in cluttered environments with many buildings, etc. This issue is dealt with in more detail in Section 67.5. The issues associated with greater base station spacing along highways are also considered in Section 67.5.
4. Low-complexity signal processing to minimize power consumption. Complexity one-tenth that of digital cellular or high-tier PCS technologies is required [29]. With only several tens of milliwatts (or less under power control) required for transmitter power, signal processing power becomes significant.
5. Low cochannel interference and high coverage area design criteria. In order to provide high-quality service over a large region, at least 99% of any covered area must receive good or better coverage and be below acceptable cochannel interference limits. This implies less than 1% of a region will receive marginal service. This is an order-of-magnitude higher service requirement than the 10% of a region permitted to receive marginal service in vehicular cellular system (high-tier PCS) design criteria.
6. Four-level phase modulation with coherent detection to maximize radio link performance and capacity with low complexity.
7. Frequency division duplexing to relax the requirement for synchronizing base station transmissions over a large region. PHS uses time division duplexing and requires base station synchronization. In first deployments, one provider did not implement this synchronization. The expected serious performance degradation prompted system upgrades to provide the needed synchronization. While this is not a big issue, it does add complexity to the system and decreases the overall robustness.

8. As noted previously, experience with PHS and CT-2 phone point have emphasized the need for vehicular speed handoff in these low-tier PCS systems. Such handoff is readily implemented in PACS and has been demonstrated in the field [51]. This issue is discussed in more detail later in this section.

Such technologies and systems have been designed, prototyped, and laboratory and field tested and evaluated for several years [7,23,24,25,26,27,28,29,31,32,50]. The viewpoint expressed here is consistent with the progress in the Joint Technical Committee (JTC) of the U.S. standards bodies, Telecommunications Industry Association (TIA) and Committee T1 of the Alliance for Telecommunications Industry Solutions (ATIS). Many technologies and systems were submitted to the JTC for consideration for wireless PCS in the new 1.9-GHz frequency bands for use in the United States [20]. Essentially all of the technologies and systems listed in Table 67.1, and some others, were submitted in late 1993. It was evident that there were at least two and perhaps three distinctly different classes of submissions. No systems optimized for packet data were submitted, but some of the technologies are optimized for voice.

One class of submissions was the group labeled high-power systems, digital cellular (high-tier PCS) in Table 67.1. These are the technologies discussed previously in this chapter. They are highly optimized for low-bit-rate voice and, therefore, have somewhat limited capability for serving packet-data applications. Since it is clear that wireless services to wide ranging high-speed mobiles will continue to be needed, and that the technology already described for low-tier PCS may not be optimum for such services, Fig. 67.1 shows a continuing evolution and need in the future for high-tier PCS systems that are the equivalent of today's cellular radio. There are more than 100 million vehicles in the United States alone. In the future, most, if not all, of these will be equipped with high-tier cellular mobile phones. Therefore, there will be a continuing and rapidly expanding market for high-tier systems.

Another class of submissions to the JTC [20] included the Japanese personal handyphone system (PHS) and a technology and system originally developed at Bellcore but carried forward to prototypes and submitted to the JTC by Motorola and Hughes Network Systems. This system was known as wireless access communications systems (WACS).[2] These two submissions were so similar in their design objectives and system characteristics that, with the agreement of the delegations from Japan and the United States, the PHS and WACS submissions were combined under a new name, personal access communication systems (PACS), that was to incorporate the best features of both. This advanced, low-power wireless access system, PACS, was to be known as low-tier PCS. Both WACS/PACS and Handyphone (PHS) are shown in Table 67.1 as low-tier PCS and represent the evolution to low-tier PCS in Fig. 67.1. The WACS/PACS/UDPC system and technology are discussed in [7,23,24,25,26,28,29,31,32,50].

In the JTC, submissions for PCS of DECT and CT-2 and their variations were also lumped under the class of low-tier PCS, even though these advanced digital cordless telephone technologies were somewhat more limited in their ability to serve all of the low-tier PCS needs. They are included under digital cordless technologies in Table 67.1. Other technologies and systems were also submitted to the JTC for high-tier and low-tier applications, but they have not received widespread industry support.

One wireless access application discussed earlier that is not addressed by either high-tier or low-tier PCS is the high-speed WLAN application. Specialized high-speed WLANs also are likely to find a place in the future. Therefore, their evolution is also continued in Fig. 67.1. The figure also recognizes that widespread low-tier PCS can support data at several hundred kilobits per second and, thus, can satisfy many of the needs of WLAN users.

It is not clear what the future roles are for paging/messaging, cordless telephone appliances, or wide-area packet-data networks in an environment with widespread contiguous coverage by low-tier and high-tier PCS. Thus, their extensions into the future are indicated with a question mark in Fig. 67.1.

Those who may object to the separation of wireless PCS into high-tier and low-tier should review this section again, and note that we have two tiers of PCS now. On the voice side there is cellular radio, i.e., high-tier PCS, and cordless telephone, i.e., an early form of low-tier PCS. On the data side there is wide-area data,

[2]WACS was known previously as Universal Digital Portable Communications (UDPC).

i.e., high-tier data PCS, and WLANs, i.e., perhaps a form of low-tier data PCS. In their evolutions, these all have the trajectories discussed and shown in Fig. 67.1 that point surely toward low-tier PCS. It is this low-tier PCS that marketing studies continue to project is wanted by more than half of the U.S. households or by half of the people, a potential market of over 100 million subscribers in the United States alone. Similar projections have been made worldwide.

PACS technology [6] has been prototyped by several manufacturers. In 1995 field demonstrations were run in Boulder, CO at a U.S. West test site using radio ports (base stations) and radio port control units made by NEC. "Handset" prototypes made by Motorola and Panasonic were trialed. The handsets and ports were brought together for the first time in Boulder. The highly successful trial demonstrated the ease of integrating the subsystems of the low-complexity PACS technology and the overall advantages of PACS from a user's perspective as noted throughout this chapter. Effective vehicular speed operation was demonstrated in these tests. Also, Hughes Network Systems (HNS) has developed and tested many sets of PACS infrastructure technology with different handsets in several settings and has many times demonstrated highly reliable high vehicular speed (in excess of 70 mi/hr) operation and handoff among several radio ports. Motorola also has demonstrated PACS equipment in several settings at vehicular speeds as well as for wireless loop applications. Highly successful demonstrations of PACS prototypes have been conducted from Alaska to Florida, from New York to California, and in China and elsewhere.

A PACS deployment in China using NEC equipment started to provide service in 1998. The U.S. service provider, 21st Century Telesis, is poised to begin a PACS deployment in several states in the U.S. using infrastructure equipment from HNS and handsets and switching equipment from different suppliers. Perhaps, with a little more support of these deployments, the public will finally be able to obtain the benefits of low-tier PCS.

67.5 Comparisons with Other Technologies

Complexity/Coverage Area Comparisons

Experimental research prototypes of radio ports and subscriber sets [64,66] have been constructed to demonstrate the technical feasibility of the radio link requirements in [7]. These WACS prototypes generally have the characteristics and parameters previously noted, with the exceptions that (1) the portable transmitter power is lower (10 mW average, 100 mW peak), (2) dynamic power control and automatic time slot transfer are not implemented, and (3) a rudimentary automatic link-transfer implementation is based only on received power. The experimental base stations transmit near 2.17 GHz; the experimental subscriber sets transmit near 2.12 GHz. Both operated under a Bellcore experimental license. The experimental prototypes incorporate application-specific, very large-scale integrated circuits[3] fabricated to demonstrate the feasibility of the low-complexity high-performance digital signal-processing techniques [63,64] for symbol timing and coherent bit detection. These techniques permit the efficient short TDMA bursts having only 100 b that are necessary for low-delay TDMA implementations. Other digital signal-processing functions in the prototypes are implemented in programmable logic devices. All of the digital signal-processing functions combined require about 1/10 of the logic gates that are required for digital signal processing in vehicular digital cellular mobile implementations [42,62,63]; that is, this low-complexity PCS implementation having no delay-dispersion-compensating circuits and no forward error-correction decoding and is about 1/10 as complex as the digital cellular implementations that include these functions.[4] The 32 kb/s ADPCM speech-encoding in the low-complexity PCS implementation is also about 1/10 as complex as the less than 10-kb/s speech encoding used in digital cellular implementations.

[3]Applications specific integrated circuits (ASIC), very large-scale integration (VLSI).

[4]Some indication of VLSI complexity can be seen by the number of people required to design the circuits. For the low-complexity TDMA ASIC set, only one person part time plus a student part time were required; the complex CDMA ASIC has six authors on the paper alone.

This significantly lower complexity will continue to translate into lower power consumption and cost. It is particularly important for low-power pocket personal communicators with power control in which the DC power expended for radio frequency transmitting can be only tens of milliwatts for significant lengths of time.

The experimental radio links have been tested in the laboratory for detection sensitivity [bit error rate (BER) vs SNR] [18,61,66] and for performance against cochannel interference [1] and intersymbol interference caused by multipath delay spread [66]. These laboratory tests confirm the performance of the radio-link techniques. In addition to the laboratory tests, qualitative tests have been made in several PCS environments to compare these experimental prototypes with several United States CT-1 cordless telephones at 50 MHz, with CT-2 cordless telephones at 900 MHz, and with DCT-900 cordless telephones at 900 MHz. Some of these comparisons have been reported [8,71,84,85]. In general, depending on the criteria, e.g., either no degradation or limited degradation of circuit quality, these WACS experimental prototypes covered areas inside buildings that ranged from 1.4 to 4 times the areas covered by the other technologies. The coverage areas for the experimental prototypes were always substantially limited in two or three directions by the outside walls of the buildings. These area factors could be expected to be even larger if the coverage were not limited by walls, i.e., once all of a building is covered in one direction, no more area can be covered no matter what the radio link margin. The earlier comparisons [8,84,85] were made with only two-branch uplink diversity before subscriber-set transmitting antenna switching was implemented and, with only one radio port before automatic radio-link transfer was implemented. The later tests [71] included these implementations. These reported comparisons agree with similar unreported comparisons made in a Bellcore Laboratory building. Similar coverage comparison results have been noted for a 900-MHz ISM-band cordless telephone compared to the 2-GHz experimental prototype. The area coverage factors (e.g., $\times 1.4$ to $\times 4$) could be expected to be even greater if the cordless technologies had also been operated at 2 GHz since attenuation inside buildings between similar small antennas is about 7 dB greater at 2 GHz than at 900 MHz [35,36] and the 900 MHz handsets transmitted only 3 dB less average power than the 2-GHz experimental prototypes. The greater area coverage demonstrated for this technology is expected because of the different compromises noted earlier; the following, in particular.

1. Coherent detection of QAM provides more detection sensitivity than noncoherent detection of frequency-shift modulations [17].
2. Antenna diversity mitigates bursts of errors from multipath fading [66].
3. Error detection and blanking of TDMA bursts having errors significantly improves perceived speech quality [72]. (Undetected errors in the most significant bit cause sharp audio pops that seriously degrade perceived speech quality.)
4. Robust symbol timing and burst and frame synchronization reduce the number of frames in error due to imperfect timing and synchronization [66].
5. Transmitting more power from the radio port compared to the subscriber set offsets the less sensitive subscriber set receiver compared to the port receiver that results from power and complexity compromises made in a portable set.

Of course, as expected, the low-power (10-mW) radio links cover less area than high-power (0.5-W) cellular mobile pocketphone radio links because of the 17-dB transmitter power difference resulting from the compromises discussed previously. In the case of vehicular mounted sets, even more radio-link advantage accrues to the mobile set because of the higher gain of vehicle-mounted antennas and higher transmitter power (3 W).

The power difference between a low-tier PACS handset and a high-tier PCS or cellular pocket handset is not as significant in limiting range as is often portrayed. Other differences in deployment scenarios for low-tier and high-tier systems are as large or even larger factors, e.g., base station antenna height and antenna gain. This can be seen by considering using the same antennas and receiver noise figures at base stations and looking at the range of high-tier and low-tier handsets. High-tier handsets typically transmit a maximum average power of 0.5 watt. The PACS handset average transmit power is 25 milliwatts

(peak power is higher for TDMA, but equal comparisons can be made considering average power and equivalent receiver sensitivities). This power ratio of ×20 translates to approximately a range reduction of a factor of about 0.5 for an environment with a distance dependence of $1/(d)^4$ or a factor of about 0.4 for a $1/(d)^{3.5}$ environment. These represent typical values of distance dependence for PCS and cellular environments. Thus, if the high-tier handset would provide a range of 5 miles in some environment, the low-tier handset would provide a range of 2 to 2.5 miles in the same environment, if the base station antennas and receiver noise figures were the same. This difference in range is no greater than the difference in range between high-tier PCS handsets used in 1.9 GHz systems and cellular handsets with the same power used in 800 MHz systems. Range factors are discussed further in the next section.

Coverage, Range, Speed, and Environments

Interest has been expressed in having greater range for low-tier PCS technology for low-population-density areas. One should first note that the range of a wireless link is highly dependent on the amount of clutter or obstructions in the environment in which it is operated. For example, radio link calculations that result in a 1400-ft base station (radio-port) separation at 1.9 GHz contain over 50-dB margin for shadowing from obstructions and multipath effects [25,37]. Thus, in an environment without obstructions, e.g., along a highway, the base station separation can be increased at least by a factor of 4 to over a mile, i.e., 25 dB for an attenuation characteristic of d^{-4}, while providing the same quality of service, without any changes to the base station or subscriber transceivers, and while still allowing over 25-dB margin for multipath and some shadowing. This remaining margin allows for operation of a handset inside an automobile. In such an unobstructed environment, multipath RMS delay spread [21,33] will still be less than the 0.5 μs in which PACS was designed to operate [28].

Operation at still greater range along unobstructed highways or at a range of a mile along more obstructed streets can be obtained in several ways. Additional link gain of 6 dB can be obtained by going from omnidirectional antennas at base stations to 90° sectored antennas (four sectors). Another 6 dB can be obtained by raising base station antennas by a factor of 2 from 27 ft to 55 ft in height. This additional 12 dB will allow another factor of 2 increase in range to 2-mile base station separation along highways, or to about 3000-ft separation in residential areas. Even higher-gain and taller antennas could be used to concentrate coverage along highways, particularly in rural areas. Of course, range could be further increased by increasing the power transmitted.

As the range of the low-tier PACS technology is extended in cluttered areas by increasing link gain, increased RMS delay spread is likely to be encountered. This will require increasing complexity in receivers. A factor of 2 in tolerance of delay spread can be obtained by interference-canceling signal combining [76,77,78,79,80] from two antennas instead of the simpler selection diversity combining originally used in PACS. This will provide adequate delay-spread tolerance for most suburban environments [21,33].

The PACS downlink contains synchronization words that could be used to train a conventional delay-spread equalizer in subscriber set receivers. Constant-modulus (blind) equalization will provide greater tolerance to delay spread in base station receivers on the uplink [45,46,47,48] than can be obtained by interference-cancellation combining from only two antennas. The use of more base-station antennas and receivers can also help mitigate uplink delay spread. Thus, with some added complexity, the low-tier PACS technology can work effectively in the RMS delay spreads expected in cluttered environments for base station separations of 2 miles or so.

The guard time in the PACS TDMA uplink is adequate for 1-mile range, i.e., 2-mile separation between base station and subscriber transceivers. A separation of up to 3 miles between transceivers could be allowed if some statistical outage were accepted for the few times when adjacent uplink timeslots are occupied by subscribers at the extremes of range (near–far). With some added complexity in assigning timeslots, the assignment of subscribers at very different ranges to adjacent timeslots could be avoided, and the base station separation could be increased to several miles without incurring adjacent slot interference. A simple alternative in low-density (rural) areas, where lower capacity could be acceptable

and greater range could be desirable, would be to use every other timeslot to ensure adequate guard time for range differences of many tens of miles. Also, the capability of transmitter time advance has been added to PACS standard in order to increase the range of operation. Such time advance is applied in the cellular TDMA technologies.

The synchronization, carrier recovery, and detection in the low-complexity PACS transceivers will perform well at highway speeds. The two-receiver diversity used in uplink transceivers also will perform well at highway speeds. The performance of the single-receiver selection diversity used in the low-complexity PACS downlink transceivers begins to deteriorate at speeds above about 30 mi/h. However, at any speed, the performance is always at least as good as that of a single transceiver without the low-complexity diversity. Also, fading in the relatively uncluttered environment of a highway is likely to have a less severe Ricean distribution, so diversity will be less needed for mitigating the fading. Cellular handsets do not have diversity. Of course, more complex two-receiver diversity could be added to downlink transceivers to provide two-branch diversity performance at highway speeds. It should be noted that the very short 2.5-ms TDMA frames incorporated into PACS to provide low transmission delay (for high speech quality) also make the technology significantly less sensitive to high-speed fading than the longer-frame-period cellular technologies. The short frame also facilitates the rapid coordination needed to make reliable high-speed handoffs between base stations. Measurements on radio links to potential handoff base stations can be made rapidly, i.e., a measurement on at least one radio link every 2.5 ms. Once a handoff decision is made, signalling exchanges every 2.5 ms ensure that the radio link handoff is completed quickly. In contrast, the long frame periods in the high-tier (cellular) technologies prolong the time it takes to complete a handoff. As noted earlier, high speed handoff has been demonstrated many times with PACS technology and at speeds over 70 mi/h.

67.6 Quality, Capacity, and Economic Issues

Although the several trajectories toward low-tier PCS discussed in the preceding section are clear, it does not fit the existing wireless communications paradigms. Thus, low-tier PCS has attracted less attention than the systems and technologies that are compatible with the existing paradigms. Some examples are cited in the following paragraphs.

The need for intense interaction with an intelligent network infrastructure in order to manage mobility is not compatible with the cordless telephone appliance paradigm. In that paradigm, independence of network intelligence and base units that mimic wireline telephones are paramount.

Wireless data systems often do not admit to the dominance of wireless voice communications and, thus, do not take advantage of the economics of sharing network infrastructure and base station equipment. Also, wireless voice systems often do not recognize the importance of data and messaging and, thus, only add them in as bandaids to systems.

The need for a dense collection of many low-complexity, low-cost, low-tier PCS base stations interconnected with inexpensive fixed-network facilities (copper or fiber based) does not fit the cellular high-tier paradigm that expects sparsely distributed $1 million cell sites. Also, the need for high transmission quality to compete with wireline telephones is not compatible with the drive toward maximizing users-per-cell-site and per megahertz to minimize the number of expensive cell sites. These concerns, of course, ignore the hallmark of frequency-reusing cellular systems. That hallmark is the production of almost unlimited overall system capacity by reducing the separation between base stations. The cellular paradigm does not recognize the fact that almost all houses in the U.S. have inexpensive copper wires connecting telephones to the telephone network. The use of low-tier PCS base stations that concentrate individual user services before backhauling in the network will result in less fixed interconnecting facilities than exist now for wireline telephones. Thus, inexpensive techniques for interconnecting many low-tier base stations are already deployed to provide wireline telephones to almost all houses. The cost of backhaul to many base stations (radio ports) in a low-tier system is often cited as an economic disadvantage that cannot be overcome. However, this perception is based on existing tariffs for T1 digital lines which are

TABLE 67.5 Comparison of Cellular (IS-54/IS-95) and Low-Tier PCS (WACS/PACS). Capacity Comparisons Made without Regard to Quality Factors, Complexity, and Cost per Base Station Are not Meaningful

Parameter	Cellular (High-Tier)	Low-Tier PCS	Capacity Factor
Speech Coding, kb/s	8 (MOS 3.4) No tandem coding	32 (MOS 4.1) 3 or 4 tandem	×4
Speech activity	Yes (MOS 3.2)	No (MOS 4.1)	×2.5
Percentage of good areas, %	90	99	×2
Propagation σ, dB	8	10	×1.5
Total: trading quality for capacity			×30

excessive considering current digital subscriber line technology. These tariffs were established many years ago when digital subscriber line electronics were very expensive. With modern low-cost high-rate digital subscriber line (HDSL) electronics, the cost of backhaul could be greatly reduced. If efforts were made to revise tariffs for digital line backhaul based on low cost electronics and copper loops like residential loops, the resulting backhaul costs would more nearly approach the cost of residential telephone lines. As it is now, backhaul costs are calculated based on antiquated high T1 line tariffs that were established for "antique" high cost electronics.

This list could be extended, but the preceding examples are sufficient, along with the earlier sections of the paper, to indicate the many complex interactions among circuit quality, spectrum utilization, complexity (circuit and network), system capacity, and economics that are involved in the design compromises for a large, high-capacity wireless-access system. Unfortunately, the tendency has been to ignore many of the issues and focus on only one, e.g., the focus on cell site capacity that drove the development of digital-cellular high-tier systems in the United States. Interactions among circuit quality, complexity, capacity, and economics are considered in the following sections.

Capacity, Quality, and Complexity

Although capacity comparisons frequently are made without regard to circuit quality, complexity, or cost per base station, such comparisons are not meaningful. An example in Table 67.5 compares capacity factors for U.S. cellular or high-tier PCS technologies with the low-tier PCS technology, PACS/WACS. The mean opinion scores (MOS) (noted in Table 67.5) for speech coding are discussed later. Detection of speech activity and turning off the transmitter during times of no activity is implemented in IS-95. Its impact on MOS also is noted later. A similar technique has been proposed as E-TDMA for use with IS-54 and is discussed with respect to TDMA system in [29]. Note that the use of low-bit-rate speech coding combined with speech activity degrades the high-tier system's quality by nearly one full MOS point on the five-point MOS scale when compared to 32 kb/s ADPCM. Tandem encoding is discussed in a later section. These speech quality degrading factors alone provide a base station capacity increasing factor of $\times 4 \times 2.5 = \times 10$ over the high-speech-quality low-tier system! Speech coding, of course, directly affects base station capacity and, thus, overall system capacity by its effect on the number of speech channels that can fit into a given bandwidth.

The allowance of extra system margin to provide coverage of 99% of an area for low-tier PCS versus 90% coverage for high-tier is discussed in the previous section and [29]. This additional quality factor costs a capacity factor of ×2. The last item in Table 67.5 does not change the actual system, but only changes the way that frequency reuse is calculated. The additional 2-dB margin in standard deviation σ, allowed for coverage into houses and small buildings for low-tier PCS, costs yet another factor of ×1.5 in calculation only. Frequency reuse factors affect the number of sets of frequencies required and, thus, the bandwidth available for use at each base station. Thus, these factors also affect the base station capacity and the overall system capacity.

For the example in Table 67.5, significant speech and coverage quality has been traded for a factor of ×30 in base station capacity! Whereas base station capacity affects overall system capacity directly, it

should be remembered that overall system capacity can be increased arbitrarily by decreasing the spacing between base stations. Thus, if the PACS low-tier PCS technology were to start with a base station capacity of ×0.5 of AMPS cellular[5] (a much lower figure than the ×0.8 sometimes quoted [20]), and then were degraded in quality as described above to yield the ×30 capacity factor, it would have a resulting capacity of ×15 of AMPS! Thus, it is obvious that making such a base station capacity comparison without including quality is not meaningful.

Economics, System Capacity, and Coverage Area Size

Claims are sometimes made that low-tier PCS cannot be provided economically, even though it is what the user wants. These claims are often made based on economic estimates from the cellular paradigm. These include the following.

- Very low estimates of market penetration, much less than cordless telephones, and often even less than cellular.
- High estimates of base station costs more appropriate to high-complexity, high-cost cellular technology than to low-complexity, low-cost, low-tier technology.
- Very low estimates of circuit usage time more appropriate to cellular than to cordless/wireline telephone usage, which is more likely for low-tier PCS.
- Backhaul costs based on existing T1 line tariffs that are based on "antique" high cost digital loop electronics. (See discussion in fourth paragraph at start of Section 67.6.)

Such economic estimates are often done by making absolute economic calculations based on very uncertain input data. The resulting estimates for low-tier and high-tier are often closer together than the large uncertainties in the input data. A perhaps more realistic approach for comparing such systems is to vary only one or two parameters while holding all others fixed and then looking at relative economics between high-tier and low-tier systems. This is the approach used in the following examples.

Example 67.1

In the first example (see Table 67.6), the number of channels per megahertz is held constant for cellular and for low-tier PCS. Only the spacing is varied between base stations, e.g., cell sites for cellular and radio ports for low-tier PCS, to account for the differences in transmitter power, antenna height, etc. In this example, overall system capacity varies directly as the square of base station spacing, but base station capacity is the same for both cellular and low-tier PCS. For the typical values in the example, the resulting low-tier system capacity is ×400 greater, only because of the closer base station spacing. If the two systems were to cost the same, the equivalent low-tier PCS base stations would have to cost less than $2,500.

This cost is well within the range of estimates for such base stations, including equivalent infrastructure. These low-tier PCS base stations are of comparable or lower complexity than cellular vehicular subscriber sets, and large-scale manufacture will be needed to produce the millions that will be required. Also, land, building, antenna tower and legal fees for zoning approval, or rental of expensive space on top of commercial buildings, represent large expenses for cellular cell sites. Low-tier PCS base stations that are mounted on utility poles and sides of buildings will not incur such large additional expenses. Therefore, costs of the order of magnitude indicated seem reasonable in large quantities. Note that, with these estimates, the per-wireless-circuit cost of the low-tier PCS circuits would be only $14/circuit compared to $5,555/circuit for the high-tier circuits. Even if there were a factor of 10 error in cost estimates, or a reduction of channels per radio port of a factor of 10, the per-circuit cost of low-tier PCS would still be only $140/circuit, which is still much less than the per-circuit cost of high-tier.

[5]Note that the ×0.5 factor is an arbitrary factor taken for illustrating this example. The so-called ×AMPS factors are only with regard to base station capacity, although they are often misused as system capacity.

TABLE 67.6 System Capacity/Coverage Area Size/Economics

Example 67.1

Assume channels/MHz are the same for cellular and PCS
Cell site: spacing = 20.000 ft cost $ = 1 M
PCS port: spacing = 1,000 ft
PCS system capacity is $(20000/1000)^2 = 400 \times$ cellular capacity

Then, for the system costs to be the same
Port cost = ($1 M/400) $2,500 a reasonable figure

If, cell site and port each have 180 channels
Cellular cost/circuit = $1 M/180 = $5,555/circuit
PCS cost/circuit = $2500/180 = $14/circuit

Example 67.2

Assume equal cellular and PCS system capacity
Cell site: spacing = 20,000 ft
PCS port: spacing = 1,000 ft

If, a cell site has 180 channels
then, for equal system capacity, a PCS port needs
180/400 < 1 channel/port

Example 67.3

Quality/cost trade
Cell site: Spacing = 20,000 ft cost = $1 M channels = 180
PCS port: Spacing = 1,000 ft cost = $2,500

Cellular to PCS, base station spacing capacity factor = ×400
PCS to cellular quality reduction factors:

32 to 8 kb/s speech	×4
Voice activity (buying)	×2
99–90% good areas	×2
Both in same environment (same σ)	×1
Capacity factor traded	×16

180 ch/16 = 11.25 channels/port then, $2500/11.25 = $222/circuit and remaining is ×400/16 = ×25 system capacity of PCS over cellular

Example 67.2

In the second example (see Table 67.6), the overall system capacity is held constant, and the number of channels/port, i.e., channels/(base station) is varied. In this example, less than 1/2 channel/port is needed, again indicating the tremendous capacity that can be produced with close-spaced low-complexity base stations.

Example 67.3

Since the first two examples are somewhat extreme, the third example (see Table 67.6) uses a more moderate, intermediate approach. In this example, some of the cellular high-tier channels/(base station) are traded to yield higher quality low-tier PCS as in the previous subsection. This reduces the channels/port to 11+, with an accompanying increase in cost/circuit up to $222/circuit, which is still much less than the $5,555/circuit for the high-tier system. Note, also, that the low-tier system still has ×25 the capacity of the high-tier system!

Low-tier base station (Port) cost would have to exceed $62,500 for the low-tier per-circuit cost to exceed that of the high-tier cellular system. Such a high port cost far exceeds any existing realistic estimate of low-tier system costs.

It can be seen from these examples, and particularly Example 67.3, that the circuit economics of low-tier PCS are significantly better than for high-tier PCS, if the user demand and density is sufficient to make use of the large system capacity. Considering the high penetration of cordless telephones, the rapid growth of cellular handsets, and the enormous market projections for wireless PCS noted earlier in this

chapter, filling such high capacity in the future would appear to be certain. The major problem is providing rapidly the widespread coverage (buildout) required by the FCC in the United States. If this unrealistic regulatory demand can be overcome, low-tier wireless PCS promises to provide the wireless personal communications that everyone wants.

Loop Evolution and Economics

It is interesting to note that several wireless loop applications are aimed at reducing cost by replacing parts of wireline or CATV loops with wireless links between transceivers. The economics of these applications are driven by the replacing of labor-intensive wireline and cable technologies with mass-produced solid-state electronics in transceivers.

Consider first a cordless telephone base unit. The cordless base-unit transceiver usually serves one or, at most, two handsets at the end of one wireline loop. Now consider moving such a base unit back along the copper-wire-pair loop end a distance that can be reliably covered by a low-power wireless link [25,31], i.e., several hundred to a thousand feet or so, and mounting it on a utility pole or a street light pole. This replaces the copper loop end with the wireless link. Many additional copper loop ends to other subscribers will be contained within a circle around the pole having a maximum usable radius of this wireless link. Replace all of the copper loop ends within the circle with cordless base units on the same pole. Note that this process replaces the most expensive parts of these many loops, i.e., the many individual loop ends, with the wireless links from cordless handsets to "equivalent" cordless base units on a pole. Of course, being mounted outside will require somewhat stronger enclosures and means of powering the base units, but these additional costs are considerably more than offset by eliminating the many copper wire drops.

It is instructive to consider how many subscribers could be collected at a pole containing base units. Consider, as an example, a coverage square of 1400 ft on a side (PACS will provide good coverage over this range, i.e., for base unit pole separations of about 1400 ft, at 1.9 GHz). Within this square will be 45 houses for a 1 house/acre density typical of low-housing-density areas, or 180 houses for 4 house/acre density more typical of high-density single-family housing areas. These represent significant concentration of traffic at a pole.

Because of the trunking advantage of the significant number of subscribers concentrated at a pole, they can share a smaller number of base unit, i.e., wireless base unit transceivers, than there are wireless subscriber sets. Therefore, the total cost compared with having a cordless base unit per subscriber also is reduced by the concentration of users.

A single PACS transceiver will support simultaneously eight TDMA channels or circuits at 32 kb/s (or 16 at 16 kb/s or 32 at 8 kb/s) [56]. Of these, one channel is reserved for system control. The cost of such moderate-rate transceivers is relatively insensitive to the number of channels supported; i.e., the cost of such an 8-channel (or 16 or 32) transceiver will be significantly less than twice the cost of a similar one-channel transceiver. Thus, another economic advantage accrues to this wireless loop approach from using time-multiplexed (TDMA) transceivers instead of single-channel-per-transceiver cordless telephone base units.

For an offered traffic of 0.06 Erlang, a typical busy-hour value for a wireline subscriber, a seven-channel transceiver could serve about 40 subscribers at 1% blocking, based on the Erlang B queuing discipline. From the earlier example, such a transceiver could serve most of the 45 houses within a 1400-ft square. Considering partial penetration, the transceiver capacity is more than adequate for the low-density housing.[6]

Considering the high-density example of 4 houses/acre, a seven-channel transceiver could serve only about 20% of the subscribers within a 1400-ft square. If the penetration became greater than about 20%, either additional transceivers, perhaps those of other service providers, or closer transceiver spacing would be required.

[6]The range could be extended by using higher base unit antennas, by using higher-gain directional (sectored) antennas, and/or by increasing the maximum power that can be transmitted.

Another advantageous economic factor for wireless loops results when considering time-multiplexed transmission in the fixed distribution facilities. For copper or fiber digital subscriber loop carrier (SLC), e.g., T1 or high-rate digital subscriber line (HDSL), a demultiplexing/multiplexing terminal and drop interface are required at the end of the time-multiplexed SLC line to provide the individual circuits for each subscriber loop-end circuit, i.e., for each drop. The most expensive part of such an SLC terminating unit is the subscriber line cards that provide per-line interfaces for each subscriber drop. Terminating a T1 or HDSL line on a wireless loop transceiver eliminates all per-line interfaces, i.e., all line cards, the most expensive part of a SLC line termination. Thus, the greatly simplified SLC termination can be incorporated within a TDMA wireless loop transceiver, resulting in another cost savings over the conventional copper-wire-pair telephone loop end.

The purpose of the previous discussions is not to give an exact system design or economic analysis, but to illustrate the inherent economic advantages of low-power wireless loops over copper loop ends and over copper loop ends with cordless telephone base units. Some economic analyses have found wireless loop ends to be more economical than copper loop ends when subscribers use low-power wireless handsets. Rizzo and Sollenberger [56] have also discussed the advantageous economics of PACS wireless loop technology in the context of low-tier PCS.

The discussions in this section can be briefly summarized as follows. Replacing copper wire telephone loop ends with low-complexity wireless loop technology like PACS can produce economic benefits in at least four ways. These are:

1. Replacing the most expensive part of a loop, the per-subscriber loop-end, with a wireless link.
2. Taking advantage of trunking in concentrating many wireless subscriber loops into a smaller number of wireless transceiver channels.
3. Reducing the cost of wireless transceivers by time multiplexing (TDMA) a few (7, 15, or 31) wireless loop circuits (channels) in each transceiver.
4. Eliminating per-line interface cards in digital subscriber line terminations by terminating time-multiplexed subscriber lines in the wireless loop transceivers.

67.7 Other Issues

Several issues in addition to those addressed in the previous two sections continue to be raised with respect to low-tier PCS. These are treated in this section.

Improvement of Batteries

Frequently, the suggestion is made that battery technology will improve so that high-power handsets will be able to provide the desired 5 or 6 hours of talk time in addition to 10 or 12 hours of standby time, and still weigh less than one-fourth of the weight of today's smallest cellular handset batteries. This hope does not take into account the maturity of battery technology, and the long history (many decades) of concerted attempts to improve it. Increases in battery capacity have come in small increments, a few percent, and very slowly over many years, and the shortfall is well over a factor of 10. In contrast, integrated electronics and radio frequency devices needed for low-power low-tier PCS continue to improve and to decrease in cost by factors of greater than 2 in time spans on the order of a year or so. It also should be noted that, as the energy density of a battery is increased, the energy release rate per volume must also increase in order to supply the same amount of power. If energy storage density and release rate are increased significantly, the difference between a battery and a bomb become indistinguishable! The likelihood of a $\times 10$ improvement in battery capacity appears to be essentially zero. If even a modest improvement in battery capacity were possible, many people would be driving electric vehicles.

As noted in the addition to the Reality Check section, new lithium batteries have become the batteries of choice for the smallest cellular/high-tier PCS handsets. While these lithium batteries have higher energy density than earlier nickel cadmium batteries, they still fall far short of the factor of 10 improvement that was needed to make long talk time, small size, and low weight possible. With the much larger advance

in electronics, the battery is even more dominant in the size and weight of the newest cellular handsets. The introduction of these batteries incurred considerable startup pain because of the greater fire and explosive hazard associated with lithium materials, i.e., closer approach to a bomb. Further attempts in this direction will be even more hazardous.

People Only Want One Handset

This issue is often raised in support of high-tier cellular handsets over low-tier handsets. Whereas the statement is likely true, the assumption that the handset must work with high-tier cellular is not. Such a statement follows from the current large usage of cellular handsets; but such usage results because that is the only form of widespread wireless service currently available, not because it is what people want. The statement assumes inadequate coverage of a region by low-tier PCS, and that low-tier handsets will not work in vehicles. The only way that high-tier handsets could serve the desires of people discussed earlier would be for an unlikely breakthrough in battery technology to occur. A low-tier system, however, can cover economically any large region having some people in it. (It will not cover rural or isolated areas but, by definition, there is essentially no one there to want communications anyway.)

Low-tier handsets will work in vehicles on village and city streets at speeds up to 30 or 40 mi/h, and the required handoffs make use of computer technology that is rapidly becoming inexpensive. As noted earlier, vehicular speed handoff is readily accomplished with PACS. Reliable handoff has been demonstrated for PACS at speeds in excess of 70 mi/h. Highways between populated areas, and also streets within them, will need to be covered by high-tier cellular PCS, but users are likely to use vehicular sets in these cellular systems. Frequently the vehicular mobile user will want a different communications device anyway, e.g., a hands-free phone. The use of hands-free phones in vehicles is becoming a legal requirement in some places now and is likely to become a requirement in many more places in the future. Thus, handsets may not be legally usable in vehicles anyway. With widespread deployment of low-tier PCS systems, the one handset of choice will be the low-power, low-tier PCS pocket handset or voice/data communicator.

As discussed in earlier sections, it is quite feasible economically to cover highways between cities with low-tier systems, if the low-tier base stations have antennas with the same height and gain as used for cellular and high-tier PCS systems. (The range penalty for the lower power was noted earlier to be only on the order of 1/2, or about the same as the range penalty in going from 800 MHz cellular to 1.9 GHz high-tier PCS.)

There are approaches for integrating low-tier pocket phones or pocket communicators with high-tier vehicular cellular mobile telephones. The user's identity could be contained either in memory in the low-tier set or in a small smart card inserted into the set, as is a feature of the European GSM system. When entering an automobile, the small low-tier communicator or card could be inserted into a receptacle in a high-tier vehicular cellular set installed in the automobile.[7] The user's identity would then be transferred to the mobile set. "Car adapters" that have a cradle for a small cellular handset providing battery charging and connection to an outside antenna are quite common—e.g., in Sweden use of such adapters is commonplace. Thus, this concept has already evolved significantly, even for the disadvantaged cellular handsets when they are used in vehicles. The mobile set could then initiate a data exchange with the high-tier system, indicating that the user could now receive calls at that mobile set. This information about the user's location would then be exchanged between the network intelligence so that calls to the user could be correctly routed.[8] In this approach the radio sets are optimized for their specific environments, high-power, high-tier vehicular or low-power, low-tier pedestrian, as discussed earlier, and the network access and call routing is coordinated by the interworking of network intelligence. This approach does not compromise the design of either radio set or radio system. It places

[7]Inserting the small personal communicator in the vehicular set would also facilitate charging the personal communicator's battery.

[8]This is a feature proposed for FPLMTS in CCIR Rec. 687.

the burden on network intelligence technology that benefits from the large and rapid advances in computer technology.

The approach of using different communications devices for pedestrians than for vehicles is consistent with what has actually happened in other applications of technology in similarly different environments. For example, consider the case of audio cassette tape players. Pedestrians often carry and listen to small portable tape players with lightweight headsets (e.g., a Walkman).[9] When one of these people enters an automobile, he or she often removes the tape from the Walkman and inserts it into a tape player installed in the automobile. The automobile player has speakers that fill the car with sound. The Walkman is optimized for a pedestrian, whereas the vehicular-mounted player is optimized for an automobile. Both use the same tape, but they have separate tape heads, tape transports, audio preamps, etc. They do not attempt to share electronics. In this example, the tape cassette is the information-carrying entity similar to the user identification in the personal communications example discussed earlier. The main points are that the information is shared among different devices but that the devices are optimized for their environments and do not share electronics.

Similarly, a high-tier vehicular-cellular set does not need to share oscillators, synthesizers, signal processing, or even frequency bands or protocols with a low-tier pocket-size communicator. Only the information identifying the user and where he or she can be reached needs to be shared among the intelligence elements, e.g., routing logic, databases, and common channel signalling [26,29] of the infrastructure networks. This information exchange between network intelligence functions can be standardized and coordinated among infrastructure subnetworks owned and operated by different business entities (e.g., vehicular cellular mobile radio networks and intelligent low-tier PCS networks). Such standardization and coordination are the same as are required today to pass intelligence among local exchange networks and interexchange carrier networks.

Other Environments

Low-tier personal communications can be provided to occupants of airplanes, trains, and buses by installing compatible low-tier radio access ports inside these vehicles. The ports can be connected to high-power, high-tier vehicular cellular mobile sets or to special air-ground or satellite-based mobile communications sets. Intelligence between the internal ports and mobile sets could interact with cellular mobile, air-ground, or satellite networks in one direction, using protocols and spectrum allocated for that purpose, and with low-tier personal communicators in the other direction to exchange user identification and route calls to and from users inside these large vehicles. Radio isolation between the low-power units inside the large metal vehicles and low-power systems outside the vehicles can be ensured by using windows that are opaque to the radio frequencies. Such an approach also has been considered for automobiles, i.e., a radio port for low-tier personal communications connected to a cellular mobile set in a vehicle so that the low-tier personal communicator can access a high-tier cellular network. (This could be done in the United States using unlicensed PCS frequencies within the vehicle.)

Speech Quality Issues

All of the PCS and cordless telephone technologies that use CCITT standardized 32-kb/s ADPCM speech encoding can provide similar error-free speech distortion quality. This quality often is rated on a five-point subjective mean opinion score (MOS) with 5 excellent, 4 good, 3 fair, 2 poor, and 1 very poor. The error-free MOS of 32-kb/s ADPCM is about 4.1 and degrades very slightly with tandem encodings. Tandem encodings could be expected in going from a digital-radio PCS access link, through a network using analog transmission or 64-kb/s PCM, and back to another digital-radio PCS access link on the other end of the circuit. In contrast, a low-bit-rate (<10-kb/s) vocoder proposed for a digital cellular system was recently reported [54] to yield an MOS of 3.4 on an error-free link without speech-activity detection. This score

[9] Walkman is a registered trademark of Sony Corporation.

dropped to 3.2 when speech-activity detection was implemented to increase system capacity. This nearly one full point decrease on the five-point MOS score indicates significant degradation below accepted CCITT wireline speech distortion quality. Either almost half of the population must have rated it as poor or most of the population must have rated it as only fair. It should also be noted that these MOS scores may not reflect additional degradation that may occur in low-bit-rate speech encoding when the speech being encoded is combined with acoustical noise in a mobile environment, e.g., tire, wind, and engine noise in automobiles and street noise, background talking, etc., in handheld phone environments along streets and in buildings. Comments from actual users of low-bit-rate speech technology in acoustically noisy environments suggest that the MOS scores just quoted are significantly degraded in these real world environments. Waveform coders, e.g., ADPCM, are not likely to experience degradation from such background noise. In addition, the low-bit-rate speech encoding is not at all tolerant of the tandem speech encodings that will inevitably occur for PCS for many years. That is, when low-bit-rate speech is transcoded to a different encoding format, e.g., to 64 kb/s as is used in many networks or from an IS-54 phone on one end to a GSM or IS-95 phone on the other end, the speech quality deteriorates precipitously. Although this may not be a serious issue for a vehicular mobile user who has no choice other than not to communicate at all, it is likely to be a serious issue in an environment where a wireline telephone is available as an alternative. It is also less serious when there are few mobile-to-mobile calls through the network, but as wireless usage increases and digital mobile-to-mobile calls become commonplace, the marginal transcoded speech quality is likely to become a serious issue. These comments in this paragraph are generally applicable to speech encoding at rates of 13 kb/s or less. The predictions above about the seriousness of the speech coding quality issues have been proven to be true in virtually all cellular technologies [2,5]. The much touted speech quality of the original IS-95 CDMA 8 kb/s QCELP speech coder proved to be unacceptable to users.

The 8 kb/s speech coder was replaced with a 13 kb/s speech coder when IS-95 CDMA technology was deployed for high-tier PCS at 1.9 GHz. Many of the 800 MHz CDMA systems are converting or considering converting to the higher bit rate coder to improve their speech quality. This change was accompanied by a base station capacity reduction of 1/3, but the CDMA advocates seldom acknowledge this fact. The GSM effort to create a "half-rate" speech coder was redirected to develop an improved speech quality coder at 13 kb/s that would be less sensitive to acoustic background noise, bit errors, tandem encoding, etc. The quality of the new coder is better under some conditions. The 8 kb/s North American TDMA/IS-136 speech coder was redesigned to improve speech quality. Some improvement has been achieved. In all these revised coder cases, the issues discussed above in the original article still apply.

In the arena of transmission delay, the short-frame (2-ms) FDMA/TDD and TDMA technologies (e.g., CT-2 and WACS noted earlier) can readily provide single-radio-link round-trip delays of <10 ms and, perhaps, even <5 ms. The longer frame (10 ms and greater) cordless-phone TDMA technologies, e.g., DCT-900/CT-3/DECT and some ISM-band implementations, inherently have a single-link round-trip delay of at least 20 ms and can range 30–40 ms or more in some implementations. As mentioned earlier, the digital vehicular-cellular technologies with low-bit-rate speech encoding, bit interleaving, forward error-correction decoding, and relatively long frame time (~16–20 ms) result in single-link round-trip delays on the order of 200 ms, well over an order of magnitude greater than the short-frame technologies, and on the same order of magnitude as a single-direction synchronous satellite link. Even with the somewhat improved low bit rate speech coders noted previously, the transmission delay remains excessive for all of the cellular/high-tier PCS digital technologies. It should be noted that almost all United States domestic long-distance telephone circuits have been removed from such satellite links, and many international satellite links also are being replaced by undersea fiber links. These long-distance-circuit technology changes are made partially to reduce the perceptual impairment of long transmission delay.

New Technology

New technology, e.g., spread spectrum or CDMA, is sometimes offered as a solution to both the higher-tier cell site capacity and transmitter power issues. As these new technologies are pursued vigorously, however, it becomes increasingly evident that the early projections were considerably overoptimistic, that

the base station capacity will be about the same as other technologies [29], and that the high complexity will result in more, not less, power consumption.

With the continuing problems and delays in initial deployments, there is increasing concern throughout the industry as to whether CDMA is a viable technology for high-capacity cellular applications. With the passage of time, it is becoming more obvious that Viterbi was correct in his 1985 paper in which he questioned the use of spread spectrum for commercial communications [73].

The IS-95 proposal is considerably more technically sophisticated than earlier spread spectrum proposals. It includes fast feedback control of mobile transmitter power, heavy forward error correction, speech detection and speech-encoding rate adjustment to take advantage of speech inactivity, and multiple receiver correlators to latch onto and track resolvable multipath maxima [57]. The spreading sequence rate is 1.23 MHz.

The near–far problem is addressed directly and elegantly on the uplink by a combination of the fast-feedback power control and a technique called soft handoff that permits the instantaneous selection of the best paths between a mobile and two cell sites. Path selection is done on a frame-by-frame basis when paths between a mobile and the two cell sites are within a specified average level (perhaps 6–10 dB) for each other. This soft handoff provides a form of macroscopic diversity [10] between pairs of cell sites when it is advantageous. Increasing capacity by soft handoff requires precise time synchronization (on the order of a microsecond) among all cell sites in a system. An advantage of this proposal is that frequency coordination is not needed among cell sites since all sites can share a frequency channel. Coordination of the absolute time delays of spreading sequences among cell sites is required, however, since these sequence delays are used to distinguish different cell sites for initial access and for soft handoff. Also, handoff from one frequency to another is complicated.

Initially, the projected cell site capacity of this CDMA system, determined by mathematical analysis and computer simulation of simplified versions of the system, was ×20 to ×40 that of the analog AMPS, with a coverage criterion of 99% of the covered area [11]. However, some other early estimates [83] suggested that the factors were more likely to be ×6 to ×8 of AMPS.

A limited experiment was run in San Diego, California, during the fourth quarter of 1991 under the observation of cellular equipment vendors and service providers. This experiment had 42–62 mobile units in fewer than that many vehicles,[10] and four or five cell sites, one with three sectors. Well over half of the mobiles needed to provide the interference environment for system capacity tests were simulated by hardware noise simulation by a method not yet revealed for technical assessment. Estimates of the cell site capacity from this CDMA experiment center around ×10 that of AMPS [12,54][11] with coverage criteria <99%, perhaps 90–95%, and with other capacity estimates ranging between ×8 and ×15.

This experiment did not exercise several potential capacity-reducing factors; the following, for example:

1. Only four cells participated in capacity tests. The test mobiles were all located in a relatively limited area and had limited choices of cell sites with which to communicate for soft handoffs. This excludes the effects of selecting a strong cell-site downlink for soft handoff that does not have the lowest uplink attenuation because of uncorrelated uplink and downlink multipath fading at slow vehicle speeds [3].
2. The distribution of time-dispersed energy in hilly environments like San Diego usually is more concentrated around one or two delays than is the dispersed energy scattered about in heavily built-up urban areas like downtown Manhattan [13,14], or Chicago. Energy concentrated at one or two delays is more fully captured by the limited number of receiver correlators than is energy more evenly dispersed in time.
3. Network delay in setting up soft handoff channels can result in stronger paths to other cell sites than to the one controlling uplink transmitter power. This effect can be more pronounced when coming out of shadows of tall buildings at intersections in heavily built-up areas. The effect will not occur as frequently in a system with four or five cell sites as it will in a large, many cell site system.

[10]Some vehicles contained more than one mobile unit.

[11]AT&T stated that the San Diego data supported a capacity improvement over analog cellular of at least ×10.

All of these effects and others [4] will increase the interference in a large system, similarly to the increase in interference that results from additional mobiles and, thus, will decrease cell site capacity significantly below that estimated in the San Diego trial. Factors like these have been shown to reduce the San Diego estimate of ×10 to an expected CDMA capacity of ×5 or ×6 of analog AMPS [3,4,13]. This further reduction in going from a limited experiment to a large-scale system in a large metropolitan area is consistent with the reduction already experienced in going from simplified theoretical estimates to the limited experiment in a restricted environment. There are many "small" effects that decrease the performance and, thus, the base station capacity of actual mobile CDMA systems compared to calculations of performance based on ideal assumptions—e.g., stationary channel and additive white Gaussian noise analysis. Some of these effects, in addition to those noted above, included inaccuracy in assigning correlators ("fingers") to time varying statistically nonstationary multiple "paths" at different delays when many "paths" exist, and lack of diversity because many environments have delay spreads less than the resolution of correlators.

The San Diego trial also indicated a higher rate of soft(er) handoffs [54] between antenna sectors at a single cell site than expected for sectors well isolated by antenna patterns. This result suggests a lower realizable sectorization gain because of reflected energy than would be expected from more idealized antennas and locations. This could further reduce the estimated cell site capacity of a large-scale system.

Even considering the aforementioned factors, capacity increases of ×5 or ×6 are significant. These estimates, however, are consistent with the factor of ×3 obtained from low-bit-rate (<10-kb/s) speech coding and the ×2 to ×2.5 obtained by taking advantage of speech pauses. These factors result in an expected increase of ×6–7.5, with none of these speech-processing-related contributions being directly attributable to the spread-spectrum processing in CDMA. These results are consistent with the factor of ×6–8 estimate made earlier [83] and are not far from the factor of ×8 quoted recently [86].

Thus, it is clear that new high-complexity, high-tier technology will not be a substitute for low-complexity, low-power, low-tier PCS.

Since the writing for the first edition of this handbook, IS-95 based CDMA systems have been deployed and are providing service in several parts of the world, e.g., the U.S.A., Korea, and Hong Kong. The "high priests" and the "disciples" of CDMA continue to espouse their "belief" in the superiority of their CDMA technology. Good hard data on system performance is still scarce and of questionable reliability. However, there is sufficient evidence from some "usually reliable sources" and from "conversations in the halls" to confirm that, not only does CDMA not perform as well as was originally promised, but also in many places it is likely to be providing performance that is inferior to the other high-tier digital technologies. In spite of the continuing dismal failures of CDMA technology to meet the early claims for it, a significant fraction of the wireless industry appears to be committed to pursuing variations of CDMA for future applications. At this time it is not obvious how many more redesign and reprogramming failures, or perhaps even failures of manufacturing companies and service providers will be needed to overcome the religiously fanatic attachment of a part of the industry to CDMA. It appears that some are heavily committed again to "sweep on to the grand fallacy." Viterbi may even have been over optimistic about CDMA in 1985 [73], but he was correct in the way it attracts advocates.

IS-95 CDMA has failed to satisfy any of the major claims made for it in the late 1980s and early 1990s. The listing in Table 67.7 and the following discussion cites specific claims that have NOT been met now that reality has set in.

Much of the original CDMA hype was focused on undefined cell site capacity. Information from many sources (e.g., [9], Benjamin and Lusignan, 1997—Benjamin is with Lucent Technologies; Lusignan with Stanford University; [40,49,69], etc.) suggest that with 13 kb/s speech coding, "fully loaded" systems may provide base station capacities on the order of AMPS ×3 or perhaps ×4. However, the meaning of fully loaded is not well defined. If it is taken to be at a very high call dropping rate (sometimes said to be as high as 40%!), perhaps even these low factors are again unrealistically inflated. A usable system loading must provide an acceptable call dropping rate, perhaps 1% or less. If loading is reduced to provide sufficient operating margin to prevent too many dropped calls, perhaps the usable base station capacity

TABLE 67.7 CDMA Scoresheet

Early Claim	Reality
Easy to install and expand (no frequency planning)	Many "parameters" to adjust; adjustment is very difficult; many software revisions required
"Capacity" of AMPS ×20 (with high expectation of AMPS ×40)	AMPS ×3 or 4; AMPS × <3? (at unspecified but likely high call dropping rate) (continuously revised downward over time)
No more dropped calls	More dropped calls than other technologies (some reports say up to 40% dropped calls when system is loaded!)
No problem with interference (CDMA "loves" interference)	Can't live with interference from AMPS or other sources
Greater range and coverage than other technologies	System coverage "breathes" as loading increases; range often is less than other technologies
"Toll quality" 8 kb/s QCELP speech	Original 8 kb/s QCELP speech quality worse than VSELP in TDMA; required change to 13 kb/s coder (change reduced "capacity")

may be even significantly less than AMPS ×3. (Note: Publicly presented CDMA claims have been consistently inflated in the past. Why should we be any more inclined to believe them now?) The capacity situation is so bad that the CDMA crowd doesn't even want to discuss this topic anymore. They now attempt to shift attention to the fact that there are operating CDMA systems, and try to brush aside the fact CDMA performance is so very poor. Low capacity has a significant negative impact on the economics of systems; high call dropping rates have a significant negative impact on service quality.

Another big early claim for CDMA was that it was easy to install and expand, that is, that it needed little engineering and planning for installation because it didn't require frequency reuse and power coordination among base stations. In reality, many parameters require adjustment and critical thresholds require setting for a system to provide even the low capacity numbers cited above. Many software revisions have been made in attempts to cope with fundamental problems that keep appearing and reappearing in new deployments and expansions. Deployments and service starts have been consistently late by many months and even years. [40,69, and others]. A partial list of late CDMA deployments that required many months or even years of "parameter adjustment" and/or "software revisions" and/or "redesign" include those in Seattle, WA, U.S.A. (800 MHz); Los Angeles, CA, U.S.A. (800 MHz); Hong Kong, China (800 MHz); Trenton, NJ, U.S.A.; Seoul, Korea (800 MHz); Houston and Dallas, TX, U.S.A. (1.9 GHz); etc. As these systems were deployed a consistent pattern developed. Announcements would concentrate on a projected future service date, e.g., Seattle, 1Q 1994. When that date approached, news focused on yet another service date elsewhere, e.g., after Seattle was Los Angeles, to divert attention from the deployment problems causing large slips in earlier announcements. Sometimes when service date slips became excessive, sizes of deployed areas were revised downward, and "trials" were substituted for service starts, e.g., Los Angeles. This pattern continued through the early deployments and service eventually was started on systems, even though their performance was significantly below earlier claims. Excuses for slips were often vague, for example: software problems (e.g., Seattle, Hong Kong, and Korea); adjusting parameters (e.g., Seattle, Houston, and Dallas); too much interference from AMPS (e.g., Los Angeles); not enough multipath! (e.g., Seattle). These sound like descriptions of attempts to overcome some of the fundamental problems in mobile CDMA systems. In any event it is obvious that CDMA is not easy to install and is likely more difficult to adjust than other technologies.

The CDMA speech quality issue was discussed earlier in the Speech Quality Issues section. It is just another example of unrealistic initial claims.

Throughout the CDMA debate over almost 10 years, when unrealistic claims were questioned and even when sound technical analysis was cited by more conservative members of the wireless technical

community, including this author, these members were immediately branded as heretics and their concerns were dismissed with rhetoric by the high priests and their disciples. (In the technical world, "never have so many been misled for so long by so few!") However, as deployments have proceeded and reality set in, even the more conservative initial assessments have, unfortunately, turned out to be somewhat optimistic. Thus, even the conservative heretics in the community now appear to be in the unfortunate position of having also been overoptimistic in their initial estimates of the performance of IS-95 CDMA.

Statistical Multiplexing, Speech Activity, CDMA, and TDMA

Factors of ×2–2.5 have been projected for the capacity increase possible by taking advantage of pauses in speech. It has been suggested that implementing statistical multiplexing is easier for CDMA systems because it is sometimes thought to be time consuming to negotiate channels for speech spurts for implementation in TDMA systems. The most negative quality-impacting factor in implementing statistical multiplexing for speech, however, is not in obtaining a channel when needed but is in the detection of the onset of speech, particularly in an acoustically noisy environment. The effect of clipping at the onset of speech is evident in the MOS scores noted for the speech-activity implementation in the United States cellular CDMA proposal discussed earlier (i.e., an MOS of 3.4 without statistical multiplexing and of 3.2 with it). The degradation in MOS can be expected to be even greater for encoding that starts with a higher MOS, e.g., 32-kb/s ADPCM.

It was noted earlier that the proposed cellular CDMA implementation was ×10 as complex as the proposed WACS wireless access for personal communications TDMA implementation. From earlier discussion, the CDMA round-trip delay approaches 200 ms, whereas the short 2-ms-frame TDMA delay is <10-ms round trip. It should be noted that the TDMA architecture could permit negotiation for time slots when speech activity is detected. Since the TDMA frames already have capability for exchange of signalling data, added complexity for statistical multiplexing of voice could readily be added within less than 200 ms of delay and less than ×10 in complexity. That TDMA implementation supports 8 circuits at 32 kb/s or 16 circuits at 16 kb/s for each frequency. These are enough circuits to gain benefit from statistical multiplexing. Even more gain could be obtained at radio ports that support two or three frequencies and, thus, have 16–48 circuits over which to multiplex.

A statistical multiplexing protocol for speech and data has been researched at Rutgers WINLAB [39]. The Rutgers packet reservation multiple access (PRMA) protocol has been used to demonstrate the feasibility of increasing capacity on TDMA radio links. These PRMA TDMA radio links are equivalent to slotted ALOHA packet-data networks. Transmission delays of less than 50 ms are realizable. The capacity increase achievable depends on the acceptable packet-dropping ratio. This increase is soft in that a small increase in users causes a small increase in packet-dropping ratio. This is analogous to the soft capacity claimed for CDMA.

Thus, for similar complexity and speech quality, there appears to be no inherent advantage of either CDMA or TDMA for the incorporation of statistical multiplexing. It is not included in the personal communications proposal but is included in cellular proposals because of the different speech-quality/complexity design compromises discussed throughout this chapter, not because of any inherent ease of incorporating it in any particular access technology.

High-Tier to Low-Tier or Low-Tier to High-Tier Dual Mode

Industry and the FCC in the United States appear willing to embrace multimode handsets for operating in very different high-tier cellular systems, e.g., analog FM AMPS, TDMA IS-54, and CDMA IS-95. Such sets incur significant penalties for dual mode operation with dissimilar air interface standards and, of course, incur the high-tier complexity penalties.

It has been suggested that multimode high-tier and low-tier handsets could be built around one air-interface standard, for example, TDMA IS-54 or GSM. When closely spaced low-power base stations

were available, the handset could turn off unneeded power-consuming circuitry, e.g., the multipath equalizer. The problem with this approach is that the handset is still encumbered with power-consuming and quality-reducing signal processing inherent in the high-tier technology, e.g., error correction decoding and low-bit-rate speech encoding and decoding.

With widespread and successful deployment of GSM, perhaps an initial dual-mode strategy of GSM and PACS could be desirable. Of all the high-tier and low-tier technologies, GSM and PACS appear most compatible for dual-mode handset and base station implementation. They are considerably more compatible than the dual-mode combinations, e.g., AMPS/IS-54/IS-136 and AMPS/IS-95, already implemented.

An alternative dual-mode low-tier, high-tier system based on a common air-interface standard can be configured around the low-tier PACS/WACS system, if such a dual-mode system is deemed desirable in spite of the discussion in this chapter. The range of PACS can readily be extended by increasing transmitter power and/or the height and gain of base station antennas. With increased range, the multipath delay spread will be more severe in some locations [21,22,33]. Two different solutions to the increased delay spread can be employed, one for the downlink and another for the uplink. The PACS radio-link architecture has a specified bit sequence, i.e., a unique word, between each data word on the TDM downlink [7,28]. This unique word can be used as a training sequence for setting the tap weights of a conventional equalizer added to subscriber sets for use in a high-tier PACS mode. Since received data can be stored digitally [62,65], tap weights can be trimmed, if necessary, by additional passes through an adaptive equalizer algorithm, e.g., a decision feedback equalizer algorithm.

The PACS TDMA uplink has no unique word. The high-tier uplink, however, will terminate on a base station that can support greater complexity but still be no more complex than the high-tier cellular technologies. Research at Stanford University has indicated that blind equalization, using constant-modulus algorithms (CMA), [58,70], can be effective for equalizing the PACS uplink. Techniques have been developed for converging the CMA equalizer on the short TDMA data burst.

The advantages of building a dual-mode high-tier, low-tier PCS system around the low-tier PACS air-interface standard follow.

1. The interface can still support small low-complexity, low-power, high-speech-quality low-tier handsets.
2. Both data and voice can be supported in a PACS personal communicator.
3. In high-tier low-tier dual mode PACS sets, circuits used for low-tier operation will also be used for high-tier operation, with additional circuits being activated only for high-tier operation.
4. The flexibility built into the PACS radio link to handle different data rates from 8 kb/s to several hundred kb/s will be available to both modes of operation.

67.8 Infrastructure Networks

It is beyond the scope of this chapter to consider the details of PCS network infrastructures. There are, however, perhaps as many network issues as there are wireless access issues discussed herein [26,32,43,81]. With the possible exception of the self-organizing WLANS, wireless PCS technologies serve as access technologies to large integrated intelligent fixed communications infrastructure networks.

These infrastructure networks must incorporate intelligence, i.e., database storage, signalling, processing and protocols, to handle both small-scale mobility, i.e., handoff from base station to base station as users move, and large-scale mobility, i.e., providing service to users who roam over large distances, and perhaps from one network to another. The fixed infrastructure networks also must provide the interconnection among base stations and other network entities, e.g., switches, databases, and control processors. Of course, existing cellular mobile networks now contain or are incorporating these infrastructure network capabilities. Existing cellular networks, however, are small compared to the expected size of future high-tier and low-tier PCS networks, e.g., 20 million cellular users in the United States compared with perhaps 100 million users or more each in the future for high-tier and low-tier PCS.

Several other existing networks have some of the capabilities needed to serve as access networks for PCS. Existing networks that could provide fixed base station interconnection include:

- Local exchange networks that could provide interconnection using copper or glass-fiber distribution facilities
- Cable TV networks that could provide interconnection using new glass-fiber and coaxial-cable distribution facilities
- Metropolitan fiber digital networks that could provide interconnection in some cities in which they are being deployed

Networks that contain intelligence, e.g., databases, control processors, and signalling that is suitable or could be readily adapted to support PCS access include:

- Local exchange networks that are equipped with signalling system 7 common channel signalling (SS7 CCS), databases, and digital control processors
- Interexchange networks that are similarly equipped

Data networks, e.g., the internet, could perhaps be adapted to provide the needed intelligence for wireless data access, but they do not have the capacity needed to support large voice/data wireless low-tier PCS access.

Many entities and standards bodies worldwide are working on the access network aspects of wireless PCS. The signalling, control processing, and database interactions required for wireless access PCS are considerably greater than those required for fixed place-to-place networks, but that fact must be accepted when considering such networks.

Low-tier PCS, when viewed from a cellular high-tier paradigm, requires much greater fixed interconnection for the much closer spaced base stations. When viewed from a cordless telephone paradigm of a base unit for every handset and, perhaps, several base units per wireline, however, the requirement is much less fixed interconnection because of the concentration of users and trunking that occurs at the multiuser base stations. One should remember that there are economical fixed wireline connections to almost all houses and business offices in the United States now. If wireless access displaces some of the wireline connections, as expected, the overall need for fixed interconnection could decrease!

67.9 Conclusion

Wireless personal communications embraces about seven relatively distinct groups of tetherless voice and data applications or services having different degrees of mobility for operation in different environments. Many different technologies and systems are evolving to provide the different perceived needs of different groups. Different design compromises are evident in the different technologies and systems. The evidence suggests that the evolutionary trajectories are aimed toward at least three large groups of applications or services, namely, high-tier PCS (current cellular radio), high-speed wireless local-area networks (WLANS), and low-tier PCS (an evolution from several of the current groups). It is not clear to what extent several groups, e.g., cordless telephones, paging, and wide-area data, will remain after some merging with the three large groups. Major considerations that separate current cellular technologies from evolving low-tier low-power PCS technologies are speech quality, complexity, flexibility of radio-link architecture, economics for serving high-user-density or low-user-density areas, and power consumption in pocket carried handsets or communicators. High-tier technologies make use of large complex expensive cell sites and have attempted to increase capacity and reduce circuit costs by increasing the capacity of the expensive cell sites. Low-tier technologies increase capacity by reducing the spacing between base stations, and achieve low circuit cost by using low-complexity, low-cost base stations. The differences between these approaches result in significantly different compromises in circuit quality and power consumption in pocket-sized handsets or communicators. These kinds of differences also can be seen in evolving wireless systems optimized for data. Advantages of the low-tier PACS/WACS

technology are reviewed in the chapter, along with techniques for using that technology in high-tier PCS systems.

References

1. Afrashteh, A., Sollenberg, N.R., and Chukurov, D.D., Signal to interference performance for a TDMA portable radio link. *IEEE VTC'91,* St. Louis, MO, 19–22, May, 1991.
2. Andrews, E.L., When digital doesn't always mean clear, *New York Times,* p. C1, 4, Jun. 26, 1995.
3. Ariyavisitakul, S., SIR-based power control in a CDMA system. *IEEE GLOBECOM'92,* Orlando, FL, Paper 26.3, Dec. 6–9, to be published in *IEEE Trans. Comm.,* 1992.
4. Ariyavisitakul, S., et al., private communications.
5. Baugh, C.R., Wireless quality: a ticking time bomb, *Telephony,* 232(9), 50–58, Mar. 3, 1997.
6. Baugh, C.R., Laborde, E., Pandey, V., and Varma, V., Personal access communications system: fixed wireless local loop and mobile configurations and services, *Proc. IEEE,* 1498–1506, July, 1998.
7. Bellcore Technical Advisories. Generic framework criteria for universal digital personal communications systems (PCS). FA-TSY-001013(1) March and FA-NWT-001013(2) Dec., and Tech. Ref. 1993. Generic criteria for version 0.1 wireless access communications systems (WACS). (1) Oct. rev. 1, Jun. 1994.
8. BellSouth Services Inc., Quarterly progress report number 3 for experimental licenses KF2XFO and KF2XFN. To the Federal Communications Commission, Nov. 25, 1991.
9. Benjamin, B. and Lusignan, B., CDMA vs. TDMA, Stanford Univ. Center for Telecom. Symposium, The Acceleration of World Wide Wireless Communication, May 21, 1997.
10. Bernhardt, R.C., Macroscopic diversity in frequency reuse radio systems. *IEEE J. Sel. Areas in Comm.,* SAC-5, 862–870, Jun. 1987.
11. Cellular Telecommunications Industries Association. CDMA digital cellular technology open forum. Jun. 6, 1989.
12. Cellular Telecommunications Industries Association. CTIA Meeting, Washington, D.C., 5–6, Dec. 1991.
13. Chang, L.F., Dispersive fading effects in CDMA radio systems. *Elect. Lett.,* 28(19), 1801–1802, 1992.
14. Chang, L.F. and Ariyavisitakul, S., Performance of a CDMA radio communications system with feedback power control and multipath dispersion. *IEEE GLOBECOM'91,* Phoenix, AZ, 1017–1021, Dec. 1991.
15. Chuang, J.C.-I., Performance issues and algorithms for dynamic channel assignment. *IEEE JSAC,* Aug. 1993.
16. Chuang, J.C.-I., Performance limitations of TDD wireless personal communications with asynchronous radio ports. *Electron. Lett.,* 28, 532–533, Mar. 1992.
17. Chuang, J.C.-I., Comparison of coherent and differential detection of BPSK and QPSK in a quasistatic fading channel. *IEEE Trans. Comm.,* 565–576, May 1990.
18. Chuang, J.C.-I. and Sollenberger, N.R., Burst coherent detection with robust frequency and timing estimation for portable radio communications. *IEEE GLOBECOM'88,* Hollywood, FL, 28–30, Nov. 1988.
19. Chuang, J.C.-I., Sollenberger, N.R., and Cox, D.C., A pilot based dynamic channel assignment scheme for wireless access TDMA/FDMA systems. *Proc. IEEE ICUPC'93,* Ottawa, Canada, Oct. 12–15, 706–712, 1994. *Intl. J. Wireless Inf. Net.,* 1(1), 37–48, 1993.
20. Cook, C.I., Development of air interface standards for PCS. *IEEE Personal Comm.,* 4th Quarter, 30–34, 1994.
21. Cox, D.C., Delay-Doppler characteristics of multipath propagation at 910 MHz in a suburban mobile radio environment. *IEEE Trans. on Antennas and Propagation,* 625–635, Sep. 1972.
22. Cox, D.C., Multipath delay spread and path loss correlation for 910 MHz urban mobile radio propagation. *IEEE Trans. on Veh. Tech.,* 340–344, Nov. 1977.

23. Cox, D.C., Universal portable radio communications. *IEEE Trans. on Veh. Tech.,* 117–121, Aug. 1985.
24. Cox, D.C., Research toward a wireless digital loop. *Bellcore Exchange,* 2, 2–7, Nov./Dec. 1986.
25. Cox, D.C., Universal digital portable radio communications. *Proc. IEEE,* 75, 436–477, Apr. 1987.
26. Cox, D.C., Portable digital radio communications—An approach to tetherless access. *IEEE Comm. Mag.,* 30–40, Jul. 1989.
27. Cox, D.C., Personal communications—A viewpoint. *IEEE Comm. Mag.,* 8–20, Nov. 1990.
28. Cox, D.C., A radio system proposal for widespread low-power tetherless communications. *IEEE Trans. on Comm.,* 324–335, Feb. 1991.
29. Cox, D.C., Wireless network access for personal communications. *IEEE Comm. Mag.,* 96–115, Dec. 1992.
30. Cox, D.C., Wireless loops: what are they. *Intl. J. Wireless Inf. Net.,* 3(3), Plenum Press, 1996.
31. Cox, D.C., Arnold, H.W., and Porter, P.T., Universal digital portable communications—a system perspective. *IEEE JSAC,* JSAC-5, 764–773, Jun. 1987.
32. Cox, D.C., Gifford, W.G., and Sherry, H., Low-power digital radio as a ubiquitous subscriber loop. *IEEE Comm. Mag.,* 92–95, Mar. 1991.
33. Devasirvatham, D.M.J., Radio propagation studies in a small city for universal portable communications. *IEEE VTC'88,* Conference Record. Philadelphia, PA, 100–104, Jun. 15–17, 1988.
34. Devasirvatham, D.M.J., Radio propagation studies in a small city for universal portable communications, *IEEE VTC'88,* Philadelphia, PA, 100–104, Jun. 15–17, 1988.
35. Devasirvatham, D.M.J., et al., Radio propagation measurements at 850 MHz, 1.7 GHz and 4 GHz inside two dissimilar office buildings. *Elect. Lett.,* 26(7), 445–447, 1990a.
36. Devasirvatham, D.M.J., et al., Multi-frequency radiowave propagation measurements in the portable radio environment. *IEEE ICC'90,* 1334–1340, Apr. 1990b.
37. Devasirvatham, D.M.J., Murray, R.R., Arnold, H.W., and Cox, D.C., Four-frequency CW measurements in residential environments for personal communications, symposium on personal indoor and mobile radio communications (PIMRC'93), Yokohama, Japan, 201–205, Sep. 9–11, 1993.
38. Goldhirsh, J. and Vogel, W., Mobile satellite system fade and statistics for shadowing and multipath from roadside trees at UFH and L-band, *IEEE Transactions on Antennas and Propagation,* 37, Apr. 1989.
39. Goodman, D.J., Trends in cellular and cordless communications. *IEEE Comm. Mag.,* 31–40, Jun. 1991.
40. Hardy, Q., Jacob's Patter, The Wall Street Journal, Friday, Sep. 6, 1996.
41. Hattori, T., PHS system deployment success story, Stanford Univ. Center for Telecom. Symposium, May 21, 1997.
42. Hinderling, J., et al., CDMA mobile station modem ASIC. *IEEE CICC 92,* Boston, MA, May. *Intl. J. Wireless Info. Net.,* 1(1), 37–48, Jan. 1994.
43. Jabbari, B., et al., Network issues for wireless personal communications. *IEEE Comm. Mag.,* 88–98, Jan. 1995.
44. Katz, R.H., Adaptation and mobility in wireless information systems. *IEEE Personal Comm.,* 1st Quarter, 6–17, 1994.
45. Kim, B.-J. and Cox, D.C., Blind equalization for short burst communications over frequency selective wireless channels, 1997 *IEEE VTC'97,* Phoenix, AZ, 2, 544–548, May 4–7, 1997a.
46. Kim, B.-J. and Cox, D.C., Blind sequence estimation for short burst communications over time-varying wireless channels, *IEEE ICUPC'97,* San Diego, CA, 2, 713–717, Oct. 12–16, 1997b.
47. Kim, B.-J. and Cox, D.C., Blind diversity combining equalization for short burst wireless communications, *IEEE Globecom'97,* Phoenix, AZ, 3, 1163–1167, Nov. 3–8, 1997c.
48. Kim, B.-J., Blind equalization for short burst wireless communications, Stanford University Ph.D. Thesis, Jan. 1998.
49. Korea Mobile Telecom, Article on CDMA, Mobile Communications International, Feb. 1997.
50. Padgett, J.E., Hattori, T., and Gunther, C., Overview of wireless personal communications. *IEEE Comm. Mag.,* 28–41, Jan. 1995.

51. Patel, V., Personal Access Communications System, Stanford Univ. Center for Telecom. Symposium, May 21, 1997.
52. Phillips, B.W., Broadband in the Local Loop, Telecommunications, 37–42, Nov. 1994.
53. PHS, World Congress, Singapore, Nov. 18–20, 1996.
54. Qualcomm Technology Forum, Open meeting on status of CDMA technology and review of San Diego experiment. San Diego, CA, Jan. 16–17, 1992.
55. Ranade, A., Local access radio interference due to building reflections, *IEEE Transaction on Communications,* 70–74, Jan. 1989.
56. Rizzo, J.F. and Sollenberger, N.R., Multiuser Wireless Access, *IEEE Personal Communications Magazine,* 18–30, Jun. 1995.
57. Salmasi, A. and Gilhousen, K.S., On the system design aspects of code division multiple access (CDMA) applied to digital cellular and personal communications networks. *IEEE VTC'91,* St. Louis, MO, 57–62, May 1991.
58. Sato, Y., A method of self-recovering equalization for multilevel amplitude modulation systems. *IEEE Trans. on Comm.,* 679–682, Jun. 1975.
59. Schneideman, R., Spread spectrum gains wireless applications. *Microwaves and RF,* 31–42, May 1992.
60. Seidel, S.Y. and Arnold, H.W., Propagation measurements of 28 GHz to investigate the performance of local multipoint distribution services (LMDS), *IEEE Globecom'95,* 1995.
61. Sollenberger, N.R. and Chuang, J.C-I., Low overhead symbol timing and carrier recovery for TDMA portable radio systems. Third nordic sem. digital land mobile radio comm., Copenhagen, Denmark, 13–15, Sep. 1988.
62. Sollenberger, N.R., An experimental VLSI implementation of low-overhead symbol timing and frequency offset estimation for TDMA portable radio applications. *IEEE Globecom'90,* San Diego, CA, 1701–1711, Dec. 1990.
63. Sollenberger, N.R., An experimental TDMA modulation/demodulation CMOS VLSI chip-set. *IEEE CICC'91,* San Diego, CA, May 12–15, 1991.
64. Sollenberger, N.R. and Afrashteh, A., An experimental low-delay TDMA portable radio link. *Wireless'91,* Calgary, Canada, Jul. 8–10, 1991.
65. Sollenberger, N.R. and Chuang, J.C.-I., Low-overhead symbol timing and carrier recovery for TDMA portable radio systems. *IEEE Trans. on Comm.,* 1886–1892, Oct. 1990.
66. Sollenberger, N.R., et al., Architecture and implementation of an efficient and robust TDMA frame structure for digital portable communications. *IEEE Trans. Veh. Tech.,* 40, 250–260, Feb. 1991.
67. Steele, R., Deploying personal communications networks. *IEEE Comm. Mag.,* 12–15, Sep. 1990.
68. Steer, D.G., Coexistence and access etiquette in the United States unlicensed PCS band. *IEEE Personal Comm. Mag.,* 4th Quarter, 36–43, 1994.
69. Titch, S., Blind Faith, Telephony, Sep. 8, 1997 and Global Telephony, Intertec Publishing Corp., Oct. 30, 1997.
70. Treichler, J.R. and Agee, B.G., A new approach to multipath correction of constant modulus signals, *IEEE Trans. on Acoustics, Speech and Signal Processing,* 459–472, Apr. 1983.
71. Tuthill, J.P., Granger, B.S., and Wurtz, J.L., Request for a pioneer's preference. Before the Federal Communications Commission, Pacific Bell submission for FCC General Docket No. 90–314, RM-7140, and RM-7175, May 4, 1992.
72. Varma, V.K., et al., Performance of sub-band and RPE coders in the portable communication environment. *Fourth Intl. Conf. on Land Mobile Radio,* Coventry, UK, 221–227, Dec. 14–17, 1987.
73. Viterbi, A.J., When not to spread spectrum—A sequel. *IEEE Comm. Mag.,* 12–17, Apr. 1985.
74. Weseloh, C., Solutions for Wireless Communications, Stanford University Center for Telecommunications, Symposium, Stanford, CA, May 16, 1995.
75. Wong, A., Regulating public wireless networks. Workshop on lightwave, wireless and networking technologies, Chinese Univ. of Hong Kong, Hong Kong, Aug. 24, 1994.
76. Wong, P., Low-power, low-complexity diversity combining algorithm and VLSI circuit chip for handheld PCS receivers, Stanford University Ph.D. Thesis, Apr. 1997.

77. Wong, P. and Cox, D.C., Low-complexity cochannel interference cancellation and macroscopic diversity for high capacity PCS, *IEEE ICC'95 Digest,* Seattle, WA, 852–857, Jun. 19–22, 1995.
78. Wong, P. and Cox, D.C., Low-complexity diversity combining algorithms and circuit architectures for cochannel interference cancellation and frequency selective fading mitigation, *IEEE ICC'96 Digest,* Dallas TX, Jun. 23–27, 1996a.
79. Wong, P. and Cox, D.C., Low-complexity diversity combining algorithms and circuit architectures for cochannel interference cancellation and frequency-selective fading mitigation, *IEEE Trans. on Comm.,* 44(9), 1107–1116, Sep. 1996b.
80. Wong, P. and Cox, D.C., Low-complexity interference cancellation and macroscopic diversity for high capacity PCS, *IEEE Trans. on Veh. Tech.,* 124–132, Feb. 1998.
81. Zaid, M., Personal mobility in PCS. *IEEE Personal Comm.,* 4th Quarter, 12–16, 1994.
82. Special issue on advanced mobile phone service (AMPS). *Bell System Tech. J.,* (BSTS). 58, Jan. 1979.
83. CDMA capacity seen as less than advertised. *Adv. Wireless Comm.,* (AWC). 1–2, Feb. 6, 1991.
84. Bellcore's Wireless Prototype. *Microcell News,* (MN). 5–6, Jan. 25, 1992.
85. Bellcore PCS phone shines in Bellsouth test; others fall short. *Adv. Wireless Comm.,* (AWC). 2–3, Feb. 19, 1992a.
86. TDMA accelerates, but CDMA could still be second standard. *Adv. Wireless Comm.,* (AWC). 6, Mar. 4, 1992b.
87. Special issue on wireless personal communications. *IEEE Comm. Mag.,* Jan. 1995.

68

Modulation Methods

Gordon L. Stüber
Georgia Institute of Technology

68.1 Introduction

Modulation is the process where the message information is added to the radio carrier. Most first generation cellular systems such as the advanced mobile telephone system (AMPS) use analog **frequency modulation** (**FM**), because analog technology was very mature when these systems were first introduced. Digital modulation schemes, however, are the obvious choice for future wireless systems, especially if data services such as wireless multimedia are to be supported. Digital modulation can also improve spectral efficiency, because digital signals are more robust against channel impairments. Spectral efficiency is a key attribute of wireless systems that must operate in a crowded radio frequency spectrum.

To achieve high spectral efficiency, modulation schemes must be selected that have a high **bandwidth efficiency** as measured in units of bits per second per Hertz of bandwidth. Many wireless communication systems, such as cellular telephones, operate on the principle of frequency reuse, where the carrier frequencies are reused at geographically separated locations. The link quality in these systems is limited by cochannel interference. Hence, modulation schemes must be identified that are both bandwidth efficient and capable of tolerating high levels of cochannel interference. More specifically, digital modulation techniques are chosen for wireless systems that satisfy the following properties.

Compact Power Density Spectrum: To minimize the effect of adjacent channel interference, it is desirable that the power radiated into the adjacent channel be 60–80 dB below that in the desired channel. Hence, modulation techniques with a narrow main lobe and fast rolloff of sidelobes are desirable.

Good Bit-Error-Rate Performance: A low-bit-error probability should be achieved in the presence of cochannel interference, adjacent channel interference, thermal noise, and other channel impairments, such as fading and intersymbol interference.

Envelope Properties: Portable and mobile applications typically employ nonlinear (class C) power amplifiers to minimize battery drain. Nonlinear amplification may degrade the bit-error-rate performance of modulation schemes that transmit information in the amplitude of the carrier. Also, spectral shaping is usually performed prior to up-conversion and nonlinear amplification. To prevent the regrowth

of spectral sidelobes during nonlinear amplification, the input signal must have a relatively constant envelope.

A variety of digital modulation techniques are currently being used in wireless communication systems. Two of the more widely used digital modulation techniques for cellular mobile radio are $\pi/4$ phase-shifted quadrature **phase shift keying** ($\pi/4$-QPSK) and **Gaussian minimum shift keying** (**GMSK**). The former is used in the North American IS-54 digital cellular system and Japanese Personal Digital Cellular (PDC), whereas the latter is used in the global system for mobile communications (GSM system). This chapter provides a discussion of these and other modulation techniques that are employed in wireless communication systems.

68.2 Basic Description of Modulated Signals

With any modulation technique, the bandpass signal can be expressed in the form

$$s(t) = \mathrm{Re}\{v(t)e^{j2\pi f_c t}\} \tag{68.1}$$

where $v(t)$ is the complex envelope, f_c is the carrier frequency, and $\mathrm{Re}\{z\}$ denotes the real part of z. For digital modulation schemes $v(t)$ has the general form

$$v(t) = A\sum_k b(t-kT, \boldsymbol{x}_k) \tag{68.2}$$

where A is the amplitude of the carrier $\boldsymbol{x}_k = (x_k, x_{k-1}, \ldots, x_{k-K})$ is the data sequence, T is the symbol or baud duration, and $b(t, \boldsymbol{x}_i)$ is an equivalent shaping function usually of duration T. The precise form of $b(t, \boldsymbol{x}_i)$ and the memory length K depends on the type of modulation that is employed. Several examples are provided in this chapter where information is transmitted in the amplitude, phase, or frequency of the bandpass signal.

The **power spectral density** of the bandpass signal $S_{ss}(f)$ is related to the power spectral density of the complex envelope $S_{vv}(f)$ by

$$S_{ss}(f) = \frac{1}{2}[S_{vv}(f-f_c) + S_{vv}(f+f_c)] \tag{68.3}$$

The power density spectrum of the complex envelope for a digital modulation scheme has the general form

$$S_{vv}(f) = \frac{A^2}{T}\sum_m S_{b,m}(f)e^{-j2\pi fmT} \tag{68.4}$$

where

$$S_{b,m}(f) = \frac{1}{2}E[B(f,\boldsymbol{x}_m)B^*(f,\boldsymbol{x}_0)] \tag{68.5}$$

$B(f, \boldsymbol{x}_m)$ is the Fourier transform of $b(t, \boldsymbol{x}_m)$, and $E[\cdot]$ denotes the expectation operator. Usually symmetric signal sets are chosen so that the complex envelope has zero mean, i.e., $E[b(t, \boldsymbol{x}_0)] = 0$. This implies that the power density spectrum has no discrete components. If, in addition, $\boldsymbol{x}_m$ and $\boldsymbol{x}_0$ are independent for $|m| > K$, then

$$S_{vv}(f) = \frac{A^2}{T}\sum_{|m|<K} S_{b,m}(f)e^{-j2\pi fmT} \tag{68.6}$$

68.3 Analog Frequency Modulation

With analog frequency modulation the complex envelope is

$$v(t) = A\exp\left[j2\pi k_f \int_0^t m(\tau)\,\mathrm{d}\tau\right] \tag{68.7}$$

where $m(t)$ is the modulating waveform and k_f in Hz/v is the frequency sensitivity of the FM modulator. The bandpass signal is

$$s(t) = A\cos\left[2\pi f_c t + 2\pi k_f \int_0^t m(t)\,\mathrm{d}t\right] \tag{68.8}$$

The instantaneous frequency of the carrier $f_i(t) = f_c + k_f m(t)$ varies linearly with the waveform $m(t)$, hence, the name frequency modulation. Notice that FM has a constant envelope making it suitable for nonlinear amplification. However, the complex envelope is a nonlinear function of the modulating waveform $m(t)$ and, therefore, the spectral characteristics of $v(t)$ cannot be obtained directly from the spectral characteristics of $m(t)$.

With the sinusoidal modulating waveform $m(t) = A_m \cos(2\pi f_m t)$ the instantaneous carrier frequency is

$$f_i(t) = f_c + \Delta_f \cos(2\pi f_m t) \tag{68.9}$$

where $\Delta_f = k_f A_m$ is the peak frequency deviation. The complex envelope becomes

$$\begin{aligned} v(t) &= \exp\left[2\pi \int_0^t f_i(t)\,\mathrm{d}t\right] \\ &= \exp[2\pi f_c t + \beta \sin(2\pi f_m t)] \end{aligned} \tag{68.10}$$

where $\beta = \Delta_f / f_m$ is called the modulation index. The bandwidth of $v(t)$ depends on the value of β. If $\beta < 1$, then narrowband FM is generated, where the spectral widths of $v(t)$ and $m(t)$ are about the same, i.e., $2f_m$. If $\beta \gg 1$, then wideband FM is generated, where the spectral occupancy of $v(t)$ is slightly greater than $2\Delta_f$. In general, the approximate bandwidth of an FM signal is

$$W \approx 2\Delta_f + 2f_m = 2\Delta_f\left(1 + \frac{1}{\beta}\right) \tag{68.11}$$

which is a relation known as Carson's rule. Unfortunately, typical analog cellular radio systems use a modulation index in the range $1 \lesssim \beta \lesssim 3$ where Carson's rule is not accurate. Furthermore, the message waveform $m(t)$ is not a pure sinusoid so that Carson's rule does not directly apply.

In analog cellular systems the waveform $m(t)$ is obtained by first companding the speech waveform and then hard limiting the resulting signal. The purpose of the limiter is to control the peak frequency deviation Δ_f. The limiter introduces high-frequency components that must be removed with a low-pass filter prior to modulation. To estimate the bandwidth occupancy, we first determine the ratio of the frequency deviation Δ_f corresponding to the maximum amplitude of $m(t)$, and the highest frequency component B that is present in $m(t)$. These two conditions are the most extreme cases, and the resulting ratio, $D = \Delta_f / B$, is called the *deviation ratio*. Then replace β by D and f_m by B in Carson's rule, giving

$$W \approx 2\Delta_f + 2B = 2\Delta_f\left(1 + \frac{1}{D}\right) \tag{68.12}$$

This approximation will overestimate the bandwidth requirements. A more accurate estimate of the bandwidth requirements must be obtained from simulation or measurements.

68.4 Phase Shift Keying (PSK) and π/4-QPSK

With **phase shift keying** (PSK), the equivalent shaping function in Eq. (68.2) has the form

$$b(t, \boldsymbol{x}_k) = \psi_T(t) \exp\left[j\frac{\pi}{M} x_k h_s(t)\right], \quad \boldsymbol{x}_k = x_k \tag{68.13}$$

where $h_s(t)$ is a phase shaping pulse, $\psi_T(t)$ an amplitude shaping pulse, and M the size of the modulation alphabet. Notice that the phase varies linearly with the symbol sequence $\{x_k\}$, hence the name phase shift keying. For a modulation alphabet size of M, $x_k \in \{\pm 1, \pm 3, \ldots, \pm(M-1)\}$. Each symbol x_k is mapped onto $\log_2 M$ source bits. A QPSK signal is obtained by using $M = 4$, resulting in a transmission rate of 2 b/symbol.

Usually, the phase shaping pulse is chosen to be the rectangular pulse $h_s(t) = u_T(t) \stackrel{\Delta}{=} u(t) - u(t-T)$, where $u(t)$ is the unit step function. The amplitude shaping pulse is very often chosen to be a square root raised cosine pulse, where the Fourier transform of $\psi_T(t)$ is

$$\Psi_T(f) = \begin{cases} \sqrt{T} & 0 \le |f| \le (1-\beta)/2T \\ \sqrt{\frac{T}{2}\left[1 - \sin\frac{\pi T}{\beta}\left(f - \frac{1}{2T}\right)\right]} & (1-\beta)/2T \le |f| \le (1+\beta)/2T \end{cases} \tag{68.14}$$

The receiver implements the same filter $\Psi_R(f) = \Psi_T(f)$ so that the overall pulse has the raised cosine spectrum $\Psi(f) = \Psi_R(f)\Psi_T(f) = |\Psi_T(f)|^2$. If the channel is affected by flat fading and additive white Gaussian noise, then this partitioning of the filtering operations between the transmitter and receiver will optimize the signal to noise ratio at the output of the receiver filter at the sampling instants. The rolloff factor β usually lies between 0 and 1 and defines the **excess bandwidth** $100\beta\%$. Using a smaller β results in a more compact power density spectrum, but the link performance becomes more sensitive to errors in the symbol timing. The IS-54 system uses $\beta = 0.35$, while PDC uses $\beta = 0.5$.

The time domain pulse corresponding to Eq. (68.14) can be obtained by taking the inverse Fourier transform, resulting in

$$\psi_T(t) = 4\beta \frac{\cos[(1+\beta)\pi t/T] + \sin[(1-\beta)\pi t/T](4\beta t/T)^{-1}}{\pi\sqrt{T}[1 - 16\beta^2 t^2/T^2]} \tag{68.15}$$

A typical square root raised cosine pulse with a rolloff factor of $\beta = 0.5$ is shown in Fig. 68.1. Strictly speaking the pulse $\psi_T(t)$ is noncausal, but in practice a truncated time domain pulse is used. For example, in Fig. 68.1 the pulse is truncated to $6T$ and time shifted by $3T$ to yield a causal pulse.

Unlike conventional QPSK that has four possible transmitted phases, π/4-QPSK has eight possible transmitted phases. Let $\theta(n)$ be the transmitted carrier phase for the nth epoch, and let $\Delta\theta(n) = \theta(n) - \theta(n-1)$ be the differential carrier phase between epochs n and $n-1$. With π/4-QPSK, the transmission rate is 2 b/symbol and the differential phase is related to the symbol sequence $\{x_n\}$ through the mapping

$$\Delta\theta(n) = \begin{cases} -3\pi/4, & x_n = -3 \\ -\pi/4, & x_n = -1 \\ \pi/4, & x_n = +1 \\ 3\pi/4, & x_n = +3 \end{cases} \tag{68.16}$$

Since the symbol sequence $\{x_n\}$ is random, the mapping in Eq. (68.16) is arbitrary, except that the phase differences must be $\pm\pi/4$ and $\pm 3\pi/4$. The phase difference with the given mapping can be written in the convenient algebraic form

$$\Delta\theta(n) = x_n \frac{\pi}{4} \tag{68.17}$$

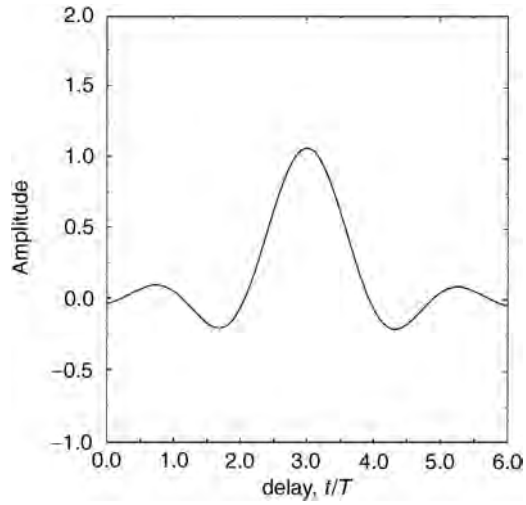


FIGURE 68.1 Square root raised cosine pulse with rolloff factor $\beta = 0.5$.

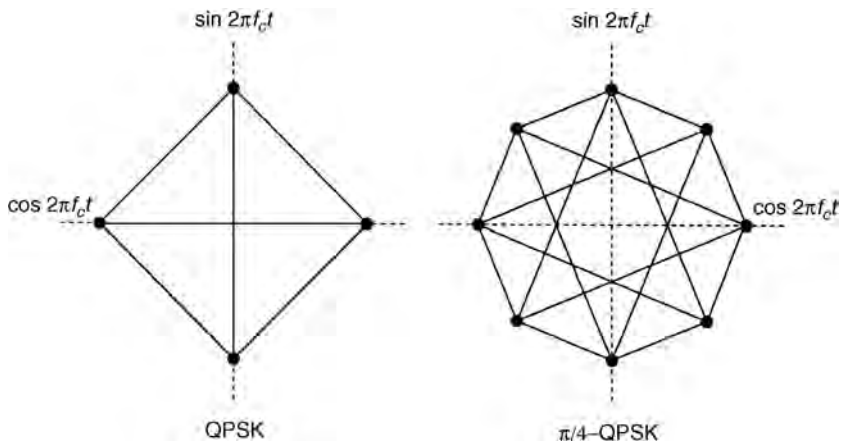


FIGURE 68.2 Signal-space constellations for QPSK and $\pi/4$-DQPSK.

which allows us to write the equivalent shaping function of the $\pi/4$-QPSK signal as

$$\begin{aligned} b(t, \underline{x}_k) &= \psi(t) \exp\left\{ j\left[\theta(k-1) + x_k \frac{\pi}{4} \right] \right\} \\ &= \psi_T(t) \exp\left[j\frac{\pi}{4} \left(\sum_{n=-\infty}^{k-1} x_n + x_k \right) \right] \end{aligned} \tag{68.18}$$

The summation in the exponent represents the accumulated carrier phase, whereas the last term is the phase change due to the kth symbol. Observe that the phase shaping function is the rectangular pulse $u_T(t)$. The amplitude shaping function $\psi_T(t)$ is usually the square root raised cosine pulse in Eq. (68.15).

The phase states of QPSK and $\pi/4$-QPSK signals can be summarized by the signal space diagram in Fig. 68.2 that shows the phase states and allowable transitions between the phase states. However, it does

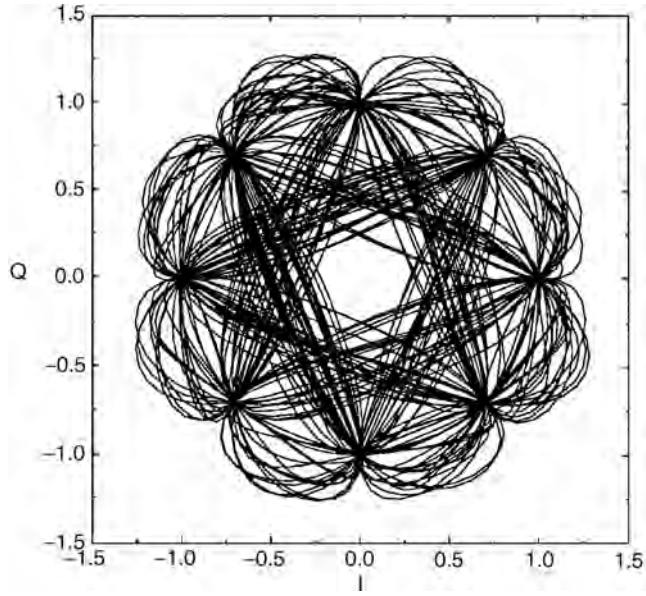


FIGURE 68.3 Phase diagram of $\pi/4$-QPSK with square root raised cosine pulse; $\beta = 0.5$.

not describe the actual phase trajectories. A typical diagram showing phase trajectories with square root raised cosine pulse shaping is shown in Fig. 68.3. Note that the phase trajectories do not pass through the origin. This reduces the envelope fluctuations of the signal making it less susceptible to amplifier nonlinearities and reduces the dynamic range required of the power amplifier.

The power density spectrum of QPSK and $\pi/4$-QPSK depends on both the amplitude and phase shaping pulses. For the rectangular phase shaping pulse $h_s(t) = u_T(t)$, the power density spectrum of the complex envelope is

$$S_{vv}(f) = \frac{A^2}{T}|\Psi_T(f)|^2 \tag{68.19}$$

With square root raised cosine pulse shaping, $\Psi_T(f)$ has the form defined in Eq. (68.14). The power density spectrum of a pulse $\tilde{\psi}_T(t)$ that is obtained by truncating $\psi_T(t)$ to length τ can be obtained by writing $\tilde{\psi}_T(t) = \psi_T(t)\,\mathrm{rect}(t/\tau)$. Then $\tilde{\Psi}_T(f) = \Psi_T(f) * \tau\,\mathrm{sinc}(f\tau)$, where $*$ denotes the operation of convolution, and the power density spectrum is again obtained by applying Eq. (68.19). Truncation of the pulse will regenerate some side lobes, thus causing adjacent channel interference. Figure 68.4 illustrates the power density spectrum of a truncated square root raised cosine pulse for various truncation lengths τ.

68.5 Continuous Phase Modulation (CPM) and MSK

Continuous phase modulation (**CPM**) refers to a broad class of frequency modulation techniques where the carrier phase varies in a continuous manner. A comprehensive treatment of CPM is provided in [1]. CPM schemes are attractive because they have constant envelope and excellent spectral characteristics. The complex envelope of any CPM signal is

$$v(t) = A\exp\left[j2\pi k_f \int_{-\infty}^{t} \sum_n x_n h_s(\tau - nT)\,\mathrm{d}\tau\right] \tag{68.20}$$

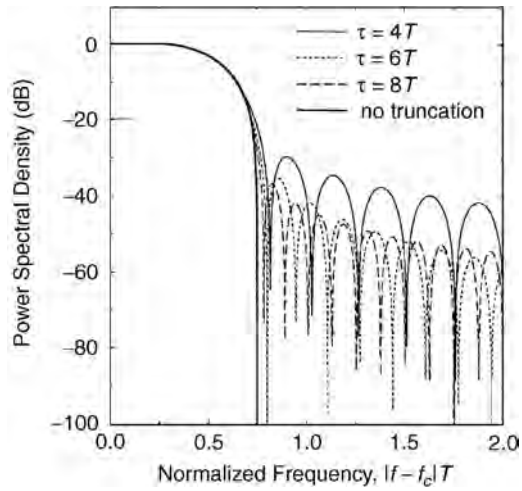


FIGURE 68.4 Power density spectrum of truncated square root raised cosine pulse with various truncation lengths; $\beta = 0.5$.

The instantaneous frequency deviation from the carrier is

$$f_{\text{dev}}(t) = k_f \sum_n x_n h_s(t - nT) \tag{68.21}$$

where k_f is the peak frequency deviation. If the frequency shaping pulse $h_s(t)$ has duration T, then the equivalent shaping function in Eq. (68.2) has the form

$$b(t, \boldsymbol{x}_k) = \exp\left\{ j\left[\beta(T) \sum_{n=-\infty}^{k-1} x_n + x_k \beta(t) \right] \right\} u_T(t) \tag{68.22}$$

where

$$\beta(T) = \begin{cases} 0, & t < 0 \\ \dfrac{\pi h}{\int_0^T h_s(\tau)\,\mathrm{d}\tau} \displaystyle\int_0^t h_s(\tau)\,\mathrm{d}\tau, & 0 \le t \le T \\ \pi h, & t \ge T \end{cases} \tag{68.23}$$

is the phase shaping pulse, and $h = \beta(T)/\pi$ is called the modulation index.

Minimum shift keying (**MSK**) is a special form of binary CPM ($x_k \in \{-1, +1\}$) that is defined by a rectangular frequency shaping pulse $h_s(t) = u_T(t)$, and a modulation index $h = 1/2$ so that

$$\beta(T) = \begin{cases} 0, & t < 0 \\ \pi t/2T, & 0 \le t \le T \\ \pi/2, & t \ge T \end{cases} \tag{68.24}$$

Therefore, the complex envelope is

$$v(t) = A \exp\left(j\frac{\pi}{2} \sum_{n=-\infty}^{k-1} x_n + \frac{\pi}{2} x_k \frac{t - kT}{T} \right) \tag{68.25}$$

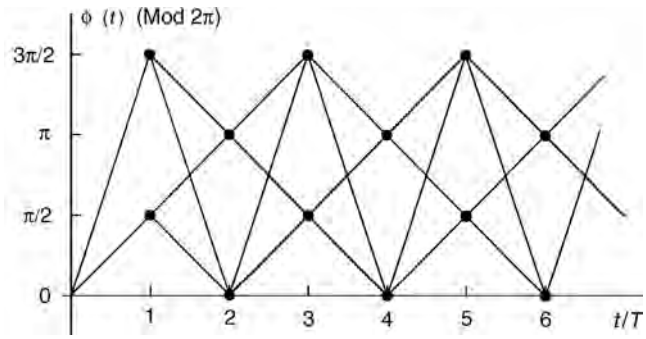


FIGURE 68.5 Phase-trellis diagram for MSK.

A MSK signal can be described by the phase trellis diagram shown in Fig. 68.5 which plots the time behavior of the phase

$$\theta(t) = \frac{\pi}{2}\sum_{n=-\infty}^{k-1} x_n + \frac{\pi}{2}x_k\frac{t-kT}{T} \tag{68.26}$$

The MSK bandpass signal is

$$\begin{aligned} s(t) &= A\cos\left(2\pi f_c t + \frac{\pi}{2}\sum_{n=-\infty}^{k-1} x_n + \frac{\pi}{2}x_k\frac{t-kT}{T}\right) \\ &= A\cos\left[2\pi\left(f_c + \frac{x_k}{4T}\right)t - \frac{k\pi}{2}x_k + \frac{\pi}{2}\sum_{n=-\infty}^{k-1} x_n\right] \quad kT \le t \le (k+1)T \end{aligned} \tag{68.27}$$

From Eq. (68.27) we observe that the MSK signal has one of two possible frequencies $f_L = f_c - 1/4T$ or $f_U = f_c + 1/4T$ during each symbol interval. The difference between these frequencies is $f_U - f_L = 1/2T$. This is the minimum frequency difference between two sinusoids of duration T that will ensure orthogonality with coherent demodulation [7], hence, the name minimum shift keying. By applying various trigonometric identities to Eq. (68.27) we can write

$$s(t) = A[x_k^I\psi(t-k2T)\cos(2\pi f_c t) - x_k^Q\psi(t-k2T-T)\sin(2\pi f_c t)], \quad kT \le t \le (k+1)T \tag{68.28}$$

where

$$\begin{aligned} x_k^I &= -x_{k-1}^Q x_{2k-1} \\ x_k^Q &= x_k^I x_{2k} \\ \psi(t) &= \cos\left(\frac{\pi t}{2T}\right), \quad -T \le t \le T \end{aligned}$$

Note that the x_k^I and x_k^Q are independent binary symbols that take on elements from the set $\{-1, +1\}$, and the half-sinusoid amplitude shaping pulse $\psi(t)$ has duration $2T$ and $\psi(t - T) = \sin(\pi t/2T)$, $0 \le t \le 2T$. Therefore, MSK is equivalent to offset quadrature amplitude shift keying (OQASK) with a half-sinusoid amplitude shaping pulse.

To obtain the power density spectrum of MSK, we observe from Eq. (68.28) that the equivalent shaping function of MSK has the form

$$b(t, \boldsymbol{x}_k) = x_k^I \psi(t) + j x_k^Q \psi(t - T) \tag{68.29}$$

The Fourier transform of Eq. (68.29) is

$$B(f, \boldsymbol{x}_k) = (x_k^I + j x_k^Q e^{-j2\pi f T}) \Psi(f) \tag{68.30}$$

Since the symbols x_k^I and x_k^Q are independent and zero mean, it follows from Eqs. (68.5) and (68.6) that

$$S_{vv}(f) = \frac{A^2 |\Psi(f)|^2}{2T} \tag{68.31}$$

Therefore, the power density spectrum of MSK is determined solely by the Fourier transform of the half-sinusoid amplitude shaping pulse $\psi(t)$, resulting in

$$S_{vv}(f) = \frac{16A^2 T}{\pi^2} \left[\frac{\cos 2\pi f T}{1 - 16 f^2 T^2} \right]^2 \tag{68.32}$$

The power spectral density of MSK is plotted in Fig. 68.8. Observe that an MSK signal has fairly large sidelobes compared to $\pi/4$-QPSK with a truncated square root raised cosine pulse (c.f., Fig. 68.4).

68.6 Gaussian Minimum Shift Keying

MSK signals have all of the desirable attributes for mobile radio, except for a compact power density spectrum. This can be alleviated by filtering the modulating signal $x(t) = \Sigma_n x_n u_T(t - nT)$ with a low-pass filter prior to frequency modulation, as shown in Fig. 68.6. Such filtering removes the higher frequency components in $x(t)$ and, therefore, yields a more compact spectrum. The low-pass filter is chosen to have (1) narrow bandwidth and a sharp transition band, (2) low-overshoot impulse response, and (3) preservation of the output pulse area to ensure a phase shift of $\pi/2$.

GMSK uses a low-pass filter with the following transfer function:

$$H(f) = A \exp\left\{ -\left(\frac{f}{B}\right)^2 \frac{\ln 2}{2} \right\} \tag{68.33}$$

where B is the 3-dB bandwidth of the filter and A a constant. It is apparent that $H(f)$ is bell shaped about $f = 0$, hence the name Gaussian MSK. A rectangular pulse $\mathrm{rect}(t/T) = u_T(t + T/2)$ transmitted through this filter yields the frequency shaping pulse

$$h_s(t) = A\sqrt{\frac{2\pi}{\ln 2}}(BT) \int_{t/T-1/2}^{t/T+1/2} \exp\left\{ -\frac{2\pi^2 (BT)^2 x^2}{\ln 2} \right\} dx \tag{68.34}$$

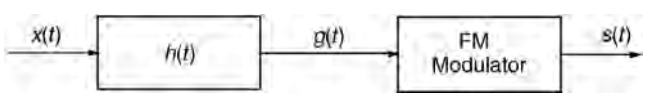


FIGURE 68.6 Premodulation filtered MSK.

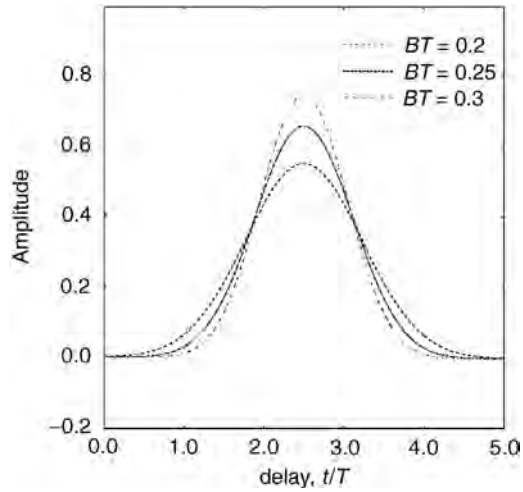


FIGURE 68.7 GMSK frequency shaping pulse for various normalized filter bandwidths *BT*.

The phase change over the time interval from $-T/2 \le t \le T/2$ is

$$\theta\left(\frac{T}{2}\right) - \theta\left(\frac{-T}{2}\right) = x_0\beta_0(T) + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} x_n\beta_n(T) \tag{68.35}$$

where

$$\beta_n(T) = \frac{\pi h}{\int_{-\infty}^{\infty} h_s(v)\,\mathrm{d}v} \int_{-T/2-nT}^{T/2-nT} h_s(v)\,\mathrm{d}v \tag{68.36}$$

The first term in Eq. (68.35) is the desired term, and the second term is the intersymbol interference (ISI) introduced by the premodulation filter. Once again, with GMSK $h = 1/2$ so that a total phase shift of $\pi/2$ is maintained.

Notice that the pulse $h_s(t)$ is noncausal so that a truncated pulse must be used in practice. Figure 68.7 plots a GMSK frequency shaping pulse that is truncated to $\tau = 5T$ and time shifted by $2.5T$, for various normalized filter bandwidths *BT*. Notice that the frequency shaping pulse has a duration greater than *T* so that ISI is introduced. As *BT* decreases, the induced ISI is increased. Thus, whereas a smaller value of *BT* results in a more compact power density spectrum, the induced ISI will degrade the bit-error-rate performance. Hence, there is a tradeoff in the choice of *BT*. Some studies have indicated that $BT = 0.25$ is a good choice for cellular radio systems [6].

The power density spectrum of GMSK is quite difficult to obtain, but can be computed by using published methods [3]. Figure 68.8 plots the power density spectrum for $BT = 0.2$, 0.25, and 0.3, obtained from Wesolowski [8]. Observe that the spectral sidelobes are greatly reduced by the Gaussian low-pass filter.

68.7 Orthogonal Frequency Division Multiplexing (OFDM)

Orthogonal frequency division multiplexing (**OFDM**) is a modulation technique that has been recently suggested for use in cellular radio [2], digital audio broadcasting [4], and digital video broadcasting. The basic idea of OFDM is to transmit blocks of symbols in parallel by employing a (large) number of orthogonal subcarriers. With block transmission, *N* serial source symbols each with period T_s are converted into a block of *N* parallel modulated symbols each with period $T = NT_s$. The block length *N* is chosen so that $NT_s \gg \sigma_\tau$, where σ_τ is the rms delay spread of the channel. Since the symbol rate on each subcarrier is much less than the serial source rate, the effects of delay spread are greatly reduced. This has practical advantages because it may reduce or even eliminate the need for equalization. Although the

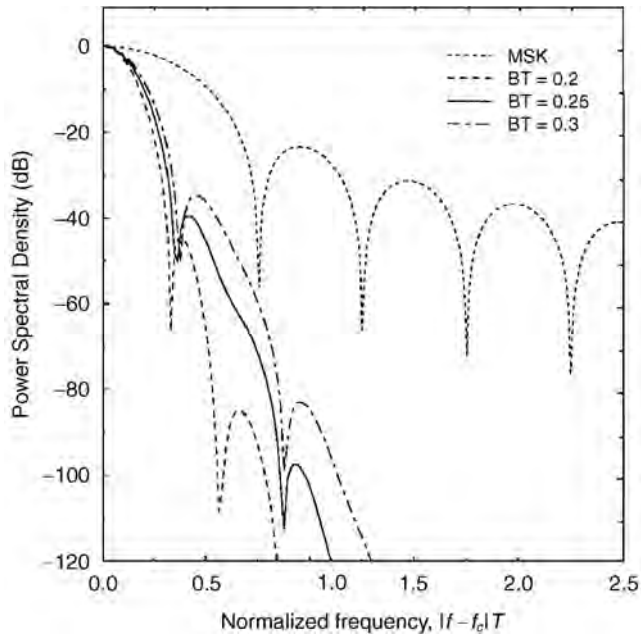


FIGURE 68.8 Power density spectrum of MSK and GMSK.

block length N is chosen so that $NT_s \gg \sigma_\tau$, the channel dispersion will still cause consecutive blocks to overlap. This results in some residual ISI that will degrade the performance. This residual ISI can be eliminated at the expense of channel capacity by using guard intervals between the blocks that are at least as long as the effective channel impulse response.

The complex envelope of an OFDM signal is described by

$$v(t) = A\sum_{k}\sum_{n=0}^{N-1} x_{k,n}\phi_n(t-kT) \tag{68.37}$$

where

$$\phi_n(t) = \exp\left\{j\frac{2\pi\left(n-\frac{N-1}{2}\right)t}{T}\right\}U_T(t), \quad n = 0,1,\ldots,N-1 \tag{68.38}$$

are orthogonal waveforms and $U_T(t)$ is a rectangular shaping function. The frequency separation of the subcarriers, $1/T$, ensures that the subcarriers are orthogonal and phase continuity is maintained from one symbol to the next, but is twice the minimum required for orthogonality with coherent detection. At epoch k, N-data symbols are transmitted by using the N distinct pulses. The data symbols $x_{k,n}$ are often chosen from an M-ary **quadrature amplitude modulation** (M-QAM) constellation, where $x_{k,n} = x^I_{k,n} + jx^Q_{k,n}$ with $x^I_{k,n}, x^Q_{k,n} \in \{\pm1, \pm3,\ldots,\pm(N-1)\}$ and $N = \sqrt{M}$.

A key advantage of using OFDM is that the modulation can be achieved in the discrete domain by using either an inverse discrete Fourier transform (IDFT) or the more computationally efficient inverse fast Fourier transform (IFFT). Considering the data block at epoch $k = 0$ and ignoring the frequency offset $\exp\{-j[2\pi(N-1)t/2T]\}$, the complex low-pass OFDM signal has the form

$$v(t) = \sum_{n=0}^{N-1} x_{0,n}\exp\left\{\frac{j2\pi nt}{NT_s}\right\}, \quad 0 \le t \le T \tag{68.39}$$

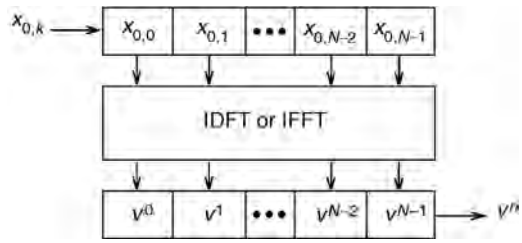


FIGURE 68.9 Block diagram of OFDM transmitter using IDFT or IFFT.

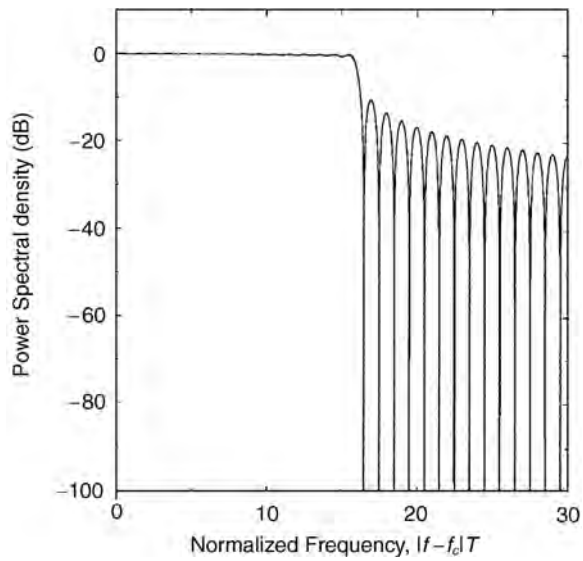


FIGURE 68.10 Power density spectrum of OFDM with $N = 32$.

If this signal is sampled at epochs $t = kT_s$, then

$$v^k = v(kT_s) = \sum_{n=0}^{N-1} x_{0,n} \exp\left\{\frac{j2\pi nk}{N}\right\}, \quad k = 0,1,\ldots,N-1 \tag{68.40}$$

Observe that the sampled OFDM signal has duration N and the samples $v^0, v^1,\ldots,v^{N-1}$ are just the IDFT of the data block $x_{0,0}, x_{0,1},\ldots,x_{0,N-1}$. A block diagram of an OFDM transmitter is shown in Fig. 68.9.

The power spectral density of an OFDM signal can be obtained by treating OFDM as independent modulation on subcarriers that are separated in frequency by $1/T$. Because the subcarriers are only separated by $1/T$, significant spectral overlap results. Because the subcarriers are orthogonal, however, the overlap improves the spectral efficiency of the scheme. For a signal constellation with zero mean and the waveforms in Eq. (68.38), the power density spectrum of the complex envelope is

$$S_{vv}(f) = \frac{A^2}{T}\sigma_x^2 \sum_{n=0}^{N-1} \left|\operatorname{sinc}\left[fT - \left(n - \frac{N-1}{2}\right)\right]\right|^2 \tag{68.41}$$

where $\sigma_x^2 = \frac{1}{2}E[|x_{k,n}|^2]$ is the variance of the signal constellation. For example, the complex envelope power spectrum of OFDM with $N = 32$ subcarriers is shown in Fig. 68.10.

68.8 Conclusions

A variety of modulation schemes are employed in wireless communication systems. Wireless modulation schemes must have a compact power density spectrum, while at the same time providing a good bit-error-rate performance in the presence of channel impairments such as cochannel interference and fading. The most popular digital modulation techniques employed in wireless systems are GMSK in the European GSM system, $\pi/4$-QPSK in the North American IS-54 and Japanese PDC systems, and OFDM in digital audio broadcasting systems.

Defining Terms

Bandwidth efficiency: Transmission efficiency of a digital modulation scheme measured in units of bits per second per Hertz of bandwidth.

Continuous phase modulation: Frequency modulation where the phase varies in a continuous manner.

Excess bandwidth: Percentage of bandwidth that is in excess of the minimum of $1/2T$ (T is the baud or symbol duration) required for data communication.

Frequency modulation: Modulation where the instantaneous frequency of the carrier varies linearly with the data signal.

Gaussian minimum shift keying: MSK where the data signal is prefiltered with a Gaussian filter prior to frequency modulation.

Minimum shift keying: A special form of continuous phase modulation having linear phase trajectories and a modulation index of 1/2.

Orthogonal frequency division multiplexing: Modulation by using a collection of low-bit-rate orthogonal subcarriers.

Phase shift keying: Modulation where the instantaneous phase of the carrier varies linearly with the data signal.

Power spectral density: Relative power in a modulated signal as a function of frequency.

Quadrature amplitude modulation: Modulation where information is transmitted in the amplitude of the cosine and sine components of the carrier.

References

1. Anderson, J.B., Aulin, T., and Sundberg, C.-E., *Digital Phase Modulation,* Plenum Press, New York, 1986.
2. Birchler, M.A. and Jasper, S.C., A 64 kbps digital land mobile radio system employing M-16QAM. *Proc. 5th Nordic Sem. Dig. Mobile Radio Commun.,* 237–241, Dec. 1992.
3. Garrison, G.J., A power spectral density analysis for digital FM, *IEEE Trans. Commun.,* COM-23, 1228–1243, Nov. 1975.
4. Le Floch, B., Halbert-Lassalle, R., and Castelain, D., Digital sound broadcasting to mobile receivers, *IEEE Trans. Consum. Elec.,* 35, Aug. 1989.
5. Murota, K. and Hirade, K., GMSK modulation for digital mobile radio telephony, *IEEE Trans. Commun.,* COM-29, 1044–1050, Jul. 1981.
6. Murota, K., Kinoshita, K., and Hirade, K., Spectral efficiency of GMSK land mobile radio. *Proc. ICC'81,* 23.8.1, Jun. 1981.
7. Proakis, J.G., *Digital Communications,* 2nd ed., McGraw-Hill, New York, 1989.
8. Wesolowski, K., Private Communication, 1994.

Further Information

A good discussion of digital modem techniques is presented in *Advanced Digital Communications,* edited by K. Feher, Prentice-Hall, 1987.

Proceedings of various IEEE conferences such as the Vehicular Technology Conference, International Conference on Communications, and Global Telecommunications Conference, document the lastest developments in the field of wireless communications each year.

Journals such as the *IEEE Transactions on Communications* and *IEEE Transactions on Vehicular Technology* report advances in wireless modulation.

69

Access Methods

Bernd-Peter Paris
George Mason University

69.1 Introduction

The radio channel is fundamentally a broadcast communication medium. Therefore, signals transmitted by one user can potentially be received by all other users within range of the transmitter. Although this high connectivity is very useful in some applications, like broadcast radio or television, it requires stringent access control in wireless communication systems to avoid, or at least to limit, interference between transmissions. Throughout, the term wireless communication systems is taken to mean communication systems that facilitate two-way communication between a portable radio communication terminal and the fixed network infrastructure. Such systems range from mobile cellular systems through personal communication systems (PCS) to cordless telephones.

The objective of wireless communication systems is to provide communication channels on demand between a portable radio station and a radio port or base station that connects the user to the fixed network infrastructure. Design criteria for such systems include **capacity**, cost of implementation, and quality of service. All of these measures are influenced by the method used for providing multiple-access capabilities. However, the opposite is also true: the access method should be chosen carefully in light of the relative importance of design criteria as well as the system characteristics.

Multiple access in wireless radio systems is based on insulating signals used in different connections from each other. The support of parallel transmissions on the uplink and downlink, respectively, is called multiple access, whereas the exchange of information in both directions of a connection is referred to as **duplexing**. Hence, multiple access and duplexing are methods that facilitate the sharing of the broadcast communication medium. The necessary insulation is achieved by assigning to each transmission different components of the domains that contain the signals. The signal domains commonly used to provide multiple access capabilities include the following.

Spatial domain: All wireless communication systems exploit the fact that radio signals experience rapid attenuation during propagation. The propagation exponent ρ on typical radio channels lies between

$\rho = 2$ and $\rho = 6$ with $\rho = 4$ a typical value. As signal strength decays inversely proportional to the ρth power of the distance, far away transmitters introduce interference that is negligible compared to the strength of the desired signal. The cellular design principle is based on the ability to reuse signals safely if a minimum reuse distance is maintained. Directional antennas can be used to enhance the insulation between signals. We will not focus further on the spatial domain in this treatment of access methods.

Frequency domain: Signals which occupy nonoverlapping frequency bands can be easily separated using appropriate bandpass filters. Hence, signals can be transmitted simultaneously without interfering with each other. This method of providing multiple access capabilities is called **frequency-division multiple access (FDMA)**.

Time domain: Signals can be transmitted in nonoverlapping time slots in a round-robin fashion. Thus, signals occupy the same frequency band but are easily separated based on their time of arrival. This multiple access method is called **time-division multiple access (TDMA)**.

Code domain: In **code-division multiple access (CDMA)** different users employ signals that have very small cross-correlation. Thus, correlators can be used to extract individual signals from a mixture of signals even though they are transmitted simultaneously and in the same frequency band. The term code-division multiple-access is used to denote this form of channel sharing. Two forms of CDMA are most widely employed and will be described in detail subsequently, frequency hopping (FH) and direct sequence (DS).

System designers have to decide in favor of one, or a combination, of the latter three domains to facilitate multiple access. The three access methods are illustrated in Fig. 69.1. The principal idea in all three of these access methods is to employ signals that are orthogonal or nearly orthogonal. Then, correlators that project the received signal into the subspace of the desired signal can be employed to extract a signal without interference from other transmissions.

Preference for one access method over another depends largely on overall system characteristics, as we will see in the sequel. No single access method is universally preferable, and system considerations should be carefully weighed before the design decision is made. Before going into the detailed description

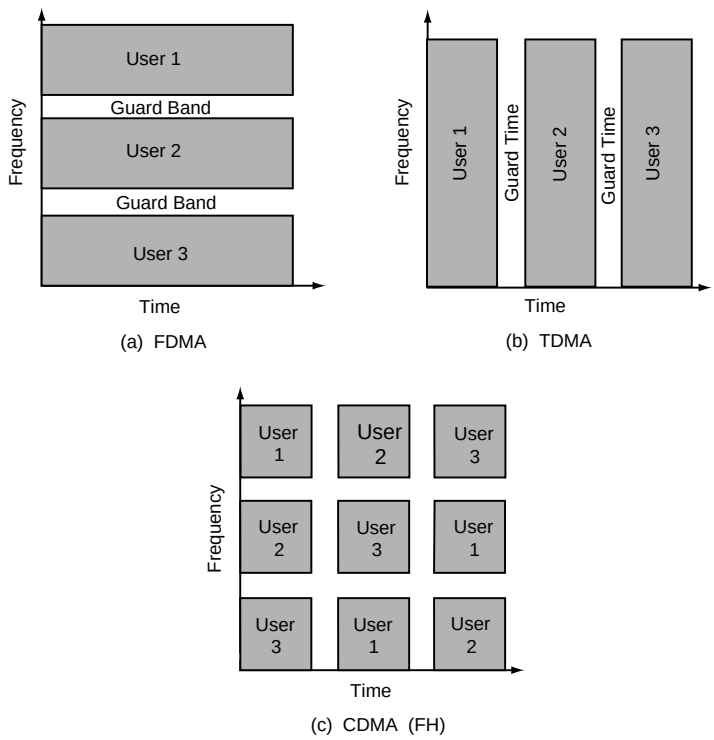


FIGURE 69.1 Multiple-access methods for wireless communication systems.

TABLE 69.1 Summary of Relevant Characteristics of Cordless Telephone and Cellular Mobile Radio

Characteristic or Parameter	Cordless Telephone	Cellular Radio
Speech quality	Toll quality	Varying with channel quality; possibly decreased by speech pause exploitation
Transmission range	<100 m	100 m–30 km
Transmit power	Milliwatts	Approx. 1 W
Base station antenna height	Approx. 1 m	Tens of meters
Delay spread	Approx. 1 μs	Approx. 10 μs
Complexity of base station	Low	High
Complexity of user set	Low	High

of the different access methods, we will discuss briefly the salient features of some wireless communication systems. This will allow us later to assess the relative merits of the access methods in different scenarios.

69.2 Relevant Wireless Communication System Characteristics

Modern wireless radio systems range from relatively simple cordless telephones to mobile cellular systems and personal communication systems (PCS). It is useful to consider such diverse systems as cordless telephone and mobile cellular radio to illustrate some of the fundamental characteristics of wireless communication systems [2].

A summary of the relevant parameters and characteristics for cordless telephone and cellular radio is given in Table 69.1. As evident from that table, the fundamental differences between the two systems are speech quality and the area covered by a base station. The high speech quality requirement in the cordless application is the consequence of the availability of tethered access in the home and office and the resulting direct competition with wire-line telephone services. In the mobile cellular application, the user has no alternative to the wireless access and may be satisfied with lower, but still acceptable, quality of service.

In cordless telephone applications the transmission range is short because the base station can simply be moved to a conveniently located wire-line access point (wall jack) to provide wireless network access where desired. In contrast, the mobile cellular base station must provide access for users throughout a large geographical area of up to approximately 30 km (20 mi) around the base station. This large coverage area is necessary to economically meet the promise of uninterrupted service to roaming users.

The different range requirements directly affect the transmit power and antenna height for the two systems. High-power transmitters used in mobile cellular user sets consume far more power than even complex signal processing hardware. Hence, sophisticated signal processing, including speech compression, voice activity detection, error correction and detection, and adaptive equalization, can be employed without substantial impact on the battery life in portable hand sets. Furthermore, such techniques are consistent with the goals of increased range and support of large numbers of users with a single, expensive base station. On the other hand, the high mobile cellular base station antennas introduce delay spreads that are one or two orders of magnitude larger than those commonly observed in cordless telephone applications.

Clearly, the two systems just considered are at extreme ends of the spectrum of wireless communications systems. Most notably, the PCS systems fall somewhere between the two. However, the comparison above highlights some of the system characteristics that should be considered when discussing access methods for wireless communication systems.

69.3 Frequency Division Multiple Access

As mentioned in Section 69.1, in FDMA nonoverlapping frequency bands are allocated to different users on a continuous time basis. Hence, signals assigned to different users are clearly orthogonal, at least ideally. In practice, out-of-band spectral components can not be completely suppressed leaving signals not quite orthogonal. This necessitates the introduction of guard bands between frequency bands to

reduce adjacent channel interference, i.e., inference from signals transmitted in adjacent frequency bands; see also Fig. 69.1(a).

It is advantageous to combine FDMA with time-division duplexing (TDD) to avoid simultaneous reception and transmission that would require insulation between receive and transmit antennas. In this scenario, the base station and portable take turns using the same frequency band for transmission. Nevertheless, combining FDMA and frequency division duplex is possible in principle, as is evident from the analog FM-based systems deployed throughout the world since the early 1980s.

Channel Considerations

In principle there exists the well-known duality between TDMA and FDMA; see [1, p. 113]. In the wireless environment, however, propagation related factors have a strong influence on the comparison between FDMA and TDMA. Specifically, the duration of a transmitted symbol is much longer in FDMA than in TDMA. As an immediate consequence, an equalizer is typically not required in an FDMA-based system because the delay spread is small compared to the symbol duration.

To illustrate this point, consider a hypothetical system that transmits information at a constant rate of 50 kb/s. This rate would be sufficient to support 32-kb/s adaptive differential pulse code modulation (ADPCM) speech encoding, some coding for error protection, and control overhead. If we assume further that some form of QPSK modulation is employed, the resulting symbol duration is 40 μs. In relation to delay spreads of approximately 1 μs in the cordless application and 10 μs in cellular systems, this duration is large enough that only little intersymbol interference is introduced. In other words, the channel is frequency nonselective, i.e., all spectral components of the signal are affected equally by the channel. In the cordless application an equalizer is certainly not required; cellular receivers may require equalizers capable of removing intersymbol interference between adjacent bits. Furthermore, it is well known that intersymbol interference between adjacent bits can be removed without loss in SNR by using maximum-likelihood sequence estimation; e.g., [8, p. 622].

Hence, rather simple receivers can be employed in FDMA systems at these data rates. However, there is a flip side to the argument. Recall that the Doppler spread, which characterizes the rate at which the channel impulse response changes, is given approximately by $B_d = v/cf_c$, where v denotes the speed of the mobile user, c is the propagation speed of the electromagnetic waves carrying the signal, and f_c is the carrier frequency. Thus, for systems operating in the vicinity of 1 GHz, B_d will be less than 1 Hz in the cordless application and typically about 100 Hz for a mobile traveling on a highway. In either case, the signal bandwidth is much larger than the Doppler spread B_d, and the channel can be characterized as slowly fading. Whereas this allows tracking of the carrier phase and the use of coherent receivers, it also means that fade durations are long in comparison to the symbol duration and can cause long sequences of bits to be subject to poor channel conditions. The problem is compounded by the fact that the channel is frequency nonselective because it implies that the entire signal is affected by a fade.

To overcome these problems either time diversity, frequency diversity, or spatial diversity could be employed. Time diversity can be accomplished by a combination of coding and interleaving if the fading rate is sufficiently large. For very slowly fading channels, such as the cordless application, the necessary interleaving depth would introduce too much delay to be practical. Frequency diversity can be introduced simply by slow frequency hopping, a technique that prescribes users to change the carrier frequency periodically. Frequency hopping is a form of spectrum spreading because the bandwidth occupied by the resulting signal is much larger than the symbol rate. In contrast to direct sequence spread spectrum discussed subsequently, however, the instantaneous bandwidth is not increased. The jumps between different frequency bands effectively emulate the movement of the portable and, thus, should be combined with the just described time-diversity methods. Spatial diversity is provided by the use of several receive or transmit antennas. At carrier frequencies exceeding 1 GHz, antennas are small and two or more antennas can be accommodated even in the hand set. Furthermore, if FDMA is combined with time-division duplexing, multiple antennas at the base station can provide diversity on both uplink and downlink. This is possible because the channels for the two links are virtually identical,

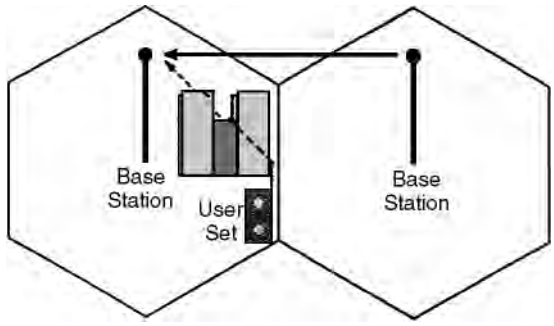


FIGURE 69.2 High base station antennas lead to stronger propagation paths between base stations than between a user set and its base stations.

and the base station, using channel information gained from observing the portable's signal, can transmit signals at each antenna such that they combine coherently at the portable's antenna. Thus, signal processing complexity is moved to the base station extending the portable's battery life.

Influence of Antenna Height

In the cellular mobile environment base station antennas are raised considerably to increase the coverage area. Antennas mounted on towers and rooftops are a common sight, and antenna heights of 50 m above ground are no exceptions. Besides increasing the coverage area, this has the additional effect that frequently there exists a better propagation path between two base station antennas than between a mobile and the base station; see Fig. 69.2.

Assuming that FDMA is used in conjunction with TDD as specified at the beginning of this section, then base stations and mobiles transmit on the same frequency. Now, unless there is tight synchronization between all base stations, signals from other base stations will interfere with the reception of signals from portables at the base station. To keep the interference at acceptable levels, it is necessary to increase the reuse distance, i.e., the distance between cells using the same frequencies. In other words, sufficient insulation in the spatial domain must be provided to facilitate the separation of signals. Note that these comments apply equally to cochannel and adjacent channel interference.

This problem does not arise in cordless applications. Base station antennas are generally of the same height as user sets. Hence, interference created by base stations is subject to the same propagation conditions as signals from user sets. Furthermore, in cordless telephone applications there are frequently attenuating obstacles, such as walls, between base stations that reduce intracell interference further. Note that this reduction is vital for the proper functioning of cordless telephones since there is typically no network planning associated with installing a cordless telephone. As a safety feature, to overcome intercell interference, adaptive channel management strategies based on sensing interference levels can be employed.

Example 69.1: CT2

The CT2 standard was originally adopted in 1987 in Great Britain and improved with a common air interface (CAI) in 1989. The CAI facilitates interoperability between equipment from different vendors whereas the original standard only guarantees noninterference. The CT2 standard is used in home and office cordless telephone equipment and has been used for telepoint applications [5].

CT2 operates in the frequency band 864–868 MHz and uses carriers spaced at 100 kHz. FDMA with time division duplexing is employed. The combined gross bit rate is 72 kb/s, transmitted in frames of 2-ms duration of which the first-half carries downlink and the second-half carries uplink information. This setup supports a net bit rate of 32 kb/s of user data (32-kb/s ADPCM encoded speech) and 2-kb/s control information in each direction. The CT2 modulation technique is binary frequency shift keying.

Further Remarks

From the preceding discussion it is obvious that FDMA is a good candidate for applications like cordless telephone. In particular, the simple signal processing makes it a good choice for inexpensive implementation in the benign cordless environment. The possibility of concentration of signal processing functions in the base station strengthens this aspect.

In the cellular application, on the other hand, FDMA is inappropriate because of the lack of built-in diversity and the potential for severe intercell interference between base stations. A further complication arises from the difficulty of performing handovers if base stations are not tightly synchronized.

For PCS the decision is not as obvious. Depending on whether the envisioned PCS application resembles more a cordless private branch exchange (PBX) than a cellular system, FDMA may be an appropriate choice. We will see later that it is probably better to opt for a combined TDMA/FDMA or a CDMA-based system to avoid the pitfalls of pure FDMA systems and still achieve moderate equipment complexities.

Finally, there is the problem of channel assignment. Clearly, it is not reasonable to assign a unique frequency to each user as there are not sufficient frequencies and the spectral resource would be unused whenever the user is idle. Instead, methods that allocate channels on demand can make much more efficient use of the spectrum. Such methods will be discussed further during the description of TDMA systems.

69.4 Time Division Multiple Access

In TDMA systems users share the same frequency band by accessing the channel in non-overlapping time intervals in a round-robin fashion [3]. Since the signals do not overlap, they are clearly orthogonal, and the signal of interest is easily extracted by switching the receiver on only during the transmission of the desired signal. Hence, the receiver filters are simply windows instead of the bandpass filters required in FDMA. As a consequence, the guard time between transmissions can be made as small as the synchronization of the network permits. Guard times of 30–50 μs between time slots are commonly used in TDMA-based systems. As a consequence, all users must be synchronized with the base station to within a fraction of the guard time. This is achievable by distributing a master clock signal on one of the base station's broadcast channels.

TDMA can be combined with TDD or frequency-division duplexing (FDD). The former duplexing scheme is used, for example, in the Digital European Cordless Telephone (DECT) standard and is well suited for systems in which base-to-base and mobile-to-base propagation paths are similar, i.e., systems without extremely high base station antennas. Since both the portable and the base station transmit on the same frequency, some signal processing functions for the downlink can be implemented in the base station, as discussed earlier for FDMA/TDD systems.

In the cellular application, the high base station antennas make FDD the more appropriate choice. In these systems, separate frequency bands are provided for uplink and downlink communication. Note that it is still possible and advisable to stagger the uplink and downlink transmission intervals such that they do not overlap, to avoid the situation that the portable must transmit and receive at the same time. With FDD the uplink and downlink channel are not identical and, hence, signal processing functions can not be implemented in the base station; antenna diversity and equalization have to be realized in the portable.

Propagation Considerations

In comparison to a FDMA system supporting the same user data rate, the transmitted data rate in a TDMA system is larger by a factor equal to the number of users sharing the frequency band. This factor is eight in the pan-European global system for mobile communications (GSM) and three in the advanced mobile phone service (D-AMPS) system. Thus, the symbol duration is reduced by the same factor and severe intersymbol interference results, at least in the cellular environment.

To illustrate, consider the earlier example where each user transmits 25 K symbols per second. Assuming eight users per frequency band leads to a symbol duration of 5 μs. Even in the cordless application with delay spreads of up to 1 μs, an equalizer may be useful to combat the resulting interference between adjacent symbols. In cellular systems, however, the delay spread of up to 20 μs introduces severe intersymbol interference spanning up to 5 symbol periods. As the delay spread often exceeds the symbol duration, the channel can be classified as frequency selective, emphasizing the observation that the channel affects different spectral components differently.

The intersymbol interference in cellular TDMA systems can be so severe that linear equalizers are insufficient to overcome its negative effects. Instead, more powerful, nonlinear decision feedback or maximum-likelihood sequence estimation equalizers must be employed [9]. Furthermore, all of these equalizers require some information about the channel impulse response that must be estimated from the received signal by means of an embedded training sequence. Clearly, the training sequence carries no user data and, thus, wastes valuable bandwidth.

In general, receivers for cellular TDMA systems will be fairly complex. On the positive side of the argument, however, the frequency selective nature of the channel provides some built-in diversity that makes transmission more robust to channel fading. The diversity stems from the fact that the multipath components of the received signal can be resolved at a resolution roughly equal to the symbol duration, and the different multipath components can be combined by the equalizer during the demodulation of the signal. To further improve robustness to channel fading, coding and interleaving, slow frequency hopping and antenna diversity can be employed as discussed in connection with FDMA.

Initial Channel Assignment

In both FDMA and TDMA systems, channels should not be assigned to a mobile on a permanent basis. A fixed assignment strategy would either be extremely wasteful of precious bandwidth or highly susceptible to cochannel interference. Instead, channels must be assigned on demand. Clearly, this implies the existence of a separate uplink channel on which mobiles can notify the base station of their need for a traffic channel. This uplink channel is referred to as the **random-access channel** because of the type of strategy used to regulate access to it.

The successful procedure for establishing a call that originates from the mobile station is outlined in Fig. 69.3. The mobile initiates the procedure by transmitting a request on the random-access channel. Since this channel is shared by all users in range of the base station, a random access protocol, like the ALOHA protocol, has to be employed to resolve possible collisions. Once the base station has received the mobile's request, it responds with an immediate assignment message that directs the mobile to tune to a dedicated control channel for the ensuing call setup. Upon completion of the call setup negotiation, a traffic channel, i.e., a frequency in FDMA systems or a time slot in TDMA systems, is assigned by the

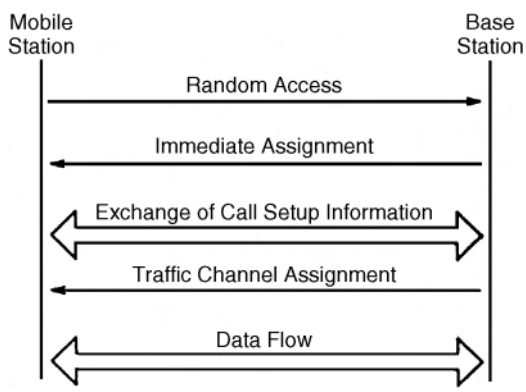


FIGURE 69.3 Mobile-originating call establishment.

base station and all future communication takes place on that channel. In the case of a mobile-terminating call request, the sequence of events is preceded by a paging message alerting the base station of the call request.

Example 69.2: GSM

Named after the organization that created the system standards (Groupe Speciale Mobile) this pan-European digital cellular system has been deployed in Europe since the early 1990s [6]. GSM uses combined TDMA and FDMA with frequency-division duplex for access. Carriers are spaced at 200 kHz and support eight TDMA time slots each. For the uplink the frequency band 890–915 MHz is allocated, whereas the downlink uses the band 935–960 MHz. Each time slot is of duration 577 μs, which corresponds to 156.26-b periods, including a guard time of 8.25-b periods. Eight consecutive time slots form a GSM frame of duration 4.62 ms.

The GSM modulation is Gaussian minimum shift keying with time-bandwidth product of 0.3, i.e., the modulator bandpass has a cutoff frequency of 0.3 times the bit rate. At the bit rate of 270.8 kb/s, severe intersymbol interference arises in the cellular environment. To facilitate coherent detection, a 26-bit training sequence is embedded into every time slot. Time diversity is achieved by interleaving over 8 frames for speech signals and 20 frames for data communication. Sophisticated error-correction coding with varying levels of protection for different outputs of the speech coder is provided. Note that the round-trip delay introduced by the interleaver is on the order of 80 ms for speech signals. GSM provides slow frequency hopping as a further mechanism to improve the efficiency of the interleaver.

Further Remarks

In cellular systems, such as GSM or the North-American D-AMPS, TDMA is combined with FDMA. Different frequencies are used in neighboring cells to provide orthogonal signalling without the need for tight synchronization of base stations. Furthermore, channel assignment can then be performed in each cell individually. Within a cell, one or more frequencies are shared by users in the time domain.

From an implementation standpoint TDMA systems have the advantage that common radio and signal processing equipment at the base station can be shared by users communicating on the same frequency. A somewhat more subtle advantage of TDMA systems arises from the possibility of monitoring surrounding base stations and frequencies for signal quality to support mobile assisted handovers.

69.5 Code Division Multiple Access

CDMA systems employ wideband signals with good cross-correlation properties [7]. That means the output of a filter matched to one user's signal is small when a different user's signal is input. A large body of work exists on spreading sequences that lead to signal sets with small cross correlations [10]. Because of their noise-like appearance such sequences are often referred to as pseudonoise (PN) sequences, and because of their wideband nature CDMA systems are often called spread-spectrum systems.

Spectrum spreading can be achieved mainly in two ways: through frequency hopping as explained earlier or through direct sequence spreading. In direct sequence spread spectrum, a high-rate, antipodal pseudorandom spreading sequence modulates the transmitted signal such that the bandwidth of the resulting signal is roughly equal to the rate of the spreading sequence. The cross-correlation of the signals is then largely determined by the cross-correlation properties of the spreading signals. Clearly, CDMA signals overlap in both time and frequency domains but are separable based on their spreading waveforms.

An immediate consequence of this observation is that CDMA systems do not require tight synchronization between users as do TDMA systems. By the same token, frequency planning and management are not required as frequencies are reused throughout the coverage area.

Propagation Considerations

Spread spectrum is well suited for wireless communication systems because of its built-in frequency diversity. As discussed before, in cellular systems the delay spread measures several microseconds and, hence, the coherence bandwidth of the channel is smaller than 1 MHz. Spreading rates can be chosen to exceed the coherence bandwidth such that the channel becomes frequency selective, i.e., different spectral components are affected unequally by the channel and only parts of the signal are affected by fades. Expressing the same observation in time domain terms, multipath components are resolvable at a resolution equal to the chip period and can be combined coherently, for example, by means of a RAKE receiver [8]. An estimate of the channel impulse response is required for the coherent combination of multipath components. This estimate can be gained from a training sequence or by means of a so-called pilot signal.

Even for cordless telephone systems, operating in environments with submicrosecond delay spread and corresponding coherence bandwidths of a few megahertz, the spreading rate can be chosen large enough to facilitate multipath diversity. If the combination of multipath components already described is deemed too complex, a simpler, but less powerful, form of diversity can be used that decorrelates only the strongest received multipath component and relies on the suppression of other path components by the matched filter.

Multiple-Access Interference

If it is possible to control the relative timing of the transmitted signals, such as on the downlink, the transmitted signals can be made perfectly orthogonal, and if the channel only adds white Gaussian noise, matched filter receivers are optimal for extracting a signal from the superposition of waveforms. If the channel is dispersive because of multipath, the signals arriving at the receiver will no longer be orthogonal and will introduce some multiple-access interference, i.e., signal components from other signals that are not rejected by the matched filter.

On the uplink, extremely tight synchronization between users to within a fraction of a chip period, which is defined as the inverse of the spreading rate, is generally not possible, and measures to control the impact of multiple-access interference must be taken. Otherwise, the near–far problem, i.e., the problem of very strong undesired users' signals overwhelming the weaker signal of the desired user, can severely decrease performance. Two approaches are proposed to overcome the near–far problem: power control with soft handovers and multiuser detection.

Power control attempts to ensure that signals from all mobiles in a cell arrive at the base station with approximately equal power levels. To be effective, power control must be accurate to within about 1 dB and fast enough to compensate for channel fading. For a mobile moving at 55 mph and transmitting at 1 GHz, the Doppler bandwidth is approximately 100 Hz. Hence, the channel changes its characteristic drastically about 100 times per second and on the order of 1000 b/s must be sent from base station to mobile for power control purposes. As different mobiles may be subject to vastly different fading and shadowing conditions, a large dynamic range of about 80 dB must be covered by power control. Notice, that power control on the downlink is really only necessary for mobiles that are about equidistant from two base stations, and even then neither the update rate nor the dynamic range of the uplink is required.

The interference problem that arises at the cell boundaries where mobiles are within range of two or more base stations can be turned into an advantage through the idea of soft handover. On the downlink, all base stations within range can transmit to the mobile, which in turn can combine the received signals to achieve some gain from the antenna diversity. On the uplink, a similar effect can be obtained by selecting the strongest received signal from all base stations that received a user's signal. The base station that receives the strongest signal will also issue power control commands to minimize the transmit power of the mobile. Note, however, that soft handover requires fairly tight synchronization between base stations, and one of the advantages of CDMA over TDMA is lost.

Multiuser detection is still an emerging technique. It is probably best used in conjunction with power control. The fundamental idea behind this technique is to model multiple-access interference explicitly

and devise receivers that reject or cancel the undesired signals. A variety of techniques have been proposed ranging from optimum maximum-likelihood sequence estimation via multistage schemes, reminiscent of decision feedback algorithms, to linear decorrelating receivers. An excellent survey of the theory and practice of multiuser detection is given by [11].

Further Remarks

CDMA systems work well in conjunction with frequency division duplexing. This arrangement decouples the power control problem on the uplink and downlink, respectively.

Signal quality enhancing methods, such as time diversity through coding and interleaving, can be applied just as with the other access methods. In spread spectrum systems, however, coding can be built into the spreading process, avoiding the loss of bandwidth associated with error protection. Additionally, CDMA lends itself naturally to the exploitation of speech pauses that make up more than half the time of a connection. If no signals are transmitted during such pauses, then the instantaneous interference level is reduced and the total number of users supportable by the system can be approximately doubled.

69.6 Comparison and Outlook

The question of which of the access methods is best does not have a single answer. Based on the preceding discussion FDMA is only suited for applications such as cordless telephone with very small cells and submicrosecond delay spreads. In cellular systems and for most versions of personal communication systems, the choice reduces to TDMA vs. CDMA.

In terms of complexity, TDMA receivers require adaptive, nonlinear equalizers when operating in environments with large delay spreads. CDMA systems, in turn, need RAKE receivers and sophisticated power control algorithms. In the future, some form of multiple-access interference rejection is likely to be implemented as well. Time synchronization is required in both systems, albeit for different reasons. The additional complexity for coding and interleaving is comparable for both access methods.

An often quoted advantage of CDMA systems is the fact that the performance will degrade gracefully as the load increases. In TDMA systems, in turn, requests will have to be blocked once all channels in a cell are in use. Hence, there is a hard limit on the number of channels per cell. There are proposals for extended TDMA systems, however, that incorporate reassignment of channels during speech pauses. Not only would such extended TDMA systems match the advantage of the exploitation of speech pauses of CDMA systems, they would also lead to a soft limit on the system capacity. The extended TDMA proposals would implement the statistical multiplexing of the user data, e.g., by means of the packet reservation multiple access protocol [4]. The increase in capacity depends on the acceptable packet loss rate; in other words, small increases in the load lead to small increases in the packet loss probability.

Many comparisons in terms of capacity between TDMA and CDMA can be found in the recent literature. Such comparisons, however, are often invalidated by making assumptions that favor one access method over the other. An important exception constitutes the recent paper by Wyner [12]. Under a simplified model that nevertheless captures the essence of cellular systems, he computes the Shannon capacity. Highlights of his results include the following.

- TDMA is distinctly suboptimal in cellular systems.
- When the signal-to-noise-ratio is large, CDMA appears to achieve twice the capacity of TDMA.
- Multiuser detectors are essential to realize near-optimum performance in CDMA systems.
- Intercell interference in CDMA systems has a detrimental effect when the signal-to-noise ratio is large, but it can be exploited via diversity combining to increase capacity when the signal-to-noise ratio is small.

More research along this avenue is necessary to confirm the validity of the results. In particular, incorporation of realistic channel models into the analysis is required. However, this work represents a substantial step toward quantifying capacity increases achievable with CDMA.

Defining Terms

Capacity: Shannon originally defined capacity as the maximum data rate which permits error-free communication in a given environment. A looser interpretation is normally employed in wireless communication systems. Here capacity denotes the traffic density supported by the system under consideration normalized with respect to bandwidth and coverage area.

Code-division multiple access (CDMA): Systems use signals with very small cross-correlations to facilitate sharing of the broadcast radio channel. Correlators are used to extract the desired user's signal while simultaneously suppressing interfering, parallel transmissions.

Duplexing: Refers to the exchange of messages in both directions of a connection.

Frequency-division multiple access (FDMA): Simultaneous access to the radio channel is facilitated by assigning nonoverlapping frequency bands to different users.

Multiple access: Denotes the support of simultaneous transmissions over a shared communication channel.

Random-access channel: This uplink control channel is used by mobiles to request assignment of a traffic channel. A random access protocol is employed to arbitrate access to this channel.

Time-division multiple access (TDMA): Systems assign nonoverlapping time slots to different users in a round-robin fashion.

References

1. Bertsekas, D. and Gallager, R., *Data Networks,* Prentice-Hall, Englewood Cliffs, NJ, 1987.
2. Cox, D.C., Wireless network access for personal communications, *IEEE Comm. Mag.,* 96–115, 1992.
3. Falconer, D.D., Adachi, F., and Gudmundson, B. Time division multiple access methods for wireless personal communications. *IEEE Comm. Mag.,* 33(1), 50–57, 1995.
4. Goodman, D., Trends in cellular and cordless communications. *IEEE Comm. Mag.,* 31–40, 1991a.
5. Goodman, D.J., Second generation wireless information networks. *IEEE Trans. on Vehicular Tech.,* 40(2), 366–374, 1991b.
6. Hodges, M.R.L., The GSM radio interface. *Br. Telecom Tech. J.,* 8(1), 31–43, 1990.
7. Kohno, R., Meidan, R., and Milstein, L.B., Spread spectrum access methods for wireless communications. *IEEE Comm. Mag.,* 33(1), 58, 1995.
8. Proakis, J.G., *Digital Communications.* 2nd ed., McGraw-Hill, New York, 1989.
9. Proakis, J.G., Adaptive equalization for TDMA digital mobile radio. *IEEE Trans. on Vehicular Tech.,* 40(2), 333–341, 1991.
10. Sarwate, D.V. and Pursley, M.B., Crosscorrelation properties of pseudorandom and related sequences. *Proceedings of the IEEE,* 68(5), 593–619, 1980.
11. Verdu, S., Multi-user detection. In *Advances in Statistical Signal Processing—Vol. 2: Signal Detection,* JAI Press, Greenwich, CT, 1992.
12. Wyner, A.D., Shannon-theoretic approach to a Gaussian cellular multiple-access channel. *IEEE Trans. on Information Theory,* 40(6), 1713–1727, 1994.

Further Information

Several of the IEEE publications, including the *Transactions on Communications, Journal on Selected Areas in Communications, Transactions on Vehicular Technology, Communications Magazine*, and *Personal Communications Magazine* contain articles on the subject of access methods on a regular basis.

70

Rayleigh Fading Channels[1]

Bernard Sklar
Communications Engineering Services

[1]A version of this chapter has appeared as two papers in the *IEEE Communications Magazine*, September 1997, under the titles "Rayleigh Fading Channels in Mobile Digital Communication Systems, Part I: Characterization" and "Part II: Mitigation."

70.1 Introduction

When the mechanisms of fading channels were first modeled in the 1950s and 1960s, the ideas were primarily applied to over-the-horizon communications covering a wide range of frequency bands. The 3–30 MHz high-frequency (HF) band is used for ionospheric communications, and the 300 MHz–3 GHz ultra-high-frequency (UHF) and 3–30 GHz super-high-frequency (SHF) bands are used for tropospheric scatter. Although the fading effects in a mobile radio system are somewhat different from those in ionospheric and tropospheric channels, the early models are still quite useful to help characterize fading effects in mobile digital communication systems. This chapter addresses Rayleigh fading, primarily in the UHF band, that affects mobile systems such as cellular and personal communication systems (PCS). The chapter itemizes the fundamental fading manifestations, types of degradation, and methods to mitigate the degradation. Two particular mitigation techniques are examined: the Viterbi equalizer implemented in the Global System for Mobile Communication (GSM), and the Rake receiver used in CDMA systems built to meet Interim Standard-95 (IS-95).

70.2 The Challenge of a Fading Channel

In the study of communication systems, the classical (ideal) additive-white-Gaussian-noise (AWGN) channel, with statistically independent Gaussian noise samples corrupting data samples free of intersymbol interference (ISI), is the usual starting point for understanding basic performance relationships. The primary source of performance degradation is thermal noise generated in the receiver. Often, external interference received by the antenna is more significant than the thermal noise. This external interference can sometimes be characterized as having a broadband spectrum and quantified by a parameter called antenna temperature [1]. The thermal noise usually has a flat power spectral density over the signal band and a zero-mean Gaussian voltage probability density function (pdf). When modeling practical systems, the next step is the introduction of bandlimiting filters. The filter in the transmitter usually serves to satisfy some regulatory requirement on spectral containment. The filter in the receiver often serves the purpose of a classical "matched filter" [1,2]. Due to the bandlimiting and phase-distortion properties of filters, special signal design and equalization techniques may be required to mitigate the filter-induced ISI.

If a radio channel's propagating characteristics are not specified, one usually infers that the signal attenuation vs. distance behaves as if propagation takes place over ideal free space. The model of free space treats the region between the transmit and receive antennas as being free of all objects that might absorb or reflect radio frequency (RF) energy. It also assumes that, within this region, the atmosphere behaves as a perfectly uniform and nonabsorbing medium. Furthermore, the earth is treated as being infinitely far away from the propagating signal (or, equivalently, as having a reflection coefficient that is negligible). Basically, in this idealized free-space model, the attenuation of RF energy between the transmitter and receiver behaves according to an inverse-square law. The received power expressed in terms of transmitted power is attenuated by a factor, $L_s(d)$, where this factor is called **path loss** or **free space loss**. When the receiving antenna is isotropic, this factor is expressed as [1]:

$$L_s(d) = \left(\frac{4\pi d}{\lambda}\right)^2 \tag{70.1}$$

In Eq. (70.1), d is the distance between the transmitter and the receiver, and λ is the wavelength of the propagating signal. For this case of idealized propagation, received signal power is very predictable.

For most practical channels, where signal propagation takes place in the atmosphere and near the ground, the free-space propagation model is inadequate to describe the channel and predict system performance. In a wireless mobile communication system, a signal can travel from transmitter to receiver over multiple reflective paths; this phenomenon is referred to as **multipath propagation**. The effect can cause fluctuations in the received signal's amplitude, phase, and angle of arrival, giving rise to the terminology **multipath fading**. Another name, **scintillation**, having originated in radio astronomy, is used to describe the multipath fading caused by physical changes in the propagating medium, such as variations in the density of ions in the ionospheric layers that reflect high frequency (HF) radio signals. Both names, fading and scintillation, refer to a signal's random fluctuations or fading due to multipath propagation. The main difference is that scintillation involves mechanisms (e.g., ions) that are much smaller than a wavelength. The end-to-end modeling and design of systems that mitigate the effects of fading are usually more challenging than those whose sole source of performance degradation is AWGN.

70.3 Mobile-Radio Propagation: Large-Scale Fading and Small-Scale Fading

Figure 70.1 represents an overview of fading channel manifestations. It starts with two types of fading effects that characterize mobile communications: large-scale fading and small-scale fading. Large-scale fading represents the average signal power attenuation or the path loss due to motion over large areas. In Fig. 70.1, the large-scale fading manifestation is shown in blocks 1, 2, and 3. This phenomenon is affected by prominent terrain contours (e.g., hills, forests, billboards, clumps of buildings, etc.) between the transmitter and receiver. The receiver is often represented as being "shadowed" by such prominences. The statistics of large-scale fading provide a way of computing an estimate of path loss as a function of distance. This is described in terms of a mean-path loss (nth-power law) and a log-normally distributed variation about the mean. Small-scale fading refers to the dramatic changes in signal amplitude and phase that can be experienced as a result of small changes (as small as a half-wavelength) in the spatial separation between a receiver and transmitter. As indicated in Fig. 70.1, blocks 4, 5, and 6, small-scale fading manifests itself in two mechanisms, namely, time-spreading of the signal (or signal dispersion) and time-variant behavior of the channel. For mobile-radio applications, the channel is time-variant because motion between the transmitter and receiver results in propagation path changes. The rate of change of these propagation conditions accounts for the fading rapidity (rate of change of the fading impairments). Small-scale fading is also called **Rayleigh fading** because if the multiple reflective paths are large in number and there is no line-of-sight signal component, the envelope of the received signal is statistically described by a Rayleigh pdf. When there is a dominant nonfading signal component present, such as a line-of-sight propagation path, the small-scale fading envelope is described by a Rician pdf [3]. A mobile radio roaming over a large area must process signals that experience both types of fading: small-scale fading superimposed on large-scale fading.

There are three basic mechanisms that impact signal propagation in a mobile communication system. They are reflection, diffraction, and scattering [3].

- Reflection occurs when a propagating electromagnetic wave impinges upon a smooth surface with very large dimensions compared to the RF signal wavelength (λ).
- Diffraction occurs when the radio path between the transmitter and receiver is obstructed by a dense body with large dimensions compared to λ, causing secondary waves to be formed behind the obstructing body. Diffraction is a phenomenon that accounts for RF energy travelling from transmitter to receiver without a line-of-sight path between the two. It is often termed **shadowing** because the diffracted field can reach the receiver even when shadowed by an impenetrable obstruction.

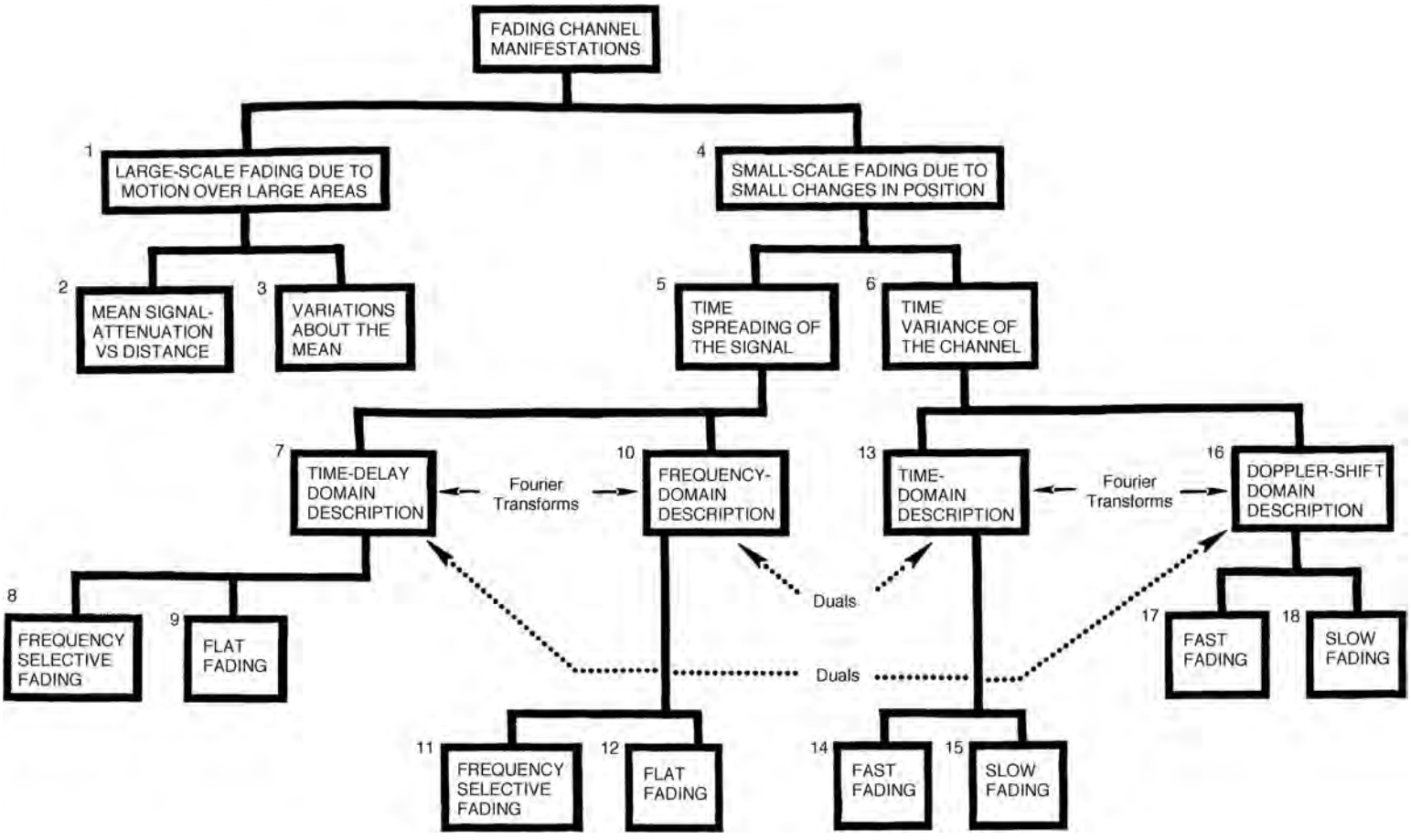


FIGURE 70.1 Fading channel manifestations.

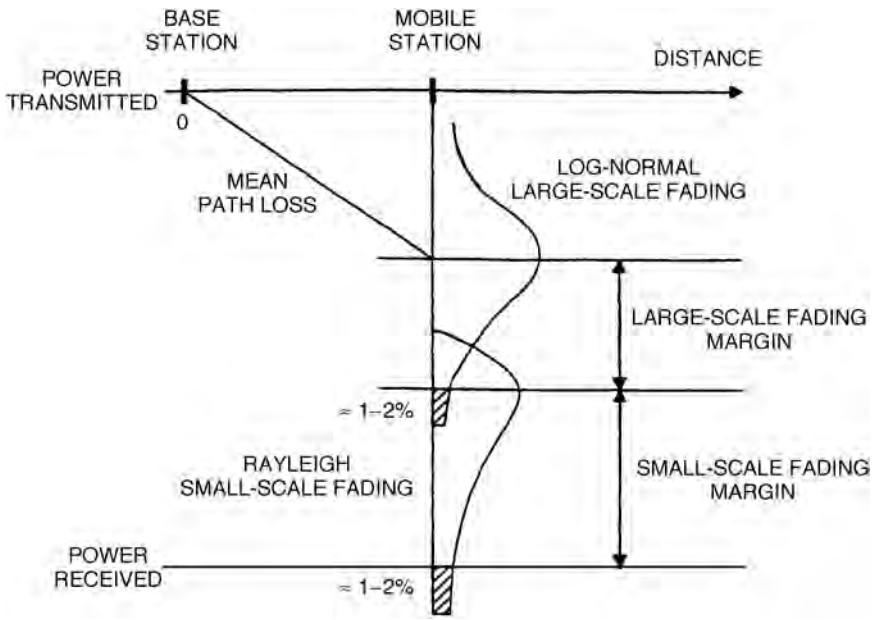


FIGURE 70.2 Link-budget considerations for a fading channel.

- Scattering occurs when a radio wave impinges on either a large rough surface or any surface whose dimensions are on the order of λ or less, causing the reflected energy to spread out (scatter) in all directions. In an urban environment, typical signal obstructions that yield scattering are lampposts, street signs, and foliage.

Figure 70.1 may serve as a table of contents for the sections that follow. We will examine the two manifestations of small-scale fading: signal time-spreading (signal dispersion) and the time-variant nature of the channel. These examinations will take place in two domains: time and frequency, as indicated in Fig. 70.1, blocks 7, 10, 13, and 16. For signal dispersion, we categorize the fading degradation types as being frequency-selective or frequency-nonselective (flat), as listed in blocks 8, 9, 11, and 12. For the time-variant manifestation, we categorize the fading degradation types as fast-fading or slow-fading, as listed in blocks 14, 15, 17, and 18. The labels indicating Fourier transforms and duals will be explained later.

Figure 70.2 illustrates the various contributions that must be considered when estimating path loss for a link budget analysis in a cellular application [4]. These contributions are:

- Mean path loss as a function of distance, due to large-scale fading
- Near-worst-case variations about the mean path loss (typically 6–10 dB) or large-scale fading margin
- Near-worst-case Rayleigh or small-scale fading margin (typically 20–30 dB)

In Fig. 70.2, the annotations "≈1–2%" indicate a suggested area (probability) under the tail of each pdf as a design goal. Hence, the amount of margin indicated is intended to provide adequate received signal power for approximately 98–99% of each type of fading variation (large- and small-scale).

A received signal, is generally described in terms of a transmitted signal $s(t)$ convolved with the impulse response of the channel $h_c(t)$. Neglecting the degradation due to noise, we write:

$$r(t) = s(t) * h_c(t) \tag{70.2}$$

where $*$ denotes convolution. In the case of mobile radios, $r(t)$ can be partitioned in terms of two component random variables, as follows [5]:

$$r(t) = m(t) \times r_0(t) \tag{70.3}$$

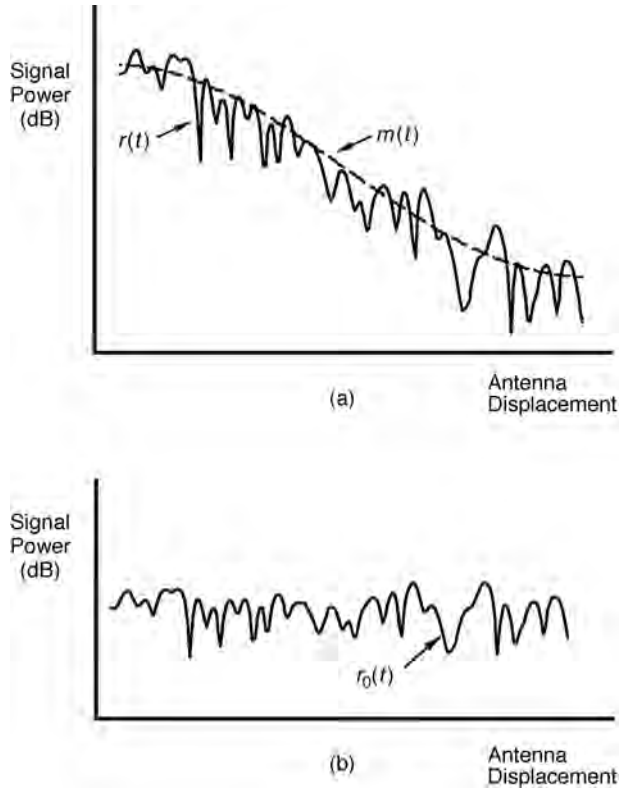


FIGURE 70.3 Large-scale fading and small-scale fading.

where $m(t)$ is called the large-scale-fading component, and $r_0(t)$ is called the small-scale-fading component. $m(t)$ is sometimes referred to as the **local mean** or **log-normal fading** because the magnitude of $m(t)$ is described by a log-normal pdf (or, equivalently, the magnitude measured in decibels has a Gaussian pdf). $r_0(t)$ is sometimes referred to as multipath or Rayleigh fading. Figure 70.3 illustrates the relationship between large-scale and small-scale fading. In Fig. 70.3(a), received signal power $r(t)$ vs. antenna displacement (typically in units of wavelength) is plotted for the case of a mobile radio. Small-scale fading superimposed on large-scale fading can be readily identified. The typical antenna displacement between the small-scale signal nulls is approximately a half wavelength. In Fig. 70.3(b), the large-scale fading or local mean, $m(t)$, has been removed in order to view the small-scale fading, $r_0(t)$, about some average constant power.

In the sections that follow, we enumerate some of the details regarding the statistics and mechanisms of large-scale and small-scale fading.

Large-Scale Fading: Path-Loss Mean and Standard Deviation

For the mobile radio application, Okumura [6] made some of the earlier comprehensive path-loss measurements for a wide range of antenna heights and coverage distances. Hata [7] transformed Okumura's data into parametric formulas. For the mobile radio application, the mean path loss, $\overline{L_p}(d)$, as a function of distance, d, between the transmitter and receiver is proportional to an nth-power of d relative to a reference distance d_0 [3].

$$\overline{L_p}(d) \propto \left(\frac{d}{d_0}\right)^n \tag{70.4}$$

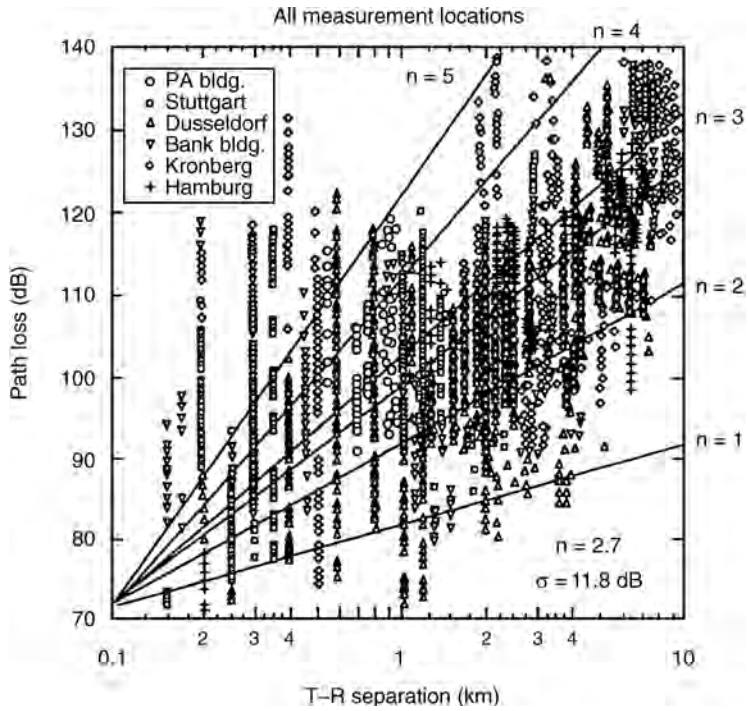


FIGURE 70.4 Path loss vs. distance measured in several German cities.

$\overline{L_p}(d)$ is often stated in decibels, as shown below.

$$\overline{L_p}(d)\ \ (\text{dB}) = L_s(d_0)\ \ (\text{dB}) + 10n\log\left(\frac{d}{d_0}\right) \tag{70.5}$$

The reference distance, d_0, corresponds to a point located in the far field of the antenna. Typically, the value of d_0 is taken to be 1 km for large cells, 100 m for microcells, and 1 m for indoor channels. $\overline{L_p}(d)$ is the average path loss (over a multitude of different sites) for a given value of d. Linear regression for a minimum mean-squared estimate (MMSE) fit of $\overline{L_p}(d)$ vs. d on a log-log scale (for distance greater than d_0) yields a straight line with a slope equal to 10 n dB/decade. The value of the exponent n depends on the frequency, antenna heights, and propagation environment. In free space, $n = 2$, as seen in Eq. (70.1). In the presence of a very strong guided wave phenomenon (like urban streets), n can be lower than 2. When obstructions are present, n is larger. The path loss $L_s(d_0)$ to the reference point at a distance d_0 from the transmitter is typically found through field measurements or is calculated using the free-space path loss given by Eq. (70.1). Figure 70.4 shows a scatter plot of path loss vs. distance for measurements made in several German cities [8]. Here, the path loss has been measured relative to the free-space reference measurement at $d_0 = 100$ m. Also shown are straight-line fits to various exponent values.

The path loss vs. distance expressed in Eq. (70.5) is an average, and therefore not adequate to describe any particular setting or signal path. It is necessary to provide for variations about the mean since the environment of different sites may be quite different for similar transmitter-receiver separations. Figure 70.4 illustrates that path-loss variations can be quite large. Measurements have shown that for any value of d, the path loss $L_p(d)$ is a random variable having a log-normal distribution about the mean distant-dependent value $\overline{L_p}(d)$ [9]. Thus, path loss $L_p(d)$ can be expressed in terms of $\overline{L_p}(d)$ plus a random variable X_σ, as follows [3].

$$L_p(d)\ \ (\text{dB}) = L_s(d_0)\ \ (\text{dB}) + 10n\ \log_{10}\left(\frac{d}{d_0}\right) + X_\sigma\ \ (\text{dB}) \tag{70.6}$$

where X_σ denotes a zero-mean, Gaussian random variable (in decibels) with standard deviation σ (also in decibels). X_σ is site and distance dependent. The choice of a value for X_σ is often based on measurements; it is not unusual for X_σ to take on values as high as 6–10 dB or greater. Thus, the parameters needed to statistically describe path loss due to large-scale fading for an arbitrary location with a specific transmitter-receiver separation are:

- The reference distance d_0
- The path-loss exponent n
- The standard deviation σ of X_σ

There are several good references dealing with the measurement and estimation of propagation path loss for many different applications and configurations [3], [7]–[11].

Small-Scale Fading: Statistics and Mechanisms

When the received signal is made up of multiple reflective rays plus a significant line-of-sight (nonfaded) component, the envelope amplitude due to small-scale fading has a Rician pdf, and is referred to as **Rician fading** [3]. The nonfaded component is called the **specular component**. As the amplitude of the specular component approaches zero, the Rician pdf approaches a Rayleigh pdf, expressed as:

$$p(r) = \begin{cases} \dfrac{r}{\sigma^2}\exp\left[-\dfrac{r^2}{2\sigma^2}\right] & \text{for } r \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{70.7}$$

where r is the envelope amplitude of the received signal, and $2\sigma^2$ is the predetection mean power of the multipath signal. The Rayleigh faded component is sometimes called the **random**, **scatter**, or **diffuse component**. The Rayleigh pdf results from having no specular component of the signal; thus for a single link it represents the pdf associated with the worst case of fading per mean received signal power. For the remainder of this chapter, it will be assumed that loss of signal-to-noise ratio (SNR) due to fading follows the Rayleigh model described. It will also be assumed that the propagating signal is in the UHF band, encompassing present-day cellular and personal communications services (PCS) frequency allocations—nominally 1 GHz and 2 GHz, respectively.

As indicated in Fig. 70.1, blocks 4, 5, and 6, small-scale fading manifests itself in two mechanisms:

- Time-spreading of the underlying digital pulses within the signal
- A time-variant behavior of the channel due to motion (e.g., a receive antenna on a moving platform)

Figure 70.5 illustrates the consequences of both manifestations by showing the response of a multipath channel to a narrow pulse vs. delay, as a function of antenna position (or time, assuming a constant velocity of motion). In Fig. 70.5, we distinguish between two different time references—delay time τ and transmission or observation time t. Delay time refers to the time-spreading manifestation which results from the fading channel's nonoptimum impulse response. The transmission time, however, is related to the antenna's motion or spatial changes, accounting for propagation path changes that are perceived as the channel's time-variant behavior. Note that, for constant velocity, as is assumed in Fig. 70.5, either antenna position or transmission time can be used to illustrate this time-variant behavior. Figures 70.5(a)–(c) show the sequence of received pulse-power profiles as the antenna moves through a succession of equally spaced positions. Here, the interval between antenna positions is 0.4 λ, where λ is the wavelength of the carrier frequency. For each of the three cases shown, the response-pattern differs significantly in the delay time of the largest signal component, the number of signal copies, their magnitudes, and the total received power (area) in the received power profile. Figure 70.6 summarizes these two small-scale fading mechanisms, the two domains (time or time-delay and frequency or Doppler shift) for viewing each mechanism and the degradation categories each mechanism can exhibit.

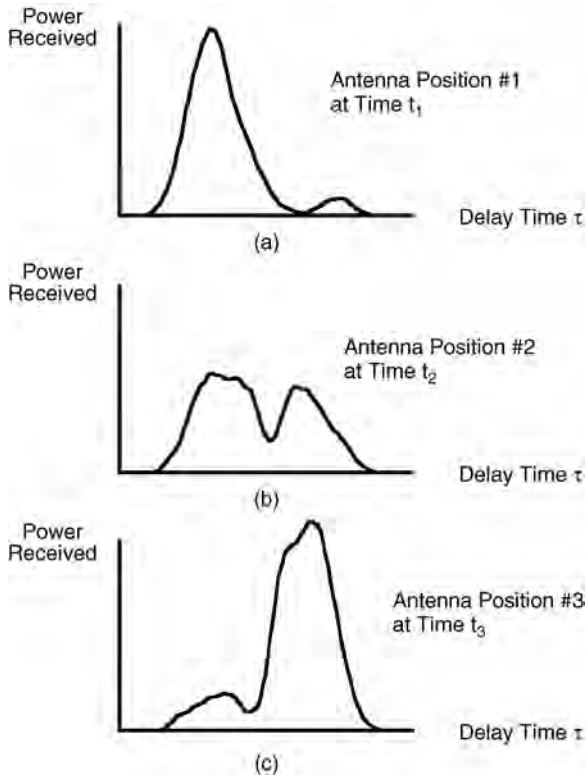


FIGURE 70.5 Response of a multipath channel to a narrow pulse vs. delay, as a function of antenna position.

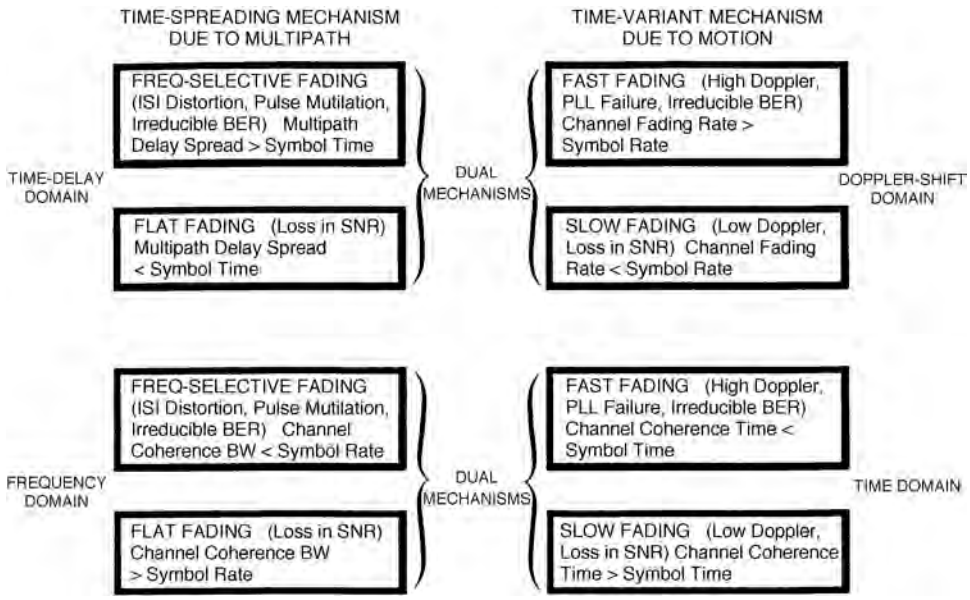


FIGURE 70.6 Small-scale fading: mechanisms, degradation categories, and effects.

Note that any mechanism characterized in the time domain can be characterized equally well in the frequency domain. Hence, as outlined in Fig. 70.6, the time-spreading mechanism will be characterized in the time-delay domain as a multipath delay spread and in the frequency domain as a channel coherence bandwidth. Similarly, the time-variant mechanism will be characterized in the time domain as a channel coherence time and in the Doppler-shift (frequency) domain as a channel fading rate or Doppler spread. These mechanisms and their associated degradation categories will be examined in greater detail in the sections that follow.

70.4 Signal Time-Spreading Viewed in the Time-Delay Domain: Figure 70.1, Block 7—The Multipath Intensity Profile

A simple way to model the fading phenomenon was introduced by Bello [13] in 1963; he proposed the notion of wide-sense stationary uncorrelated scattering (WSSUS). The model treats signal variations arriving with different delays as uncorrelated. It can be shown [4,13] that such a channel is effectively WSS in both the time and frequency domains. With such a model of a fading channel, Bello was able to define functions that apply for all time and all frequencies. For the mobile channel, Fig. 70.7 contains four functions that make up this model [4], [13]–[16]. We will examine these functions, starting with Fig. 70.7(a) and proceeding counter-clockwise toward Fig. 70.7(d).

In Fig. 70.7(a), a **multipath-intensity profile**, $S(\tau)$ vs. time delay τ is plotted. Knowledge of $S(\tau)$ helps answer the question, "For a transmitted impulse, how does the average received power vary as a function of time delay, τ?" The term "time delay" is used to refer to the excess delay. It represents the signal's

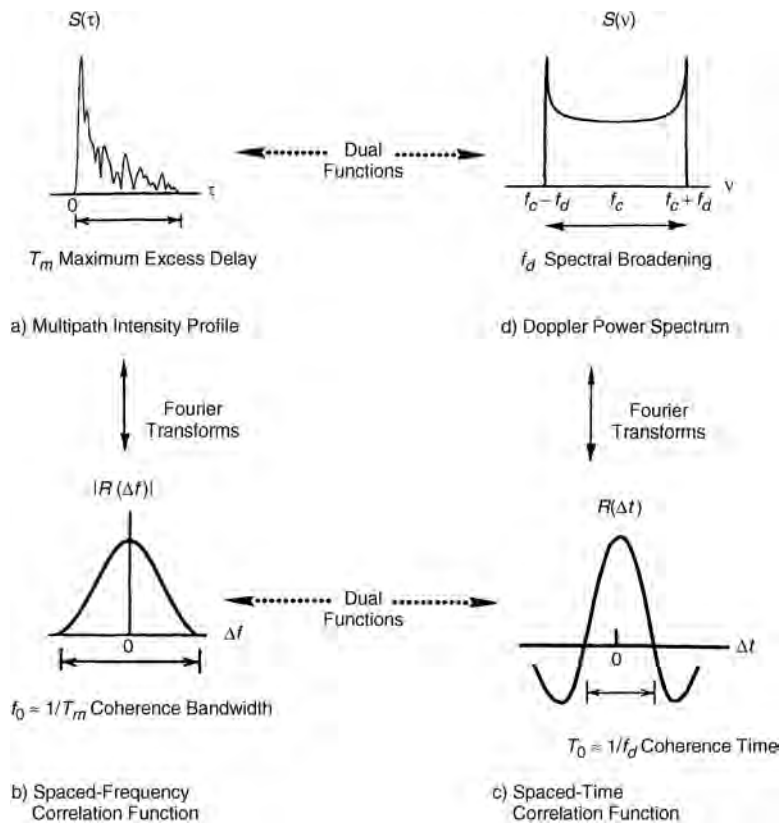


FIGURE 70.7 Relationships among the channel correlation functions and power density functions.

propagation delay that exceeds the delay of the first signal arrival at the receiver. For a typical wireless radio channel, the received signal usually consists of several discrete multipath components, sometimes referred to as fingers. For some channels, such as the tropospheric scatter channel, received signals are often seen as a continuum of multipath components [14,16]. For making measurements of the multipath intensity profile, wideband signals (impulses or spread spectrum) need to be used [16]. For a single transmitted impulse, the time, T_m, between the first and last received component represents the **maximum excess delay**, during which the multipath signal power falls to some threshold level below that of the strongest component. The threshold level might be chosen at 10 dB or 20 dB below the level of the strongest component. Note, that for an ideal system (zero excess delay), the function $S(\tau)$ would consist of an ideal impulse with weight equal to the total average received signal power.

Degradation Categories due to Signal Time-Spreading Viewed in the Time-Delay Domain

In a fading channel, the relationship between maximum excess delay time, T_m, and symbol time, T_s, can be viewed in terms of two different degradation categories, **frequency-selective fading** and **frequency nonselective** or **flat fading**, as indicated in Fig. 70.1, blocks 8 and 9, and Fig. 70.6. A channel is said to exhibit frequency-selective fading if $T_m > T_s$. This condition occurs whenever the received multipath components of a symbol extend beyond the symbol's time duration. Such multipath dispersion of the signal yields the same kind of ISI distortion that is caused by an electronic filter. In fact, another name for this category of fading degradation is **channel-induced ISI**. In the case of frequency-selective fading, mitigating the distortion is possible because many of the multipath components are resolvable by the receiver. Later, several such mitigation techniques are described.

A channel is said to exhibit frequency nonselective or flat fading if $T_m < T_s$. In this case, all of the received multipath components of a symbol arrive within the symbol time duration; hence, the components are not resolvable. Here, there is no channel-induced ISI distortion, since the signal time spreading does not result in significant overlap among neighboring received symbols. There is still performance degradation since the unresolvable phasor components can add up destructively to yield a substantial reduction in SNR. Also, signals that are classified as exhibiting flat fading can sometimes experience frequency-selective distortion. This will be explained later when viewing degradation in the frequency domain, where the phenomenon is more easily described. For loss in SNR due to flat fading, the mitigation technique called for is to improve the received SNR (or reduce the required SNR). For digital systems, introducing some form of signal diversity and using error-correction coding is the most efficient way to accomplish this.

70.5 Signal Time-Spreading Viewed in the Frequency Domain: Figure 70.1, Block 10—The Spaced-Frequency Correlation Function

A completely analogous characterization of signal dispersion can begin in the frequency domain. In Fig. 70.7(b), the function $|R(\Delta f)|$ is seen, designated a **spaced-frequency correlation function**; it is the Fourier transform of $S(\tau)$. $R(\Delta f)$ represents the correlation between the channel's response to two signals as a function of the frequency difference between the two signals. It can be thought of as the channel's frequency transfer function. Therefore, the time-spreading manifestation can be viewed as if it were the result of a filtering process. Knowledge of $R(\Delta f)$ helps answer the question, "What is the correlation between received signals that are spaced in frequency $\Delta f = f_1 - f_2$?" $R(\Delta f)$ can be measured by transmitting a pair of sinusoids separated in frequency by Δf, cross-correlating the two separately received signals, and repeating the process many times with ever-larger separation Δf. Therefore, the measurement of $R(\Delta f)$ can be made with a sinusoid that is swept in frequency across the band of interest (a wideband signal). The **coherence bandwidth**, f_0, is a statistical measure of the range of frequencies over which the channel passes all spectral components with approximately equal gain and linear phase. Thus, the

coherence bandwidth represents a frequency range over which frequency components have a strong potential for amplitude correlation. That is, a signal's spectral components in that range are affected by the channel in a similar manner, as for example, exhibiting fading or no fading. Note that f_0 and T_m are reciprocally related (within a multiplicative constant). As an approximation, it is possible to say that

$$f_0 \approx \frac{1}{T_m} \tag{70.8}$$

The maximum excess delay, T_m, is not necessarily the best indicator of how any given system will perform on a channel because different channels with the same value of T_m can exhibit very different profiles of signal intensity over the delay span. A more useful measurement of delay spread is most often characterized in terms of the root mean squared (rms) delay spread, σ_τ, where

$$\sigma_\tau = \sqrt{\overline{\tau^2} - (\bar{\tau})^2} \tag{70.9}$$

τ is the mean excess delay, $(\tau)^2$ is the mean squared, $\overline{\tau^2}$ is the second moment, and σ_τ is the square root of the second central moment of $S(\tau)$ [3].

An exact relationship between coherence bandwidth and delay spread does not exist, and must be derived from signal analysis (usually using Fourier techniques) of actual signal dispersion measurements in particular channels. Several approximate relationships have been described. If coherence bandwidth is defined as the frequency interval over which the channel's complex frequency transfer function has a correlation of at least 0.9, the coherence bandwidth is approximately [17]

$$f_0 \approx \frac{1}{50\sigma_\tau} \tag{70.10}$$

For the case of a mobile radio, an array of radially uniformly spaced scatterers, all with equal-magnitude reflection coefficients but independent, randomly occurring reflection phase angles [18,19] is generally accepted as a useful model for urban surroundings. This model is referred to as the **dense-scatterer channel model**. With the use of such a model, coherence bandwidth has similarly been defined [18] for a bandwidth interval over which the channel's complex frequency transfer function has a correlation of at least 0.5 to be

$$f_0 = \frac{0.276}{\sigma_\tau} \tag{70.11}$$

The ionospheric-effects community employs the following definition

$$f_0 = \frac{1}{2\pi\sigma_\tau} \tag{70.12}$$

A more popular approximation of f_0 corresponding to a bandwidth interval having a correlation of at least 0.5 is [3]

$$f_0 \approx \frac{1}{5\sigma_\tau} \tag{70.13}$$

Degradation Categories due to Signal Time-Spreading Viewed in the Frequency Domain

A channel is referred to as frequency-selective if $f_0 < 1/T_s \approx W$, where the symbol rate $1/T_s$ is nominally taken to be equal to the signal bandwidth W. In practice, W may differ from $1/T_s$ due to system filtering or data modulation type (quaternary phase shift keying, QPSK, minimum shift keying, MSK, etc.) [21]. Frequency-selective fading distortion occurs whenever a signal's spectral components are not all affected equally by the channel. Some of the signal's spectral components, falling outside the coherence bandwidth,

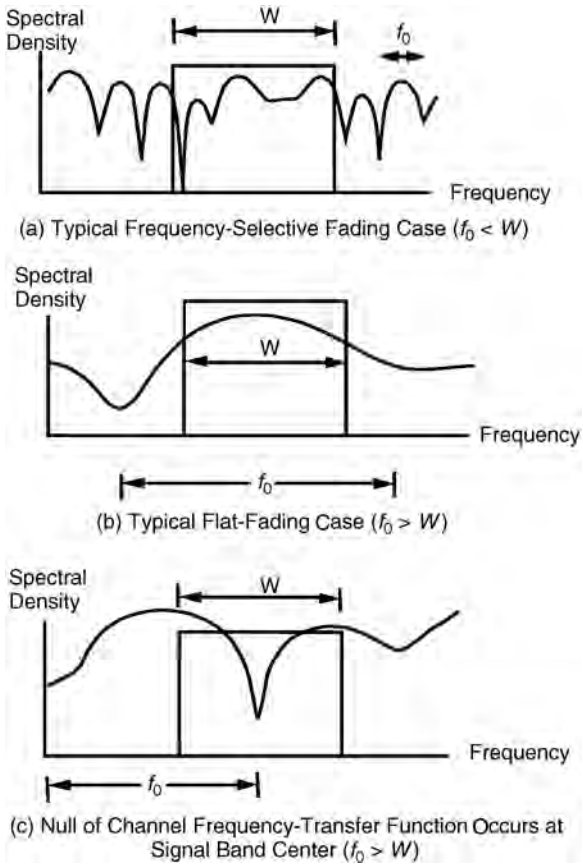


FIGURE 70.8 Relationships between the channel frequency-transfer function and a signal with bandwidth W.

will be affected differently (independently) compared to those components contained within the coherence bandwidth. This occurs whenever $f_0 < W$ and is illustrated in Fig. 70.8(a).

Frequency-nonselective or flat fading degradation occurs whenever $f_0 > W$. Hence, all of the signal's spectral components will be affected by the channel in a similar manner (e.g., fading or no fading); this is illustrated in Fig. 70.8(b). Flat-fading does not introduce channel-induced ISI distortion, but performance degradation can still be expected due to the loss in SNR whenever the signal is fading. In order to avoid channel-induced ISI distortion, the channel is required to exhibit flat fading by ensuring that

$$f_0 > W \approx \frac{1}{T_s} \tag{70.14}$$

Hence, the channel coherence bandwidth f_0 sets an upper limit on the transmission rate that can be used without incorporating an equalizer in the receiver.

For the flat-fading case, where $f_0 > W$ (or $T_m < T_s$), Fig. 70.8(b) shows the usual flat-fading pictorial representation. However, as a mobile radio changes its position, there will be times when the received signal experiences frequency-selective distortion even though $f_0 > W$. This is seen in Fig. 70.8(c), where the null of the channel's frequency transfer function occurs at the center of the signal band. Whenever this occurs, the baseband pulse will be especially mutilated by deprivation of its DC component. One consequence of the loss of DC (zero mean value) is the absence of a reliable pulse peak on which to establish the timing synchronization, or from which to sample the carrier phase carried by the pulse [18]. Thus, even though a channel is categorized as flat fading (based on rms relationships), it can still manifest

frequency-selective fading on occasions. It is fair to say that a mobile radio channel, classified as having flat-fading degradation, cannot exhibit flat fading all of the time. As f_0 becomes much larger than W (or T_m becomes much smaller than T_s), less time will be spent in conditions approximating Fig. 70.8(c). By comparison, it should be clear that in Fig. 70.8(a) the fading is independent of the position of the signal band, and frequency-selective fading occurs all the time, not just occasionally.

70.6 Typical Examples of Flat Fading and Frequency-Selective Fading Manifestations

Figure 70.9 shows some examples of flat fading and frequency-selective fading for a direct-sequence spread-spectrum (DS/SS) system [20,22]. In Fig. 70.9, there are three plots of the output of a pseudonoise (PN) code correlator vs. delay as a function of time (transmission or observation time). Each amplitude vs. delay plot is akin to $S(\tau)$ vs. τ shown in Fig. 70.7(a). The key difference is that the amplitudes shown in Fig. 70.9 represent the output of a correlator; hence, the waveshapes are a function not only of the impulse response of the channel, but also of the impulse response of the correlator. The delay time is expressed in units of chip durations (chips), where the chip is defined as the spread-spectrum minimal-duration keying element. For each plot, the observation time is shown on an axis perpendicular to the amplitude vs. time-delay plane. Figure 70.9 is drawn from a satellite-to-ground communications link exhibiting

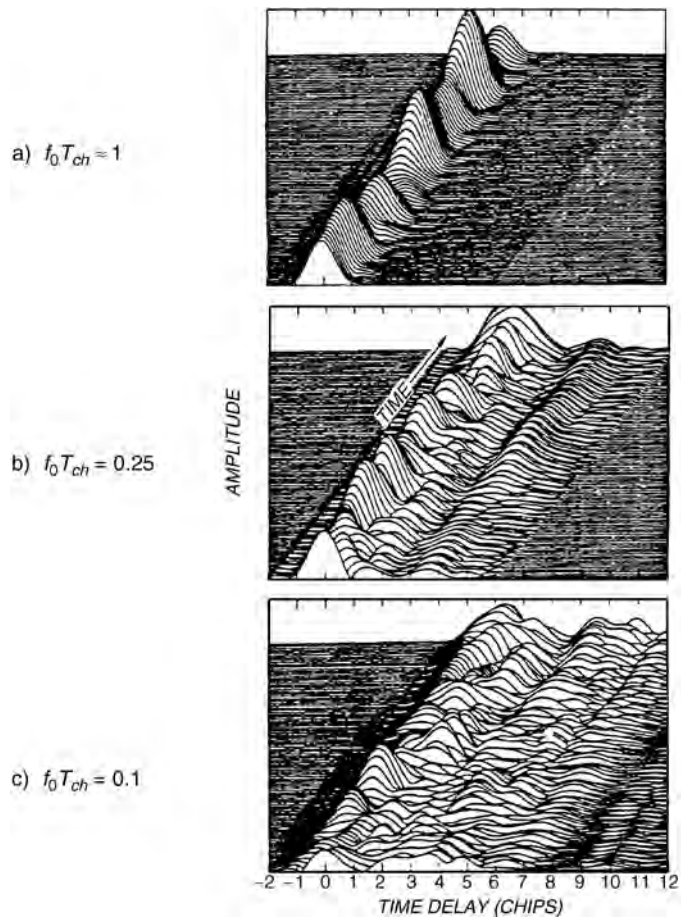


FIGURE 70.9 DS/SS matched-filter output time-history examples for three levels of channel conditions, where T_{ch} is the time duration of a chip.

scintillation because of atmospheric disturbances. However, Fig. 70.9 is still a useful illustration of three different channel conditions that might apply to a mobile radio situation. A mobile radio that moves along the observation-time axis is affected by changing multipath profiles along the route, as seen in the figure. The scale along the observation-time axis is also in units of chips. In Fig. 70.9(a), the signal dispersion (one "finger" of return) is on the order of a chip time duration, T_{ch}. In a typical DS/SS system, the spread-spectrum signal bandwidth is approximately equal to $1/T_{ch}$; hence, the normalized coherence bandwidth f_0T_{ch} of approximately unity in Fig. 70.9(a) implies that the coherence bandwidth is about equal to the spread-spectrum bandwidth. This describes a channel that can be called frequency-nonselective or slightly frequency-selective. In Fig. 70.9(b), where $f_0T_{ch} = 0.25$, the signal dispersion is more pronounced. There is definite interchip interference, and the coherence bandwidth is approximately equal to 25% of the spread-spectrum bandwidth. In Fig. 70.9(c), where $f_0T_{ch} = 0.1$, the signal dispersion is even more pronounced, with greater interchip-interference effects, and the coherence bandwidth is approximately equal to 10% of the spread-spectrum bandwidth. The channels of Figs. 70.9(b) and (c) can be categorized as moderately and highly frequency-selective, respectively, with respect to the basic signalling element, the chip. Later, we show that a DS/SS system operating over a frequency-selective channel at the chip level does not necessarily experience frequency-selective distortion at the symbol level.

70.7 Time Variance Viewed in the Time Domain: Figure 70.1, Block 13—The Spaced-Time Correlation Function

Until now, we have described signal dispersion and coherence bandwidth, parameters that describe the channel's time-spreading properties in a local area. However, they do not offer information about the time-varying nature of the channel caused by relative motion between a transmitter and receiver, or by movement of objects within the channel. For mobile-radio applications, the channel is time variant because motion between the transmitter and receiver results in propagation-path changes. Thus, for a transmitted continuous wave (CW) signal, as a result of such motion, the radio receiver sees variations in the signal's amplitude and phase. Assuming that all scatterers making up the channel are stationary, then whenever motion ceases, the amplitude and phase of the received signal remain constant; that is, the channel appears to be time invariant. Whenever motion begins again, the channel appears time variant. Since the channel characteristics are dependent on the positions of the transmitter and receiver, time variance in this case is equivalent to spatial variance.

Figure 70.7(c) shows the function $R(\Delta t)$, designated the **spaced-time correlation function**; it is the autocorrelation function of the channel's response to a sinusoid. This function specifies the extent to which there is correlation between the channel's response to a sinusoid sent at time t_1 and the response to a similar sinusoid sent at time t_2, where $\Delta t = t_2 - t_1$. The **coherence time**, T_0, is a measure of the expected time duration over which the channel's response is essentially invariant. Earlier, we made measurements of signal dispersion and coherence bandwidth by using wideband signals. Now, to measure the time-variant nature of the channel, we use a narrowband signal. To measure $R(\Delta t)$ we can transmit a single sinusoid ($\Delta f = 0$) and determine the autocorrelation function of the received signal. The function $R(\Delta t)$ and the parameter T_0 provide us with knowledge about the fading rapidity of the channel. Note that for an ideal **time-invariant channel** (e.g., a mobile radio exhibiting no motion at all), the channel's response would be highly correlated for all values of Δt, and $R(\Delta t)$ would be a constant function. When using the dense-scatterer channel model described earlier, with constant velocity of motion, and an unmodulated CW signal, the normalized $R(\Delta t)$ is described as

$$R(\Delta t) = J_0(kV\Delta t) \tag{70.15}$$

where $J_0(\cdot)$ is the zero-order Bessel function of the first kind, V is velocity, $V\Delta t$ is distance traversed, and $k = 2\pi/\lambda$ is the free-space phase constant (transforming distance to radians of phase). Coherence time can be measured in terms of either time or distance traversed (assuming some fixed velocity of motion).

Amoroso described such a measurement using a CW signal and a dense-scatterer channel model [18]. He measured the statistical correlation between the combination of received magnitude and phase sampled at a particular antenna location x_0, and the corresponding combination sampled at some displaced location $x_0 + \zeta$, with displacement measured in units of wavelength λ. For a displacement ζ of 0.38λ between two antenna locations, the combined magnitudes and phases of the received CW are statistically uncorrelated. In other words, the state of the signal at x_0 says nothing about the state of the signal at $x_0 + \zeta$. For a given velocity of motion, this displacement is readily transformed into units of time (coherence time).

The Concept of Duality

Two operators (functions, elements, or systems) are dual when the behavior of one with reference to a time-related domain (time or time-delay) is identical to the behavior of the other with reference to the corresponding frequency-related domain (frequency or Doppler shift).

In Fig. 70.7, we can identify functions that exhibit similar behavior across domains. For understanding the fading channel model, it is useful to refer to such functions as duals. For example, $R(\Delta f)$ in Fig. 70.7(b), characterizing signal dispersion in the frequency domain, yields knowledge about the range of frequency over which two spectral components of a received signal have a strong potential for amplitude and phase correlation. $R(\Delta t)$ in Fig. 70.7(c), characterizing fading rapidity in the time domain, yields knowledge about the span of time over which two received signals have a strong potential for amplitude and phase cor- relation. We have labeled these two correlation functions as duals. This is also noted in Fig. 70.1 as the duality between blocks 10 and 13, and in Fig. 70.6 as the duality between the time-spreading mechanism in the frequency domain and the time-variant mechanism in the time domain.

Degradation Categories due to Time Variance Viewed in the Time Domain

The time-variant nature of the channel or fading rapidity mechanism can be viewed in terms of two degradation categories as listed in Fig. 70.6: **fast fading** and **slow fading**. The terminology "fast fading" is used for describing channels in which $T_0 < T_s$, where T_0 is the channel coherence time and T_s is the time duration of a transmission symbol. Fast fading describes a condition where the time duration in which the channel behaves in a correlated manner is short compared to the time duration of a symbol. Therefore, it can be expected that the fading character of the channel will change several times during the time that a symbol is propagating, leading to distortion of the baseband pulse shape. Analogous to the distortion previously described as channel-induced ISI, here distortion takes place because the received signal's components are not all highly correlated throughout time. Hence, fast fading can cause the baseband pulse to be distorted, resulting in a loss of SNR that often yields an irreducible error rate. Such distorted pulses cause synchronization problems (failure of phase-locked-loop receivers), in addition to difficulties in adequately defining a matched filter.

A channel is generally referred to as introducing slow fading if $T_0 > T_s$. Here, the time duration that the channel behaves in a correlated manner is long compared to the time duration of a transmission symbol. Thus, one can expect the channel state to virtually remain unchanged during the time in which a symbol is transmitted. The propagating symbols will likely not suffer from the pulse distortion described above. The primary degradation in a slow-fading channel, as with flat fading, is loss in SNR.

70.8 Time Variance Viewed in the Doppler-Shift Domain: Figure 70.1, Block 16—The Doppler Power Spectrum

A completely analogous characterization of the time-variant nature of the channel can begin in the Doppler-shift (frequency) domain. Figure 70.7(d) shows a **Doppler power spectral density**, $S(\nu)$, plotted as a function of Doppler-frequency shift, ν. For the case of the dense-scatterer model, a vertical receive antenna with constant azimuthal gain, a uniform distribution of signals arriving at all arrival angles throughout the range

$(0, 2\pi)$, and an unmodulated CW signal, the signal spectrum at the antenna terminals is [19]

$$S(\nu) = \frac{1}{\pi f_d \sqrt{1 - \left(\frac{\nu - f_c}{f_d}\right)^2}} \tag{70.16}$$

The equality holds for frequency shifts of ν that are in the range $\pm f_d$ about the carrier frequency f_c and would be zero outside that range. The shape of the RF Doppler spectrum described by Eq. (70.16) is classically bowl-shaped, as seen in Fig. 70.7(d). Note that the spectral shape is a result of the dense-scatterer channel model. Equation (70.16) has been shown to match experimental data gathered for mobile radio channels [23]; however, different applications yield different spectral shapes. For example, the dense-scatterer model does not hold for the indoor radio channel; the channel model for an indoor area assumes $S(\nu)$ to be a flat spectrum [24].

In Fig. 70.7(d), the sharpness and steepness of the boundaries of the Doppler spectrum are due to the sharp upper limit on the Doppler shift produced by a vehicular antenna traveling among the stationary scatterers of the dense scatterer model. The largest magnitude (infinite) of $S(\nu)$ occurs when the scatterer is directly ahead of the moving antenna platform or directly behind it. In that case the magnitude of the frequency shift is given by

$$f_d = \frac{V}{\lambda} \tag{70.17}$$

where V is relative velocity and λ is the signal wavelength. f_d is positive when the transmitter and receiver move toward each other and negative when moving away from each other. For scatterers directly broadside of the moving platform, the magnitude of the frequency shift is zero. The fact that Doppler components arriving at exactly 0° and 180° have an infinite power spectral density is not a problem, since the angle of arrival is continuously distributed and the probability of components arriving at exactly these angles is zero [3,19].

$S(\nu)$ is the Fourier transform of $R(\Delta t)$. We know that the Fourier transform of the autocorrelation function of a time series is the magnitude squared of the Fourier transform of the original time series. Therefore, measurements can be made by simply transmitting a sinusoid (narrowband signal) and using Fourier analysis to generate the power spectrum of the received amplitude [16]. This Doppler power spectrum of the channel yields knowledge about the spectral spreading of a transmitted sinusoid (impulse in frequency) in the Doppler-shift domain. As indicated in Fig. 70.7, $S(\nu)$ can be regarded as the dual of the multipath intensity profile, $S(\tau)$, since the latter yields knowledge about the time spreading of a transmitted impulse in the time-delay domain. This is also noted in Fig. 70.1 as the duality between blocks 7 and 16, and in Fig. 70.6 as the duality between the time-spreading mechanism in the time-delay domain and the time-variant mechanism in the Doppler-shift domain.

Knowledge of $S(\nu)$ allows us to glean how much spectral broadening is imposed on the signal as a function of the rate of change in the channel state. The width of the Doppler power spectrum is referred to as the **spectral broadening** or **Doppler spread**, denoted by f_d, and sometimes called the **fading bandwidth** of the channel. Equation (70.16) describes the Doppler frequency shift. In a typical multipath environment, the received signal arrives from several reflected paths with different path distances and different angles of arrival, and the Doppler shift of each arriving path is generally different from that of another path. The effect on the received signal is seen as a Doppler spreading or spectral broadening of the transmitted signal frequency, rather than a shift. Note that the Doppler spread, f_d, and the coherence time, T_0, are reciprocally related (within a multiplicative constant). Therefore, we show the approximate relationship between the two parameters as

$$T_0 \approx \frac{1}{f_d} \tag{70.18}$$

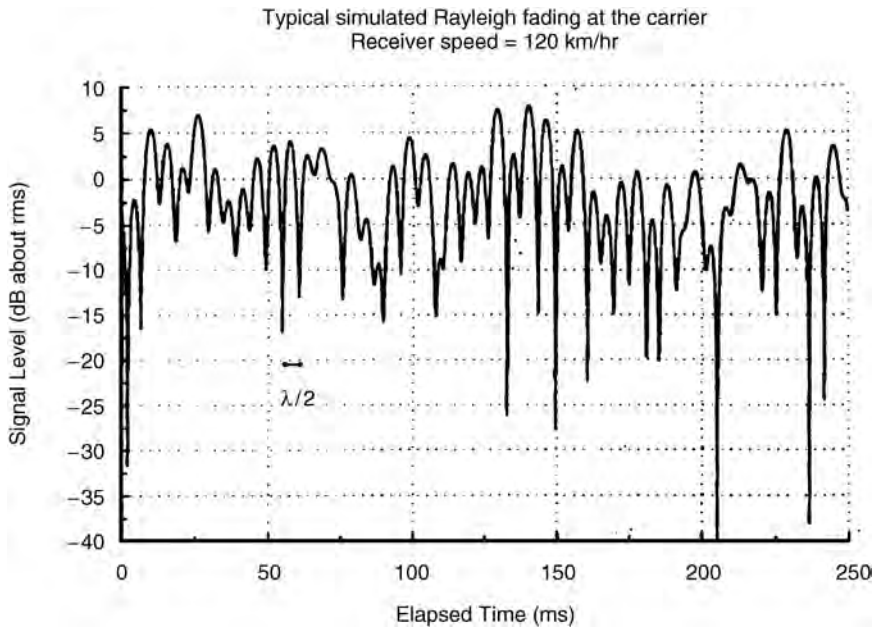


FIGURE 70.10 A typical Rayleigh fading envelope at 900 MHz.

Hence, the Doppler spread f_d or $1/T_0$ is regarded as the typical **fading rate** of the channel. Earlier, T_0 was described as the expected time duration over which the channel's response to a sinusoid is essentially invariant. When T_0 is defined more precisely as the time duration over which the channel's response to a sinusoid has a correlation of at least 0.5, the relationship between T_0 and f_d is approximately [4]

$$T_0 \approx \frac{9}{16\pi f_d} \tag{70.19}$$

A popular "rule of thumb" is to define T_0 as the geometric mean of Eqs. (70.18) and (70.19). This yields

$$T_0 = \sqrt{\frac{9}{16\pi f_d^2}} = \frac{0.423}{f_d} \tag{70.20}$$

For the case of a 900 MHz mobile radio, Fig. 70.10 illustrates the typical effect of Rayleigh fading on a signal's envelope amplitude vs. time [3]. The figure shows that the distance traveled by the mobile in the time interval corresponding to two adjacent nulls (small-scale fades) is on the order of a half-wavelength ($\lambda/2$) [3]. Thus, from Fig. 70.10 and Eq. (70.17), the time (approximately, the coherence time) required to traverse a distance $\lambda/2$ when traveling at a constant velocity, V, is:

$$T_0 \approx \frac{\lambda/2}{V} = \frac{0.5}{f_d} \tag{70.21}$$

Thus, when the interval between fades is taken to be $\lambda/2$, as in Fig. 70.10, the resulting expression for T_0 in Eq. (70.21) is quite close to the rule-of-thumb shown in Eq. (70.20). Using Eq. (70.21), with the parameters shown in Fig. 70.10 (velocity = 120 km/hr, and carrier frequency = 900 MHz), it is straightforward to compute that the coherence time is approximately 5 ms and the Doppler spread (channel fading rate) is approximately 100 Hz. Therefore, if this example represents a voice-grade channel with a typical transmission rate of 10^4 symbols/s, the fading rate is considerably less than the symbol rate. Under such conditions, the channel would manifest slow-fading effects. Note that if the abscissa of Fig. 70.10 were labeled in units of wavelength instead of time, the figure would look the same for any radio frequency and any antenna speed.

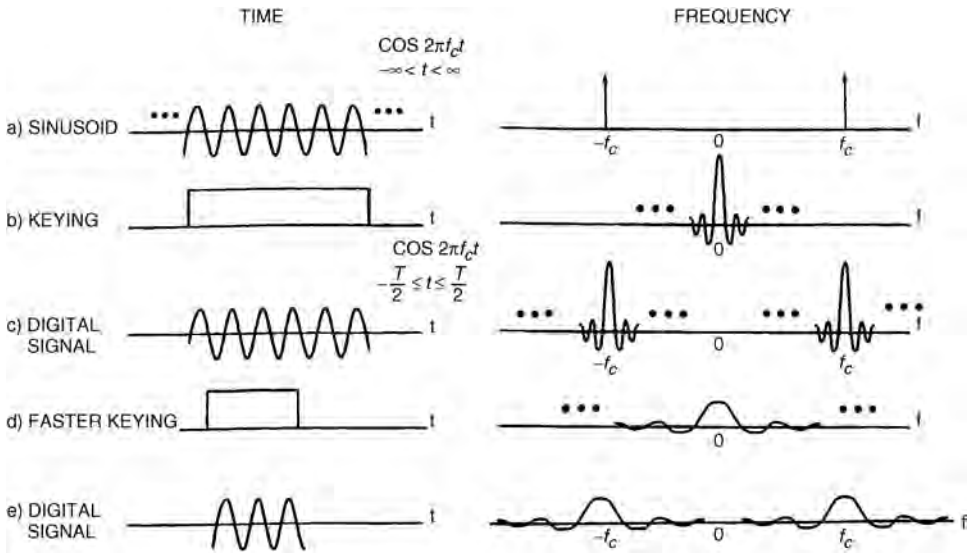


FIGURE 70.11 Analogy between spectral broadening in fading and spectral broadening in keying a digital signal.

70.9 Analogy Between Spectral Broadening in Fading Channels and Spectral Broadening in Digital Signal Keying

Help is often needed in understanding why spectral broadening of the signal is a function of fading rate of the channel. Figure 70.11 uses the keying of a digital signal (such as amplitude-shift-keying or frequency-shift-keying) to illustrate an analogous case. Figure 70.11(a) shows that a single tone, $\cos 2\pi f_c t$ $(-\infty < t < \infty)$ that exists for all time is characterized in the frequency domain in terms of impulses (at $\pm f_c$). This frequency domain representation is ideal (i.e., zero bandwidth), since the tone is pure and neverending. In practical applications, digital signalling involves switching (keying) signals on and off at a required rate. The keying operation can be viewed as multiplying the infinite-duration tone in Fig. 70.11(a) by an ideal rectangular (switching) function in Fig. 70.11(b). The frequency-domain description of the ideal rectangular function is of the form $(\sin f)/f$. In Fig. 70.11(c), the result of the multiplication yields a tone, $\cos 2\pi f_c t$, that is time-duration limited in the interval $-T/2 < t < T/2$. The resulting spectrum is obtained by convolving the spectral impulses in part (a) with the $(\sin f)/f$ function in part (b), yielding the broadened spectrum in part (c). It is further seen that, if the signalling occurs at a faster rate characterized by the rectangle of shorter duration in part (d), the resulting spectrum of the signal in part (e) exhibits greater spectral broadening. The changing state of a fading channel is somewhat analogous to the keying on and off of digital signals. The channel behaves like a switch, turning the signal "on" and "off." The greater the rapidity of the change in the channel state, the greater the spectral broadening of the received signals. The analogy is not exact because the on and off switching of signals may result in phase discontinuities, but the typical multipath-scatterer environment induces phase-continuous effects.

70.10 Degradation Categories due to Time Variance, Viewed in the Doppler-Shift Domain

A channel is referred to as fast fading if the symbol rate, $1/T_s$ (approximately equal to the signalling rate or bandwidth W) is less than the fading rate, $1/T_0$ (approximately equal to f_d); that is, fast fading is characterized by

$$W < f_d \tag{70.22a}$$

or

$$T_s > T_0 \tag{70.22b}$$

Conversely, a channel is referred to as slow fading if the signalling rate is greater than the fading rate. Thus, in order to avoid signal distortion caused by fast fading, the channel must be made to exhibit slow fading by ensuring that the signalling rate must exceed the channel fading rate. That is

$$W > f_d \tag{70.23a}$$

or

$$T_s < T_0 \tag{70.23b}$$

In Eq. (70.14), it was shown that due to signal dispersion, the coherence bandwidth, f_0, sets an upper limit on the signalling rate which can be used without suffering frequency-selective distortion. Similarly, Eq. (70.23) shows that due to Doppler spreading, the channel fading rate, f_d, sets a lower limit on the signalling rate that can be used without suffering fast-fading distortion. For HF communicating systems, when teletype or Morse-coded messages were transmitted at a low data rate, the channels were often fast fading. However, most present-day terrestrial mobile-radio channels can generally be characterized as slow fading.

Equation (70.23) doesn't go far enough in describing what we desire of the channel. A better way to state the requirement for mitigating the effects of fast fading would be that we desire $W \gg f_d$ (or $T_s \ll T_0$). If this condition is not satisfied, the random frequency modulation (FM) due to varying Doppler shifts will limit the system performance significantly. The Doppler effect yields an irreducible error rate that cannot be overcome by simply increasing E_b/N_0 [25]. This irreducible error rate is most pronounced for any modulation that involves switching the carrier phase. A single specular Doppler path, without scatterers, registers an instantaneous frequency shift, classically calculated as $f_d = V/\lambda$. However, a combination of specular and multipath components yields a rather complex time dependence of instantaneous frequency which can cause much larger frequency swings than $\pm V/\lambda$ when detected by an instantaneous frequency detector (a nonlinear device) [26]. Ideally, coherent demodulators that lock onto and track the information signal should suppress the effect of this FM noise and thus cancel the impact of Doppler shift. However, for large values of f_d, carrier recovery becomes a problem because very wideband (relative to the data rate) phase-lock loops (PLLs) need to be designed. For voice-grade applications with bit-error rates of 10^{-3} to 10^{-4}, a large value of Doppler shift is considered to be on the order of $0.01 \times W$. Therefore, to avoid fast-fading distortion and the Doppler-induced irreducible error rate, the signalling rate should exceed the fading rate by a factor of 100 to 200 [27]. The exact factor depends on the signal modulation, receiver design, and required error-rate [3], [26]–[29]. Davarian [29] showed that a frequency-tracking loop can help lower, but not completely remove, the irreducible error rate in a mobile system when using differential minimum-shift keyed (DMSK) modulation.

70.11 Mitigation Methods

Figure 70.12, subtitled "The Good, The Bad, and The Awful," highlights three major performance categories in terms of bit-error probability, P_B, vs. E_b/N_0. The leftmost exponentially-shaped curve represents the performance that can be expected when using any nominal modulation type in AWGN. Observe that with a reasonable amount of E_b/N_0, good performance results. The middle curve, referred to as the **Rayleigh limit**, shows the performance degradation resulting from a loss in SNR that is characteristic of flat fading or slow fading when there is no line-of-sight signal component present. The curve is a function of the reciprocal of E_b/N_0 (an inverse-linear function), so for reasonable values of SNR, performance will generally be "bad." In the case of Rayleigh fading, parameters with overbars are often introduced to indicate that a mean is being taken over the "ups" and "downs" of the fading experience. Therefore, one often sees such bit-error probability plots with mean parameters denoted by

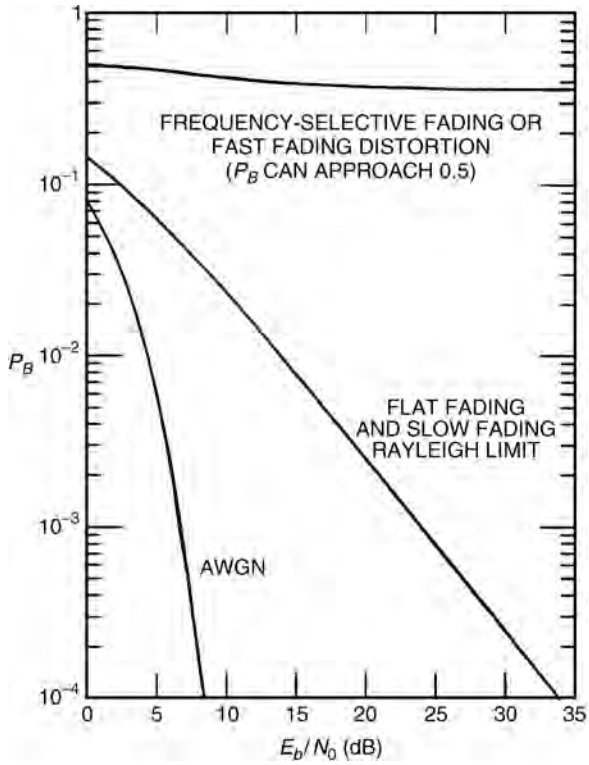


FIGURE 70.12 Error performance: The good, the bad, and the awful.

$\overline{P_B}$ and $\overline{E_b}/N_0$. The curve that reaches an irreducible level, sometimes called an **error floor**, represents "awful" performance, where the bit-error probability can approach the value of 0.5. This shows the severe distorting effects of frequency-selective fading or fast fading.

If the channel introduces signal distortion as a result of fading, the system performance can exhibit an irreducible error rate; when larger than the desired error rate, no amount of E_b/N_0 will help achieve the desired level of performance. In such cases, the general approach for improving performance is to use some form of mitigation to remove or reduce the distortion. The mitigation method depends on whether the distortion is caused by frequency-selective fading or fast fading. Once the distortion has been mitigated, the P_B vs. E_b/N_0 performance should have transitioned from the "awful" bottoming out curve to the merely "bad" Rayleigh limit curve. Next, we can further ameliorate the effects of fading and strive to approach AWGN performance by using some form of diversity to provide the receiver with a collection of uncorrelated samples of the signal, and by using a powerful error-correction code.

In Fig. 70.13, several mitigation techniques for combating the effects of both signal distortion and loss in SNR are listed. Just as Figs. 70.1 and 70.6 serve as a guide for characterizing fading phenomena and their effects, Fig. 70.13 can similarly serve to describe mitigation methods that can be used to ameliorate the effects of fading. The mitigation approach to be used should follow two basic steps: first, provide distortion mitigation; second, provide diversity.

Mitigation to Combat Frequency-Selective Distortion

- Equalization can compensate for the channel-induced ISI that is seen in frequency-selective fading. That is, it can help move the operating point from the error-performance curve that is "awful" in Fig. 70.12 to the one that is "bad." The process of equalizing the ISI involves some method of gathering the dispersed symbol energy back together into its original time interval. In effect,

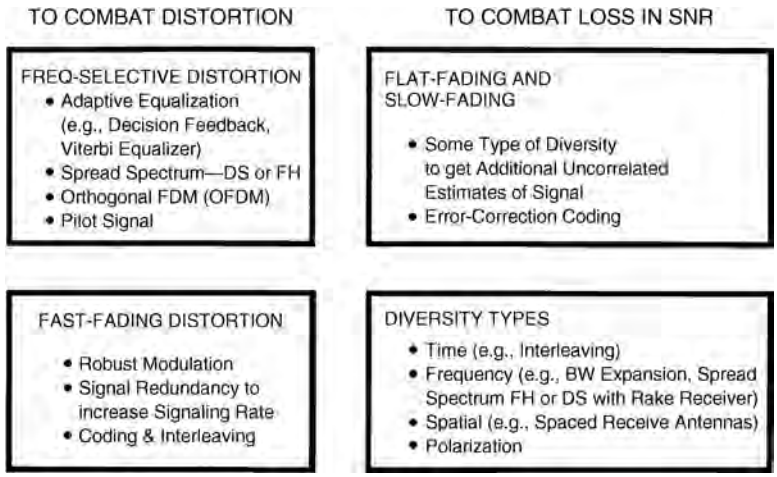


FIGURE 70.13 Basic mitigation types.

equalization involves insertion of a filter to make the combination of channel and filter yield a flat response with linear phase. The phase linearity is achieved by making the equalizer filter the complex conjugate of the time reverse of the dispersed pulse [30]. Because in a mobile system the channel response varies with time, the equalizer filter must also change or adapt to the time-varying channel. Such equalizer filters are, therefore, called adaptive equalizers. An equalizer accomplishes more than distortion mitigation; it also provides diversity. Since distortion mitigation is achieved by gathering the dispersed symbol's energy back into the symbol's original time interval so that it doesn't hamper the detection of other symbols, the equalizer is simultaneously providing each received symbol with energy that would otherwise be lost.

- The decision feedback equalizer (DFE) has a feedforward section that is a linear transversal filter [30] whose length and tap weights are selected to coherently combine virtually all of the current symbol's energy. The DFE also has a feedback section which removes energy that remains from previously detected symbols [14], [30]–[32]. The basic idea behind the DFE is that once an information symbol has been detected, the ISI that it induces on future symbols can be estimated and subtracted before the detection of subsequent symbols.
- The maximum-likelihood sequence estimation (MLSE) equalizer tests all possible data sequences (rather than decoding each received symbol by itself) and chooses the data sequence that is the most probable of the candidates. The MLSE equalizer was first proposed by Forney [33] when he implemented the equalizer using the Viterbi decoding algorithm [34]. The MLSE is optimal in the sense that it minimizes the probability of a sequence error. Because the Viterbi decoding algorithm is the way in which the MLSE equalizer is typically implemented, the equalizer is often referred to as the **Viterbi equalizer**. Later in this chapter, we illustrate the adaptive equalization performed in the Global System for Mobile Communications (GSM) using the Viterbi equalizer.
- Spread-spectrum techniques can be used to mitigate frequency-selective ISI distortion because the hallmark of any spread-spectrum system is its capability to reject interference, and ISI is a type of interference. Consider a direct-sequence spread-spectrum (DS/SS) binary phase shift keying (PSK) communication channel comprising one direct path and one reflected path. Assume that the propagation from transmitter to receiver results in a multipath wave that is delayed by τ_k compared to the direct wave. If the receiver is synchronized to the waveform arriving via the direct path, the received signal, $r(t)$, neglecting noise, can be expressed as

$$r(t) = Ax(t)g(t)\cos(2\pi f_c t) + \alpha Ax(t-\tau_k)g(t-\tau_k)\cos(2\pi f_c t + \theta) \tag{70.24}$$

where $x(t)$ is the data signal, $g(t)$ is the pseudonoise (PN) spreading code, and τ_k is the differential time delay between the two paths. The angle θ is a random phase, assumed to be uniformly distributed in the range $(0, 2\pi)$, and α is the attenuation of the multipath signal relative to the direct path signal. The receiver multiplies the incoming $r(t)$ by the code $g(t)$. If the receiver is synchronized to the direct path signal, multiplication by the code signal yields

$$Ax(t)g^2(t)\cos(2\pi f_c t) + \alpha Ax(t-\tau_k)g(t)g(t-\tau_k)\cos(2\pi f_c t + \theta) \tag{70.25}$$

where $g^2(t) = 1$, and if τ_k is greater than the chip duration, then,

$$\left|\int g^*(t)g(t-\tau_k)dt\right| \ll \int g^*(t)g(t)dt \tag{70.26}$$

over some appropriate interval of integration (correlation), where $*$ indicates complex conjugate, and τ_k is equal to or larger than the PN chip duration. Thus, the spread spectrum system effectively eliminates the multipath interference by virtue of its code-correlation receiver. Even though channel-induced ISI is typically transparent to DS/SS systems, such systems suffer from the loss in energy contained in all the multipath components not seen by the receiver. The need to gather up this lost energy belonging to the received chip was the motivation for developing the Rake receiver [35]–[37]. The Rake receiver dedicates a separate correlator to each multipath component (finger). It is able to coherently add the energy from each finger by selectively delaying them (the earliest component gets the longest delay) so that they can all be coherently combined.

- Earlier, we described a channel that could be classified as flat fading, but occasionally exhibits frequency-selective distortion when the null of the channel's frequency transfer function occurs at the center of the signal band. The use of DS/SS is a good way to mitigate such distortion because the wideband SS signal would span many lobes of the selectively faded frequency response. Hence, a great deal of pulse energy would then be passed by the scatterer medium, in contrast to the nulling effect on a relatively narrowband signal [see Fig. 70.8(c)] [18].
- Frequency-hopping spread-spectrum (FH/SS) can be used to mitigate the distortion due to frequency-selective fading, provided the hopping rate is at least equal to the symbol rate. Compared to DS/SS, mitigation takes place through a different mechanism. FH systems avoid the degradation effects due to multipath by rapidly changing the carrier-frequency band in the transmitter. Interference is avoided by similarly changing the band position in the receiver before the arrival of the multipath signal.
- Orthogonal frequency-division multiplexing (OFDM) can be used in frequency-selective fading channels to avoid the use of an equalizer by lengthening the symbol duration. The signal band is partitioned into multiple subbands, each one exhibiting a lower symbol rate than the original band. The subbands are then transmitted on multiple orthogonal carriers. The goal is to reduce the symbol rate (signalling rate), $W \approx 1/T_s$, on each carrier to be less than the channel's coherence bandwidth f_0. OFDM was originally referred to as Kineplex. The technique has been implemented in the U.S. in mobile radio systems [38], and has been chosen by the European community under the name Coded OFDM (COFDM), for high-definition television (HDTV) broadcasting [39].
- Pilot signal is the name given to a signal intended to facilitate the coherent detection of waveforms. Pilot signals can be implemented in the frequency domain as an in-band tone [40], or in the time domain as a pilot sequence, which can also provide information about the channel state and thus improve performance in fading [41].

Mitigation to Combat Fast-Fading Distortion

- For fast fading distortion, use a robust modulation (noncoherent or differentially coherent) that does not require phase tracking, and reduces the detector integration time [20].
- Increase the symbol rate, $W \approx 1/T_s$, to be greater than the fading rate, $f_d \approx 1/T_0$, by adding signal redundancy.

- Error-correction coding and interleaving can provide mitigation because instead of providing more signal energy, a code reduces the required E_b/N_0. For a given E_b/N_0, with coding present, the error floor will be lowered compared to the uncoded case.
- An interesting filtering technique can provide mitigation in the event of fast-fading distortion and frequency-selective distortion occurring simultaneously. The frequency-selective distortion can be mitigated by the use of an OFDM signal set. Fast fading, however, will typically degrade conventional OFDM because the Doppler spreading corrupts the orthogonality of the OFDM subcarriers. A polyphase filtering technique [42] is used to provide time-domain shaping and duration extension to reduce the spectral sidelobes of the signal set and thus help preserve its orthogonality. The process introduces known ISI and adjacent channel interference (ACI) which are then removed by a post-processing equalizer and canceling filter [43].

Mitigation to Combat Loss in SNR

After implementing some form of mitigation to combat the possible distortion (frequency-selective or fast fading), the next step is to use some form of diversity to move the operating point from the error-performance curve labeled as "bad" in Fig. 70.12 to a curve that approaches AWGN performance. The term "diversity" is used to denote the various methods available for providing the receiver with uncorrelated renditions of the signal. Uncorrelated is the important feature here, since it would not help the receiver to have additional copies of the signal if the copies were all equally poor. Listed below are some of the ways in which diversity can be implemented.

- Time diversity—Transmit the signal on L different time slots with time separation of at least T_0. Interleaving, often used with error-correction coding, is a form of time diversity.
- Frequency diversity—Transmit the signal on L different carriers with frequency separation of at least f_0. Bandwidth expansion is a form of frequency diversity. The signal bandwidth, W, is expanded to be greater than f_0, thus providing the receiver with several independently fading signal replicas. This achieves frequency diversity of the order $L = W/f_0$. Whenever W is made larger than f_0, there is the potential for frequency-selective distortion unless we further provide some mitigation such as equalization. Thus, an expanded bandwidth can improve system performance (via diversity) only if the frequency-selective distortion the diversity may have introduced is mitigated.
- Spread spectrum is a form of bandwidth expansion that excels at rejecting interfering signals. In the case of direct-sequence spread-spectrum (DS/SS), it was shown earlier that multipath components are rejected if they are delayed by more than one chip duration. However, in order to approach AWGN performance, it is necessary to compensate for the loss in energy contained in those rejected components. The Rake receiver (described later) makes it possible to coherently combine the energy from each of the multipath components arriving along different paths. Thus, used with a Rake receiver, DS/SS modulation can be said to achieve path diversity. The Rake receiver is needed in phase-coherent reception, but in differentially coherent bit detection, a simple delay line (one bit long) with complex conjugation will do the trick [44].
- Frequency-hopping spread-spectrum (FH/SS) is sometimes used as a diversity mechanism. The GSM system uses slow FH (217 hops/s) to compensate for those cases where the mobile user is moving very slowly (or not at all) and happens to be in a spectral null.
- Spatial diversity is usually accomplished through the use of multiple receive antennas, separated by a distance of at least 10 wavelengths for a base station (much less for a mobile station). Signal processing must be employed to choose the best antenna output or to coherently combine all the outputs. Systems have also been implemented with multiple spaced transmitters; an example is the Global Positioning System (GPS).
- Polarization diversity [45] is yet another way to achieve additional uncorrelated samples of the signal.

- Any diversity scheme may be viewed as a trivial form of repetition coding in space or time. However, there exist techniques for improving the loss in SNR in a fading channel that are more efficient and more powerful than repetition coding. Error-correction coding represents a unique mitigation technique, because instead of providing more signal energy it reduces the required E_b/N_0 in order to accomplish the desired error performance. Error-correction coding coupled with interleaving [20], [46]–[51] is probably the most prevalent of the mitigation schemes used to provide improved performance in a fading environment.

70.12 Summary of the Key Parameters Characterizing Fading Channels

We summarize the conditions that must be met so that the channel does not introduce frequency-selective distortion and fast-fading distortion. Combining the inequalities of Eqs. (70.14) and (70.23), we obtain

$$f_0 > W > f_d \tag{70.27a}$$

or

$$T_m < T_s < T_0 \tag{70.27b}$$

In other words, we want the channel coherence bandwidth to exceed our signalling rate, which in turn should exceed the fading rate of the channel. Recall that without distortion mitigation, f_0 sets an upper limit on signalling rate, and f_d sets a lower limit on it.

Fast-Fading Distortion: Example #1

If the inequalities of Eq. (70.27) are not met and distortion mitigation is not provided, distortion will result. Consider the fast-fading case where the signalling rate is less than the channel fading rate, that is,

$$f_0 > W < f_d \tag{70.28}$$

Mitigation consists of using one or more of the following methods. (See Fig. 70.13).

- Choose a modulation/demodulation technique that is most robust under fast-fading conditions. That means, for example, avoiding carrier recovery with PLLs since the fast fading could keep a PLL from achieving lock conditions.
- Incorporate sufficient redundancy so that the transmission symbol rate exceeds the channel fading rate. As long as the transmission symbol rate does not exceed the coherence bandwidth, the channel can be classified as flat fading. However, even flat-fading channels will experience frequency-selective distortion whenever a channel null appears at the band center.

Since this happens only occasionally, mitigation might be accomplished by adequate error-correction coding and interleaving.

- The above two mitigation approaches should result in the demodulator operating at the Rayleigh limit [20] (see Fig. 70.12). However, there may be an irreducible floor in the error-performance vs. E_b/N_0 curve due to the FM noise that results from the random Doppler spreading. The use of an in-band pilot tone and a frequency-control loop can lower this irreducible performance level.
- To avoid this error floor caused by random Doppler spreading, increase the signalling rate above the fading rate still further (100–200 × fading rate) [27]. This is one architectural motive behind time-division multiple access (TDMA) mobile systems.
- Incorporate error-correction coding and interleaving to lower the floor and approach AWGN performance.

Frequency-Selective Fading Distortion: Example #2

Consider the frequency-selective case where the coherence bandwidth is less than the symbol rate; that is,

$$f_0 < W > f_d \tag{70.29}$$

Mitigation consists of using one or more of the following methods. (See Fig. 70.13).

- Since the transmission symbol rate exceeds the channel-fading rate, there is no fast-fading distortion. Mitigation of frequency-selective effects is necessary. One or more of the following techniques may be considered:
- Adaptive equalization, spread spectrum (DS or FH), OFDM, pilot signal. The European GSM system uses a midamble training sequence in each transmission time slot so that the receiver can learn the impulse response of the channel. It then uses a Viterbi equalizer (explained later) for mitigating the frequency-selective distortion.
- Once the distortion effects have been reduced, introduce some form of diversity and error-correction coding and interleaving in order to approach AWGN performance. For direct-sequence spread-spectrum (DS/SS) signalling, the use of a Rake receiver (explained later) may be used for providing diversity by coherently combining multipath components that would otherwise be lost.

Fast-Fading and Frequency-Selective Fading Distortion: Example #3

Consider the case where the coherence bandwidth is less than the signalling rate, which in turn is less than the fading rate. The channel exhibits both fast-fading and frequency-selective fading which is expressed as

$$f_0 < W < f_d \tag{70.30a}$$

or

$$f_0 < f_d \tag{70.30b}$$

Recalling from Eq. (70.27) that f_0 sets an upper limit on signalling rate and f_d sets a lower limit on it, this is a difficult design problem because, unless distortion mitigation is provided, the maximum allowable signalling rate is (in the strict terms of the above discussion) less than the minimum allowable signalling rate. Mitigation in this case is similar to the initial approach outlined in example #1.

- Choose a modulation/demodulation technique that is most robust under fast-fading conditions.
- Use transmission redundancy in order to increase the transmitted symbol rate.
- Provide some form of frequency-selective mitigation in a manner similar to that outlined in example #2.
- Once the distortion effects have been reduced, introduce some form of diversity and error-correction coding and interleaving in order to approach AWGN performance.

70.13 The Viterbi Equalizer as Applied to GSM

Figure 70.14 shows the GSM time-division multiple access (TDMA) frame, having a duration of 4.615 ms and comprising 8 slots, one assigned to each active mobile user. A normal transmission burst occupying one slot of time contains 57 message bits on each side of a 26-bit midamble called a **training** or **sounding sequence**. The slot-time duration is 0.577 ms (or the slot rate is 1733 slots/s). The purpose of the midamble is to assist the receiver in estimating the impulse response of the channel in an adaptive way (during the time duration of each 0.577 ms slot). In order for the technique to be effective, the fading

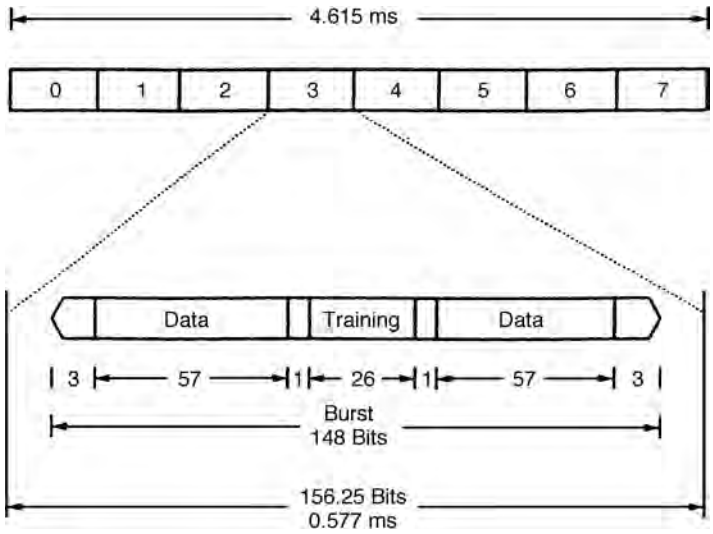


FIGURE 70.14 The GSM TDMA frame and time-slot containing a normal burst.

behavior of the channel should not change appreciably during the time interval of one slot. In other words, there should not be any fast-fading degradation during a slot time when the receiver is using knowledge from the midamble to compensate for the channel's fading behavior. Consider the example of a GSM receiver used aboard a high-speed train, traveling at a constant velocity of 200 km/hr (55.56 m/s). Assume the carrier frequency to be 900 MHz (the wavelength is $\lambda = 0.33$ m). From Eq. (70.21), we can calculate that a half-wavelength is traversed in approximately the time (coherence time)

$$T_0 \approx \frac{\lambda/2}{V} \approx 3 \text{ ms} \tag{70.31}$$

Therefore, the channel coherence time is over 5 times greater than the slot time of 0.577 ms. The time needed for a significant change in fading behavior is relatively long compared to the time duration of one slot. Note, that the choices made in the design of the GSM TDMA slot time and midamble were undoubtedly influenced by the need to preclude fast fading with respect to a slot-time duration, as in this example.

The GSM symbol rate (or bit rate, since the modulation is binary) is 271 kilosymbols/s and the bandwidth is $W = 200$ kHz. If we consider that the typical rms delay spread in an urban environment is on the order of $\sigma = 2\mu$s, then using Eq. (70.13) the resulting coherence bandwidth is $f_0 \approx 100$ kHz. It should therefore be apparent that since $f_0 < W$, the GSM receiver must utilize some form of mitigation to combat frequency-selective distortion. To accomplish this goal, the Viterbi equalizer is typically implemented.

Figure 70.15 illustrates the basic functional blocks used in a GSM receiver for estimating the channel impulse response, which is then used to provide the detector with channel-corrected reference waveforms [52]. In the final step, the Viterbi algorithm is used to compute the MLSE of the message. As stated in Eq. (70.2), a received signal can be described in terms of the transmitted signal convolved with the impulse response of the channel, $h_c(t)$. We show this below, using the notation of a received training sequence, $r_{tr}(t)$, and the transmitted training sequence, $s_{tr}(t)$, as follows:

$$r_{tr}(t) = s_{tr}(t) * h_c(t) \tag{70.32}$$

where $*$ denotes convolution. At the receiver, $r_{tr}(t)$ is extracted from the normal burst and sent to a filter having impulse response, $h_{mf}(t)$, that is matched to $s_{tr}(t)$. This matched filter yields at its output an

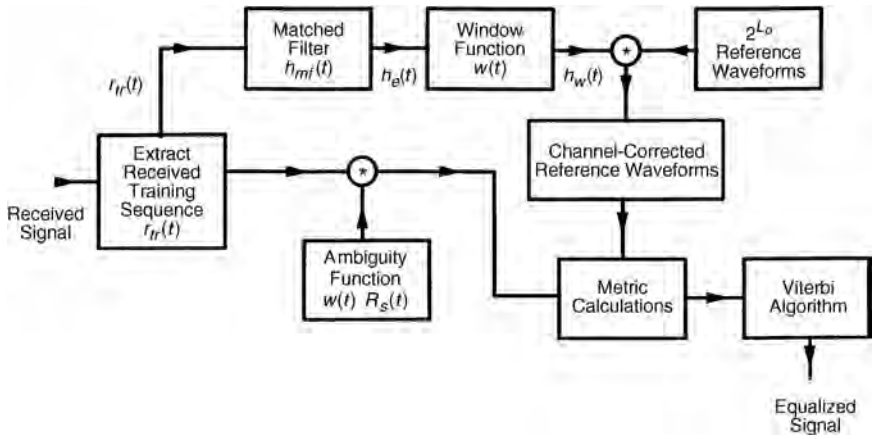


FIGURE 70.15 The Viterbi equalizer as applied to GSM.

estimate of $h_c(t)$, denoted $h_e(t)$, developed from Eq. (70.32) as follows.

$$\begin{aligned} h_e(t) &= r_{tr}(t) * h_{mf}(t) \\ &= s_{tr}(t) * h_c(t) * h_{mf}(t) \\ &= R_s(t) * h_c(t) \end{aligned} \tag{70.33}$$

where $R_s(t)$ is the autocorrelation function of $s_{tr}(t)$. If $R_s(t)$ is a highly peaked (impulse-like) function, then $h_e(t) \approx h_c(t)$.

Next, using a windowing function, $w(t)$, we truncate $h_e(t)$ to form a computationally affordable function, $h_w(t)$. The window length must be large enough to compensate for the effect of typical channel-induced ISI. The required observation interval L_0 for the window can be expressed as the sum of two contributions. The interval of length L_{CISI} is due to the controlled ISI caused by Gaussian filtering of the baseband pulses, which are then MSK modulated. The interval of length L_C is due to the channel-induced ISI caused by multipath propagation; therefore, L_0 can be written as

$$L_0 = L_{CISI} + L_C \tag{70.34}$$

The GSM system is required to provide mitigation for distortion due to signal dispersions of approximately 15–20 μs. The bit duration is 3.69 μs. Thus, the Viterbi equalizer used in GSM has a memory of 4–6 bit intervals. For each L_0-bit interval in the message, the function of the Viterbi equalizer is to find the most likely L_0-bit sequence out of the 2^{L_0} possible sequences that might have been transmitted. Determining the most likely L_0-bit sequence requires that 2^{L_0} meaningful reference waveforms be created by modifying (or disturbing) the 2^{L_0} ideal waveforms in the same way that the channel has disturbed the transmitted message. Therefore, the 2^{L_0} reference waveforms are convolved with the windowed estimate of the channel impulse response, $h_w(t)$ in order to derive the disturbed or channel-corrected reference waveforms. Next, the channel-corrected reference waveforms are compared against the received data waveforms to yield metric calculations. However, before the comparison takes place, the received data waveforms are convolved with the known windowed autocorrelation function $w(t)R_s(t)$, transforming them in a manner comparable to that applied to the reference waveforms. This filtered message signal is compared to all possible 2^{L_0} channel-corrected reference signals, and metrics are computed as required by the Viterbi decoding algorithm (VDA). The VDA yields the maximum likelihood estimate of the transmitted sequence [34].

70.14 The Rake Receiver Applied to Direct-Sequence Spread-Spectrum (DS/SS) Systems

Interim Specification 95 (IS-95) describes a DS/SS cellular system that uses a Rake receiver [35]–[37] to provide path diversity. In Fig. 70.16, five instances of chip transmissions corresponding to the code sequence 1 0 1 1 1 are shown, with the transmission or observation times labeled t_{-4} for the earliest transmission and t_0 for the latest. Each abscissa shows three "fingers" of a signal that arrive at the receiver with delay times τ_1, τ_2, and τ_3. Assume that the intervals between the t_i transmission times and the intervals between the τ_i delay times are each one chip long. From this, one can conclude that the finger arriving at the receiver at time t_{-4}, with delay τ_3, is time coincident with two other fingers, namely the fingers arriving at times t_{-3} and t_{-2} with delays τ_2 and τ_1, respectively. Since, in this example, the delayed components are separated by exactly one chip time, they are *just* resolvable. At the receiver, there must be a sounding device that is dedicated to estimating the τ_i delay times. Note that for a terrestrial mobile radio system, the fading rate is relatively slow (milliseconds) or the channel coherence time large compared to the chip time ($T_0 > T_{ch}$). Hence, the changes in τ_i occur slowly enough so that the receiver can readily adapt to them.

Once the τ_i delays are estimated, a separate correlator is dedicated to processing each resolvable multipath finger. In this example, there would be three such dedicated correlators, each one processing a delayed version of the same chip sequence 1 0 1 1 1. In Fig. 70.16, each correlator receives chips with power profiles represented by the sequence of fingers shown along a diagonal line. Each correlator attempts to match these arriving chips with the same PN code, similarly delayed in time. At the end of a symbol interval (typically there may be hundreds or thousands of chips per symbol), the outputs of the correlators are coherently combined, and a symbol detection is made. At the chip level, the Rake receiver resembles an equalizer, but its real function is to provide diversity.

The interference-suppression nature of DS/SS systems stems from the fact that a code sequence arriving at the receiver merely one chip time late, will be approximately orthogonal to the particular PN code with which the sequence is correlated. Therefore, any code chips that are delayed by one or more chip times will be suppressed by the correlator. The delayed chips only contribute to raising the noise floor (correlation sidelobes). The mitigation provided by the Rake receiver can be termed path diversity, since it allows the energy of a chip that arrives via multiple paths to be combined coherently. Without the Rake receiver, this energy would be transparent and therefore lost to the DS/SS system. In Fig. 70.16, looking

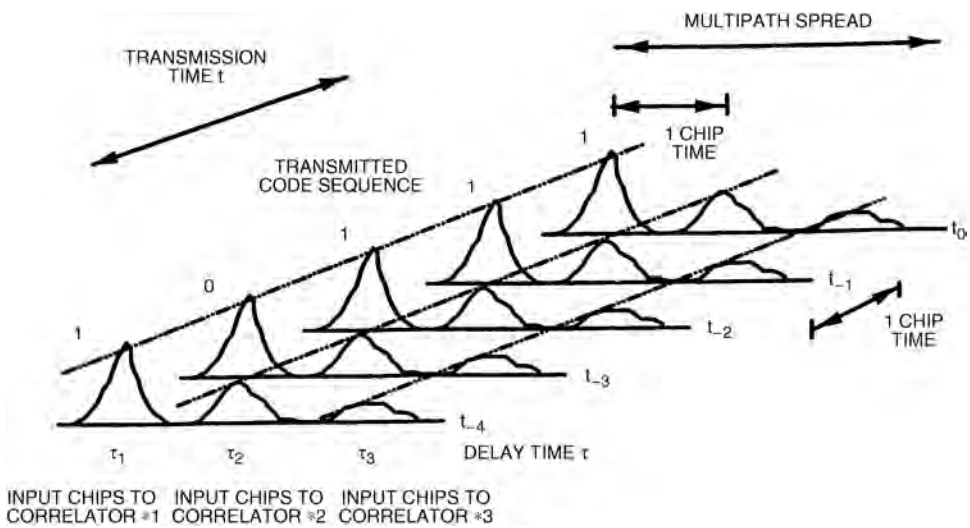


FIGURE 70.16 Example of received chips seen by a 3-finger Rake receiver.

vertically above point τ_3, it is clear that there is interchip interference due to different fingers arriving simultaneously. The spread-spectrum processing gain allows the system to endure such interference at the chip level. No other equalization is deemed necessary in IS-95.

70.15 Conclusion

In this chapter, the major elements that contribute to fading in a communication channel have been characterized. Figure 70.1 was presented as a guide for the characterization of fading phenomena. Two types of fading, large-scale and small-scale, were described. Two manifestations of small-scale fading (signal dispersion and fading rapidity) were examined, and the examination involved two views, time and frequency. Two degradation categories were defined for dispersion: frequency-selective fading and flat-fading. Two degradation categories were defined for fading rapidity: fast and slow. The small-scale fading degradation categories were summarized in Fig. 70.6. A mathematical model using correlation and power density functions was presented in Fig. 70.7. This model yields a nice symmetry, a kind of "poetry" to help us view the Fourier transform and duality relationships that describe the fading phenomena. Further, mitigation techniques for ameliorating the effects of each degradation category were treated, and these techniques were summarized in Fig. 70.13. Finally, mitigation methods that have been implemented in two system types, GSM and CDMA systems meeting IS-95, were described.

References

1. Sklar, B., *Digital Communications: Fundamentals and Applications,* 2nd ed., Prentice-Hall, Englewood Cliffs, NJ, Ch. 4, 2001.
2. Van Trees, H.L., *Detection, Estimation, and Modulation Theory, Part I,* John Wiley & Sons, New York, Ch. 4, 1968.
3. Rappaport, T.S., *Wireless Communications,* Prentice-Hall, Upper Saddle River, New Jersey, Chs. 3 and 4, 1996.
4. Greenwood, D. and Hanzo, L., Characterisation of Mobile Radio Channels, *Mobile Radio Communications,* Steele, R., Ed., Pentech Press, London, Ch. 2, 1994.
5. Lee, W.C.Y., Elements of cellular mobile radio systems, *IEEE Trans. Vehicular Technol.,* V-35(2), 48–56, May 1986.
6. Okumura, Y. et al., Field strength and its variability in VHF and UHF land mobile radio service, *Rev. Elec. Comm. Lab.,* 16(9-10), 825–873, 1968.
7. Hata, M., Empirical formulæ for propagation loss in land mobile radio services, *IEEE Trans. Vehicular Technol.,* VT-29(3), 317–325, 1980.
8. Seidel, S.Y. et al., Path loss, scattering and multipath delay statistics in four European cities for digital cellular and microcellular radiotelephone, *IEEE Trans. Vehicular Technol.,* 40(4), 721–730, Nov. 1991.
9. Cox, D.C., Murray, R., and Norris, A., 800 MHz Attenuation measured in and around suburban houses, *AT&T Bell Laboratory Technical Journal,* 673(6), 921–954, Jul.-Aug. 1984.
10. Schilling, D.L. et al., Broadband CDMA for personal communications systems, *IEEE Commun. Mag.,* 29(11), 86–93, Nov. 1991.
11. Andersen, J.B., Rappaport, T.S., and Yoshida, S., Propagation measurements and models for wireless communications channels, *IEEE Commun. Mag.,* 33(1), 42–49, Jan. 1995.
12. Amoroso, F., Investigation of signal variance, bit error rates and pulse dispersion for DSPN signalling in a mobile dense scatterer ray tracing model, *Intl. J. Satellite Commun.,* 12, 579–588, 1994.
13. Bello, P.A., Characterization of randomly time-variant linear channels, *IEEE Trans. Commun. Syst.,* 360–393, Dec. 1963.
14. Proakis, J.G., *Digital Communications,* McGraw-Hill, New York, Ch. 7, 1983.
15. Green, P.E., Jr., Radar astronomy measurement techniques, *MIT Lincoln Laboratory,* Lexington, MA, Tech. Report No. 282, Dec. 1962.

16. Pahlavan, K. and Levesque, A.H., *Wireless Information Networks,* John Wiley & Sons, New York, Chs. 3 and 4, 1995.
17. Lee, W.Y.C., *Mobile Cellular Communications,* McGraw-Hill, New York, 1989.
18. Amoroso, F., Use of DS/SS signalling to mitigate Rayleigh fading in a dense scatterer environment, *IEEE Personal Commun.,* 3(2), 52–61, Apr. 1996.
19. Clarke, R.H., A statistical theory of mobile radio reception, *Bell Syst. Tech. J.,* 47(6), 957–1000, Jul.-Aug. 1968.
20. Bogusch, R.L., *Digital Communications in Fading Channels: Modulation and Coding,* Mission Research Corp., Santa Barbara, California, Report No. MRC-R-1043, Mar. 11, 1987.
21. Amoroso, F., The bandwidth of digital data signals, *IEEE Commun. Mag.,* 18(6), 13–24, Nov. 1980.
22. Bogusch, R.L. et al., Frequency selective propagation effects on spread-spectrum receiver tracking, *Proc. IEEE,* 69(7), 787–796, Jul. 1981.
23. Jakes, W.C., Ed., *Microwave Mobile Communications,* John Wiley & Sons, New York, 1974.
24. *Joint Technical Committee of Committee T1 R1P1.4 and TIA TR46.3.3/TR45.4.4 on Wireless Access,* Draft Final Report on RF Channel Characterization, Paper No. JTC(AIR)/94.01.17-238R4, Jan. 17, 1994.
25. Bello, P.A. and Nelin, B.D., The influence of fading spectrum on the binary error probabilities of incoherent and differentially coherent matched filter receivers, *IRE Trans. Commun. Syst.,* CS-10, 160–168, Jun. 1962.
26. Amoroso, F., Instantaneous frequency effects in a Doppler scattering environment, *IEEE International Conference on Communications,* 1458–1466, Jun. 7–10, 1987.
27. Bateman, A.J. and McGeehan, J.P., Data transmission over UHF fading mobile radio channels, *IEEE Proc.,* 131, Pt. F(4), 364–374, Jul. 1984.
28. Feher, K., *Wireless Digital Communications,* Prentice-Hall, Upper Saddle River, NJ, 1995.
29. Davarian, F., Simon, M., and Sumida, J., DMSK: A Practical 2400-bps Receiver for the Mobile Satellite Service, Jet Propulsion Laboratory Publication 85-51 (MSAT-X Report No. 111), Jun. 15, 1985.
30. Rappaport, T.S., *Wireless Communications,* Prentice-Hall, Upper Saddle River, NJ, Ch. 6, 1996.
31. Bogusch, R.L., Guigliano, F.W., and Knepp, D.L., Frequency-selective scintillation effects and decision feedback equalization in high data-rate satellite links, *Proc. IEEE,* 71(6), 754–767, Jun. 1983.
32. Qureshi, S.U.H., Adaptive equalization, *Proc. IEEE,* 73(9), 1340–1387, Sept. 1985.
33. Forney, G.D., The Viterbi algorithm, *Proc. IEEE,* 61(3), 268–278, Mar. 1978.
34. Sklar, B., *Digital Communications: Fundamentals and Applications,* Prentice-Hall, Englewood Cliffs, NJ, Ch. 6, 1988.
35. Price, R. and Green, P.E., Jr., A communication technique for multipath channels, *Proc. IRE,* 555–570, Mar. 1958.
36. Turin, G.L., Introduction to spread-spectrum antimultipath techniques and their application to urban digital radio, *Proc. IEEE,* 68(3), 328–353, Mar. 1980.
37. Simon, M.K., Omura, J.K., Scholtz, R.A., and Levitt, B.K., *Spread Spectrum Communications Handbook,* McGraw-Hill, New York, 1994.
38. Birchler, M.A. and Jasper, S.C., A 64 kbps Digital Land Mobile Radio System Employing M-16QAM, *Proceedings of the 1992 IEEE Intl. Conference on Selected Topics in Wireless Communications,* Vancouver, British Columbia, 158–162, Jun. 25–26, 1992.
39. Sari, H., Karam, G., and Jeanclaude, I., Transmission techniques for digital terrestrial TV broadcasting, *IEEE Commun. Mag.,* 33(2), 100–109, Feb. 1995.
40. Cavers, J.K., The performance of phase locked transparent tone-in-band with symmetric phase detection, *IEEE Trans. Commun.,* 39(9), 1389–1399, Sept. 1991.
41. Moher, M.L. and Lodge, J.H., TCMP—A modulation and coding strategy for Rician fading channel, *IEEE J. Selected Areas Commun.,* 7(9), 1347–1355, Dec. 1989.
42. Harris, F., On the Relationship Between Multirate Polyphase FIR Filters and Windowed, Overlapped FFT Processing, *Proceedings of the Twenty Third Annual Asilomar Conference on Signals, Systems, and Computers,* Pacific Grove, California, 485–488, Oct. 30 to Nov. 1, 1989.

43. Lowdermilk, R.W. and Harris, F., Design and Performance of Fading Insensitive Orthogonal Frequency Division Multiplexing (OFDM) using Polyphase Filtering Techniques, *Proceedings of the Thirtieth Annual Asilomar Conference on Signals, Systems, and Computers,* Pacific Grove, California, Nov. 3–6, 1996.
44. Kavehrad, M. and Bodeep, G.E., Design and experimental results for a direct-sequence spread-spectrum radio using differential phase-shift keying modulation for indoor wireless communications, *IEEE JSAC,* SAC-5(5), 815–823, Jun. 1987.
45. Hess, G.C., *Land-Mobile Radio System Engineering,* Artech House, Boston, 1993.
46. Hagenauer, J. and Lutz, E., Forward error correction coding for fading compensation in mobile satellite channels, *IEEE JSAC,* SAC-5(2), 215–225, Feb. 1987.
47. McLane, P.I. et al., PSK and DPSK trellis codes for fast fading, shadowed mobile satellite communication channels, *IEEE Trans. Commun.,* 36(11), 1242–1246, Nov. 1988.
48. Schlegel, C. and Costello, D.J., Jr., Bandwidth efficient coding for fading channels: code construction and performance analysis, *IEEE JSAC,* 7(9), 1356–1368, Dec. 1989.
49. Edbauer, F., Performance of interleaved trellis-coded differential 8–PSK modulation over fading channels, *IEEE J. Selected Areas Commun.,* 7(9), 1340–1346, Dec. 1989.
50. Soliman, S. and Mokrani, K., Performance of coded systems over fading dispersive channels, *IEEE Trans. Commun.,* 40(1), 51–59, Jan. 1992.
51. Divsalar, D. and Pollara, F., Turbo Codes for PCS Applications, *Proc. ICC'95,* Seattle, Washington, 54–59, Jun. 18–22, 1995.
52. Hanzo, L. and Stefanov, J., The Pan-European Digital Cellular Mobile Radio System—known as GSM, *Mobile Radio Communications.* Steele, R., Ed., Pentech Press, London, Ch. 8, 1992.

71

Space-Time Processing

Arogyaswami J. Paulraj
Stanford University

71.1 Introduction

Mobile radio signal processing includes modulation and demodulation, channel coding and decoding, equalization and diversity. Current cellular modems mainly use temporal signal processing. Use of spatio-temporal signal processing can improve average signal power, mitigate fading, and reduce cochannel and intersymbol interference. This can significantly improve the capacity, coverage, and quality of wireless networks.

A space-time processing radio operates simultaneously on multiple antennas by processing signal samples both in space and time. In receive, space-time (ST) processing can increase array gain, spatial and temporal diversity and reduce cochannel interference and intersymbol interference. In transmit, the spatial dimension can enhance array gain, improve diversity, reduce generation of cochannel and intersymbol interference.

71.2 The Space-Time Wireless Channel

Multipath Propagation

Multipath scattering gives rise to a number of propagation effects described below.

Scatterers Local to Mobile

Scattering local to the mobile is caused by buildings/other scatterers in the vicinity of the mobile (a few tens of meters). Mobile motion and local scattering give rise to Doppler spread which causes time-selective fading. For a mobile traveling at 65 mph, the Doppler spread is about 200 Hz in the 1900 MHz band. While local scatterers contribute to Doppler spread, the delay spread they contribute is usually insignificant because of the small scattering radius. Likewise, the angle spread induced at the base station is also small.

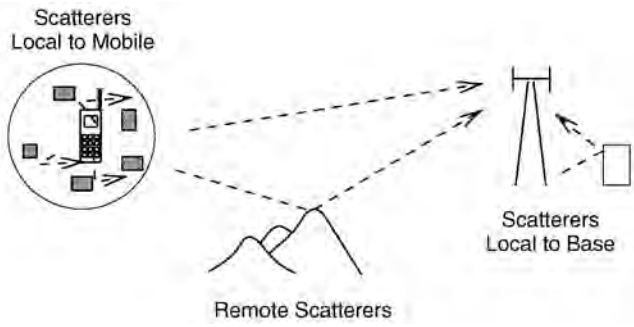


FIGURE 71.1 Multipath propagation in macrocells.

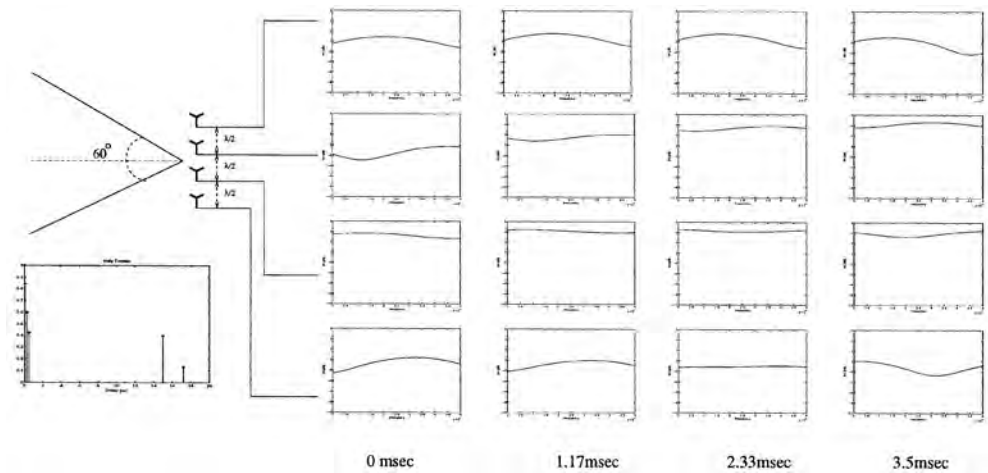


FIGURE 71.2 ST channel.

Remote Scatterers

The emerging wavefront from the local scatterers may then travel directly to the base or may be scattered toward the base by remote dominant scatterers, giving rise to specular multipaths. These remote scatterers can be either terrain features or high-rise building complexes. Remote scattering can cause significant delay and angle spreads.

Scatterers Local to Base

Once these multiple wavefronts reach the base station, they may be scattered further by local structures such as buildings or other structures in the vicinity of the base. Such scattering will be more pronounced for low elevation and below roof-top antennas. Scattering local to the base can cause severe angle spread which in turn causes space-selective fading. See Figure 71.1 for a depiction of different types of scattering.

The forward link channel is affected in similar ways by these scatterers, but in a reverse order.

Space-Time Channel Model

The effect of delay, Doppler and angle spreads makes the channel selective in frequency, time, and space. Figure 71.2 shows plots of the frequency response at each branch of a four-antenna receiver operating with a 200 Khz bandwidth. We can see that the channel is highly frequency-selective since the delay spread reaches 10 to 15 μs. Also, an angle spread of 30∞ causes variations in the channel from antenna to antenna. The channel variation in time depends upon the Doppler spread. As expected, the plots show

negligible channel variation between adjacent time slots, despite the high velocity of the mobile (100 kph). Use of longer time slots such as in IS-136 will result in significant channel variations over the slot period. Therefore, space-time processing should address the effect of the three spreads on the signal.

71.3 Signal Models

We develop signal models for nonspread modulation used in time division multiple access (TDMA) systems.

Signal Model at Base Station (Reverse Link)

We assume that antenna arrays are used at the base station only and that the mobile has a single omni antenna. The mobile transmits a channel coded and modulated signal which does not incorporate any spatial (or indeed any special temporal) processing. See Figure 71.3.

The baseband signal $x_i(t)$ received by the base station at the ith element of an m element antenna array is given by

$$x_i(t) = \sum_{l=1}^{L} a_i(\theta_l)\alpha_l^R(t)u(t-t_l) + n_i(t) \tag{71.1}$$

where L is the number of multipaths, $a_i(\theta_l)$ is the response of the ith element for the lth path from direction q_l, $a_l^R(t)$ is the complex path fading, τ_l is the path delay, $n_i(t)$ is the additive noise and $u(\cdot)$ is the transmitted signal that depends on the modulation waveform and the information data stream.

For a linear modulation, the baseband transmitted signal is given by

$$u(t) = \sum_k g(t-kT)s(k) \tag{71.2}$$

where $g(\cdot)$ is the pulse shaping waveform and $s(k)$ represents the information bits.

In the above model we have assumed that the inverse signal bandwidth is large compared to the travel time across the array. Therefore, the complex envelopes of the signals received by different antennas from a given path are identical except for phase and amplitude differences that depend on the path angle-of-arrival, array geometry and the element pattern. This angle-of-arrival dependent phase and amplitude response at the ith element is $a_i(\theta_l)$.

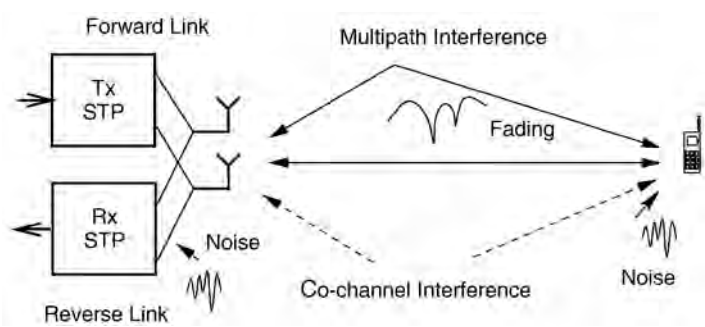


FIGURE 71.3 ST Processing Model.

We collect all the element responses to a path arriving from angle θ_l into an m-dimensional vector, called the *array response vector* defined as

$$\begin{aligned} \mathbf{a}(\theta_l) &= [a_1(\theta_l)a_2(\theta_l)\ldots a_m(\theta_l)]^t \\ \mathbf{x}(t) &= \sum_{l=1}^{L}\mathbf{a}(\theta_l)\alpha_l^R(t)u(t-t_l) + \mathbf{n}(t) \end{aligned} \tag{71.3}$$

where $\mathbf{x}(t)$ and $\mathbf{n}(t)$ are m-dimensional complex vectors. The fading $|\alpha^R(t)|$ is Rayleigh or Rician distributed depending on the propagation model.

Signal Model at Mobile (Forward Link)

In this model, the base station transmits different signals from each antenna with a defined relationship between them. In the case of a two element array, some examples of transmitted signals $u_i(t)$, $i = 1, 2$ can be: (a) delay diversity: $u_2(t) = u_1(t - T)$ where T is the symbol period; (b)Doppler diversity: $u_2(t) = u_1(t)e^{j\omega t}$ where ω is differential carrier offset; (c) beamforming: $u_2(t) = w_2u_1(t)$ where w_2 is complex scalar; and (d) space-time coding: $u_1(t) = \sum_k g(t - kT)s^1(k)$, $u_2(t) = \sum_k g(t - kT)s^2(k)$ where $s^1(k)$ and $s^2(k)$ are related to the symbol sequence $s(k)$ through coding.

The received signal at the mobile is then given by

$$x(t) = \sum_{i=1}^{m}\sum_{l=1}^{L}a_i(\theta_l)\alpha_l^F(t)u_i(t-\tau_l) + n(t) \tag{71.4}$$

where the path delay τ_l and angle parameters θ_l are the same as those of the reverse link. $\alpha_l^F(t)$ is the complex fading on the forward link. In (fast) **TDD** systems $\alpha_l^F(t)$ will be identical to the reverse link complex fading $\alpha_l^R(t)$. In a **FDD** system $\alpha_l^F(t)$ and $\alpha_l^R(t)$ will usually have the same statistics but will in general be uncorrelated with each other. We assume $a_i(\theta_l)$ is the same for both links. This is only approximately true in FDD systems.

If simple beamforming alone is used in transmit, the signals radiated from the antennas are related by a complex scalar and result in a directional transmit beam which may selectively couple into the multipath environment and differentially scale the power in each path.

The signal received by the mobile in this case can be written as

$$x(t) = \sum_{l=1}^{L}\mathbf{w}^H\mathbf{a}(\theta_l)\alpha_l^F(t)u(t-\tau_l) + n(t) \tag{71.5}$$

where $\mathbf{w}$ is the beamforming vector.

Discrete Time Signal Model

The channel model described above uses physical path parameters such as path gain, delay, and angle of arrival. In practice these are not known and the discrete time received signal uses a more convenient discretized "symbol response" channel model.

We derive a discrete-time signal model at the base station antenna array. Let the continuous-time output from the receive antenna array $\mathbf{x}(t)$ be sampled at the symbol rate at instants $t = t_o + kT$. Then the vector array output may be written as

$$\mathbf{x}(k) = \mathbf{H}^R\mathbf{s}(k) + n(k) \tag{71.6}$$

where $\mathbf{H}^R$ is the reverse link symbol response channel (a $m \times N$ matrix) that captures the effects of the array response, symbol waveform and path fading. m is the number of antennas, N is the channel length in symbol periods and $\mathbf{n}(k)$ the sampled vector of additive noise. Note that $\mathbf{n}(k)$ may be colored in space and time, as discussed later. $\mathbf{H}^R$ is assumed to be time invariant. $\mathbf{s}(k)$ is a vector of N consecutive elements of the data sequence and is defined as

$$\mathbf{s}(k) = \begin{bmatrix} s(k) \\ \vdots \\ s(k-N+1) \end{bmatrix} \tag{71.7}$$

Note that we have assumed a sampling rate of one sample per symbol. Higher sampling rates may be used. Also, $\mathbf{H}^R$ is given by

$$\mathbf{H}^R = \sum_{l=1}^{L} \mathbf{a}(\theta_l)\alpha_l^R \mathbf{g}^T(\tau_l) \tag{71.8}$$

where $\mathbf{g}(\tau_l)$ is a vector defined by T spaced sampling of the pulse shaping function $g(\cdot)$ with an offset of τ_l. Likewise the forward discrete signal model at the mobile is given by

$$x(k) = \sum_{i=1}^{m} \mathbf{h}_i^F \mathbf{s}(k) + n(k) \tag{71.9}$$

where $\mathbf{h}_i^F$ is a $1 \times N$ composite channel from the symbol sequence via the ith antenna to the mobile receiver which includes the effect transmit ST processing at the base station.

In the case of two antenna delay diversity, $\mathbf{h}_i^F$ is given by

$$\mathbf{h}_1^F = \sum_{l=1}^{L} a_1(\theta_l)\alpha_l^F \mathbf{g}(\tau_l) \tag{71.10}$$

and

$$\mathbf{h}_2^F = \sum_{l=1}^{L} a_2(\theta_l)\alpha_l^F \mathbf{g}(\tau_l - T) \tag{71.11}$$

If spatial beamforming alone is used, the signal model becomes

$$x(k) = \sum_{l=1}^{L} \mathbf{w}^H \mathbf{H}^F \mathbf{s}(k) + n(k) \tag{71.12}$$

where $\mathbf{H}^F$ is the *intrinsic* forward (F) channel given by

$$\mathbf{H}^F = \begin{bmatrix} \mathbf{h}_1^F \\ \mathbf{h}_2^F \end{bmatrix} = \sum_{l=1}^{L} \mathbf{a}(\theta_l)\alpha_l^F \mathbf{g}^T(\tau_l) \tag{71.13}$$

Signal-Plus-Interference Model

The overall received signal-plus-interference-and-noise model at the base station antenna array can be written as

$$\mathbf{x}(k) = \mathbf{H}_s^R \mathbf{s}_s(k) + \sum_{q=1}^{Q-1} \mathbf{H}_q^R \mathbf{s}_q(k) + \mathbf{n}(k) \tag{71.14}$$

where $\mathbf{H}_s^R$ and $\mathbf{H}_q^R$ are channels for signal and **CCI**, respectively, while $\mathbf{s}_s$ and $\mathbf{s}_q$ are the corresponding data sequences. Note that Eq. (71.14) appears to suggest that the signal and interference are baud synchronous. However, this can be relaxed and the time offsets can be absorbed into the channel $\mathbf{H}_q^R$.

Similarly, the signal at the mobile can also be extended to include CCI. Note that in this case, the source of interference is from other base stations (in TDMA) and the channel is between the interfering base station and the desired mobile. It is often convenient to handle signals in blocks. Therefore, we may collect M consecutive snapshots of $\mathbf{x}(\cdot)$ corresponding to time instants $k,\ldots,k + M - 1$, (and dropping subscripts for a moment), we get

$$\mathbf{X}(k) = \mathbf{H}^R \mathbf{S}(k) + \mathbf{N}(k) \tag{71.15}$$

where $\mathbf{X}(k)$, $\mathbf{S}(k)$ and $\mathbf{N}(k)$ are defined appropriately. Similarly the received signal at the mobile in the forward link has a block representation using a row vector.

71.4 ST Receive Processing (Base)

The base station receiver receives the signals from the array antenna which consist of the signals from the desired mobile and the cochannel signals along with associated intersymbol interference and fading. The task of the receiver is to maximize signal power and mitigate fading, CCI and **ISI**. There are two broad approaches for doing this—one is multiuser detection wherein we demodulate both the cochannel and desired signals jointly, the other is to cancel CCI. The structure of the receiver depends on the nature of the channel estimates available and the tolerable receiver complexity. There are a number of options and we discuss only a few salient cases. Before discussing the receiver processing, we discuss how receiver channel is estimated.

Receive Channel Estimation (Base)

In many mobile communications standards, such as GSM and IS-54, explicit training signals are inserted inside the TDMA data bursts.

Let $\mathbf{T}$ be the training sequence arranged in a matrix form ($\mathbf{T}$ is arranged to be a Toeplitz matrix). Then, during the training burst, the received data is given by

$$\mathbf{X} = \mathbf{H}^R \mathbf{T} + \mathbf{N} \tag{71.16}$$

Clearly $\mathbf{H}^R$ can be estimated using least squares as

$$\mathbf{H}^R = \mathbf{X}\mathbf{T}^\dagger \tag{71.17}$$

where $\mathbf{T}^\dagger = \mathbf{T}^H(\mathbf{T}\mathbf{T}^H)^{-1}$.

The use of training consumes spectrum resource. In GSM, for example, about 20% of the bits are dedicated to training. Moreover, in rapidly varying mobile channels, we may have to retrain frequently, resulting in even poorer spectral efficiency. There is, therefore, increased interest in blind methods that can estimate a channel without an explicit training signal.

Multiuser ST Receive Algorithms

In multiuser (MU) algorithms, we address the problem of jointly demodulating the multiple signals. Recall the received signal is given by

$$\mathbf{X} = \mathbf{H}^R\mathbf{S} + \mathbf{N} \tag{71.18}$$

where $\mathbf{H}^R$ and $\mathbf{S}$ are suitably defined to include multiple users and are of dimensions $m \times NQ$ and $NQ \times M$, respectively.

If the channels for all the arriving signals are known, then we jointly demodulate all the user data sequences using multiuser **maximum likelihood sequence estimation** (MLSE). Starting with the data model in Eq. (71.18), we can then search for multiple user data sequences that minimize the ML cost function

$$\min_{\mathbf{S}} \left\|\mathbf{X} - \mathbf{H}^R\mathbf{S}\right\|_F^2 \tag{71.19}$$

The multiuser MLSE will have a large number of states in the trellis. Efficient techniques for implementing this complex receiver are needed. Multiuser MLSE detection schemes outperform all other receivers.

Single-User ST Receive Algorithms

In this scheme we only demodulate the desired user and cancel the CCI. Therefore, after CCI cancellation we can use MLSE receivers to handle diversity and ISI. In this method there is potential conflict between CCI mitigation and diversity maximization. We are forced to allocate the available degrees of freedom (antennas) to the competing requirements.

One approach is to cancel CCI by a space-time filter followed by an MLSE receiver to handle ISI. We do this by reformulating the MLSE criterion to arrive at a joint solution for the ST-MMSE filter and the effective channel for the scalar MLSE.

Another approach is to use a ST-MMSE receiver to handle both CCI and ISI. In a space-time filter (equalizer-beamformer), $\mathbf{W}$ has the following form

$$\mathbf{W}(k) = \begin{bmatrix} w_{11}(k) & \dots & w_{1M}(k) \\ \vdots & \dots & \vdots \\ w_{m1}(k) & \dots & w_{mM}(k) \end{bmatrix} \tag{71.20}$$

In order to obtain a convenient formulation for the space-time filter output, we introduce the quantities $W(k)$ and $X(k)$ as follows

$$\begin{aligned} X(k) &= vec(\mathbf{X}(k)) \quad (mM \times 1) \\ W(k) &= vec(\mathbf{W}(k)) \quad (mM \times 1) \end{aligned} \tag{71.21}$$

where the operator $vec(\cdot)$is defined as:

$$vec\left([\mathbf{v}_1 \dots \mathbf{v}_M]\right) = \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_M \end{bmatrix}$$

The ST-MMSE filter chooses the filter weights to achieve the minimum mean square error. The ST-MMSE filter takes the familiar form

$$W = \mathbf{R}_{XX}^{-1}\overline{H^R} \tag{71.22}$$

where $\overline{H^R}$ is one column of $vec(\mathbf{H}^R)$. In ST-MMSE the CCI and spatial diversity conflict for the spatial degrees of freedom. Likewise, temporal diversity and ISI cancellation conflict for the temporal degrees of freedom.

71.5 ST Transmit Processing (Base)

The goal in ST transmit processing is to maximize the average signal power and diversity at the receiver as well as minimize cochannel generation to other mobiles. Note that the base station transmission cannot directly affect the CCI seen by its intended mobile. In transmit the space-time processing needs channel knowledge, but since it is carried out prior to transmission and, therefore, before the signal encounters the channel, this is different from the reverse link where the space-time processing is carried out after the channel has affected the signal. Note that the mobile receiver will, of course, need to know the channel for signal demodulation, but since it sees the signal after transmission through the channel, it can estimate the forward link channel using training signals transmitted from the individual transmitter antennas.

Transmit Channel Estimation (Base)

The transmit channel estimation at the base of the vector forward channel can be done via feedback by use of reciprocity principles. In a TDD system, if the duplexing time is small compared to the coherence time of the channel, both channels are the same and the base station can use its estimate of the reverse channel as the forward channel; i.e., $\mathbf{H}^F = \mathbf{H}^R$, where $\mathbf{H}^R$ is the reverse channel (we have added superscript R to emphasize the receive channel). In FDD systems, the forward and reverse channels can potentially be very different. This arises from differences in instantaneous complex path gains $\alpha^R \pi \alpha^F$. The other channel components $\mathbf{a}(\theta_l)$ and $\mathbf{g}(\tau_l)$ are very nearly equal.

A direct approach to estimating the forward channel is to feed back the signal from the mobile unit and then estimate the channel. We can do this by transmitting orthogonal training signals through each base station antenna. We can feed back from the mobile to the base the received signal for each transmitted signal and thus estimate the channel.

ST Transmit Processing

The primary goals at the transmitter are to maximize diversity in the link and to reduce CCI generation to other mobiles. The diversity maximization depends on the inherent diversity at the antenna array and cannot be created at the transmitter. The role of ST processing is limited to maximizing the exploitability of this diversity at the receiver. This usually leads to use of orthogonal or near orthogonal signalling at each antenna: Ú$u_1(t)u_2(t)dt$ ª 0. Orthogonality ensures that the transmitted signals are separable at the mobile which can now combine these signals after appropriate weighting to attain maximum diversity.

In order to minimize CCI, our goal is to use the beamforming vector $\mathbf{w}$ to steer the radiated energy and therefore minimize the interference at the other mobiles while maximizing the signal level at one's own mobile. Note that the CCI at the reference mobile is not controlled by its own base station but is generated by other base stations. Reducing CCI at one's own mobile requires the cooperation of the other base stations.

Therefore we choose $\mathbf{w}$ such that

$$\max_{\mathbf{w}} \frac{E(\mathbf{w}^H\mathbf{H}^F\mathbf{s}(k)\mathbf{s}(k)^H\mathbf{H}^{FH}\mathbf{w})}{\sum_{q=1}^{Q-1}\mathbf{w}^H\mathbf{H}_q^F\mathbf{H}_q^{FH}\mathbf{w}} \tag{71.23}$$

where $Q - 1$ is the number of susceptible outer cell mobiles. $\mathbf{H}_q^F$ is the channel from the base station to the qth outer cell mobile. In order to solve the above equation, we need to know the forward link channel $\mathbf{H}^F$ to the reference mobile and $\mathbf{H}_q^F$ to cochannel mobiles. In general, such complete channel knowledge may not be available and suboptimum receivers must be designed. Furthermore, we need to find a receiver that harmonizes maximization of diversity and reduction of CCI. Use of transmit ST processing affects $\mathbf{H}^F$ and thus can be incorporated.

Forward Link Processing at the Mobile

The mobile will receive the composite signal from all the base station transmit antennas and will need to demodulate the signal to estimate the symbol sequence. In doing so it usually needs to estimate the individual channels from each base station antenna to itself. This is usually done via the use of training signals on each transmit antenna. Note that as the number of transmit antennas increases, there is a greater burden of training requirements. The use of transmit ST processing reduces the CCI power observed by the mobile as well enhances the diversity available.

71.6 Summary

Use of space-time processing can significantly improve average signal power, mitigate fading, and reduce cochannel and intersymbol interference in wireless networks. This can in turn result in significantly improved capacity, coverage, and quality of wireless networks.

In this chapter we have discussed applications of ST processing to TDMA systems. The applications to CDMA systems follow similar principles, but differences arise due to the nature of the signal and interference models.

Defining Terms

CCI: Cochannel interference arises from neighboring cells where the frequency channel is reused.

ISI: Intersymbol intereference is caused by multipath propagation where one symbol interferes with other symbols.

Maximum Likelihood Sequence Estimation: A technique for channel equalization based on determining the best symbol sequence that matches the received signal.

References

1. Lindskog, E. and Paulraj, A., A taxonomy of space-time signal processing, *IEE Trans. Radar and Sonar,* 25–31, Feb. 1998.
2. Ng, B.C. and Paulraj, A., Space-time processing for PCS, *IEEE PCS Magazine,* 5(1), 36–48, Feb.1998.
3. Paulraj, A. and Papadias, C.B., Space-time processing for wireless communications, *IEEE Signal Processing Magazine,* 14(5), 49–83, Nov. 1997.
4. Paulraj, A., Papadias, C., Reddy, V.U., and Van der Veen, A., *A Review of Space-Time Signal Processing for Wireless Communications,* in *Signal Processing for Wireless Communications,* V. Poor, Ed., Prentice Hall, 179–210, Dec. 1997.

72

Location Strategies for Personal Communications Services

Ravi K. Jain
Telcordia Technologies

Yi-Bing Lin
National Chaio Tung University

Seshadri Mohan
Comverse Network Systems

72.1 Introduction

The vision of nomadic personal communications is the ubiquitous availability of services to facilitate exchange of information (voice, data, video, image, etc.) between nomadic end users independent of time, location, or access arrangements. To realize this vision, it is necessary to locate users that move

[1]" 1996 by Bell Communications Research, Inc. Used with permission. The material in this chapter appeared originally in the following IEEE publications: S. Mohan and R. Jain. 1994. Two user location strategies for personal communications services, *IEEE Personal Communications: The Magazine of Nomadic Communications and Computing*, pp. 42–50, Feb., and R. Jain, C.N. Lo, and S. Mohan, 1994. A caching strategy to reduce network impacts of PCS, J-SAC Special Issue on Wireless and Mobile Networks, Aug.

from place to place. The strategies commonly proposed are two-level hierarchical strategies, which maintain a system of mobility databases, home location registers (HLR) and visitor location resisters (VLR), to keep track of user locations. Two standards exist for carrying out two-level hierarchical strategies using HLRs and VLRs. The standard commonly used in North America is the EIA/TIA Interim Standard 41 (IS 41) [6] and in Europe the Global System for Mobile Communications (GSM) [15,18]. In this chapter, we refer to these two strategies as *basic* location strategies.

We introduce these two strategies for locating users and provide a tutorial on their usage. We then analyze and compare these basic location strategies with respect to load on mobility databases and signalling network. Next we propose an auxiliary strategy, called the *per-user caching* or, simply, the *caching* strategy, that augments the basic location strategies to reduce the signalling and database loads.

The outline of this chapter is as follows. In Section 72.2 we discuss different forms of mobility in the context of personal communications services (PCS) and describe a reference model for a PCS architecture. In Sections 72.3 and 72.4, we describe the user location strategies specified in the IS-41 and GSM standards, respectively, and in Section 72.5, using a simple example, we present a simplified analysis of the database loads generated by each strategy. In Section 72.6, we briefly discuss possible modifications to these protocols that are likely to result in significant benefits by either reducing query and update rate to databases or reducing the signalling traffic or both. Section 72.7 introduces the caching strategy followed by an analysis in the next two sections. This idea attempts to exploit the spatial and temporal locality in calls received by users, similar to the idea of exploiting locality of file access in computer systems [20]. A feature of the caching location strategy is that it is useful only for certain classes of PCS users, those meeting certain call and mobility criteria. We encapsulate this notion in the definition of the user's call-to-mobility ratio (CMR), and local CMR (LCMR), in Section 72.8. We then use this definition and our PCS network reference architecture to quantify the costs and benefits of caching and the threshold LCMR for which caching is beneficial, thus characterizing the classes of users for which caching should be applied. In Section 72.9, we describe two methods for estimating users' LCMR and compare their effectiveness when call and mobility patterns are fairly stable, as well as when they may be variable. In Section 72.10, we briefly discuss alternative architectures and implementation issues of the strategy proposed and mention other auxiliary strategies that can be designed. Section 72.11 provides some conclusions and discussion of future work.

The choice of platforms on which to realize the two location strategies (IS-41 and GSM) may vary from one service provider to another. In this paper, we describe a possible realization of these protocols based on the advanced intelligent network (AIN) architecture (see [2,5]), and signalling system 7 (SS7). It is also worthwhile to point out that several strategies have been proposed in the literature for locating users, many of which attempt to reduce the signalling traffic and database loads imposed by the need to locate users in PCS.

72.2 An Overview of PCS

This section explains different aspects of mobility in PCS using an example of two nomadic users who wish to communicate with each other. It also describes a reference model for PCS.

Aspects of Mobility—Example 72.1

PCS can involve two possible types of mobility, terminal mobility and personal mobility, that are explained next.

Terminal Mobility: This type of mobility allows a terminal to be identified by a unique terminal identifier independent of the point of attachment to the network. Calls intended for that terminal can therefore be delivered to that terminal regardless of its network point of attachment. To facilitate terminal mobility, a network must provide several functions, which include those that locate, identify, and validate a terminal and provide services (e.g., deliver calls) to the terminal based on the location information. This implies that the network must store and maintain the location information of the terminal based on a unique

identifier assigned to that terminal. An example of a terminal identifier is the IS-41 EIA/TIA cellular industry term mobile identification number (MIN), which is a North American Numbering Plan (NANP) number that is stored in the terminal at the time of manufacture and cannot be changed. A similar notion exists in GSM (see Section 72.4).

Personal Mobility: This type of mobility allows a PCS user to make and receive calls independent of both the network point of attachment and a specific PCS terminal. This implies that the services that a user has subscribed to (stored in that user's service profile) are available to the user even if the user moves or changes terminal equipment. Functions needed to provide personal mobility include those that identify (authenticate) the end user and provide services to an end user independent of both the terminal and the location of the user. An example of a functionality needed to provide personal mobility for voice calls is the need to maintain a user's location information based on a unique number, called the universal personal telecommunications (UPT) number, assigned to that user. UPT numbers are also NANP numbers. Another example is one that allows end users to define and manage their service profiles to enable users to tailor services to suit their needs. In Section 72.4, we describe how GSM caters to personal mobility via smart cards.

For the purposes of the example that follows, the terminal identifiers (TID) and UPT numbers are NANP numbers, the distinction being TIDs address terminal mobility and UPT numbers address personal mobility. Though we have assigned two different numbers to address personal and terminal mobility concerns, the same effect could be achieved by a single identifier assigned to the terminal that varies depending on the user that is currently utilizing the terminal. For simplicity we assume that two different numbers are assigned.

Figure 72.1 illustrates the terminal and personal mobility aspects of PCS, which will be explained via an example. Let us assume that users Kate and Al have, respectively, subscribed to PCS services from PCS service provider (PSP) A and PSP B. Kate receives the UPT number, say, 500 111 4711, from PSP A. She also owns a PCS terminal with TID 200 777 9760. Al too receives his UPT number 500 222 4712

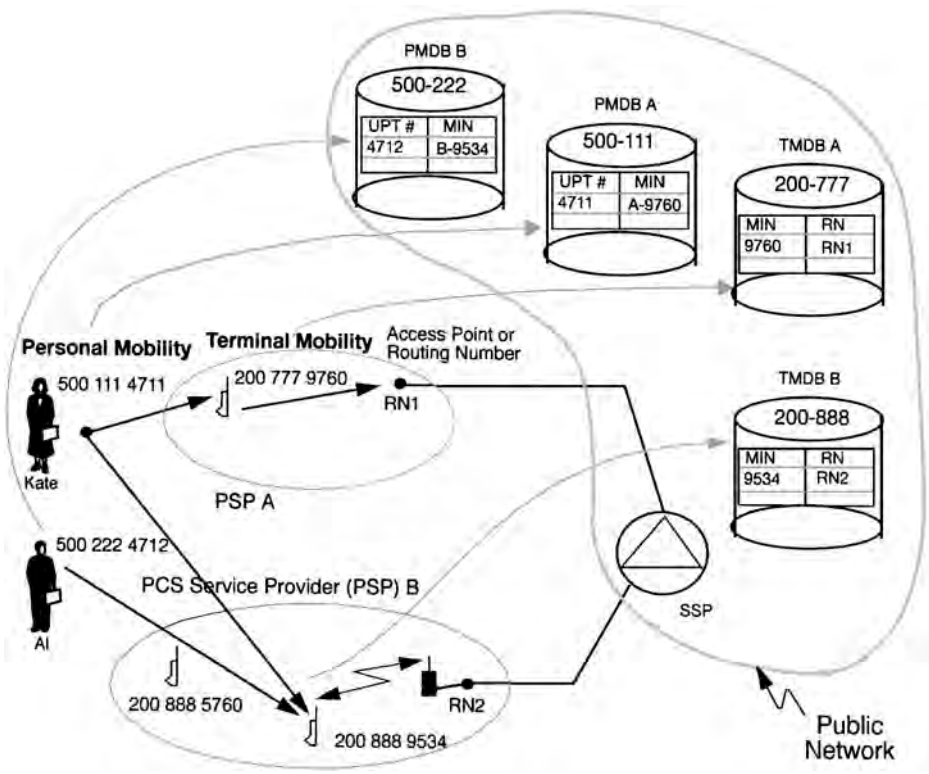


FIGURE 72.1 Illustrating terminal and personal mobility.

from PSP B, and he owns a PCS terminal with TID 200 888 5760. Each has been provided a personal identification number (PIN) by their respective PSP when subscription began. We assume that the two PSPs have subscribed to PCS access services from a certain network provider such as, for example, a local exchange carrier (LEC). (Depending on the capabilities of the PSPs, the access services provided may vary. Examples of access services include translation of UPT number to a routing number, terminal and personal registration, and call delivery. Refer to Bellcore [3] for further details.) When Kate plugs in her terminal to the network, or when she activates it, the terminal registers itself with the network by providing its TID to the network. The network creates an entry for the terminal in an appropriate database, which, in this example, is entered in the terminal mobility database (TMDB) A. The entry provides a mapping of her terminal's TID, 200 777 9760, to a routing number (RN), RN1. All of these activities happen without Kate being aware of them. After activating her terminal, Kate registers herself at that terminal by entering her UPT number (500 111 4711) to inform the network that all calls to her UPT number are to be delivered to her at the terminal. For security reasons, the network may want to authenticate her and she may be prompted to enter her PIN number into her terminal. (Alternatively, if the terminal is equipped with a smart card reader, she may enter her smart card into the reader. Other techniques, such as, for example, voice recognition, may be employed.) Assuming that she is authenticated, Kate has now registered herself. As a result of personal registration by Kate, the network creates an entry for her in the personal mobility database (PMDB) A that maps her UPT number to the TID of the terminal at which she registered. Similarly, when Al activates his terminal and registers himself, appropriate entries are created in TMDB B and PMDB B. Now Al wishes to call Kate and, hence, he dials Kate's UPT number (500 111 4711). The network carries out the following tasks.

1. The switch analyzes the dialed digits and recognizes the need for AIN service, determines that the dialed UPT number needs to be translated to a RN by querying PMDB A and, hence, it queries PMDB A.
2. PMDB A searches its database and determines that the person with UPT number 500 111 4711 is currently registered at terminal with TID 200 777 9760.
3. PMDB A then queries TMDB A for the RN of the terminal with TID 200 777 9760. TMDB A returns the RN (RN1).
4. PMDB A returns the RN (RN1) to the originating switch.
5. The originating switch directs the call to the switch RN1, which then alerts Kate's terminal. The call is completed when Kate picks up her terminal.

Kate may take her terminal wherever she goes and perform registration at her new location. From then on, the network will deliver all calls for her UPT number to her terminal at the new location. In fact, she may actually register on someone else's terminal too. For example, suppose that Kate and Al agree to meet at Al's place to discuss a school project they are working on together. Kate may register herself on Al's terminal (TID 200 888 9534). The network will now modify the entry corresponding to 4711 in PMDB A to point to B 9534. Subsequent calls to Kate will be delivered to Al's terminal.

The scenario given here is used only to illustrate the key aspects of terminal and personal mobility; an actual deployment of these services may be implemented in ways different from those suggested here. We will not discuss personal registration further. The analyses that follow consider only terminal mobility but may easily be modified to include personal mobility.

A Model for PCS

Figure 72.2 illustrates the reference model used for the comparative analysis. The model assumes that the HLR resides in a service control point (SCP) connected to a regional signal transfer point (RSTP). The SCP is a storehouse of the AIN service logic, i.e., functionality used to perform the processing required to provide advanced services, such as speed calling, outgoing call screening, etc., in the AIN architecture (see Bellcore [2] and Berman and Brewster [5]). The RSTP and the local STP (LSTP) are packet switches, connected together by various links such A links or D links, that perform the signalling

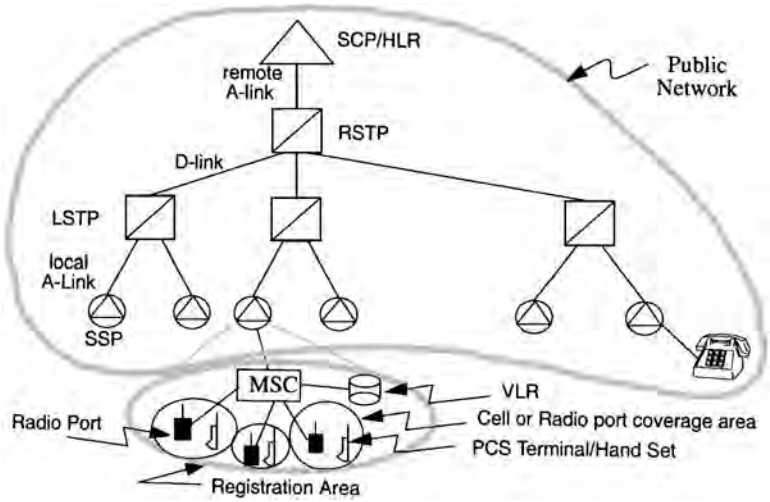


FIGURE 72.2 Example of a reference model for a PCS.

functions of the SS7 network. Such functions include, for example, global title translation for routing messages between the AIN switching system, which is also referred to as the service switching point (SSP), and SCP and IS-41 messages [6]. Several SSPs may be connected to an LSTP.

The reference model in Fig. 72.2 introduces several terms which are explained next. We have tried to keep the terms and discussions fairly general. Wherever possible, however, we point to equivalent cellular terms from IS-41 or GSM.

For our purposes, the geographical area served by a PCS system is partitioned into a number of radio port coverage areas (or cells, in cellular terms) each of which is served by a radio port (or, equivalently, base station) that communicates with PCS terminals in that cell. A registration area (also known in the cellular world as location area) is composed of a number of cells. The base stations of all cells in a registration area are connected by wireline links to a mobile switching center (MSC). We assume that each registration area is served by a single VLR. The MSC of a registration area is responsible for maintaining and accessing the VLR and for switching between radio ports. The VLR associated with a registration area is responsible for maintaining a subset of the user information contained in the HLR.

Terminal registration process is initiated by terminals whenever they move into a new registration area. The base stations of a registration area periodically broadcast an identifier associated with that area. The terminals periodically compare an identifier they have stored with the identifier to the registration area being broadcast. If the two identifiers differ, the terminal recognizes that it has moved from one registration area to another and will, therefore, generate a registration message. It also replaces the previous registration area identifier with that of the new one. Movement of a terminal within the same registration area will not generate registration messages. Registration messages may also be generated when the terminals are switched on. Similarly, messages are generated to deregister them when they are switched off.

PCS services may be provided by different types of commercial service vendors. Bellcore [3] describes three different types of PSPs and the different access services that a public network may provide to them. For example, a PSP may have full network capabilities with its own switching, radio management, and radio port capabilities. Certain others may not have switching capabilities, and others may have only radio port capabilities. The model in Fig. 72.2 assumes full PSP capabilities. The analysis in Section 72.5 is based on this model and modifications may be necessary for other types of PSPs.

It is also quite possible that one or more registration areas may be served by a single PSP. The PSP may have one or more HLRs for serving its service area. In such a situation users that move within the PSP's serving area may generate traffic to the PSP's HLR (not shown in Fig. 72.2) but not to the network's HLR (shown in Fig. 72.2). In the interest of keeping the discussions simple, we have assumed that there

is one-to-one correspondence between SSPs and MSCs and also between MSCs, registration areas, and VLRs. One impact of locating the SSP, MSC, and VLR in separate physical sites connected by SS7 signalling links would be to increase the required signalling message volume on the SS7 network. Our model assumes that the messages between the SSP and the associated MSC and VLR do not add to signalling load on the public network. Other configurations and assumptions could be studied for which the analysis may need to be suitably modified. The underlying analysis techniques will not, however, differ significantly.

72.3 IS-41 Preliminaries

We now describe the message flow for call origination, call delivery, and terminal registration, sometimes called location registration, based on the IS-41 protocol. This protocol is described in detail in EIA/TIA [6]. Only an outline is provided here.

Terminal/Location Registration

During IS-41 registration, signalling is performed between the following pairs of network elements:

- New serving MSC and the associated database (or VLR)
- New database (VLR) in the visited area and the HLR in the public network
- HLR and the VLR in former visited registration area or the old MSC serving area

Figure 72.3 shows the signalling message flow diagram for IS-41 registration activity, focusing only on the essential elements of the message flow relating to registration; for details of variations from the basic registration procedure, see Bellcore [3].

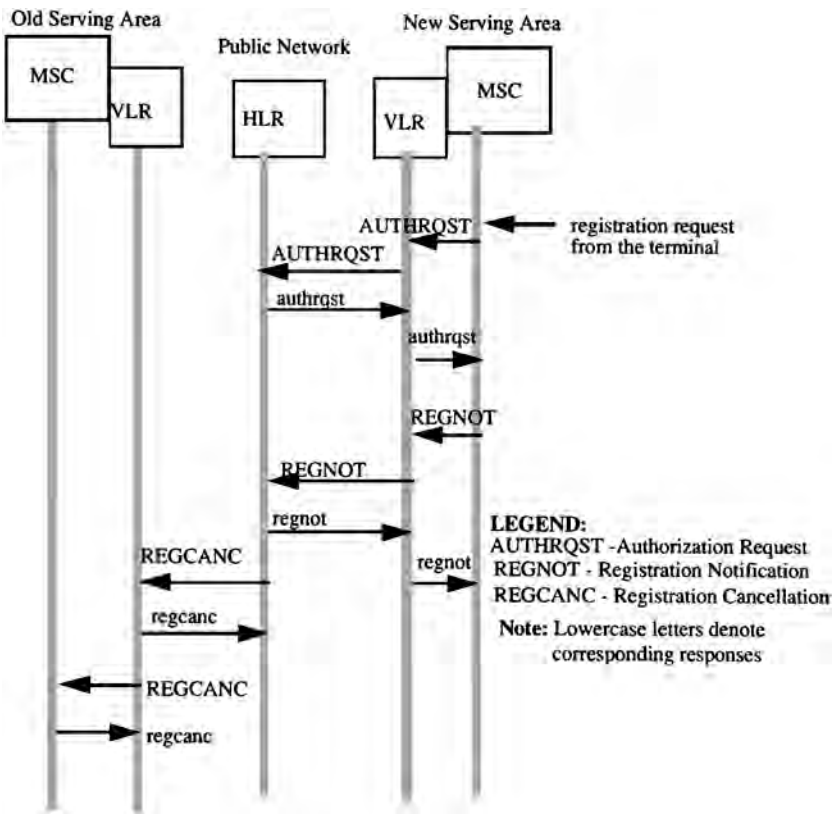


FIGURE 72.3 Signalling flow diagram for registration in IS-41.

The following steps describe the activities that take place during registration.

1. Once a terminal enters a new registration area, the terminal sends a registration request to the MSC of that area.
2. The MSC sends an authentication request (AUTHRQST) message to its VLR to authenticate the terminal, which in turn sends the request to the HLR. The HLR sends its response in the authrqst message.
3. Assuming the terminal is authenticated, the MSC sends a registration notification (REGNOT) message to its VLR.
4. The VLR in turn sends a REGNOT message to the HLR serving the terminal. The HLR updates the location entry corresponding to the terminal to point to the new serving MSC/VLR. The HLR sends a response back to the VLR, which may contain relevant parts of the user's service profile. The VLR stores the service profile in its database and also responds to the serving MSC.
5. If the user A terminal was registered previously in a different registration area, the HLR sends a registration cancellation (REGCANC) message to the previously visited VLR. On receiving this message, the VLR erases all entries for the terminal from the record and sends a REGCANC message to the previously visited MSC, which then erases all entries for the terminal from its memory.

The protocol shows authentication request and registration notification as separate messages. If the two messages can be packaged into one message, then the rate of queries to HLR may be cut in half. This does not necessarily mean that the total number of messages are cut in half.

Call Delivery

The signalling message flow diagram for IS-41 call delivery is shown in Fig. 72.4. The following steps describe the activities that take place during call delivery.

1. A call origination is detected and the number of the called terminal (for example, MIN) is received by the serving MSC. Observe that the call could have originated from within the public network

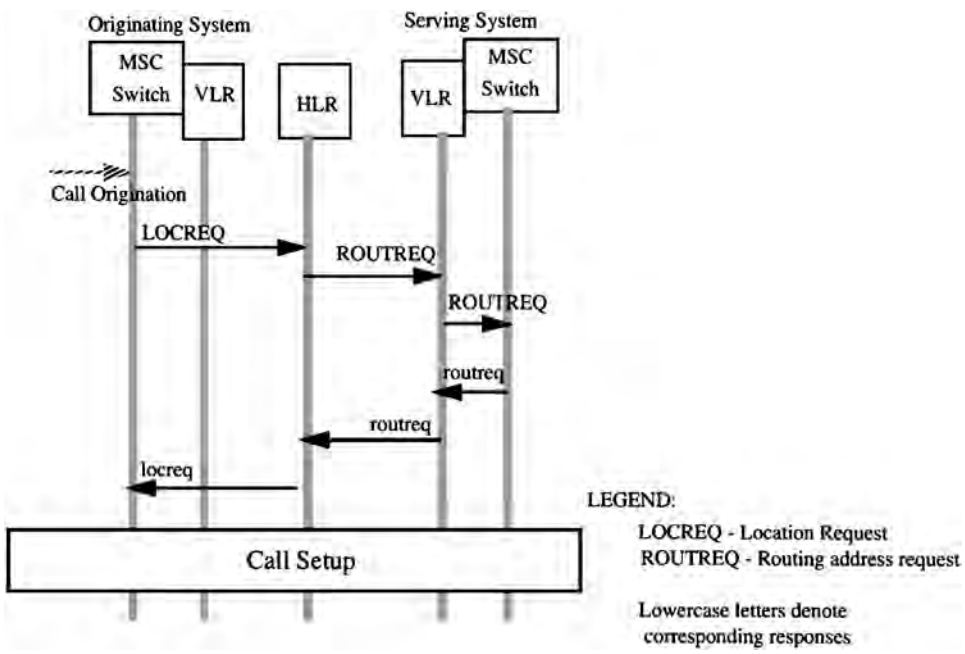


FIGURE 72.4 Signalling flow diagram for call delivery in IS-41.

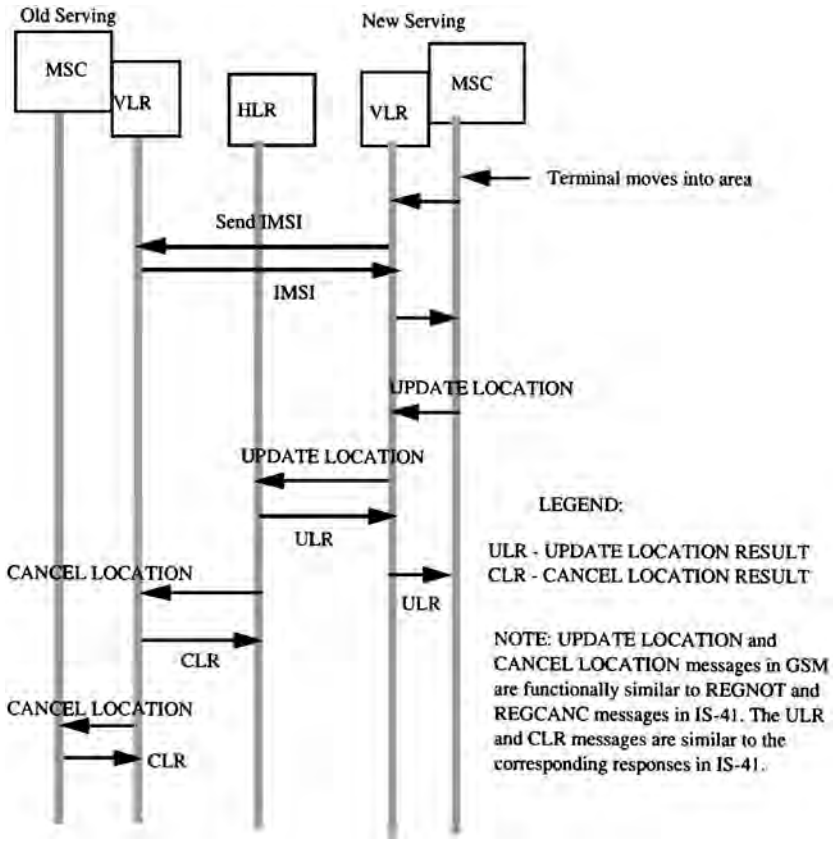


FIGURE 72.5 Flow diagram for registration in GSM.

from a wireline phone or from a wireless terminal in an MSC/VLR serving area. (If the call originated within the public network, the AIN SSP analyzes the dialed digits and sends a query to the SCP.)

2. The MSC determines the associated HLR serving the called terminal and sends a location request (LOCREQ) message to the HLR.
3. The HLR determines the serving VLR for that called terminal and sends a routing address request (ROUTEREQ) to the VLR, which forwards it to the MSC currently serving the terminal.
4. Assuming that the terminal is idle, the serving MSC allocates a temporary identifier, called a temporary local directory number (TLDN), to the terminal and returns a response to the HLR containing this information. The HLR forwards this information to the originating SSP/MSC in response to its LOCREQ message.
5. The originating SSP requests call setup to the serving MSC of the called terminal via the SS7 signalling network using the usual call setup protocols.

Similar to the considerations for reducing signalling traffic for location registration, the VLR and HLR functions could be united in a single logical database for a given serving area and collocated; further, the database and switch can be integrated into the same piece of physical equipment or be collocated. In this manner, a significant portion of the messages exchanged between the switch, HLR, and VLR as shown in Fig. 72.4 will not contribute to signalling traffic.

72.4 Global System for Mobile Communications

In this section we describe the user location strategy proposed in the European Global System for Mobile Communications (GSM) standard and its offshoot, digital cellular system 1800 (DCS1800). There has recently been increased interest in GSM in North America, since it is possible that early deployment of PCS will be facilitated by using the communication equipment already available from European manufacturers who use the GSM standard. Since the GSM standard is relatively unfamiliar to North American readers, we first give some background and introduce the various abbreviations. The reader will find additional details in Mouley and Pautet [18]. For an overview on GSM, refer to Lycksell [15].

The abbreviation GSM originally stood for Groupe Special Mobile, a committee created within the pan-European standardization body Conference Europeenne des Posts et Telecommunications (CEPT) in 1982. There were numerous national cellular communication systems and standards in Europe at the time, and the aim of GSM was to specify a uniform standard around the newly reserved 900-MHz frequency band with a bandwidth of twice 25 MHz. The phase 1 specifications of this standard were frozen in 1990. Also in 1990, at the request of the United Kingdom, specification of a version of GSM adapted to the 1800-MHz frequency, with bandwidth of twice 75 MHz, was begun. This variant is referred to as DCS1800; the abbreviation GSM900 is sometimes used to distinguish between the two variations, with the abbreviation GSM being used to encompass both GSM900 and DSC1800. The motivation for DCS1800 is to provide higher capacities in densely populated urban areas, particularly for PCS. The DCS1800 specifications were frozen in 1991, and by 1992 all major GSM900 European operators began operation.

At the end of 1991, activities concerning the post-GSM generation of mobile communications were begun by the standardization committee, using the name universal mobile telecommunications system (UMTS) for this effort. In 1992, the name of the standardization committee was changed from GSM to special mobile group (SMG) to distinguish it from the 900-MHz system itself, and the term GSM was chosen as the commercial trademark of the European 900-MHz system, where GSM now stands for global system for mobile communications.

The GSM standard has now been widely adopted in Europe and is under consideration in several other non-European countries, including the United Arab Emirates, Hong Kong, and New Zealand. In 1992, Australian operators officially adopted GSM.

Architecture

In this section we describe the GSM architecture, focusing on those aspects that differ from the architecture assumed in the IS-41 standard.

A major goal of the GSM standard was to enable users to move across national boundaries and still be able to communicate. It was considered desirable, however, that the operational network within each country be operated independently. Each of the operational networks is called a public land mobile network (PLMN) and its commercial coverage area is confined to the borders of one country (although some radio coverage overlap at national boundaries may occur), and each country may have several competing PLMNs.

A GSM customer subscribes to a single PLMN called the home PLMN, and subscription information includes the services the customer subscribes to. During normal operation, a user may elect to choose other PLMNs as their service becomes available (either as the user moves or as new operators enter the marketplace). The user's terminal [GSM calls the terminal a mobile station (MS)] assists the user in choosing a PLMN in this case, either presenting a list of possible PLMNs to the user using explicit names (e.g., DK Sonofon for the Danish PLMN) or choosing automatically based on a list of preferred PLMNs stored in the terminal's memory. This PLMN selection process allows users to choose between the services and tariffs of several competing PLMNs. Note that the PLMN selection process differs from the cell selection and handoff process that a terminal carries out automatically without any possibility of user

intervention, typically based on received radio signal strengths and, thus, requires additional intelligence and functionality in the terminal.

The geographical area covered by a PLMN is partitioned into MSC serving areas, and a registration area is constrained to be a subset of a single MSC serving area. The PLMN operator has complete freedom to allocate cells to registration areas. Each PLMN has, logically speaking, a single HLR, although this may be implemented as several physically distributed databases, as for IS-41. Each MSC also has a VLR, and a VLR may serve one or several MSCs. As for IS-41, it is interesting to consider how the VLR should be viewed in this context. The VLR can be viewed as simply a database off loading the query and signalling load on the HLR and, hence, logically tightly coupled to the HLR or as an ancillary processor to the MSC. This distinction is not academic; in the first view, it would be natural to implement a VLR as serving several MSCs, whereas in the second each VLR would serve one MSC and be physically closely coupled to it. For GSM, the MSC implements most of the signalling protocols, and at present all switch manufacturers implement a combined MSC and VLR, with one VLR per MSC [18].

A GSM mobile station is split in two parts, one containing the hardware and software for the radio interface and the other containing subscribers-specific and location information, called the subscriber identity module (SIM), which can be removed from the terminal and is the size of a credit card or smaller. The SIM is assigned a unique identity within the GSM system, called the international mobile subscriber identity (IMSI), which is used by the user location strategy as described in the next subsection. The SIM also stores authentication information, services lists, PLMN selection lists, etc., and can itself be protected by password or PIN.

The SIM can be used to implement a form of large-scale mobility called SIM roaming. The GSM specifications standardize the interface between the SIM and the terminal, so that a user carrying his or her SIM can move between different terminals and use the SIM to personalize the terminal. This capability is particularly useful for users who move between PLMNs which have different radio interfaces. The user can use the appropriate terminal for each PLMN coverage area while obtaining the personalized facilities specified in his or her SIM. Thus, SIMs address personal mobility. In the European context, the usage of two closely related standards at different frequencies, namely, GSM900 and DCS1800, makes this capability an especially important one and facilitates interworking between the two systems.

User Location Strategy

We present a synopsis of the user location strategy in GSM using call flow diagrams similar to those used to describe the strategy in IS-41.

In order to describe the registration procedure, it is first useful to clarify the different identifiers used in this procedure. The SIM of the terminal is assigned a unique identity, called the IMSI, as already mentioned. To increase confidentiality and make more efficient use of the radio bandwidth, however, the IMSI is not normally transmitted over the radio link. Instead, the terminal is assigned a temporary mobile subscriber identity (TMSI) by the VLR when it enters a new registration area. The TMSI is valid only within a given registration area and is shorter than the IMSI. The IMSI and TMSI are identifiers that are internal to the system and assigned to a terminal or SIM and should not be confused with the user's number that would be dialed by a calling party; the latter is a separate number called the mobile subscriber integrated service digital network (ISDN) number (MSISDN), and is similar to the usual telephone number in a fixed network.

We now describe the procedure during registration. The terminal can detect when it has moved into the cell of a new registration area from the system information broadcast by the base station in the new cell. The terminal initiates a registration update request to the new base station; this request includes the identity of the old registration area and the TMSI of the terminal in the old area. The request is forwarded to the MSC, which, in turn, forwards it to the new VLR. Since the new VLR cannot translate the TMSI to the IMSI of the terminal, it sends a request to the old VLR to send the IMSI of the terminal corresponding to that TMSI. In its response, the old VLR also provides the required authentication information. The new VLR then initiates procedures to authenticate the terminal. If the authentication succeeds, the VLR uses the IMSI to determine the address of the terminal's HLR.

The ensuing protocol is then very similar to that in IS-41, except for the following differences. When the new VLR receives the registration affirmation (similar to regnot in IS-41) from the HLR, it assigns a new TMSI to the terminal for the new registration area. The HLR also provides the new VLR with all relevant subscriber profile information required for call handling (e.g., call screening lists, etc.) as part of the affirmation message. Thus, in contrast with IS-41, authentication and subscriber profile information are obtained from both the HLR and old VLR and not just the HLR.

The procedure for delivering calls to mobile users in GSM is very similar to that in IS-41. The sequence of messages between the caller and called party's MSC/VLRs and the HLR is identical to that shown in the call flow diagrams for IS-41, although the names, contents and lengths of messages may be different and, hence, the details are left out. The interested reader is referred to Mouly and Pautet [18] or Lycksell [15] for further details.

72.5 Analysis of Database Traffic Rate for IS-41 and GSM

In the two subsections that follow, we state the common set of assumptions on which we base our comparison of the two strategies.

The Mobility Model for PCS Users

In the analysis that follows in the IS-41 analysis subsection, we assume a simple mobility model for the PCS users. The model, which is described in [23], assumes that PCS users carrying terminals are moving at an average velocity of v and their direction of movement is uniformly distributed over $[0, 2\pi]$. Assuming that the PCS users are uniformly populated with a density of ρ and the registration area boundary is of length L, it has been shown that the rate of registration area crossing R is given by

$$R = \frac{\rho v L}{\pi} \tag{72.1}$$

Using Eq. (72.1), we can calculate the signalling traffic due to registration, call origination, and delivery. We now need a set of assumptions so that we may proceed to derive the traffic rate to the databases using the model in Fig. 72.2.

Additional Assumptions

The following assumptions are made in performing the analysis.

- 128 total registration areas
- Square registration area size: $(7.575 \text{ km})^2 = 57.5 \text{ km}^2$, with border length $L = 30.3$ km
- Average call origination rate = average call termination (delivery) rate = 1.4/h/terminal
- Mean density of mobile terminals = $\rho = 390/\text{km}^2$
- Total number of mobile terminals = $128 \times 57.4 \times 390 = 2.87 \times 10^6$
- Average call origination rate = average call termination (delivery) rate = 1.4/h/terminal
- Average speed of a mobile, $v = 5.6$ km/h
- Fluid flow mobility model

The assumptions regarding the total number of terminals may also be obtained by assuming that a certain public network provider serves 19.15×10^6 users and that 15% (or 2.87×10^6) of the users also subscribe to PCS services from various PSPs.

Note that we have adopted a simplified model that ignores situations where PCS users may turn their handsets on and off that will generate additional registration and deregistration traffic. The model also ignores wireline registrations. These activities will increase the total number of queries and updates to HLR and VLRs.

Analysis of IS-41

Using Eq. (72.1) and the parameter values assumed in the preceding subsection, we can compute the traffic due to registration. The registration traffic is generated by mobile terminals moving into a new registration area, and this must equal the mobile terminals moving out of the registration area, which per second is

$$R_{\text{reg, VLR}} = \frac{390 \times 30.3 \times 5.6}{3600\pi} = 5.85$$

This must also be equal to the number of deregistrations (registration cancellations),

$$R_{\text{dereg, VLR}} = 5.85$$

The total number of registration messages per second arriving at the HLR will be

$$R_{\text{reg, HLR}} = R_{\text{reg, VLR}} \times \text{total No. of registration areas} = 749$$

The HLR should, therefore, be able to handle, roughly, 750 updates per second. We observe from Fig. 72.3 that authenticating terminals generate as many queries to VLR and HLR as the respective number of updates generated due to registration notification messages.

The number of queries that the HLR must handle during call origination and delivery can be similarly calculated. Queries to HLR are generated when a call is made to a PCS user. The SSP that receives the request for a call, generates a location request (LOCREQ) query to the SCP controlling the HLR. The rate per second of such queries must be equal to the rate of calls made to PCS users.

This is calculated as

$$\begin{aligned} R_{\text{CallDeliv, HLR}} &= \text{call rate per user} \times \text{total number of users} \\ &= \frac{1.4 \times 2.87 \times 10^5}{3600} \\ &= 1116 \end{aligned}$$

For calls originated from a mobile terminal by PCS users, the switch authenticates the terminal by querying the VLR. The rate per second of such queries is determined by the rate of calls originating in an SSP serving area, which is also a registration area (RA). This is given by

$$R_{\text{CallOrig, VLR}} = \frac{1116}{128} = 8.7$$

This is also the number of queries per second needed to authenticate terminals of PCS users to which calls are delivered:

$$R_{\text{CallDeliv, VLR}} = 8.7$$

Table 72.1 summarizes the calculations.

Analysis of GSM

Calculations for query and update rates for GSM may be performed in the same manner as for IS-41, and they are summarized in Table 72.2. The difference between this table and Table 72.1 is that in GSM the new serving VLR does not query the HLR separately in order to authenticate the terminal during registration and, hence, there are no HLR queries during registration. Instead, the entry (749 queries)

TABLE 72.1 IS-41 Query and Update Rates to HLR and VLR

Activity	HLR Updates/s	VLR Updates/s	HLR Queries/s	VLR Queries/s
Mobility-related activities at registration	749	5.85	749	5.85
Mobility-related activities at deregistration		5.85		
Call origination				8.7
Call delivery			1116	8.7
Total (per RA)	5.85	11.7	14.57	23.25
Total (Network)	749	1497.6	1865	2976

TABLE 72.2 GMS Query and Update Rates to HLR and VLR

Activity	HLR Updates/s	VLR Updates/s	HLR Queries/s	VLR Queries/s
Mobility-related activities at registration	749	5.85		11.7
Mobility-related activities at deregistration		5.85		
Call origination				8.7
Call delivery			1116	8.7
Total (per VLR)	749	11.7	1116	29.1
Total (Network)	749	1497.6	1116	3724.8

under HLR queries in Table 72.1, corresponding to mobility-related authentication activity at registration, gets equally divided between the 128 VLRs. Observe that with either protocol the total database traffic rates are conserved, where the total database traffic for the entire network is given by the sum of all of the entries in the last row total (Network), i.e.,

$$\text{HLR updates} + \text{VLR updates} + \text{HLR queries} + \text{VLR queries}$$

From Tables 72.1 and 72.2 we see that this quantity equals 7087.

The conclusion is independent of any variations we may provide to the assumptions in earlier in the section. For example, if the PCS penetration (the percentage of the total users subscribing to PCS services) were to increase from 15 to 30%, all of the entries in the two tables will double and, hence, the total database traffic generated by the two protocols will still be equal.

72.6 Reducing Signalling During Call Delivery

In the preceding section, we provided a simplified analysis of some scenarios associated with user location strategies and the associated database queries and updates required. Previous studies [13,16] indicate that the signalling traffic and database queries associated with PCS due to user mobility are likely to grow to levels well in excess of that associated with a conventional call. It is, therefore, desirable to study modifications to the two protocols that would result in reduced signalling and database traffic. We now provide some suggestions.

For both GSM and IS-41, delivery of calls to a mobile user involves four messages: from the caller's VLR to the called party's HLR, from the HLR to the called party's VLR, from the called party's VLR to the HLR, and from the HLR to the caller's VLR. The last two of these messages involve the HLR, whose role is to simply relay the routing information provided by the called party's VLR to the caller's VLR. An obvious modification to the protocol would be to have the called VLR directly send the routing information to the calling VLR. This would reduce the total load on the HLR and on signalling network

links substantially. Such a modification to the protocol may not be easy, of course, due to administrative, billing, legal, or security concerns. Besides, this would violate the query A response model adopted in IS-41, requiring further analysis.

A related question which arises is whether the routing information obtained from the called party's VLR could instead be stored in the HLR. This routing information could be provided to the HLR, for example, whenever a terminal registers in a new registration area. If this were possible, two of the four messages involved in call delivery could be eliminated. This point was discussed at length by the GSM standards body, and the present strategy was arrived at. The reason for this decision was to reduce the number of temporary routing numbers allocated by VLRs to terminals in their registration area. If a temporary routing number (TLDN in IS-41 or MSRN in GSM) is allocated to a terminal for the whole duration of its stay in a registration area, the quantity of numbers required is much greater than if a number is assigned on a per-call basis. Other strategies may be employed to reduce signalling and database traffic via intelligent paging or by storing user's mobility behavior in user profiles (see, for example, Tabbane [22]). A discussion of these techniques is beyond the scope of the paper.

72.7 Per-User Location Caching

The basic idea behind per-user location caching is that the volume of SS7 message traffic and database accesses required in locating a called subscriber can be reduced by maintaining local storage, or cache, of user location information at a switch. At any switch, location caching for a given user should be employed only if a large number of calls originate for that user from that switch, relative to the user's mobility. Note that the cached information is kept at the switch from which calls originate, which may or may not be the switch where the user is currently registered.

Location caching involves the storage of location pointers at the originating switch; these point to the VLR (and the associated switch) where the user is currently registered. We refer to the procedure of locating a PCS user a *FIND* operation, borrowing the terminology from Awerbuch and Peleg [1]. We define a basic *FIND*, or *BasicFIND*( ), as one where the following sequence of steps takes place.

1. The incoming call to a PCS user is directed to the nearest switch.
2. Assuming that the called party is not located within the immediate RA, the switch queries the HLR for routing information.
3. The HLR contains a pointer to the VLR in whose associated RA the subscriber is currently situated and launches a query to that VLR.
4. The VLR, in turn, queries the MSC to determine whether the user terminal is capable of receiving the call (i.e., is idle) and, if so, the MSC returns a routable address (TLDN in IS-41) to the VLR.
5. The VLR relays the routing address back to the originating switch via the HLR.

At this point, the originating switch can route the call to the destination switch. Alternately, *BasicFIND*( ) can be described by pseudocode as follows. (We observe that a more formal method of specifying PCS protocols may be desirable.)

```
BasicFIND( ){
    Call to PCS user is detected at local switch;
    if called party is in same RA then return;
    Switch queries called party's HLR;
    Called party's HLR queries called party's current VLR, V;
    V returns called party's location to HLR;
    HLR returns location to calling switch;
}
```

In the *FIND* procedure involving the use of location caching, or *CacheFIND*( ), each switch contains a local memory (cache) that stores location information for subscribers. When the switch receives a call

origination (from either a wire-line or wireless caller) directed to a PCS subscriber, it first checks its cache to see if location information for the called party is maintained. If so, a query is launched to the pointed VLR; if not, *BasicFIND*( ), as just described, is followed. If a cache entry exists and the pointed VLR is queried, two situations are possible. If the user is still registered at the RA of the pointed VLR (i.e., we have a *cache hit*), the pointed VLR returns the user's routing address. Otherwise, the pointed VLR returns a *cache miss*.

```
CacheFIND( ){
    Call to PCS user is detected at local switch;
    if called is in same RA then return;
    if there is no cache entry for called user
    then invoke BasicFIND( ) and return;
    Switch queries the VLR, V, specified in the cache entry;
    if called is at V, then
        V returns called party's location to calling switch;
    else {
        V returns "miss" to calling switch;
        Calling switch invokes BasicFIND( );
    }
}
```

When a cache hit occurs we save one query to the HLR [a VLR query is involved in both *CacheFIND*( ) and *BasicFIND*( )], and we also save traffic along some of the signalling links; instead of four message transmissions, as in *BasicFIND*( ), only two are needed. In steady-state operation, the cached pointer for any given user is updated only upon a miss.

Note that the *BasicFIND*( ) procedure differs from that specified for roaming subscribers in the IS-41 standard EIA/TIA [6]. In the IS-41 standard, the second line in the *BasicFIND*( ) procedure is omitted, i.e., every call results in a query of the called user's HLR. Thus, in fact, the procedure specified in the standard will result in an even higher network load than the *BasicFIND*( ) procedure specified here. To make a fair assessment of the benefits of *CacheFIND*( ), however, we have compared it against *BasicFIND*( ). Thus, the benefits of *CacheFIND*( ) investigated here depend specifically on the use of caching and not simply on the availability of user location information at the local VLR.

72.8 Caching Threshold Analysis

In this section we investigate the classes of users for which the caching strategy yields net reductions in signalling traffic and database loads. We characterize classes of users by their CMR. The CMR of a user is the average number of calls to a user per unit time, divided by the average number of times the user changes registration areas per unit time. We also define a LCMR, which is the average number of calls to a user from a given originating switch per unit time, divided by the average number of times the user changes registration areas per unit time.

For each user, the amount of savings due to caching is a function of the probability that the cached pointer correctly points to the user's location and increases with the user's LCMR. In this section we quantify the minimum value of LCMR for caching to be worthwhile. This caching threshold is parameterized with respect to costs of traversing signalling network elements and network databases and can be used as a guide to select the subset of users to whom caching should be applied. The analysis in this section shows that estimating user's LCMRs, preferably dynamically, is very important in order to apply the caching strategy. The next section will discuss methods for obtaining this estimate.

From the pseudocode for *BasicFIND*( ), the signalling network cost incurred in locating a PCS user in the event of an incoming call is the sum of the cost of querying the HLR (and receiving the response),

and the cost of querying the VLR which the HLR points to (and receiving the response). Let

α = cost of querying the HLR and receiving a response
β = cost of querying the pointed VLR and receiving a response

Then, the cost of *BasicFIND*( ) operation is

$$C_B = \alpha + \beta \tag{72.2}$$

To quantify this further, assume costs for traversing various network elements as follows.

A_l = cost of transmitting a location request or response message on A link between SSP and LSTP
D = cost of transmitting a location request on response message or D link
A_r = cost of transmitting a location request or response message on A link between RSTP and SCP
L = cost of processing and routing a location request or response message by LSTP
R = cost of processing and routing a location request or response message by RSTP
H_Q = cost of a query to the HLR to obtain the current VLR location
V_Q = cost of a query to the VLR to obtain the routing address

Then, using the PCS reference network architecture

$$\alpha = 2(A_l + D + A_r + L + R) + H_Q \tag{72.3}$$

$$\beta = 2(A_l + D + A_r + L + R) + V_Q \tag{72.4}$$

From Eqs. (72.2)–(72.4)

$$C_B = 4(A_l + D + A_r + L + R) + H_Q + V_Q \tag{72.5}$$

We now calculate the cost of *CacheFIND*( ). We define the *hit ratio* as the relative frequency with which the cached pointer correctly points to the user's location when it is consulted. Let

p = cache hit ratio
C_H = cost of the *CacheFIND*( ) procedure when there is a hit
C_M = cost of the *CacheFIND*( ) procedure when there is a miss

Then the cost of *CacheFIND*( ) is

$$C_C = pC_H + (1-p)C_M \tag{72.6}$$

For *CacheFIND*( ), the signalling network costs incurred in locating a user in the event of an incoming call depend on the hit ratio as well as the cost of querying the VLR, which is stored in the cache; this VLR query may or may not involve traversing the RSTP. In the following, we say a VLR is a *local* VLR if it is served by the same LSTP as the originating switch, and a *remote* VLR otherwise. Let

q = Prob (VLR in originating switch's cache is a local VLR)
δ = cost of querying a local VLR
ϵ = cost of querying a remote VLR
η = cost of updating the cache upon a miss

Then,

$$d = 4A_l + 2L + V_Q \tag{72.7}$$

$$\epsilon = 4(A_l + D + L) + 2R + V_Q \tag{72.8}$$

$$C_H = q\delta + (1-q)\epsilon \tag{72.9}$$

Since updating the cache involves an operation to a fast local memory rather than a database operation, we shall assume in the following that $\eta = 0$. Then,

$$C_M = C_H + C_B = q\delta + (1 - q)\epsilon + \alpha + \beta \tag{72.10}$$

From Eqs. (72.6), (72.9), and (72.10) we have

$$C_C = \alpha + \beta + \epsilon - p(\alpha + \beta) + q(\delta - \epsilon) \tag{72.11}$$

For net cost savings we require $C_C < C_B$, or that the hit ratio exceeds a *hit ratio threshold* p_T, derived using Eqs. (72.6), (72.9), and (72.2),

$$p > p_T = \frac{C_H}{C_B} = \frac{\epsilon + q(\delta - \epsilon)}{\alpha + \beta} \tag{72.12}$$

$$= \frac{4A_l + 4D + 4L + 2R + V_Q - q(4D + 2L + 2R)}{4A_l + 4D + 4A_r + 4L + 4R + H_Q + V_Q} \tag{72.13}$$

Equation (72.13) specifies the hit ratio threshold for a user, evaluated at a given switch, for which local maintenance of a cached location entry produces cost savings. As pointed out earlier, a given user's hit ratio may be location dependent, since the rates of calls destined for that user may vary widely across switches.

The hit ratio threshold in Eq. (72.13) is comprised of heterogeneous cost terms, i.e., transmission link utilization, packet switch processing, and database access costs. Therefore, numerical evaluation of the hit ratio threshold requires either detailed knowledge of these individual quantities or some form of simplifying assumptions. Based on the latter approach, the following two possible methods of evaluation may be employed.

1. Assume one or more cost terms dominate, and simplify Eq. (72.13) by setting the remaining terms to zero.
2. Establish a common unit of measure for all cost terms, for example, *time delay*. In this case, A_l, A_r, and D may represent transmission delays of fixed transmission speed (e.g., 56 kb/s) signalling links, L and R may constitute the sum of queueing and service delays of packet switches (i.e., STPs), and H_Q and V_Q the transaction delays for database queries.

In this section we adopt the first method and evaluate Eq. (72.13) assuming a single term dominates. (In Section 72.9 we present results using the second method.) Table 72.3 shows the hit ratio threshold required to obtain net cost savings, for each case in which one of the cost terms is dominant.

TABLE 72.3 Minimum Hit Ratios and LCMRs for Various Individual Dominant Signalling Network Cost Terms

Dominant Cost Term	Hit ratio Threshold, p_T	LCMR Threshold, $LCMR_T$	LCMR Threshold ($q = 0.043$)	LCMR Threshold ($q = 0.25$)
A_l	1	∞	∞	∞
A_r	0	0	0	0
D	1 - q	1/q - 1	22	3
L	1 - q/2	2/q - 1	45	7
R	1 - q/2	2/q - 1	45	7
H_Q	0	0	0	0
V_Q	1	∞	∞	∞

In Table 72.3 we see that if the cost of querying a VLR or of traversing a local A link is the dominant cost, caching for users who may move is never worthwhile, regardless of users' call reception and mobility patterns. This is because the caching strategy essentially distributes the functionality of the HLR to the VLRs. Thus, the load on the VLR and the local A link is always increased, since any move by a user results in a cache miss. On the other hand, for a fixed user (or telephone), caching is always worthwhile. We also observe that if the remote A links or HLR querying are the bottlenecks, caching is worthwhile even for users with very low hit ratios.

As a simple average-case calculation, consider the net network benefit of caching when HLR access and update is the performance bottleneck. Consider a scenario where $u = 50\%$ of PCS users receive $c = 80\%$ of their calls from $s = 5$ RAs where their hit ratio $p > 0$, and $s' = 4$ of the SSPs at those RAs contain sufficiently large caches. Assume that caching is applied only to this subset of users and to no other users. Suppose that the average hit ratio for these users is $p = 80\%$, so that 80% of the HLR accesses for calls to these users from these RA are avoided. Then the net saving in the accesses to the system's HLR is $H = (ucs'p)/s = 25\%$.

We discuss other quantities in Table 72.3 next. It is first useful to relate the cache hit ratio to users' calling and mobility patterns directly via the LCMR. Doing so requires making assumptions about the distribution of the user's calls and moves. We consider the steady state where the incoming call stream from an SSP to a user is a Poisson process with arrival rate λ, and the time that the user resides in an RA has a general distribution $F(t)$ with mean $1/\mu$. Thus,

$$LCMR = \frac{\lambda}{\mu} \tag{72.14}$$

Let t be the time interval between two consecutive calls from the SSP to the user and t_1 be the time interval between the first call and the time when the user moves to a new RA. From the random observer property of the arrival call stream [7], the hit ratio is

$$= Pr[t < t_1] = \int_{t=0}^{\infty} l e^{-lt} \int_{t_1=t}^{\infty} \mu[1 - F(t_1)]\mathrm{d}t_1 \mathrm{d}t$$

If $F(t)$ is an exponential distribution, then

$$p = \frac{\lambda}{\lambda + \mu} \tag{72.15}$$

and we can derive the *LCMR threshold*, the minimum LCMR required for caching to be beneficial assuming incoming calls are a Poisson process and intermove times are exponentially distributed,

$$LCMR_T = \frac{p_T}{1 - p_T} \tag{72.16}$$

Equation (72.16) is used to derive LCMR thresholds assuming various dominant costs terms, as shown in Table 72.3.

Several values for $LCMR_T$ in Table 72.3 involve the term q, i.e., the probability that the pointed VLR is a local VLR. These values may be numerically evaluated by simplifying assumptions. For example, assume that all of the SSPs in the network are uniformly distributed amongst l LSTPs. Also, assume that all of the PCS subscribers are uniformly distributed in location across all SSPs and that each subscriber exhibits the same incoming call rate at every SSP. Under those conditions, q is simply $1/l$. Consider the case of the public switched telephone network. Given that there are a total of 160 local access transport

area (LATA) across the 7 Regional Bell Operating Company (RBOC) regions [4], the average number of LATAs, or l, is 160/7 or 23. Table 72.3 shows the results with $q = 1/l$ in this case.

We observe that the assumption that all users receive calls uniformly from all switches in the network is extremely conservative. In practice, we expect that user call reception patterns would display significantly more locality, so that q would be larger and the LCMR thresholds required to make caching worthwhile would be smaller. It is also worthwhile to consider the case of a RBOC region with PCS deployed in a few LATA only, a likely initial scenario, say, 4 LATAs. In either case the value of q would be significantly higher; Table 72.3 shows the LCMR threshold when $q = 0.25$.

It is possible to quantify the net costs and benefits of caching in terms of signalling network impacts in this way and to determine the hit ratio and LCMR threshold above which users should have the caching strategy applied. Applying caching to users whose hit ratio and LCMR is below this threshold results in net increases in network impacts. It is, thus, important to estimate users' LCMRs accurately. The next section discusses how to do so.

72.9 Techniques for Estimating Users' LCMR

Here we sketch some methods of estimating users' LCMR. A simple and attractive policy is to not estimate these quantities on a per-user basis at all. For instance, if the average LCMR over all users in a PCS system is high enough (and from Table 72.3, it need not be high depending on which network elements are the dominant costs), then caching could be used at every SSP to yield net system-wide benefits. Alternatively, if it is known that at any given SSP the average LCMR over all users is high enough, a cache can be installed at that SSP. Other variations can be designed.

One possibility for deciding about caching on a per-user basis is to maintain information about a user's calling and mobility pattern at the HLR and to download it periodically to selected SSPs during off-peak hours. It is easy to envision numerous variations on this idea.

In this section we investigate two possible techniques for estimating LCMR on a per-user basis when caching is to be deployed. The first algorithm, called the *running average* algorithm, simply maintains a running average of the hit ratio for each user. The second algorithm, called the *reset-K* algorithm, attempts to obtain a measure of the hit ratio over the recent history of the user's movements. We describe the two algorithms next and evaluate their effectiveness using a stochastic analysis taking into account user calling and mobility patterns.

The Running Average Algorithm

The running average algorithm maintains, for every user that has a cache entry, the running average of the hit ratio. A running count is kept of the number of calls to a given user, and, regardless of the *FIND* procedure used to locate the user, a running count of the number of times that the user was at the same location for any two consecutive calls; the ratio of these numbers provides the measured running average of the hit ratio. We denote the measured running average of the hit ratio by p_M; in steady state, we expect that $p_M = p$. The user's previous location as stored in the cache entry is used only if the running average of the hit ratio p_M is greater than the cache hit threshold p_T. Recall that the cache scheme outperforms the basic scheme if $p > p_T = C_H/C_B$. Thus, in steady state, the running average algorithm will outperform the basic scheme when $p_M > p_T$.

We consider, as before, the steady state where the incoming call stream from an SSP to a user is a Poisson process with arrival rate λ, and the time that the user resides in an RA has an exponential distribution with mean $1/\mu$. Thus $LCMR = \lambda/\mu$ [Eq. (72.14)] and the location tracking cost at steady state is

$$C_C = \begin{cases} p_M C_H + (1 - p_M) C_B, & p_M > p_T \\ C_B, & \text{otherwise} \end{cases} \tag{72.17}$$

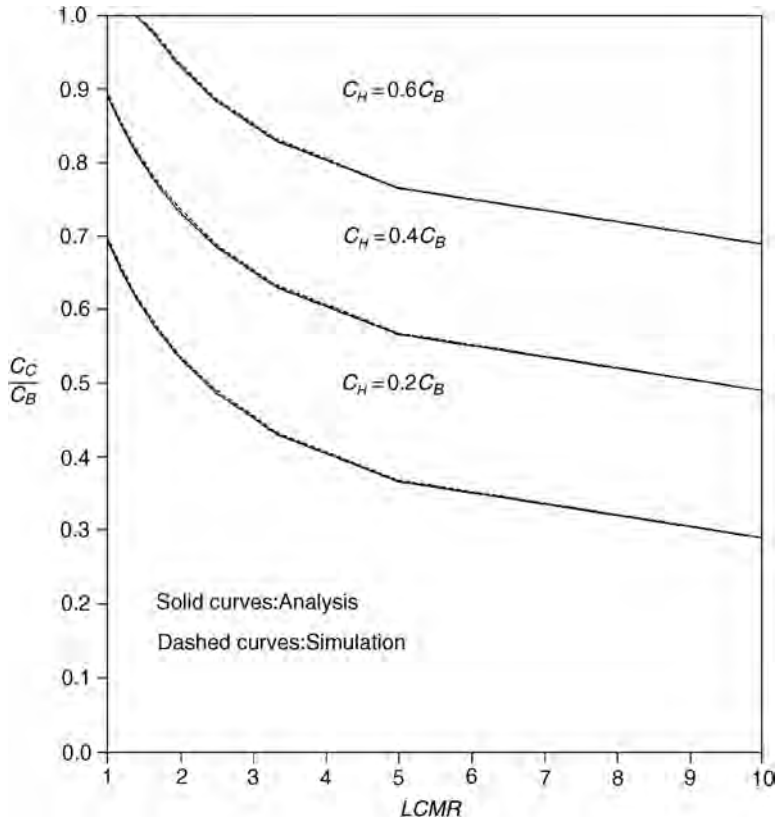


FIGURE 72.6 The location tracking cost for the running average algorithm.

Figure 72.6 plots the cost ratio C_C/C_B from Eq. (72.17) against *LCMR*. (This corresponds to assigning uniform units to all cost terms in Eq. (72.13), i.e., the second evaluation method as discussed in Section 72.8. Thus, the ratio C_C/C_B may represent the percentage reduction in user location time with the caching strategy compared to the basic strategy.) The figure indicates that in the steady state, the caching strategy with the running average algorithm for estimating LCMR can significantly outperform the basic scheme if *LCMR* is sufficiently large. For instance with *LCMR* ~ 5, caching can lead to cost savings of 20–60% over the basic strategy.

Equation (72.17) (cf., solid curves in Fig. 72.6) is validated against a simple Monte Carlo simulation (cf., dashed curves in Fig. 72.6). In the simulation, the confidence interval for the 95% confidence level of the output measure C_C/C_B is within 3% of the mean value. This simulation model will later be used to study the running average algorithm when the mean of the movement distribution changes from time to time [which cannot be modeled by using Eq. (72.17)].

One problem with the running average algorithm is that the parameter p is measured from the entire past history of the user's movement, and the algorithm may not be sufficiently dynamic to adequately reflect the recent history of the user's mobility patterns.

The Reset-*K* Algorithm

We may modify the running average algorithm such that p is measured from the recent history. Define every K incoming calls as a *cycle*. The modified algorithm, which is referred to as the reset-K algorithm, counts the number of cache hits n in a cycle. If the measured hit ratio for a user, $p_M = n/K \geq p_T$, then the cache is enabled for that user, and the cached information is always used to locate the user in the

next cycle. Otherwise, the cache is disabled for that user, and the basic scheme is used. At the beginning of a cycle, the cache hit count is reset, and a new p_M value is measured during the cycle.

To study the performance of the reset-K algorithm, we model the number of cache misses in a cycle by a Markov process. Assume as before that the call arrivals are a Poisson process with arrival rate λ and the time period the user resides in an RA has an exponential distribution with mean $1/\mu$. A pair (i, j), where $i > j$, represents the state that there are j cache misses before the first i incoming phone calls in a cycle. A pair $(i, j)^*$, where $i \geq j \geq 1$, represents the state that there are $j - 1$ cache misses before the first i incoming phone calls in a cycle, and the user moves between the ith and the $i + 1$ phone calls. The difference between (i, j) and $(i, j)^*$ is that if the Markov process is in the state (i, j) and the user moves, then the process moves into the state $(i, j + 1)^*$. On the other hand, if the process is in state $(i, j)^*$ when the user moves, the process remains in $(i, j)^*$ because at most one cache miss occurs between two consecutive phone calls.

Figure 72.7(a) illustrates the transitions for state $(i, 0)$ where $2 < i < K + 1$. The Markov process moves from $(i - 1, 0)$ to $(i, 0)$ if a phone call arrives before the user moves out. The rate is λ. The process moves from $(i, 0)$ to $(i, 1)^*$ if the user moves to another RA before the $i + 1$ call arrival. Let $\pi(i, j)$ denote the probability of the process being in state (i, j). Then the transition equation is

$$\pi(i,0) = \frac{\lambda}{\lambda+\mu}\pi(i-1,0), \quad 2 < i < K+1 \tag{72.18}$$

Figure 72.7(b) illustrates the transitions for state $(i, i - 1)$ where $1 < i < K + 1$. The only transition into the state $(i, i - 1)$ is from $(i - 1, i - 1)^*$, which means that the user always moves to another RA after a phone call. [Note that there can be no state $(i - 1, i - 1)$ by definition and, hence, no transition from such a state.] The transition rate is λ. The process moves from $(i, i - 1)$ to $(i, i)^*$ with rate μ, and moves to $(i + 1, i - 1)$ with rate λ. Let $\pi^*(i, j)$ denote the probability of the process being in state $(i, j)^*$. Then the transition equation is

$$\pi(i,i-1) = \frac{\lambda}{\lambda+\mu}\pi^*(i-1,i-1), \quad 1 < i < K+1 \tag{72.19}$$

Figure 72.7(c) illustrates the transitions for state (i, j) where $2 < i < K + 1$, $0 < j < i - 1$. The process may move into state (i, j) from two states $(i - 1, j)$ and $(i - 1, j)^*$ with rate λ, respectively. The process moves from (i, j) to $(i, j + 1)^*$ or $(i + 1, j)$ with rates μ and λ, respectively. The transition equation is

$$\pi(i,j) = \frac{\lambda}{\lambda+\mu}[\pi(i-1,j)+\pi^*(i-1,j)], \quad 2 < i < K+1, \quad 0 < j < i-1 \tag{72.20}$$

Figure 72.7(d) illustrates the transitions for state $(K + 1, j)$ where $0 < j < K + 1$. Note that if a phone call arrives when the process is in (K, j) or $(K, j)^*$, the system enters a new cycle (with rate λ), and we could represent the new state as $(1, 0)$. In our model, we introduce the state $(K + 1, j)$ instead of $(1, 0)$, where

$$\sum_{0 \leq j \leq K} \pi(K+1,j) = \pi(1,0)$$

so that the hit ratio, and thus the location tracking cost, can be derived [see Eq. (72.25)]. The process moves from $(K + 1, j)$ [i.e., $(1, 0)$] to $(1, 1)^*$ with rate μ if the user moves before the next call arrives. Otherwise, the process moves to $(2, 0)$ with rate λ. The transition equation is

$$\pi(K+1,j) = \frac{\lambda}{\lambda+\mu}[\pi(K,j)+\pi^*(K,j)], \quad 0 < j < K+1 \tag{72.21}$$

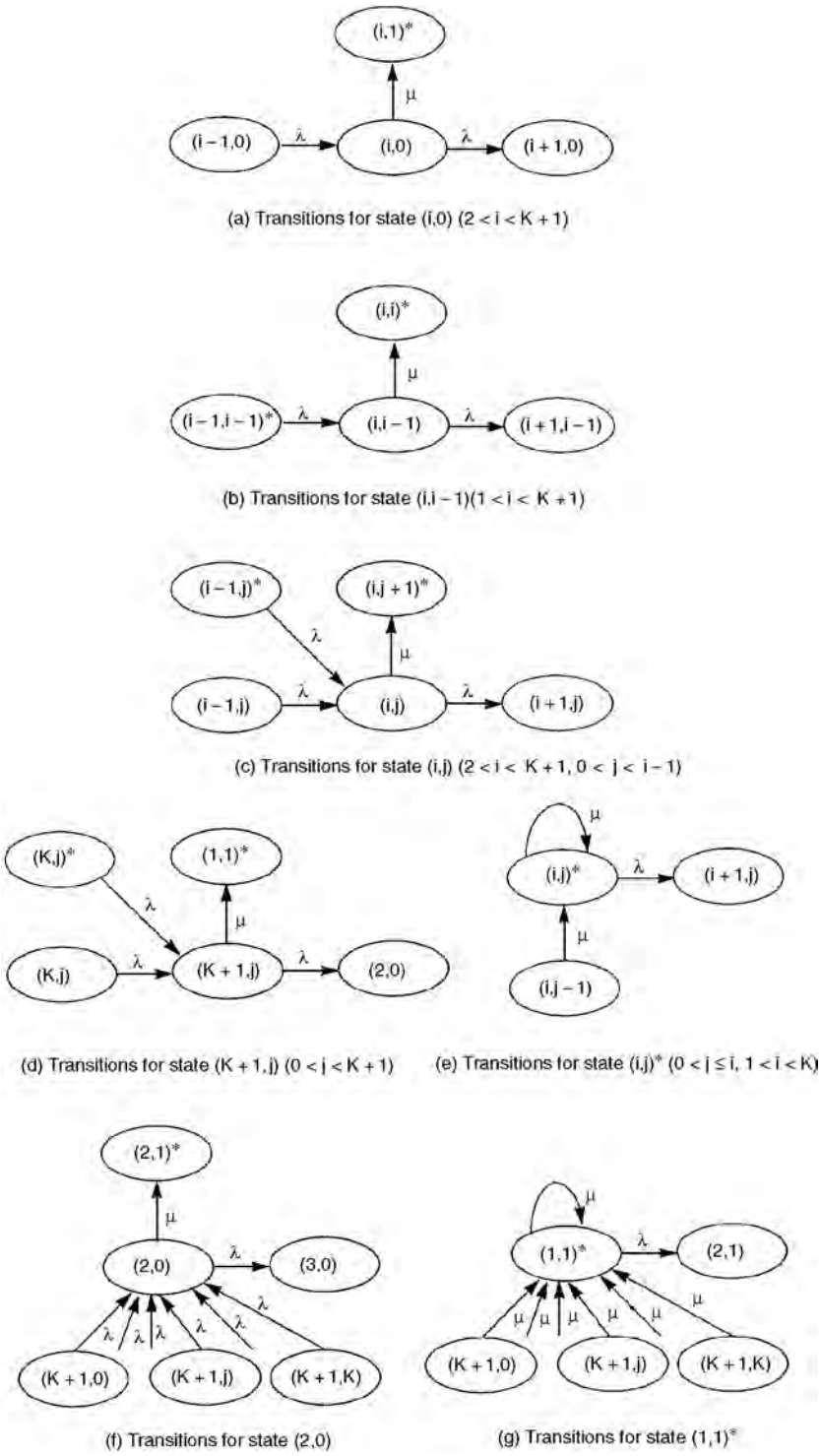


FIGURE 72.7 State transitions.

For $j = 0$, the transition from $(K, j)^*$ to $(K + 1, 0)$ should be removed in Fig. 72.7(d) because the state $(K, 0)^*$ does not exist. The transition equation for $(K + 1, 0)$ is given in Eq. (72.18). Figure 72.7(e) illustrates the transitions for state $(i, j)^*$ where $0 < j < i$, $1 < i < K + 1$. The process can only move to $(i, j)^*$ from $(i, j - 1)$ (with rate μ). From the definition of $(i, j)^*$, if the user moves when the process is in $(i, j)^*$, the process remains in $(i, j)^*$ (with rate μ). Otherwise, the process moves to $(i + 1, j)$ with rate λ. The transition equation is

$$\pi^*(i, j) = \frac{\mu}{\lambda}\pi(i, j-1), \qquad 0 < j \le i, \quad 1 < i < K+1, \quad i \ge 2 \tag{72.22}$$

The transitions for $(2, 0)$ are similar to the transitions for $(i, 0)$ except that the transition from $(1, 0)$ is replaced by $(K + 1, 0),\ldots,(K + 1, K)$ [cf., Fig. 72.7(f)]. The transition equation is

$$\pi(2, 0) = \frac{\lambda}{\lambda + \mu}\left[\sum_{0 \le j \le K} \pi(K+1, j)\right] \tag{72.23}$$

Finally, the transitions for $(1, 1)^*$ is similar to the transitions for $(i, j)^*$ except that the transition from $(1, 0)$ is replaced by $(K + 1, 0),\ldots,(K + 1, K)$ [cf., Fig. 72.7(g)]. The transition equation is

$$\pi^*(1, 1) = \frac{\mu}{\lambda}\left[\sum_{0 \le j \le K} \pi(K+1, j)\right] \tag{72.24}$$

Suppose that at the beginning of a cycle, the process is in state $(K + 1, j)$, then it implies that there are j cache misses in the previous cycle. The cache is enabled if and only if

$$p_M \ge p_T = \frac{C_H}{C_B} \Rightarrow 1 - \frac{j}{K} \ge \frac{C_H}{C_B} \Rightarrow 0 \le j \le \left[K\left(1 - \frac{C_H}{C_B}\right)\right]$$

Thus, the probability that the measured hit ratio $p_M < p_T$ in the previous cycle is

$$Pr[p_M < p_T] = \frac{\sum_{\lceil k[1-(C_H/C_B)]\rceil < j \le K} \pi(K+1, j)}{\sum_{0 \le j \le K} \pi(K+1, j)}$$

and the location tracking cost for the reset-K algorithm is

$$C_C = C_B Pr[p_M < p_T] + (1 - Pr[p_M < p_T])\left\{\sum_{0 \le j \le K}\left(\frac{(K-j)C_H}{K} + \frac{j(C_H + C_B)}{K}\right)\left[\frac{\pi(K+1, j)}{\sum_{0 \le i \le K}\pi(K+1, i)}\right]\right\} \tag{72.25}$$

The first term Eq. (72.25) represents the cost incurred when caching is disabled because the hit ratio threshold exceeds the hit ratio measured in the previous cycle. The second term is the cost when the cache is enabled and consists of two parts, corresponding to calls during which hits occur and calls during which misses occur. The ratio in square brackets is the conditional probability of being in state $\pi(K + 1, j)$ during the current cycle.

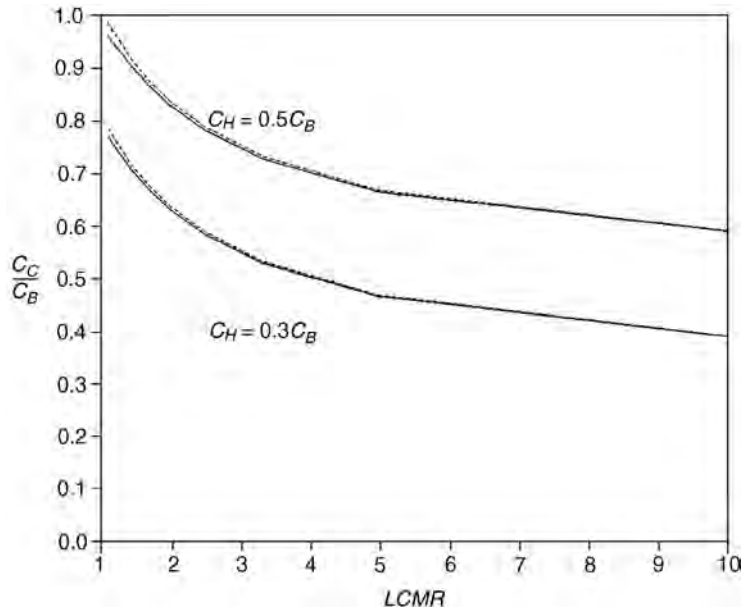


FIGURE 72.8 The location tracking costs for the reset-K algorithm; $K = 20$.

The numerical computation of $\pi(K+1, j)$ can be done as follows. First, compute $a_{i,j}$ and $b_{i,j}$ where $\pi(i,j) = a_{i,j}\pi^*(1,1)$ and $\pi^*(i,j) = b_{i,j}\pi^*(1,1)$. Note that $a_{i,j} = 0 (b_{i,j} = 0)$ if $\pi(i,j)[\pi^*(i,j)]$ is not defined in Eqs. (72.18)–(72.24). Since

$$\sum_{i,j}[\pi(i,j) + \pi^*(i,j)] = 1$$

we have

$$\pi^*(1,1) = \frac{1}{\sum_{i,j}(a_{i,j} + b_{i,j})}$$

and $\pi(K+1, j)$ can be computed and the location tracking cost for the reset-K algorithm is obtained using Eq. (72.25).

The analysis is validated by a Monte Carlo simulation. In the simulation, the confidence interval for the 98% confidence level of the output measure C_C/C_B is within 3% of the mean value. Figure 72.8 plots curves for Eq. (72.25) (the solid curves) against the simulation experiments (the dashed curves) for $K = 20$ and $C_H = 0.5C_B$ and $0.3C_B$, respectively. The figure indicates that the analysis is consistent with the simulation model.

Comparison of the LCMR Estimation Algorithms

If the distributions for the incoming call process and the user movement process never change, then we would expect the running average algorithm to outperform the reset-K algorithm (especially when K is small) because the measured hit ratio p_M in the running average algorithm approaches the true hit ratio value p in the steady state. Surprisingly, the performance for the reset-K algorithm is roughly the same as the running average algorithm even if K is as small as 10. Figure 72.9 plots the location tracking costs for the running average algorithm and the reset-K algorithm with different K values.

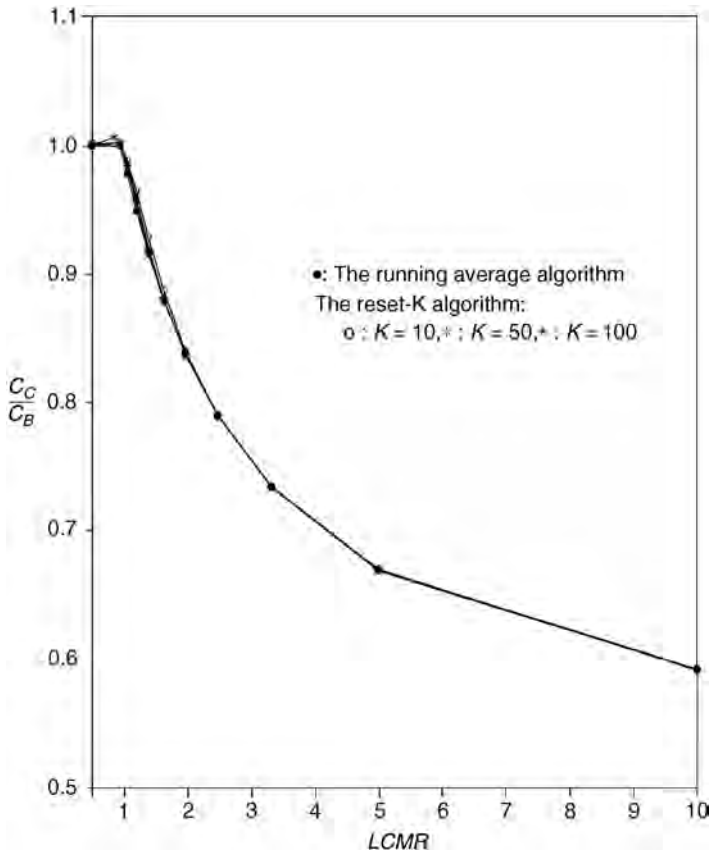


FIGURE 72.9 The location tracking costs for the running average algorithm and the reset-K algorithm; $C_H = 0.5C_B$.

The figure indicates that in steady state, when the distributions for the incoming call process and the user movement process never change, the running average algorithm outperforms reset K, and a large value of K outperforms a small K but the differences are insignificant.

If the distributions for the incoming call process or the user movement process change from time to time, we expect that the reset-K algorithm outperforms the running average algorithm. We have examined this proposition experimentally. In the experiments, 4000 incoming calls are simulated. The call arrival rate changes from 0.1 to 1.0, 0.3, and then 5.0 for every 1000 calls (other sequences have been tested and similar results are observed). For every data point, the simulation is repeated 1000 times to ensure that the confidence interval for the 98% confidence level of the output measure C_C/C_B is within 3% of the mean value. Figure 72.10 plots the location tracking costs for the two algorithms for these experiments. By changing the distributions of the incoming call process, we observe that the reset-K algorithm is better than the running average algorithm for all C_H/C_B values.

72.10 Discussion

In this section we discuss aspects of the caching strategy presented here. Caching in PCS systems raises a number of issues not encountered in traditional computer systems, particularly with respect to architecture and locality in user call and mobility patterns. In addition, several variations in our reference assumptions are possible for investigating the implementation of the caching strategies. Here we sketch some of the issues involved.

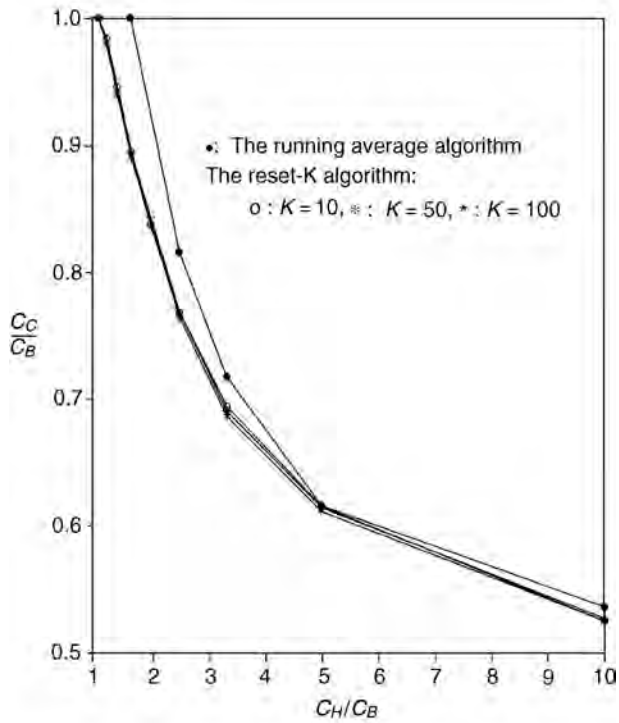


FIGURE 72.10 Comparing the running average algorithm and the reset-*K* algorithm under unstable call traffic.

Conditions When Caching Is Beneficial

We summarize the conditions for which the auxiliary strategies are worthwhile, under the assumptions of our analysis.

The caching strategy is very promising when the HLR update (or query load) or the remote A link is the performance bottleneck, since a low *LCMR*(*LCMR* > O) is required. For caching, the total database load and signalling network traffic is reduced whenever there is a cache hit. In addition, load and traffic is redistributed from the HLR and higher level SS7 network elements (RSTP, D links) to the VLRs and lower levels where excess network capacity may be more likely to exist. If the VLR is the performance bottleneck, the caching strategy is not promising, unless the VLR capacity is upgraded.

The benefits of the caching strategy depend on user call and mobility patterns when the D link, RSTP, and LSTP are the performance bottlenecks. We have used a Poisson call arrival model and exponential intermove time to estimate this dependence. Under very conservative assumptions, for caching to be beneficial requires relatively high *LCMR* (25–50); we expect that in practice this threshold could be lowered significantly (say, *LCMR* > 7). Further experimental study is required to estimate the amount of locality in user movements for different user populations to investigate this issue further. It is possible that for some classes of users data obtained from active badge location system studies (e.g., Fishman and Mazer [8]) could be useful. In general, it appears that caching could also potentially provide benefits to some classes of users even when the D link, the RSTP, or the LSTP are the bottlenecks.

We observe that more accurate models of user calling and mobility patterns are required to help resolve the issues raised in this section. We are currently engaged in developing theoretical models for user mobility and estimating their effect on studies of various aspects of PCS performance [10].

Alternative Network Architectures

The reference architecture we have assumed (Fig. 72.2) is only one of several possible architectures. It is possible to consider variations in the placement of the HLR and VLR functionality, (e.g., placing the VLR at a local SCP associated with the LSTP instead of at the SSP), the number of SSPs served by an LSTP, the number of HLRs deployed, etc. It is quite conceivable that different regional PCS service providers and telecommunications companies will deploy different signalling network architectures, as well as placement of databases for supporting PCS within their serving regions [19]. It is also possible that the number and placement of databases in a network will change over time as the number of PCS users increases.

Rather than consider many possible variations of the architecture, we have selected a reference architecture to illustrate the new auxiliary strategies and our method of calculating their costs and benefits. Changes in the architecture may result in minor variations in our analysis but may not significantly affect our qualitative conclusions.

LCMR Estimation and Caching Policy

It is possible that for some user populations estimating the LCMR may not be necessary, since they display a relatively high-average LCMR. For some populations, as we have shown in Section 72.9, obtaining accurate estimates of user LCMR in order to decide whether or not to use caching can be important in determining the net benefits of caching.

We have presented two simple distributed algorithms for estimating LCMR, based on a long-range and short-range running calculation; the former is preferable if the call and mobility pattern of users is fairly consistent. Tuning the amount of history that is used to determine whether caching should be employed for a particular user is an obvious area for further study but is outside the scope of this chapter.

An alternative approach is to utilize some user-supplied information, by requesting profiles of user movements (e.g., see Tabbane [22] and to integrate this with the caching strategy. A variation of this approach is to use some domain knowledge about user populations and their characteristics.

A related issue is that of cache size and management. In practice it is likely that the monetary cost of deploying a cache may limit its size. In that case, cache entries may not be maintained for some users; selecting these users carefully is important to maximize the benefits of caching. Note that the cache hit ratio threshold cannot necessarily be used to determine which users have cache entries, since it may be useful to maintain cache entries for some users even though their hit ratios have temporarily fallen below the threshold. A simple policy that has been found to be effective in computer systems in the least recently used (LRU) policy [20] in which cache entries that have been least recently used are discarded; LRU may offer some guidance in this context.

72.11 Conclusions

We began this chapter with an overview of the nuances of PCS, such as personal and terminal mobility, registration, deregistration, call delivery, etc. A tutorial was then provided on the two most common strategies for locating users in PCS, in North American interim standard IS-41 and the Pan-European standard GSM. A simplified analysis of the two standards was then provided to show the reader the extent to which database and signalling traffic is likely to be generated by PCS services. Suggestions were then made that are likely to result in reduced traffic.

Previous studies [12,13,14,16] of PCS-related network signalling and data management functionalities suggest a high level of utilization of the signalling network in supporting call and mobility management activities for PCS systems. Motivated by the need to evolve location strategies to reduce signalling and database loads, we then presented an auxiliary strategy, called per-user caching, to augment the basic user location strategy proposed in standards [6,18].

Using a reference system architecture for PCS, we quantified the criteria under which the caching strategy produces reductions in the network signalling and database loads in terms of users' LCMRs. We have shown that, if the HLR or the remote A link in an SS7 architecture is the performance bottleneck, caching is useful regardless of user call and mobility patterns. If the D link or STPs are the performance bottlenecks, caching is potentially beneficial for large classes of users, particularly if they display a degree of locality in their call reception patterns. Depending on the numbers of PCS users who meet these criteria, the system-wide impacts of these strategies could be significant. For instance, for users with $LCMR \sim 5$ and stable call and move patterns, caching can result in cost reduction of 20–60% over the basic user location strategy *BasicFIND*( ) under our analysis. Our results are conservative in that the *BasicFIND*( ) procedure we have used for comparison purposes already reduces the network impacts compared to the user location strategy specified in PCS standards such as IS-41.

We have also investigated in detail two simple on-line algorithms for estimating users' LCMRs and examined the call and mobility patterns for which each would be useful. The algorithms allow a system designer to tune the amount of history used to estimate a users' LCMR and, hence, to attempt to optimize the benefits due to caching. The particular values of cache hit ratios and LCMR thresholds will change with variations in the way the PCS architecture and the caching strategy is implemented, but our general approach can still be applied. There are several issues deserving further study with respect to deployment of the caching strategy, such as the effect of alternative PCS architectures, integration with other auxiliary strategies such as the use of user profiles, and effective cache management policies.

Recently, we have augmented the work reported in this paper by a simulation study in which we have compared the caching and basic user location strategies [9]. The effect of using a time-based criterion for enabling use of the cache has also been considered [11]. We have proposed elsewhere, for users with low CMRs, an auxiliary strategy involving a system of forwarding pointers to reduce the signalling traffic and database loads [10], a description of which is beyond the scope of this chapter.

Acknowledgment

We acknowledge a number of our colleagues in Bellcore who have reviewed several previous papers by the authors and contributed to improving the clarity and readability of this work.

References

1. Awerbuch, B. and Peleg, D., Concurrent online tracking of mobile users. In *Proc. SIGCOMM Symp. Comm. Arch. Prot.,* Oct. 1991.
2. Bellcore., Advanced intelligent network release 1 network and operations plan, Issue 1. Tech. Rept. SR-NPL-001623. Bell Communications Research, Morristown, NJ, Jun. 1991.
3. Bellcore., Personal communications services (PCS) network access services to PCS providers, Special Report SR-TSV-002459, Bell Communications Research, Morristown, NJ, Oct. 1993.
4. Bellcore., Switching system requirements for interexchange carrier interconnection using the integrated services digital network user part (ISDNUP). Tech. Ref. TR-NWT-000394. Bell Communications Research. Morristown, NJ, Dec. 1992.
5. Berman, R.K. and Brewster, J.H., Perspectives on the AIN architecture. *IEEE Comm. Mag.,* 1(2), 27–32, 1992.
6. Electronic Industries Association/Telecommunications Industry Association., Cellular radio telecommunications intersystem operations. Tech. Rept. IS-41. Rev. B. Jul. 1991.
7. Feller, W., *An Introduction to Probability Theory and Its Applications.* John Wiley & Sons, New York, 1966.
8. Fishman, N. and Mazer, M., Experience in deploying an active badge system. In *Proc. Globecom Workshop on Networking for Pers. Comm. Appl.,* Dec. 1992.
9. Harjono, H., Jain, R., and Mohan, S., Analysis and simulation of a cache-based auxiliary location strategy for PCS. In *Proc. IEEE Conf. Networks Pers. Comm.,* 1994.

10. Jain, R. and Lin Y.-B., An auxiliary user location strategy employing forwarding pointers to reduce network impacts of PCS. *ACM Journal on Wireless Info. Networks (WINET),* 1(2), 1995.
11. Lin, Y.-B., Determining the user locations for personal communications networks. *IEEE Trans. Vehic. Tech.,* 466–473, Aug. 1994.
12. Lo, C., Mohan, S., and Wolff, R., A comparison of data management alternatives for personal communications applications. Second Bellcore Symposium on Performance Modeling, SR-TSV-002424, Bell Communications Research, Morristown, NJ, Nov. 1993.
13. Lo, C.N., Wolff, R.S., and Bernhardt, R.C., An estimate of network database transaction volume to support personal communications services. In *Proc. Intl. Conf. Univ. Pers. Comm.,* 1992.
14. Lo, C. and Wolff, R., Estimated network database transaction volume to support wireless personal data communications applications. In *Proc. Intl. Conf. Comm.,* May 1993.
15. Lycksell, E., GSM system overview. Tech. Rept. Swedish Telecom. Admin., Jan. 1991.
16. Meier-Hellstern, K. and Alonso, E., The use of SS7 and GSM to support high density personal communications. In *Proc. Intl. Conf. Comm.,* 1992.
17. Mohan, S. and Jain, R., Two user location strategies for PCS. *IEEE Pers. Comm. Mag.,* Premiere issue. 42–50, Feb. 1994.
18. Mouly, M. and Pautet, M.B., *The GSM System for Mobile Communications.* M. Mouly, 49 rue Louise Bruneau, Palaiseau, France, 1992.
19. Russo, P., Bechard, K., Brooks, E., Corn, R.L., Honig, W.L., Gove, R., and Young, J., In rollout in the United States. *IEEE Comm. Mag.,* 56–63, Mar. 1993.
20. Silberschatz, A. and Peterson, J., *Operating Systems Concepts.* Addison-Wesley, Reading, MA, 1988.
21. Tabbane, S., Comparison between the alternative location strategy (AS) and the classical location strategy (CS). Tech. Rept. Rutgers Univ. WINLAB. Rutgers, NJ, Jul. 1992.
22. Tabbane, S., Evaluation of an alternative location strategy for future high density wireless communications systems. Tech. Rept. WINALAB-TR-51, Rutgers Univ. WINLAB. Rutgers, NJ, Jan. 1993.
23. Thomas, R., Gilbert, H., and Mazziotto, G., Influence of the mobile station on the performance of a radio mobile cellular network. In *Proc. 3rd Nordic Seminar.* Sep. 1988.

73

Cell Design Principles

Michel Daoud Yacoub
FEEC/UNICAMP

73.1 Introduction

Designing a cellular network is a challenging task that invites engineers to exercise all of their knowledge in telecommunications. Although it may not be necessary to work as an expert in all of the fields, the interrelationship among the areas involved impels the designer to naturally search for a deeper understanding of the main phenomena. In other words, the time for segregation, when radio engineers and traffic engineers would not talk to each other, at least through a common vocabulary, is probably gone.

A great many aspects must be considered in a cellular network planning. The main ones include the following.

Radio Propagation: Here the topography and the morphology of the terrain, the urbanization factor and the clutter factor of the city, and some other aspects of the target geographical region under investigation will constitute the input data for the radio coverage design.

Frequency Regulation and Planning: In most countries there is a centralized organization, usually performed by a government entity, regulating the assignment and use of the radio spectrum. The frequency planning within the assigned spectrum should then be made so that interferences are minimized and the traffic demand is satisfied.

Modulation: As far as analog systems are concerned, the narrowband FM is widely used due to its remarkable performance in the presence of fading. The North American Digital Cellular Standard IS-54 proposes the $\pi/4$ differential quadrature phase-shift keying ($\pi/4$ DQPSK) modulation, whereas the Global Standard for Mobile Communications (GSM) establishes the use of the Gaussian minimum-shift keying (GMSK).

Antenna Design: To cover large areas and for low-traffic applications omnidirectional antennas are recommended. Some systems at their inception may have these characteristics, and the utilization of

omnidirectional antennas certainly keeps the initial investment low. As the traffic demand increases, the use of some sort of capacity enhancement technique to meet the demand, such as replacing the omnidirectional by directional antennas, is mandatory.

Transmission Planning: The structure of the channels, both for signalling and voice, is one of the aspects to be considered in this topic. Other aspects include the performance of the transmission components (power capacity, noise, bandwidth, stability, etc.) and the design or specification of transmitters and receivers.

Switching Exchange: In most cases this consists of adapting the existing switching network for mobile radio communications purposes.

Teletraffic: For a given grade of service and number of channels available, how many subscribers can be accommodated into the system? What is the proportion of voice and signalling channels?

Software Design: With the use of microprocessors throughout the system there are software applications in the mobile unit, in the base station, and in the switching exchange.

Other aspects, such as human factors, economics, etc., will also influence the design.

This chapter outlines the aspects involving the basic design steps in cellular network planning. Topics, such as traffic engineering, cell coverage, and interference, will be covered, and application examples will be given throughout the section so as to illustrate the main ideas. We start by recalling the basic concepts including *cellular principles, performance measures and system requirements,* and *system expansion techniques*.

73.2 Cellular Principles

The basic idea of the cellular concept is *frequency reuse* in which the same set of channels can be reused in different geographical locations sufficiently apart from each other so that *cochannel interference* be within tolerable limits. The set of channels available in the system is assigned to a group of *cells* constituting the *cluster*. Cells are assumed to have a *regular hexagonal* shape and the number of cells per cluster determines the *repeat pattern*. Because of the hexagonal geometry only certain repeat patterns can tessellate. The number N of cells per cluster is given by

$$N = i^2 + ij + j^2 \tag{73.1}$$

where i and j are integers. From Eq. (73.1) we note that the clusters can accommodate only certain numbers of cells such as 1, 3, 4, 7, 9, 12, 13, 16, 19, 21, … , the most common being 4 and 7. The number of cells per cluster is intuitively related with system capacity as well as with transmission quality. The fewer cells per cluster, the larger the number of channels per cell (higher traffic carrying capacity) and the closer the cocells (potentially more cochannel interference). An important parameter of a cellular layout relating these entities is the D/R ratio, where D is the distance between cocells and R is the cell radius. In a hexagonal geometry it is found that

$$D/R = \sqrt{3N} \tag{73.2}$$

73.3 Performance Measures and System Requirements

Two parameters are intimately related with the grade of service of the cellular systems: carrier-to-cochannel interference ratio and blocking probability.

A high carrier-to-cochannel interference ratio in connection with a low-blocking probability is the desirable situation. This can be accomplished, for instance, in a large cluster with a low-traffic condition. In such a case the required grade of service can be achieved, although the resources may not be efficiently utilized. Therefore, a measure of efficiency is of interest. The **spectrum efficiency** η_s expressed in erlang per square meter per hertz, yields a measure of how efficiently space, frequency, and time are used, and

it is given by

$$\eta_s = \frac{\text{number of reuses}}{\text{coverage area}} \times \frac{\text{number of channels}}{\text{bandwidth available}} \times \frac{\text{time the channel is busy}}{\text{total time of the channel}}$$

Another measure of interest is the **trunking efficiency** in which the number of subscribers per channel is obtained as a function of the number of channels per cell for different values of blocking probability. As an example, assume that a cell operates with 40 channels and that the mean blocking probability is required to be 5%. Using the erlang-B formula (refer to the Traffic Engineering section of this chapter), the traffic offered is calculated as 34.6 erlang. If the traffic per subscriber is assumed to be 0.02 erl, a total of 34.6/0.02 = 1730 subscribers in the cell is found. In other words, the trunking efficiency is 1730/40 = 43.25 subscribers per channel in a 40-channel cell. Simple calculations show that the trunking efficiency decreases rapidly when the number of channels per cell falls below 20.

The basic specifications require cellular services to be offered with a fixed telephone network quality. Blocking probability should be kept below 2%. As for the transmission aspect, the aim is to provide good quality service for 90% of the time. Transmission quality concerns the following parameters:

- Signal-to-cochannel interference (S/I_c) ratio
- Carrier-to-cochannel interference ratio (C/I_c)
- Signal plus noise plus distortion-to-noise plus distortion (*SINAD*) ratio
- Signal-to-noise (S/N) ratio
- Adjacent channel interference selectivity (*ACS*)

The S/I_c is a subjective measure, usually taken to be around 17 dB. The corresponding C/I_c depends on the modulation scheme. For instance, this is around 8 dB for 25-kHz FM, 12 dB for 12.5-kHz FM, and 7 dB for GMSK, but the requirements may vary from system to system. A common figure for *SINAD* is 12 dB for 25-kHz FM. The minimum S/N requirement is 18 dB, whereas *ACS* is specified to be no less than 70 dB.

73.4 System Expansion Techniques

The obvious and most common way of permitting more subscribers into the network is by allowing a system performance degradation but within acceptable levels. The question is how to objectively define what is acceptable. In general, the subscribers are more likely to tolerate a poor quality service rather than not having the service at all. Some alternative expansion techniques, however, do exist that can be applied to increase the system capacity. The most widely known are as follows.

Adding New Channels: In general, when the system is set up not all of the channels need be used, and growth and expansion can be planned in an orderly manner by utilizing the channels that are still available.

Frequency Borrowing: If some cells become more overloaded than others, it may be possible to reallocate channels by transferring frequencies so that the traffic demand can be accommodated.

Change of Cell Pattern: Smaller clusters can be used to allow more channels to attend a bigger traffic demand at the expense of a degradation of the transmission quality.

Cell Splitting: By reducing the size of the cells, more cells per area, and consequently more channels per area, are used with a consequent increase in traffic capacity. A radius reduction by a factor of f reduces the coverage area and increases the number of base stations by a factor of f^2. Cell splitting usually takes place at the midpoint of the congested areas and is so planned in order that the old base stations are kept.

Sectorization: A cell is divided into a number of sectors, three and six being the most common arrangements, each of which is served by a different set of channels and illuminated by a directional antenna. The sector, therefore, can be considered as a new cell. The base stations can be located either at the center or at the corner of the cell. The cells in the first case are referred to as center-excited cells and in the

second as corner-excited cells. Directional antennas cut down the cochannel interference, allowing the cocells to be more closely spaced. Closer cell spacing implies smaller D/R ratio, corresponding to smaller clusters, i.e., higher capacity.

Channel Allocation Algorithms: The efficient use of channels determines the good performance of the system and can be obtained by different channel assignment techniques. The most widely used algorithm is based on fixed allocation. Dynamic allocation strategies may give better performance but are very dependent on the traffic profile and are usually difficult to implement.

73.5 Basic Design Steps

Engineering a cellular system to meet the required objectives is not a straightforward task. It demands a great deal of information, such as market demographics, area to be served, traffic offered, and other data not usually available in the earlier stages of system design. As the network evolves, additional statistics will help the system performance assessment and replanning. The main steps in a cellular system design are as follows.

Definition of the Service Area: In general, the responsibility for this step of the project lies on the operating companies and constitutes a tricky task, because it depends on the market demographics and, consequently, on how much the company is willing to invest.

Definition of the Traffic Profile: As before, this step depends on the market demographics and is estimated by taking into account the number of potential subscribers within the service area.

Choice of Reuse Pattern: Given the traffic distribution and the interference requirements a choice of the reuse pattern is carried out.

Location of the Base Stations: The location of the first base station constitutes an important step. A significant parameter to be taken into account in this is the relevance of the region to be served. The base station location is chosen so as to be at the center of or as close as possible to the target region. Data, such as available infrastructure and land, as well as local regulations are taken into consideration in this step. The cell radius is defined as a function of the traffic distribution. In urban areas, where the traffic is more heavily concentrated, smaller cells are chosen so as to attend the demand with the available channels. In suburban and in rural areas, the radius is chosen to be large because the traffic demand tends to be small. Once the placement of the first base station has been defined, the others will be accommodated in accordance with the repeat pattern chosen.

Radio Coverage Prediction: Given the topography and the morphology of the terrain, a radio prediction algorithm, implemented in the computer, can be used to predict the signal strength in the geographic region. An alternative to this relies on field measurements with the use of appropriate equipment. The first option is usually less costly and is widely used.

Design Checkup: At this point it is necessary to check whether or not the parameters with which the system has been designed satisfy the requirements. For instance, it may be necessary to re-evaluate the base station location, the antenna height, etc., so that better performance can be attained.

Field Measurements: For a better tuning of the parameters involved, field measurements (radio survey) should be included in the design. This can be carried out with transmitters and towers provisionally set up at the locations initially defined for the base station.

The cost assessment may require that a redesign of the system should be carried out.

73.6 Traffic Engineering

The starting point for engineering the traffic is the knowledge of the required grade of service. This is usually specified to be around 2% during the busy hour. The question lies on defining the busy hour. There are usually three possible definitions: (1) busy hour at the busiest cell, (2) system busy hour, and (3) system average over all hours.

The estimate of the subscriber usage rate is usually made on a demographic basis from which the traffic distribution can be worked out and the cell areas identified. Given the repeat pattern (cluster size), the cluster with the highest traffic is chosen for the initial design. The traffic A in each cell is estimated and, with the desired blocking probability $E(A, M)$, the erlang-B formula as given by Eq. (73.3) is used to determine the number of channels per cell, M

$$E(M, A) = \frac{A^M/M!}{\sum_{i=0}^{M} A^i/i!} \tag{73.3}$$

In case the total number of available channels is not large enough to provide the required grade of service, the area covered by the cluster should be reduced in order to reduce the traffic per cell. In such a case, a new study on the interference problems must be carried out. The other clusters can reuse the same channels according to the reuse pattern. Not all channels need be provided by the base stations of those cells where the traffic is supposedly smaller than that of the heaviest loaded cluster. They will eventually be used as the system grows.

The traffic distribution varies in time and space, but it is commonly bell shaped. High concentrations are found in the city center during the rush hour, decreasing toward the outskirts. After the busy hour and toward the end of the day, this concentration changes as the users move from the town center to their homes. Note that because of the mobility of the users handoffs and roaming are always occurring, reducing the channel holding times in the cell where the calls are generated and increasing the traffic in the cell where the mobiles travel. Accordingly, the erlang-B formula is, in fact, a rough approximation used to model the traffic process in this ever-changing environment. A full investigation of the traffic performance in such a dynamic system requires all of the phenomena to be taken into account, making any traffic model intricate. Software simulation packages can be used so as to facilitate the understanding of the main phenomena as well as to help system planning. This is a useful alternative to the complex modeling, typically present in the analysis of cellular networks, where closed-form solutions are not usually available.

On the other hand, conventional traffic theory, in particular, the erlang-B formula, is a handy tool widely used in cellular planning. At the inception of the system the calculations are carried out based on the best available traffic estimates, and the system capacity is obtained by grossly exaggerating the calculated figures. With the system in operation some adjustments must be made so that the requirements are met.

The approach just mentioned assumes the simplest channel assignment algorithm: the fixed allocation. It has the maximum spatial efficiency in channel reuse, since the channels are always assigned at the minimum reuse distance. Moreover, because each cell has a fixed set of channels, the channel assignment control for the calls can be distributed among the base stations.

The main problem of fixed allocation is its inability to deal with the alteration of the traffic pattern. Because of the mobility of the subscribers, some cells may experience a sudden growth in the traffic offered, with a consequent deterioration of the grade of service, whereas other cells may have free channels that cannot be used by the congested cells.

A possible solution for this is the use of dynamic channel allocation algorithms in which the channels are allocated on a demand basis. There is an infinitude of strategies using the dynamic assignment principles, but they are usually complex to implement. An interim solution can be exercised if the change of the traffic pattern is predictable. For instance, if a region is likely to have an increase of the traffic on a given day (say, a football stadium on a match day), a mobile base station can be moved toward such a region in order to alleviate the local base.

Another specific solution uses the traffic available at the boundary between cells that may well communicate with more than one base station. In this case, a call that is blocked in its own cell can be directed to the neighboring cell to be served by its base station. This strategy, called *directed retry*, is known to substantially improve the traffic capacity. On the other hand, because channels with marginally

acceptable transmission quality may be used, an increase in the interference levels, both for adjacent channel and cochannel, can be expected. Moreover, subscribers with radio access only to their own base will experience an increase in blocking probability.

73.7 Cell Coverage

The propagation of energy in a mobile radio environment is strongly influenced by several factors, including the natural and artificial relief, propagation frequency, antenna heights, and others. A precise characterization of the signal variability in this environment constitutes a hard task. Deterministic methods, such as those described by the *free space, plane earth*, and *knife-edge diffraction* propagation models, are restricted to very simple situations. They are useful, however, in providing the basic mechanisms of propagation. Empirical methods, such as those proposed by many researchers (e.g., [1,4,5,8]; and others), use curves and/or formulas based on field measurements, some of them including deterministic solutions with various correction factors to account for the propagation frequency, antenna height, polarization, type of terrain, etc. Because of the random characteristics of the mobile radio signal, however, a single deterministic treatment of this signal will certainly lead the problem to a simplistic solution. Therefore, we may treat the signal on a statistical basis and interpret the results as random events occurring with a given probability. The cell coverage area is then determined as the proportion of locations where the received signal is greater than a certain threshold considered to be satisfactory.

Suppose that at a specified distance from the base station the *mean signal strength* is considered to be known. Given this we want to determine the cell radius such that the mobiles experience a received signal above a certain threshold with a stipulated probability. The mean signal strength can be determined either by any of the prediction models or by field measurements. As for the statistics of the mobile radio signal, five distributions are widely accepted today: lognormal, Rayleigh, Suzuki [11], Rice, and Nakagami. The lognormal distribution describes the variation of the mean signal level (large-scale variations) for points having the same transmitter–receiver antennas separation, whereas the other distributions characterize the instantaneous variations (small-scale variations) of the signal. In the calculations that follow we assume a lognormal environment. The other environments can be analyzed in a like manner; although this may not be of interest if some sort of diversity is implemented, because then the effects of the small-scale variations are minimized.

Propagation Model

Define m_w and k as the mean powers at distances x and x_0, respectively, such that

$$m_w = k\left(\frac{x}{x_0}\right)^{-\alpha} \tag{73.4}$$

where α is the path loss coefficient. Expressed in decibels, $M_w = 10 \log m_w$, $K = 10 \log k$ and

$$M_w = K - 10\alpha \log\left(\frac{x}{x_0}\right) \tag{73.5}$$

Define the received power as $w = v^2/2$, where v is the received envelope. Let $p(W)$ be the probability density function of the received power W, where $W = 10 \log w$. In a lognormal environment, v has a lognormal distribution and

$$p(W) = \frac{1}{\sqrt{2\pi}\sigma_w}\exp\left(-\frac{(W - M_w)^2}{2\sigma_w^2}\right) \tag{73.6}$$

where M_W is the mean and σ_w is the standard deviation, all given in decibels. Define w_T and $W_T = 10 \log w_T$ as the threshold above which the received signal is considered to be satisfactory. The probability that the received signal is below this threshold is its *probability distribution function* $P(W_T)$, such that

$$P(W_T) = \int_{-\infty}^{W_T} p(W)\,\mathrm{d}W = \frac{1}{2} + \frac{1}{2}\,\mathrm{erf}\left[\frac{(W_T - M_W)^2}{2\sigma_w^2}\right] \tag{73.7}$$

where erf( ) is the error function defined as

$$\mathrm{erf}(y) = \frac{2}{\sqrt{\pi}} \int_0^y \exp(-t^2)\,\mathrm{d}t \tag{73.8}$$

Base Station Coverage

The problem of estimating the cell area can be approached in two different ways. In the first approach, we may wish to determine the proportion β of locations at x_0 where the received signal power w is above the threshold power w_T. In the second approach, we may estimate the proportion μ of the circular area defined by x_0 where the signal is above this threshold. In the first case, this proportion is averaged over the perimeter of the circumference (cell border); whereas in the second approach, the average is over the circular area (cell area).

The proportion β equals the probability that the signal at x_0 is greater than this threshold. Hence,

$$\beta = \mathrm{prob}(W \geq W_T) = 1 - P(W_T) \tag{73.9}$$

Using Eqs. (73.5) and (73.7) in Eq. (73.9) we obtain

$$\beta = \frac{1}{2} - \frac{1}{2}\,\mathrm{erf}\left[\frac{W_T - K + 10\alpha \log(x/x_0)}{\sqrt{2}\,\sigma_w}\right] \tag{73.10}$$

This probability is plotted in Fig. 73.1, for $x = x_0$ (cell border).

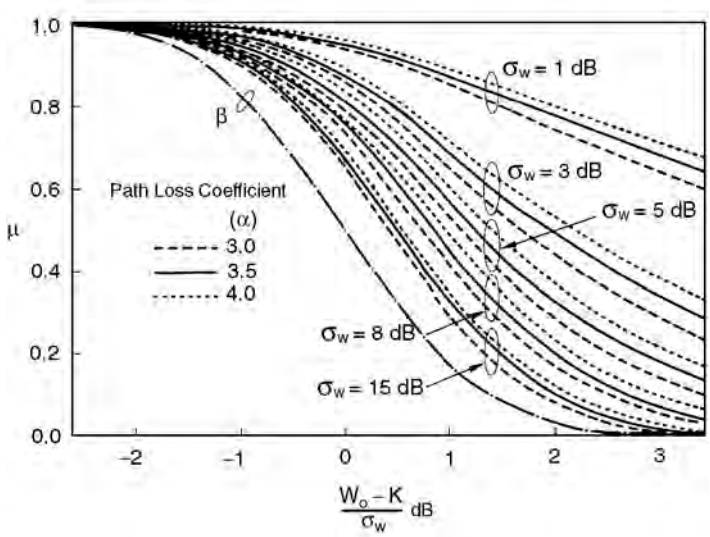


FIGURE 73.1 Proportion of locations where the received signal is above a given threshold; the dashdot line corresponds to the β approach and the other lines to the μ approach.

Let prob($W \geq W_T$) be the probability of the received power W being above the threshold W_T within an infinitesimal area dS. Accordingly, the proportion μ of locations within the circular area S experiencing such a condition is

$$\mu = \frac{1}{S}\int_S [1 - P(W_T)]\,\mathrm{d}S \tag{73.11}$$

where $S = \pi r^2$ and $\mathrm{d}S = x\,\mathrm{d}x\mathrm{d}\theta$. Note that $0 \leq x \leq x_0$ and $0 \leq \theta \leq 2\pi$. Therefore, solving for dθ, we obtain

$$\mu = 2\int_0^1 u\beta\,\mathrm{d}u \tag{73.12}$$

where $u = x/x_0$ is the normalized distance.

Inserting Eq. (73.10) in Eq. (73.12) results in

$$\mu = 0.5\left\{1 + \mathrm{erf}(a) + \exp\left(\frac{2ab+1}{b^2}\right)\left[1 - \mathrm{erf}\left(\frac{ab+1}{b}\right)\right]\right\} \tag{73.13}$$

where $a = (K - W_T)/\sqrt{2}\sigma_w$ and $b = 10\alpha\log(e)/\sqrt{2}\sigma_w$.

These probabilities are plotted in Fig. 73.1 for different values of standard deviation and path loss coefficients.

Application Examples

From the theory that has been developed it can be seen that the parameters affecting the probabilities β and μ for cell coverage are the path loss coefficient α, the standard deviation σ_w, the required threshold W_T, and a certain power level K, measured or estimated at a given distance from the base station.

The applications that follow are illustrated for two different standard deviations: $\sigma_w = 5$ dB and $\sigma_w = 8$ dB. We assume the path loss coefficient to be $\alpha = 4$ (40 dB/decade), the mobile station receiver sensitivity to be −116 dB (1 mW), and the power level estimated at a given distance from the base station as being that at the cell border, $K = -102$ dB (1 mW). The receiver is considered to operate with a *SINAD* of 12 dB for the specified sensitivity. Assuming that cochannel interference levels are negligible and given that a signal-to-noise ratio S/N of 18 dB is required, the threshold W_T will be −116 dB (1 mW) + (18 − 12) dB (1 mW) = −110 dB (1 mW).

Three cases will be explored as follows.

Case 1: We want to estimate the probabilities β and μ that the received signal exceeds the given threshold (1) at the border of the cell, probability β and (2) within the area delimited by the cell radius, probability μ.

Case 2: It may be interesting to estimate the cell radius x_0 such that the received signal be above the given threshold with a given probability (say 90%) (1) at the perimeter of the cell and (2) within the cell area. This problem implies the calculation of the mean signal strength K at the distance x_0 (the new cell border) of the base station. Given K and given that at a distance x_0 (the former cell radius) the mean signal strength M_w is known [note that in this case $M_w = -102$ dB (1 mW)], the ratio x_0/x can be estimated.

Case 3: To fulfill the coverage requirement, rather than calculating the new cell radius, as in Case 2, a signal strength at a given distance can be estimated such that a proportion of the locations at this distance, proportion β, or within the area delimited by this distance, proportion μ, will experience a received signal above the required threshold. This corresponds to calculating the value of the parameter K already carried out in Case 2 for the various situations.

The calculation procedures are now detailed for $\sigma_w = 5$ dB. Results are also shown for $\sigma_w = 8$ dB.

TABLE 73.1 Case 1 Coverage Probability

Standard Deviation, dB	β Approach (Border Coverage), %	μ Approach (Area Coverage), %
5	97	100
8	84	95

TABLE 73.2 Case 2 Normalized Radius

Standard Deviation, dB	β Approach (Border Coverage)	μ Approach (Area Coverage)
5	1.10	1.38
8	0.88	1.27

TABLE 73.3 Case 3 Signal Power

Standard Deviation dB	β Approach (Border Coverage), dB (1 mW)	μ Approach (Area Coverage), dB (1 mW)
5	−103.7	−107.6
8	−99.8	−106.2

Case 1: Using the given parameters we obtain $(W_T - K)/\sigma_w = -1.6$. With this value in Fig. 73.1, we obtain the probability that the received signal exceeds −116 dB (1 mW) for $S/N = 18$ dB given that at the cell border the mean signal power is −102 dB (1 mW) given in Table 73.1.

Note, from Table 73.1, that the signal at the cell border exceeds the receiver sensitivity with 97% probability for $\sigma_w = 5$ dB and with 84% probability for $\sigma_w = 8$ dB. If, on the other hand, we are interested in the area coverage rather than in the border coverage, then these figures change to 100% and 95%, respectively.

Case 2: From Fig. 73.1, with $\beta = 90\%$ we find $(W_T - K)/\sigma_w = -1.26$. Therefore, $K = -103.7$ dB (1 mW). Because $M_w - K = -10\alpha \log(x/x_0)$, then $x_0/x = 1.10$. Again, from Fig. 73.1, with $\mu = 90\%$ we find $(W_T - K)/\sigma_w = -0.48$, yielding $K = -107.6$ dB (1 mW). Because $M_w - K = -10\alpha \log(x/x_0)$, then $x_0/x = 1.38$. These results are summarized in Table 73.2, which shows the normalized radius of a cell where the received signal power is above −116 dB (1 mW) with 90% probability for $S/N = 18$ dB, given that at a reference distance from the base station (the cell border) the received mean signal power is −102 dB (1 mW).

Note, from Table 73.2, that in order to satisfy the 90% requirement at the cell border the cell radius can be increased by 10% for $\sigma_w = 5$ dB. If, on the other hand, for the same standard deviation the 90% requirement is to be satisfied within the cell area, rather than at the cell border, a substantial gain in power is achieved. In this case, the cell radius can be increased by a factor of 1.38. For $\sigma_w = 8$ dB and 90% coverage at the cell border, the cell radius should be reduced to 88% of the original radius. For area coverage, an increase of 27% of the cell radius is still possible.

Case 3: The values of the mean signal power K are taken from Case 2 and shown in Table 73.3, which shows the signal power at the cell border such that 90% of the locations will experience a received signal above −116 dB for $S/N = 18$ dB.

73.8 Interference

Radio-frequency interference is one of the most important issues to be addressed in the design, operation, and maintenance of mobile communication systems. Although both intermodulation and intersymbol interferences also constitute problems to account for in system planning, a mobile radio system designer is mainly concerned about adjacent channel and cochannel interferences.

Adjacent Channel Interference

Adjacent channel interference occurs due to equipment limitations, such as frequency instability, receiver bandwidth, filtering, etc. Moreover, because channels are kept very close to each other for maximum spectrum efficiency, the random fluctuation of the signal, due to fading and near–far effect, aggravates this problem.

Some simple, but efficient, strategies are used to alleviate the effects of adjacent channel interference. In narrowband systems, the total frequency spectrum is split into two halves so that the reverse channels, composing the uplink (mobile to base station) and the forward channels, composing the downlink (base station to mobile), can be separated by half of the spectrum. If other services can be inserted between the two halves, then a greater frequency separation, with a consequent improvement in the interference levels, is accomplished. Adjacent channel interference can also be minimized by avoiding the use of adjacent channels within the same cell. In the same way, by preventing the use of adjacent channels in adjacent cells a better performance is achieved. This strategy, however, is dependent on the cellular pattern. For instance, if a seven-cell cluster is chosen, adjacent channels are inevitably assigned to adjacent cells.

Cochannel Interference

Undoubtedly the most critical of all interferences that can be engineered by the designer in cellular planning is cochannel interference. It arises in mobile radio systems using cellular architecture because of the frequency reuse philosophy.

A parameter of interest to assess the system performance in this case is the carrier-to-cochannel interference ratio C/I_c. The ultimate objective of estimating this ratio is to determine the reuse distance and, consequently, the repeat pattern. The C/I_c ratio is a random variable, affected by random phenomena such as (1) location of the mobile, (2) fading, (3) cell site location, (4) traffic distribution, and others. In this subsection we shall investigate the **outage probability**, i.e., the probability of failing to achieve adequate reception of the signal due to cochannel interference. This parameter will be indicated by $p(CI)$. As can be inferred, this is intrinsically related to the repeat pattern.

Cochannel interference will occur whenever the wanted signal does not simultaneously exceed the minimum required signal level s_0 and the n interfering signals, $i_1, i_2, \ldots, i_n$, by some protection ratio r. Consequently, the conditional outage probability, given n interferers, is

$$p(CI \mid n) = 1 - \int_{s_0}^{\infty} p(s) \int_0^{s/r} p(i_1) \int_0^{(s/r)-i_1} p(i_2) \cdots \times \int_0^{(s/r)-i_1-\cdots-i_{n-1}} p(i_n) \mathrm{d}i_n \cdots \mathrm{d}i_2 \, \mathrm{d}i_1 \, \mathrm{d}s \tag{73.14}$$

The total outage probability can then be evaluated by

$$p(CI) = \sum_n p(CI \mid n) p(n) \tag{73.15}$$

where $p(n)$ is the distribution of the number of active interferers.

In the calculations that follow we shall assume an interference-only environment, i.e., $s_0 = 0$, and the signals to be Rayleigh faded. In such a fading environment the probability density function of the signal-to-noise ratio x is given by

$$p(x) = \frac{1}{x_m} \exp\left(-\frac{x}{x_m}\right) \tag{73.16}$$

where x_m is the mean signal-to-noise ratio. Note that $x = s$ and $x_m = s_m$ for the wanted signal, and $x = i_j$ and $x_m = i_{mj}$ for the interfering signal j, with s_m and i_{mj} being the mean of s and i_j, respectively.

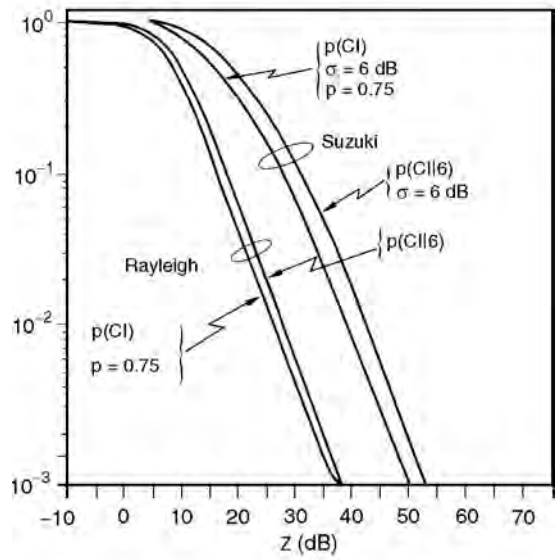


FIGURE 73.2 Conditional and unconditional outage probability for $n = 6$ interferes in a Rayleigh environment and in a Suzuki environment with $\sigma = 6$ dB.

By using the density of Eq. (73.16) in Eq. (73.14) we obtain

$$p(CI \mid n) = \sum_{j=1}^{n} \prod_{k=1}^{j} \frac{z_k}{1 + z_k} \tag{73.17}$$

where $z_k = rs_m / i_{mk}$.

If the interferers are assumed to be equal, i.e., $z_k = z$ for $k = 1, 2, \ldots, n$, then

$$p(CI \mid n) = 1 - \left(\frac{z}{1+z}\right)^n \tag{73.18}$$

Define $Z = 10 \log z$, $S_m = 10 \log s_m$, $I_m = 10 \log i_m$, and $R_r = 10 \log r$. Then, $Z = S_m - (I_m + R_r)$. Equation (73.18) is plotted in Fig. 73.2 as a function of Z for $n = 1$ and $n = 6$, for the situation in which the interferers are equal.

If the probability of finding an interferer active is p, the distribution of active interferers is given by the binomial distribution. Considering the closest surrounding cochannels to be the most relevant interferers we then have six interferers. Thus

$$p(n) = \binom{6}{n} p^n (1-p)^{6-n} \tag{73.19}$$

For equal capacity cells and an evenly traffic distribution system, the probability p is approximately given by

$$p = \sqrt[M]{B} \tag{73.20}$$

where B is the blocking probability and M is the number of channels in the cell.

Now Eqs. (73.20), (73.19), and (73.18) can be combined into Eq. (73.15) and the outage probability is estimated as a function of the parameter Z and the channel occupancy p. This is shown in Fig. 73.2 for $p = 75\%$ and $p = 100\%$.

A similar, but much more intricate, analysis can be carried out for the other fading environments. Note that in our calculations we have considered only the situation in which both the wanted signal and the interfering signals experience Rayleigh fading. For a more complete analysis we may assume the wanted signal to fade differently from the interfering signals, leading to a great number of possible combinations. A case of interest is the investigation of the influence of the standard deviation in the outage probability analysis. This is illustrated in Fig. 73.2 for the Suzuki (lognormal plus Rayleigh) environment with $\sigma = 6$ dB.

Note that by definition the parameter z is a function of the carrier-to-cochannel interference ratio, which, in turn, is a function of the reuse distance. Therefore, the outage probability can be obtained as a function of the cluster size, for a given protection ratio.

The ratio between the mean signal power s_m and the mean interfering power i_m equals the ratio between their respective distances d_s and d_i such that

$$\frac{s_m}{i_m} = \left(\frac{d_s}{d_i}\right)^{-\alpha} \tag{73.21}$$

where α is the path loss coefficient. Now, (1) let D be the distance between the wanted and interfering base stations, and (2) let R be the cell radius. The cochannel interference worst case occurs when the mobile is positioned at the boundary of the cell, i.e., $d_s = R$ and $d_i = D - R$. Then,

$$\frac{i_m}{s_m} = \left(\frac{D}{R} - 1\right)^{-\alpha} \tag{73.22a}$$

or, equivalently,

$$S_m - I_m = 10\alpha \log\left(\frac{D}{R} - 1\right) \tag{73.22b}$$

In fact, $S_m - I_m = Z + R_r$. Therefore,

$$Z + R_r = 10\alpha \log(\sqrt{3N} - 1) \tag{73.23}$$

With Eq. (73.23) and the curves of Fig. 73.2, we can compare some outage probabilities for different cluster sizes. The results are shown in Table 73.4 where we have assumed a protection ratio $R_r = 0$ dB. The protection ratio depends on the modulation scheme and varies typically from 8 dB (25-kHz FM) to 20 dB [single sideband (SSB) modulation].

Note, from Table 73.4, that the standard deviation has a great influence in the calculations of the outage probability.

TABLE 73.4 Probability of Cochannel Interference in Different Cell Clusters

		Outage Probability, %			
		Rayleigh		Suzuki $\sigma = 6$ dB	
N	$Z + R$, dB	$p = 75\%$	$p = 100\%$	$p = 75\%$	$p = 100\%$
1	−4.74	100	100	100	100
3	10.54	31	40	70	86
4	13.71	19	26	58	74
7	19.40	4.7	7	29	42
12	24.46	1	2.1	11	24
13	25.19	0.9	1.9	9	22

73.9 Conclusions

The interrelationship among the areas involved in a cellular network planning is substantial. Vocabularies belonging to topics, such as radio propagation, frequency planning and regulation, modulation schemes, antenna design, transmission, teletraffic, and others, are common to all cellular engineers.

Designing a cellular network to meet system requirements is a challenging task which can only be partially and roughly accomplished at the design desk. Field measurements play an important role in the whole process and constitute an essential step used to tune the parameters involved.

Defining Terms

Outage probability: The probability of failing to achieve adequate reception of the signal due to, for instance, cochannel interference.

Spectrum efficiency: A measure of how efficiently space, frequency, and time are used. It is expressed in erlang per square meter per hertz.

Trunking efficiency: A function relating the number of subscribers per channel and the number of channels per cell for different values of blocking probability.

References

1. Egli, J., Radio above 40 Mc over irregular terrain. *Proc. IRE.,* 45(10), 1383–1391, 1957.
2. Hata, M., Empirical formula for propagation loss in land-mobile radio services. *IEEE Trans. Vehicular Tech.,* VT-29, 317–325, 1980.
3. Ho, M.J. and Stüber, G.L., Co-channel interference of microcellular systems on shadowed Nakagami fading channels. *Proc. IEEE Vehicular Tech. Conf.,* 568–571, 1993.
4. Ibrahim, M.F. and Parsons, J.D., Signal strength prediction in built-up areas, Part I: median signal strength. *Proc. IEEE.* Pt. F. (130), 377–384, 1983.
5. Lee, W.C.Y., *Mobile Communications Design Fundamentals,* Howard W. Sams, Indianapolis, IN, 1986.
6. Leonardo, E.J. and Yacoub, M.D., A statistical approach for cell coverage area in land mobile radio systems. Proceedings of the 7th IEE. Conf. on Mobile and Personal Comm., Brighton, UK, 16–20, Dec. 1993a.
7. Leonardo, E.J. and Yacoub, M.D., (Micro) Cell coverage area using statistical methods. Proceedings of the IEEE Global Telecom. Conf. GLOBECOM'93, Houston, TX, 1227–1231, Dec. 1993b.
8. Okumura, Y., Ohmori, E., Kawano, T., and Fukuda, K., Field strength and its variability in VHF and UHF land mobile service. *Rev. Elec. Comm. Lab.,* 16, 825–873, Sept.-Oct. 1968.
9. Reudink, D.O., Large-scale variations of the average signal. In *Microwave Mobile Communications,* 79–131, John Wiley & Sons, New York, 1974.
10. Sowerby, K.W. and Williamson, A.G., Outage probability calculations for multiple cochannel interferers in cellular mobile radio systems. *IEE Proc.,* Pt. F. 135(3), 208–215, 1988.
11. Suzuki, H., A statistical model for urban radio propagation. *IEEE Trans. Comm.,* 25(7), 673–680, 1977.

Further Information

The fundamentals of mobile radio engineering in connection with many practical examples and applications as well as an overview of the main topics involved can be found in Yacoub, M.D., *Foundations of Mobile Radio Engineering,* CRC Press, Boca Raton, FL, 1993.

74

Microcellular Radio Communications

Raymond Steele
Multiple Access Communications

74.1 Introducing Microcells

In mobile radio communications, an operator will be licenced a specific bandwidth in which to operate a service. The operator will, in general, not design the mobile equipment, but purchase equipment that has been designed and standardized by others. The performance of this equipment will have a profound effect on the number of subscribers the network can support. To shoehorn us into the our subject of microcellular radio communications we will commence with the simplist of systems; namely, circuit-switched frequency division multiple access (FDMA) where each user is assigned a unique radio frequency (RF) bandwidth (B) for the duration of their call. If an operator has been assigned a bandwidth W for calls from the network to the mobile stations (MSs), i.e., handsets carried by subscribers, then the operator can deploy $N_T = W/B$ channels. The operator must also be assigned another bandwidth W in order that there are N_T channels to be used by MSs to transmit their messages to the network. The two bands of WHz must be spaced apart so that the high level transmit signals will be miniscule at the transceivers receiver ports.

Communications with MSs are made from fixed sites, known as base stations (BSs). Should a mobile travel too far from its BS, the quality of the communications link becomes unacceptable. The perimeter around the BS where acceptable communications occur is called a cell and, hence, the term cellular radio. BSs are arranged so that their radio coverage areas, or cells, overlap, and each BS may be given $N = N_T/M$ channels. This implies that there are M BS and each BS uses a different set of channels.

The number N_T is typically relatively low, suppose only 1000. As radio channels cannot operate with 100% utilization, the cluster of BSs or cells has fewer than 1000 simultaneous calls. In order to make the business viable, more users must be supported by the network. This is achieved by repeatedly reusing the channels. Clusters of BSs are tessellated with each cluster using the same N_T channels. This means that there are users in each cluster using the same frequency band at the same time, and inevitably there will be interference. This interference is known as cochannel interference. Cochannel cells, i.e., cells using

the same channels, must be spaced sufficiently far apart for the interference levels to be acceptable. An MS will therefore receive the wanted signal of power S and a total interference power of I, and the signal-to-interference ratio (SIR) is a key system design parameter.

Suppose we have large cells, a condition that occurs during the initial stages of deploying a network when coverage is important. For a given geographical area G_A we may have only one cluster of seven cells, and this may support some 800 simultaneous calls in our example. As the subscriber base grows, the number of clusters increases to, say, 100 with the area of each cluster being appropriately decreased. The network can now support some 80,000 simultaneous calls in the area G_A. As the number of subscribers continues to expand, we increase the number of clusters. The geographical area occupied by each cluster is now designed to match the number of potential users residing in that area. Consequently, the smallest clusters having the highest density of channels per area is found in the center of cities. As each cluster has the same number of channels, the smaller the clusters and, therefore, the smaller the cells, the greater the **spectral efficiency** measured in erlang per hertz per square meter. Achieving this higher spectral efficiency requires a concomitant increase in the infrastructure that connects the small cell BSs to their base station controller (BSC). The BSCs are connected to other switches and data bases required to register mobiles, page them when they are being called, or to allow them to make calls. The mobile operators usually lease high capacity trunk links from the national and international fixed network operators to support their teletraffic. The consequence is that roving MSs can be connected to other MSs, or to fixed terminals, who are anywhere in the world.

As we make the cells smaller, we change from locating the BS antennas on top of tall buildings or hills, where they produce large cells or macrocells, to the tops of small buildings or the sides of large buildings, where they form minicells, to lamp post elevations (several metres), where they form **street microcells**. Each decrease in cell size is accompanied by a reduction in the radiated power levels from the BSs and from the mobiles. As the BS antenna height is lowered, the neighboring buildings and streets increasingly control the radio propagation. This chapter is concerned with microcells and microcellular networks. We commence with the simplest type of microcells, namely, those used for highways.

74.2 Highway Microcells

Since their conception by Steele and Prabhu [1], many papers have been published on **highway microcells**, ranging from propagation measurements to teletraffic issues. Figure 74.1 shows the basic concepts for a highway microcellular system having two cells per cluster where microcells 1 use channel set 1 while microcells 2 use channel set 2. Mobile stations using channel set 1 are always separated in distance by at least the length of microcell 2, and this distance, we assume, provides an acceptable SIR for all MSs. Similar comments apply for MSs using channel set 2. The highway is partitioned into contiguous cigar-shaped segments formed by directional antennas. Omnidirectional antennas can be used at junctions, roundabouts, cloverleafs, and other road intersections. The BS antennas are sited at elevations of some 6–12 m. Figure 74.2 shows received signal levels as a function of the distance d between BS and MS for different roads [2]. The average loss in received signal level, or path loss, is approximately inversely proportional to d^4. The path loss is associated

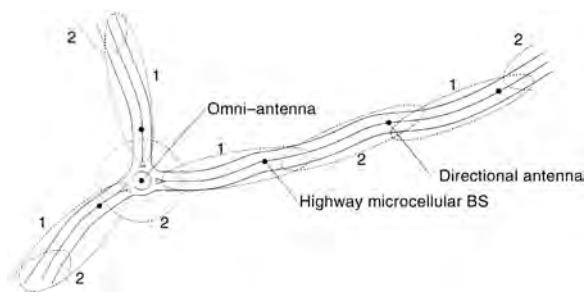


FIGURE 74.1 Highway microcellular clusters. Microcells with the same number use the same channel set.

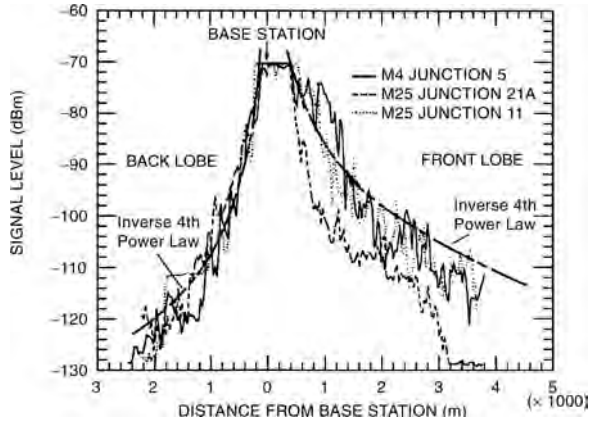


FIGURE 74.2 Overlayed received signal strength profiles of various highway cells including the inverse fourth power law curve for both the front and back antenna lobes. (*Source:* Chia et al. 1987. Propagation and bit error ratio measurements for a microcellular system. *JIERE*. 57(6), 5255–5266. With permission.)

with a slow fading component that is due to the variations in the terrain, the road curvature and cuttings, and the presence of other vehicles.

The curves in the figure are plotted for an 18-element yagi BS antenna having a gain of 15 dB and a front-to-back ratio of 25 dB. In Fig. 74.2, reference is made to junctions on different motorways, e.g., junction 5 on motorway M4. This is because the BS antennas are mounted at these road junctions with the yagi antenna pointing along the highway in order to create a cigar-shaped cell. The flat part of the curve near the BS is due to the MS receiver being saturated by high-signal levels. Notice that the curve related to M25, junction 11, decreases rapidly with distance when the MS leaves the immediate vicinity of the BS. This is due to the motorway making a sharp turn into a cutting, causing the MS to lose line of sight (LOS) with the BS. Later the path loss exponent is approximately 4. Experiments have shown that using the arrangement just described, with each BS transmitting 16 mW at 900 MHz, 16 kb/s noncoherent frequency shift keying, two-cell clusters could be formed where each cell has a length along the highway ranging from 1 to 2 km. For MSs traveling at 110 km/h the average handover rate is 1.2 per minute [2].

Spectral Efficiency of Highway Microcellular Network

Spectral efficiency is a key system parameter. The higher the efficiency, the greater will be the teletraffic carried by the network for the frequency band assigned by the regulating authorities per unit geographical area. We define the spectral efficiency in mobile radio communications in erlang per hertz per square meter as

$$h \triangleq A_{CT}/S_T W \tag{74.1}$$

although erlang per megahertz per square kilometer is often used. In this equation, A_{CT} is the total traffic carried by the microcellular network,

$$A_{CT} = CA_C \tag{74.2}$$

where C is the number of microcells in the network and A_C the carried traffic by each microcellular BS. The total area covered by the tessellated microcells is

$$S_T = CS \tag{74.3}$$

where S is the average area of a microcell, whereas the total bandwidth available is

$$W = MNB \tag{74.4}$$

whose terms M, N, and B were defined in Section 74.1. Substituting Eqs. (74.2)–(74.4) into Eq. (74.1) yields

$$\eta = \frac{\rho}{SMB} \tag{74.5}$$

where

$$r = A_C/N \tag{74.6}$$

is the utilization of each BS channel.

If the length of each microcell is L, there are n up-lanes and n down-lanes, and each vehicle occupies an effective lane length V, which is speed dependent, then the total number of vehicles in a highway microcell is

$$K = 2nL/V \tag{74.7}$$

Given that all vehicles have a mobile terminal, the maximum number of MSs in a microcell is K. In a highway microcell we are not interested in the actual area $S = 2nL$ but in how many vehicles can occupy this area, namely, the effective area K. Notice that K is largest in a traffic jam when all vehicles are stationary and V only marginally exceeds the vehicle length. Given that N is sufficiently large, η is increased when the traffic flow is decreased.

The traffic carried by a microcellular BS is

$$A_C = [\lambda_N(1 - P_{bn}) + \lambda_H(1 - P_{fhm})]\bar{T}_H \tag{74.8}$$

where P_{bn} is the probability of a new call being blocked, P_{fhm} is the probability of handover failure when mobiles enter the microcell while making a call and concurrently no channel is available, λ_N and λ_H are the new call and handover rates, respectively, and $\bar{T}_H$ is the mean channel holding time of all calls. For the simple case where no channels are exclusively reserved for handovers, $P_{bn} = P_{fhm}$, and

$$A_C = l_T\bar{T}_H(1 - P_{bn}) = A(1 - P_{bn}) \tag{74.9}$$

where

$$\lambda_T = \lambda_N + \lambda_H \tag{74.10}$$

and A is the total offered traffic. The mathematical complexity resides in calculating A and P_{bn}, and the reader is advised to consult El-Dolil, Wong, and Steele [3] and Steele and Nofal [4].

Priority schemes have been proposed whereby N channels are available for handover, but only $N - N_h$ for new calls. Thus N_h channels are exclusively reserved for handover [3,4]. While P_{bn} marginally increases, P_{fhm} decreases by orders of magnitude for the same average number of new calls per sec per microcell. This is important as people prefer to be blocked while attempting to make a call compared to having a call in progress terminated due to no channel being available on handover. An important enhancement is to use an oversailing macrocellular cluster, where each macrocell supports a microcellular cluster. The role of the macrocell is to provide channels to support microcells that are overloaded and to provide communications to users who are in areas not adequately covered by the microcells [3,4]. When vehicles are in a solid traffic jam, there are no handovers and so N_h should be zero. When traffic is flowing fast, N_h should be high. Accordingly, a useful strategy is to make N_h adaptive to the new call and handover rates [5].

74.3 City Street Microcells

We will define a city street microcell as one where the BS antenna is located below the lowest building. As a consequence, the diffraction over the buildings can be ignored, and the heights of the buildings are of no consequence. Roads and their attendant buildings form trenches or canyons through which the mobiles travel. If there is a direct line-of-sight path between the BS and a MS and a ground-reflected path, the received signal level vs. BS–MS distance is as shown in Fig. 74.3. Should there be two additional paths from rays reflected from the buildings, then the profile for this four-ray situation is also shown in Fig. 74.3. These theoretical curves show that as the MS travels from the BS the average received signal level is relatively constant and then decreases relatively rapidly. This is a desirable characteristic as it offers a good signal level within the microcell, while the interference into adjacent microcells falls off rapidly with distance.

In practice there are many paths, but there is often a dominant one. As a consequence the fading is Rician [6,7]. The Rician distribution approximates to a Gaussian one when the received signal is from a dominant path with the power in the scattered paths being negligible, to a Rayleigh one when there is no dominant path. Macrocells usually have Rayleigh fading, whereas in microcells the fading only occasionally becomes Rayleigh and is more likely to be closer to Gaussian. This means that the depth of the fades in microcells are usually significantly smaller than in macrocells enabling microcellular communications to operate closer to the receiver noise floor without experiencing error bursts and to accommodate higher cochannel interference levels. Because of the small dimensions of the microcells, the delays between the first and last significant paths is relatively small compared to the corresponding delays in macrocells. Consequently, the impulse response is generally shorter in microcells and, therefore, the transmitted bit rate can be significantly higher before intersymbol interference is experienced compared to the situation in macrocells. Microcells are, therefore, more spectrally efficient with an enhanced propagation environment.

We may classify two types of city street microcells, one for pedestrians and the other for vehicles. In general, there are more MSs carried by pedestrians than in cars. Also, as cars travel more quickly than people, street microcells for cars are accordingly larger than those for pedestrians. The handover rates for MSs in pedestrian and vehicular street microcells may be similar, and networks must be capable of handling the many handovers per call that may occur. In addition, the time available to complete a handover may be very short compared to those in macrocells.

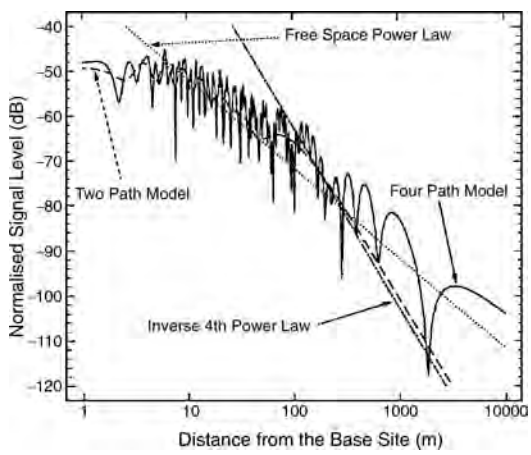


FIGURE 74.3 Signal level profiles for the two- and four-path models. Also shown are the free space and inverse fourth power laws. (*Source:* Green. 1990. Radio link design for microcellular system, *British Telecom. Tech. J.*, 8(1), 85–96. With permission.)

FIGURE 74.4 NP WorkPlace coverage plot of a city street microcell. The map gridsize is 500 meters.

City street microcells are irregular when the streets are irregular as demonstrated by the NP WorkPlace[1] coverage plot for the microcellular BS in Southampton city area displayed in Fig. 74.4. The microcell is represented by the white arrow in the plot, the buildings are in light grey, the vegetation in green, and open spaces in black. The other colours are received signal strength zone defined by the legend. Having sited the first BS and observed its signal level contours we still do not know the microcell boundary, as this boundary is not set by coverage but by the minimum acceptable SIR level for the system being deployed. For FDMA or time division multiple access (TDMA) networks [code division multiple access (CDMA) networks will be addressed later], the SIR levels cannot be computed until we have a frequency plan and hence know which other BSs are using the same channel set as assigned to our microcell. However, coverage is a pre-requisite, i.e., first we must obtain coverage and then we will decide on the "frequency plan" to yield our cell shapes. Figure 74.5 shows the highest received level at any location due to a cluster of microcells. Again we observe the white arrows signifying the microcellular BSs, and the high signal levels around each BS (represented by the colour red). Using a planning tool, like the NPWorkPlace, a frequency plan is assigned. This process is too complex to be described here, but intuitively it is apparent that if two adjacent BSs are assigned the same channel set then the SIR values are likely to be <8 dB at juctions where both BSs have a high signal level. Indeed if adjacent BSs were assigned the same channel set then the microcell for each BS would be a island around it, with areas between the BSs unusable because of the unacceptably low SIR values. However, if the cochannel BSs are sufficiently far apart, and we take care with the siting of our BSs, we can tailor the cells to suit the teletraffic requirements. In Fig. 74.5, the BSs at the top-left hand and bottom right-hand corner of the cluster could be cochanneled.

In the middle of cities, microcell clusters will be tesselated to form a mosaic of concatenated microcells. They will probably be assigned their own channel set, as distinct from the set used by macrocells. Arguments for channel set partitioning are presented in reference [8]. Channel set partitioning or not, macrocells over-sailing microcellular clusters are needed to remove radio dead spots, enables the channel utilisation of microcell channels to be high as well as providing support for MSs that cannot handover to another microcellular BS.

There are important differences between highway microcells and city microcells, which relate to their one- and two-dimensional characteristics. A similar comment applies to street microcells and hexagonal cells. Basically, the buildings have a profound effect on cochannel interference. The buildings shield much

[1]NP WorkPlace is a propriety software outdoor planning tool developed by Multiple Access Communications Ltd.

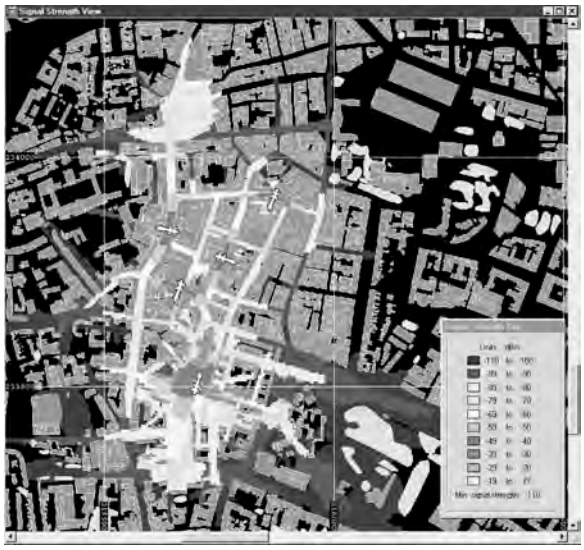


FIGURE 74.5 Cluster of microcells in a city centre. The map gridsize is 500 meters.

of the cochannel interference, and the double regression path loss law of microcells [6] also decreases interference if the break-distance constitutes the notional microcell boundary. City microcellular clusters may have as few as two microcells, but generally more are required, and in some topologies six or more may be deployed [9]. The irregularity of city streets means that some signals can find paths through building complexes to give cochannel interference where it is least expected.

In highway microcells two-cell clusters can be used with TDMA and FDMA. Street microcells have complex shapes, see Figs. 74.4 and 74.5, and when fixed channel assignment (FCA) is used with TDMA and FDMA, there is a danger that high levels of interference will be ducted through streets and cause high-interference levels in a small segment of a microcell. To accommodate this phenomenon, the system must have a rapid handover (HO) capability, with HO to either a different channel at the same BS, to the interfering cell, or to an oversailing macrocell. The irregularity of street microcells, except in regularly shaped cities, such as midtown Manhattan, suggests that FCA should not be used. If it is, it requires $M \geq 4$. Instead, dynamic channel assignment (DCA) should be employed. For example, when DCA is used with TDMA we abandon the notion of clusters of microcells. We may arrange for all microcells to have the same frequency set and design the system with accurate power control to contain the cochannel interference, and to encourage a MS to switch to another channel at the current BS or switch to a new BS directly when the SIR becomes below a threshold at either end of the duplex link. The application of DCA increases the capacity and can also deal with the situation where a MS suddenly experiences a rapid peak of cochannel interference during its travels.

TDMA/DCA is appropriate for indoor microcells, where the complexity of the DCA is easier to implement compared to street microcells. FDMA should be considered for indoor microcells where it is well suited to provide high-bit-rate services since the transmitted rate is the same as the source rate. It also benefits from the natural shielding that exists within buildings to contain the cochannel interference and the low-power levels that simplify equipment design.

Teletraffic Issues

Consider the arrangement where each microcellular cluster is overlaid by a macrocell. The macrocells are also clustered.The arrangement is shown in Fig. 74.6 [10]. The total traffic carried is

$$A_{CT} = C_m A_{cm} + C_M A_{CM} \tag{74.11}$$

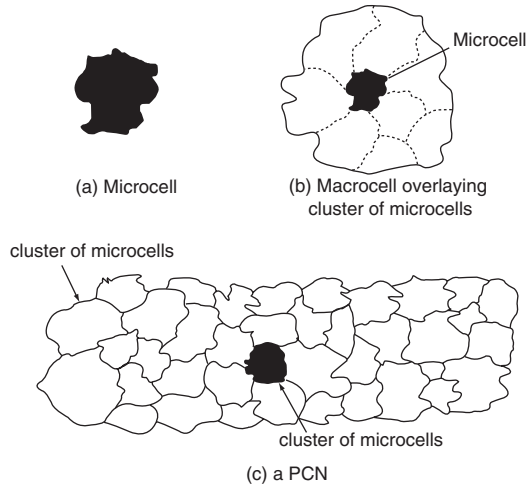


FIGURE 74.6 Microcellular clusters with oversailing macrocellular clusters. Each macrocell is associated with a particular microcellular cluster. (*Source:* Steele and Williams. 1993. Third generation PCN and the intelligent multimode mobile portable. *IEE Elect. and Comm. Eng. J.*, 5(3), 147–156. With permission.)

where C_m and C_M are number of microcells and macrocells in the network, respectively. Each microcellular BS has N channels and carries A_{cm} erlang. The corresponding values for the macrocellular BSs are N_0 and A_{CM}. The channel utilization for the network is

$$\rho_2 = \frac{C_m A_{cm} + C_M A_{CM}}{C_m N + C_M N_0} = \frac{M A_{cm} + A_{CM}}{MN + N_0} \tag{74.12}$$

where M is the number of microcells per cluster.

The spectral efficiency is found by noting that the total bandwidth is

$$B_T = B_c(MN + M_0 N_0) \tag{74.13}$$

where B_c is the effective channel bandwidth, and M_0 is the number of macrocells per macrocellular cluster. The traffic carried by a macrocellular cluster and its cluster of microcells is

$$A_M = A_{CM} M_0 + M_0(A_{cm} M) \tag{74.14}$$

over an area of

$$S_M = (S_m M) M_0 \tag{74.15}$$

where S_m is the area of each microcell.

The spectral efficiency is, therefore,

$$\begin{aligned} \eta &= \frac{A_{CM} + A_{cm} M}{B_c(MN + M_0 N_0) S_m M} \\ &= \frac{\rho_2}{B_c S_m M} \left\{ \frac{MN + N_0}{MN + M_0 N_0} \right\} \end{aligned} \tag{74.16}$$

We note that by using oversailing macrocells to assist microcells experiencing overloading we are able to operate the microcells at high levels of channel utilization. However, the channel utilization of the

macrocells must not be high if the probability of calls forced to terminate due to handover failure is to be minuscule.

74.4 Indoor Microcells

Microcellular BSs may be located within buildings to produce **indoor microcells** often called picocells, whose dimensions may extend from a small office, to part of larger offices, to a complete floor, or to a number of floors. The microcells are box-like for a single office microcell; or may contain many boxes, e.g., when a microcell contains contiguous offices. Furniture, such as bookcases, filing cabinets, and desks may represent large obstacles that will introduce shadowing effects. The signal attenuation through walls, floors, and ceilings may vary dramatically depending on the construction of the building. There is electromagnetic leakage down stairwells and through service ducting, and signals may leave the building and re-enter it after reflection and diffractions from other buildings.

Predicting the path loss in office microcells is, therefore, fraught with difficulties. At the outset it may not be easy to find out the relevant details of the building construction. Even if these are known, an estimation of the attenuation factors for walls, ceilings, and floors from the building construction is far from simple. Then there is a need to predict the effect of the furniture, effect of doors, the presence of people, and so on. Simple equations have been proposed. For example, the one by Keenan and Motley [11], who represent the path loss in decibels by

$$PL = L(V) + 20\log_{10} d + n_f a_f + n_w a_w \tag{74.17}$$

where d is the straight line distance between the BS and the MS, a_f and a_w are the attenuation of a floor and a wall, respectively, n_f and n_w are the number of floors and walls along the line d, respectively, and $L(V)$ is a so-called clutter loss, which is frequency dependent. This equation should be used with caution. Researchers have made many measurements and found that computer ray tracing techniques give disparate results. Prediction errors having a standard deviation of 8–12 dB are not unusual. Given the wide range of path loss in mobile communications, however, and the expense of making measurements, particularly when many BS locations are examined, means that there is nevertheless an important role for planning tools to play, albeit their poorer accuracy compared to street microcellular planning tools.

As might be expected, the excess path delays in buildings is relatively small. The maximum delay spread within rooms and corridors may be <200 and 300 ns, respectively [12]. The digitally enhanced cordless telecommunication (DECT) indoor system operates at 1152 kb/s without either channel coding or equalization [13]. This means that the median rms values are relatively low, and 25 ns has been measured [12]. When the delay spread becomes too high resulting in bit errors, the DECT system hops the user to a better channel using DCA.

74.5 Microcellular Infrastructure

An important requirement of first and second generation mobile networks is to restrict the number of base stations to achieve sufficient network capacity with an acceptably low probability of blocking. This approach is wise given the cost of base stations and their associated equipment, plus the cost and difficulties in renting sites. It is somewhat reminiscent of the situation faced by early electronic circuit designers who needed to minimize the number of tubes and later the number of discrete transistors in their equipment. It was the introduction of microelectronics that freed the circuit designer. We are now approaching an analogous situation where we can free the network designers of the third generation (3G) communication networks, allowing the design to have microcellular BSs in the position where they are required, without being concerned if they are rarely used, and knowing that the cost of the microcellular network is a minor one compared to the overall network cost. This approach is equivalent to installing electric lighting where we are not unduly concerned if not all the lights are switched on at any particular time, preferring to be able to provide illumination where and when it is needed.

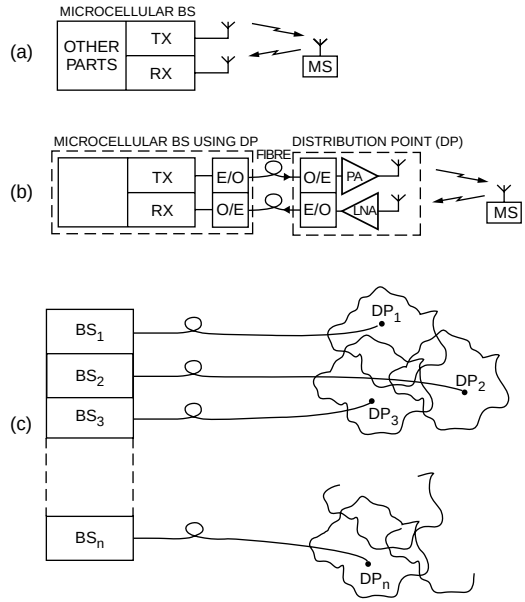


FIGURE 74.7 Creating microcells using DPs: (a) microcellular BS communicating with MS, (b) radio over fiber to distribution point, and (c) microcellular clusters using DPs.

To realize high-capacity mobile communications we need to design microcellular BSs of negligible costs, of match-box to coffee mug dimensions and to interconnect them with their infrastructure using either a fibre network or a wireless local area network (W-LAN). At one extreme we could deploy distribution points (DPs) that only have optical-to-microwave converters, microwave-to-optical converters, and linear amplifiers, with the remainder of the BS at another location. In between the miniaturized, fully functional BSs and the DPs there is a range of options that depends on how much complexity is built into the microcellular BS and how the intelligence of the network is distributed.

Radio over Fiber

The method of using DPs to form microcells is often referred to as radio over fiber (ROF) [14]. Figure 74.7(a) shows a microcellular BS transmitting to a MS. When the DP concept is evoked, the microcellular BS contains electrical-to-optical (E/O) and optical-to-electrical (O/E) converters as shown in Fig. 74.7(b). The microwave signal that would have been radiated to the mobile is now applied, after suitable attenuation, to a laser transmitter. Essentially, the microwave signal amplitude modulates the laser, and the modulated signal is conveyed over a single-mode optical fiber to the distribution point. O/E conversion ensues followed by power amplification, and the resulting signal is transmitted to the MS. Signals from the MS are low-noise amplified and applied to the laser transmitter in the DP. Optical signals are sent from the DP to the BS where O/E conversion is performed followed by radio reception.

In general, the BS transceiver will be handling multicarrier signals for many mobiles, and the DP will accordingly be transceiving signals with many mobiles whose power levels may be significantly different, even when power control is used. Care must be exercised to avoid serious intermodulation products arising in the optical components.

Figure 74.7(c) shows the cositing of n microcellular BSs for use with DPs. This cositing may be conveniently done at a mobile switching center (MSC). Shown in the figure are the DPs and their irregular shaped overlapping microcells. The DPs can be attached to lamp posts in city streets, using electrical power from the electric light supply and the same ducting as used by the electrical wiring, or local telephone ducting. The DPs can also be attached to the outside of buildings. DPs within buildings may be conveniently mounted on ceilings.

The DP concept allows small, lightweight equipment in the form of DPs to be geographically distributed to form microcells; however, there are problems. The N radio carriers cause intermodulation products (IMPs), which may be reduced by decreasing the depth of amplitude modulation for each radio channel. Unfortunately this also decreases the carrier-to-noise ratio (CNR) and the dynamic range of the link. With TDMA having many channels per carrier, we can decrease the number of radio carriers and make the IMPs more controllable. CDMA is particularly adept at coping with IMPs. The signals arriving from the MSs may have different power levels, in spite of power control. Because of the small size cells, the dynamic range of the signals arriving from MSs having power control should be <20 dB. If not, the power levels of the signals arriving at the DP from the MSs may need to be made approximately similar by individual amplification at the DP. We also must be careful to limit the length of the fiber as optical signals propagate along fibers more slowly than radio signals propagate in free space. This should not be a problem in microcells, unless the fiber makes many detours before arriving at its DP.

Miniaturized Microcellular BSs

The low-radiated power levels used by BSs and MSs in microcells have important ramifications on BS equipment design. Small fork combiners can be used along with linear amplifiers. Even FDMA BSs become simple when the high-radiated power levels are abandoned. It is the changes to the size of the RF components that enables the size of the microcellular BSs to be small.

The interesting question that next arises is, how much baseband signal processing complexity should the microcellular BS have? If the microcellular BSs are connected to nodes in an optical LAN, we can convey the final baseband signals to the BS, leaving the BS with the IF and front-end RF components. This means that processing of the baseband signals will be done at the group station (GS), which may be a MSC connected to the LAN. The GS will, therefore, transcode the signals from the ISDN into a suitable format. Using an optical LAN, however, and with powerful microelectronics, the transcoding and full BS operations could be done at each microcellular BS. Indeed, the microcellular BS may eventually execute many of the operations currently handled by the MSC.

74.6 Microcells in CDMA Networks

Cellular radio operates in an intentional interference environment where the level of interference is a compromise between an acceptable received signal quality and a sufficiently high network capacity. Because CDMA was conceived to operate in military environments where enemy jamming interference is expected, it follows that CDMA technology is well suited to handling interference in cellular radio. As a consequence of CDMA technology's robustness to interference, CDMA networks are able to operate with single cell clusters ($M = 1$). In principle, an entire CDMA network can work with the same radio frequency carrier in every sector, although the more likely strategy is to assign one carrier to, say, all indoor microcells, another to all outdoor microcells, and a further carrier to all macrocells. The problem with this scenario is that the adjacent channel interference from the macrocells to the microcells must be carefully managed. Failure to do this will result in radio deserts where few or no calls can be made. Should sufficient carriers be available then frequency planning can be used with effect in CDMA networks.

The International Telecommunications Union (ITU) is supporting a family of 3G networks to provide multimedia mobile services on a global basis. The leading 3G systems are IMT DS, also known as UTRA FDD or wideband CMDA (W-CDMA); and IMT MC which is another name for cdma2000 [9]. There are another three members of the 3G terrestrial system family, but they will not be considered here. The main point is that W-CDMA is the dominant 3G technology.

Let us consider a CDMA network of BSs in a city centre. In our discussions on FDMA and TDMA we showed a coverage plot in Fig. 74.5 of a cluster of microcells. The NPWorkPlace planning tool will plan small 3-D cells in cities [15] using 3-D map data having a building data resolution of 1m. BSs can be placed on the tops or the sides of buildings at any elevation, yielding SIR horizontal profiles at the elevations of interest. This means, for example, that the SIR values on the outside of buildings at ground

FIGURE 74.8 The minimum path loss from all sites.

and at any other floor can be predicted. As a consequence we can have both microcells in the street canyons, and oversailing minicells (i.e., small macrocells) covering buildings and their neighbouring streets below the minicell antennas.

Figure 74.8 shows an NPWorkPlace plot of minimum path loss, i.e., highest received signal level for four microcells; and for three 3-sector minicells and two 2-sector minicells whose antennas are located on top of the buildings (shown in grey). This plot, when coupled with a frequency plan, will yield the cells with suitable SIR values for a TDMA network. But, the situation is different for CDMA, and is further different from second generation TDMA systems due to the presence of multimedia services with their specific bit rates and quality-of-service (QoS) agreements. In CDMA planning we also need to consider the number and behaviour of the 3G MSs.

Consider a simple example of a single cell and only one user. The cell boundary is identified by the user travelling away from the BS in a specific direction until the quality of reception is just acceptable, and then repeating this procedure for different directions from the BS. If a second user closer to the BS than the original user now communicates with the BS, its received power will be interference to the first user, whose position on the boundary was previously just able to sustain communications. The first user will have his call dropped unless he moves toward the BS until his received power matches the received power of the second user. Notice that in CDMA each user is power controlled in an attempt to deliver the same power S as the BS. The effect of the first user needing to move towards the BS means that the presence of the second user has shrunk the cell. Further users, each delivering a powers at the BS, will cause the cell to get smaller as the interference to any user is $(N-1)S$, where N is the number of users in the cell. The cell boundary depends on the user on the boundary having sufficient radiated power to deliver the power S at the BS. The received power S is simply related to the received energy per bit E_b, and the interference power $(N-1)S$ is related via the transmission bandwidth to the power spectral density (I_o) of the total interference. The quotient E_b/I_o is a key system parameter in CDMA.

Notice that for a given number of users N, their position in the cell is unimportant, provided they can all deliver the power S at the BS. However, the profile of E_b/I_o versus mobile speed is far from constant, being low when the MS is stationary when power control is effective, and low at higher speeds when channel coding and interleaving combats errors. For intermediate speeds a higher E_b/I_o is required. If the user is allowed to increase its power the interference to other users increases. Also, even if all users move at the same speed, perfect power control is not realised. Consequently the E_b/I_o has a distribution about its target value, and the larger the standard deviation of this distribution the greater the decrease in capacity for a given data erasure rate [9].

TABLE 74.1 The Assumed Service Mix Within the Network

Service Type	Bit Rate	Penetration	Source Activity Factor	Symmetry (UL/DL)
High-rate data	384 kb/s	0.05	0.8	0.026
Medium-rate data	144 kb/s	0.15	0.8	0.026
Speech	12.2 kb/s	0.8	0.5	1

In a multicellular environment there is intercellular interference in addition to the intracellular interference just discussed. The intercellular interference is difficult to control, and generally each BS does the best it can for its own users rather than the network attempting a global optimisation of mitigating all forms of interference. The important point to appreciate is that the cell boundary is a system variable, with the cell size increasing as the interference levels decrease, and vice versa. So if there are few users, as in a start-up scenario, the cell size may be large, but as traffic builds up the cell size decreases; and as users clear-down calls, or additional users in the cell make new calls, or change their behaviour (e.g., speed), or the intercellular interference changes, then the cell size also changes.

This interference modulated cell size is difficult to manage, and matters are exacerbated by other factors. For example, in 3G networks the chip rate may be constant, but the bit rate varies with the type of service. This means the spreading factor, which is equal to the CDMA code length, decreases with increasing bit rate. Consequently, the processing gain of the CDMA receiver is decreased, and this in turn decreases the link budget. Hence increasing the bit rate results in the cell size contracting. This has another knock-on effect, namely, if the cells become islands for high bit rate services, then soft handover (SHO) that requires more than one BS involvement cannot be used, and this makes matters worse. There are other major difficulties, such as if the QoS requires a lower bit error rate, or guaranteed packet delivery, etc., then these factors also decrease the effective size of the cell for this service.

Our above comments lead us to the opinion that because cell shapes and sizes are so variable, depending as they do on a myriad of factors, then the basic concept of a cell may have become out-moded. Instead we may adopt the notion of 3-D space in which MSs roam and may be connected to one or more BSs, and that for some services there is a continuum of space for communications, and at other times only islands that will support particular services, and that these islands have continuously changing boundaries. One BS does not have one island, but more like an archipelago, small volumes of space where the service is best from this BS.

In the plots that follow we will illustrate some of these ideas. Cells will not be specifically shown. Rather areas where communications can be supported by a family of BSs. The plots were made using the MACdma planning tool,[2] which will highlight some of the key phenomena for the exemplary conditions given in Table 74.1.

The table displays the downlink (forward link) transmission rates for three services. The probability that the service will be used is given by the penetration column, and hence the sum of the numbers in this column is unity. From the penetration column we observe that 80% of all calls are for speech, while only 5% and 15% of the calls are for high-rate (384 kb/s) and medium rate (144 kb/s) data, respectively. Consequently an average user is 0.8/0.05 = 16 times more likely to make a voice call than a high-rate user. A voice call is 0.8/0.15 = 5.3 times more probable than a medium-rate call. Because most calls do not have data to transmit continuously we need to introduce the source activity factor (SAF). We see that the speech service has a SAF of 0.5, which may be interpreted as a speaker requiring the channel on average for half the time. For the remainder of the time the other person will speak. The system performance is enhanced if the SAF is low as less interference is generated to other users. The symmetry column gives the ratio of the uplink (UL) to downlink (DL) data rates. We observe that for the high-rate and medium-rate services, the DL rate is 384 kb/s and 144 kb/s. The corresponding UL rate from the MSs is also 384 kb/s and 144 kb/s. However, because these services have large asymmetry, the uplinks transmit

[2]MACdma planning tool developed by Multiple Access Communications Ltd.

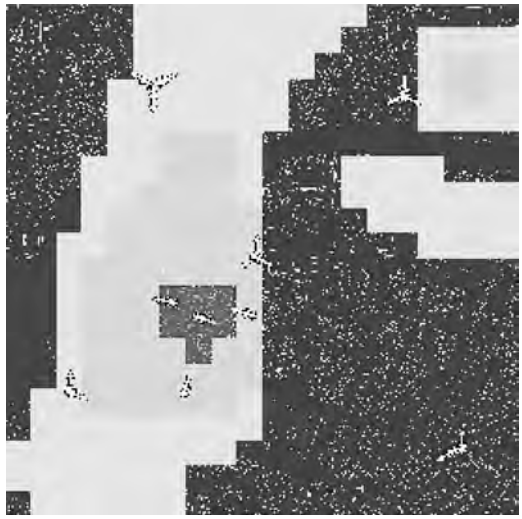


FIGURE 74.9 Variation of traffic density across the network. Red represents 128 $\mu E/m^2$, green represents 64 $\mu E/m^2$, light blue represents 32 $\mu E/m^2$, and dark blue represents 16 $\mu E/m^2$.

relatively infrequently, and the long-term average data rate is 384 × 0.026 = 9.984 kb/s and 144 × 0.026 = 3.744 kb/s, respectively.

We now need a map of how we expect the traffic density measured in E/m^2 to be distributed throughout the city. Generating such a map is difficult, but as a first approximation it could be based on 2G telephony traffic, and then modified by measurements as the network capacity grows and the popularity of different services becomes apparent. Let us utilise the arbitrary traffic distribution shown in Fig. 74.9, where the highest traffic of 128 $\mu E/m^2$ is carried by the microcellular cluster. Other parameters now need to be considered, such as the channel holding time for each service, the mean call arrival rate and its distribution, and so on. We will refrain from entering into such detail, but emphasise that the carried traffic for a specific service is independent of the bit rate.

In the plots that follow we will assume that the three services listed in Table 74.1 are being used concurrently, and that the traffic density for each service depends on where the MSs are located, i.e., in accordance with Fig. 74.9. Our first step is to assure ourselves that the plot (not shown) of downlink pilot channel energy per chip-to-interference PSD, namely, E_c/I_o, is in excess of −16 dB throughout the complete coverage area. The condition $(E_c/I_o) > -16$ dB is a number we assume the system requires in order for an MS to register on the network. Whether an MS is allowed to make a call depends on a number of factors, such as whether the BS is able to allocate sufficient transmit power to a traffic channel in the presence of many calls in progress.

At any instant we may have one to all of the service types being handled by each BS. It may be that the required MS transmit power is sufficient for a particular service as the received E_b/I_o is acceptable, given the current level of prevailing interference. Such a situation is shown in Fig. 74.10 for the speech service. Scrutiny of the plot may lead the observer to consider that the buildings are displayed in orange. But this is not so. MSs are uniformly distributed and some are therefore located within buildings. These users have to transmit sufficient power to overcome the electromagnetic shielding of the buildings they are in, and as a consequence there is a considerable difference in the transmit power from an MS within a building and another on the street outside. However, we observe that for some buildings additional transmit power is required (shown in red), while for buildings close to a cell site the transmit power is at a minimum irrespective of whether an MS is within or outside the building (see the green areas).

Consideration must also be given to the consequences of granting access to the network. To this end, call access control (CAC) is employed, and if the network considers that by granting access to an MS the level

FIGURE 74.10 The required transmit power for the speech service.

FIGURE 74.11 The probability of blocking for the speech service.

of interference will become unacceptably high to other users then the access request fails. Figure 74.11 displays the probability of blocking for the speech service. The dimensions of this plot are smaller than in the previous ones as for our traffic levels and service mix the blocking only occurs close to the microcellular BSs. This may seem surprising but it is due to the path losses between MSs and their BSs being so small that when the MSs are power controlled to their lowest specified levels they deliver a power at the cell site that is unacceptably high, causing unwanted interference to other MSs. Accordingly the system's CAC blocks these calls. Interestingly, MSs located within buildings close to the cell sites may not be blocked if the building penetration losses are sufficiently high to cause the received power at the BS to be brought close to the specified received power by the power control system. Fortunately this high level of blocking around the cell site is small, ranging from a few meters to usually less than 80 metres. We note that all of the plots in Section 74.6 relate to the same carrier being deployed at all cell sites, be they microcellular or minicellular. If an additional carrier is available then MSs in the high blocking probability areas would be handed over to the other carrier, probably used by the minicells.

To appreciate how the degree of traffic effects the MS transmit power and the probability of blocking of an access call for the speech service, we increase the variation of teletraffic density shown in Fig. 74.9 by a factor of four for all services. So the red area where the microcells reside now supports 512 $\mu E/m^2$ of traffic.

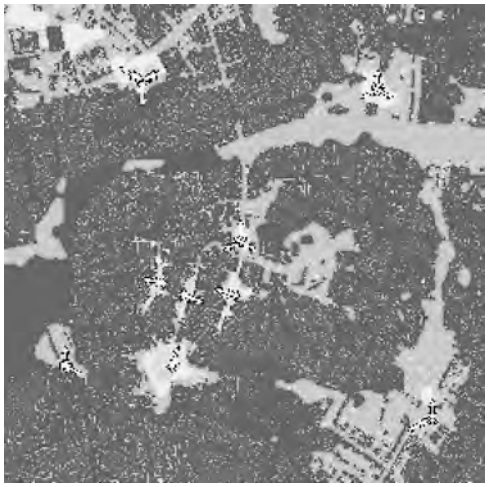


FIGURE 74.12 The required transmit power for MSs using the speech service, when the level of offered traffic has been increased by a factor of four.

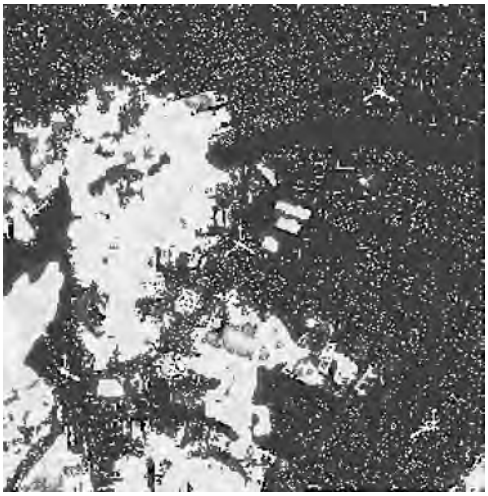


FIGURE 74.13 The probability of blocking for speech users when the offered traffic has been increased by a factor of four. For the key see Fig. 74.11.

Figures 74.12 and 74.13 apply for this fourfold increase in traffic. We observe in Fig. 74.12 that at many locations the MSs are required to transmit more power to make a speech call compared to the situation in Fig. 74.10 where the offered traffic is four times less. The MS transmit values of Fig. 74.12 (and Fig. 74.10) relate to MSs that are able to make a call. More significantly, the probability of blocking for speech users is dramatically increased as shown in Fig. 74.13. Notice that the plot size in Fig. 74.13 is much larger than in Fig. 74.11 where blocking is confined to areas close to the microcellular sites. Comparing the traffic map of Fig. 74.9 (with traffic levels now four times higher) with the probability of blocking map of Fig. 74.13 we observe that in spite of the traffic being highest in the vicinity of the microcells, the blocking probability is low (except in the immediate vicinity of the cell sites, as previously described) because of the high capacity per metre squared offered by the microcellular cluster. The minicells on the right hand side of Fig. 74.13 are able to accommodate the increase in traffic. Unacceptably high blocking levels occur on the left half of the plot where the traffic is high and the network capacity is insufficient.

We now return to the offered traffic levels shown in Fig. 74.9. An important feature of CDMA is soft handover (SHO), a form of macro-diversity, where a MS communicates with more than one BS when its received signal levels from its current and another BS differs by a sufficiently small amount (a system parameter [9]). By communicating with more than one BS, diversity combining occurs enabling the transmit power to be decreased, the overall interference power to fall, and the capacity and coverage to increase. The benefits of SHO, only applied to our speech service, are present in all of our plots. A useful plot for an operator to have is one showing SHO regions, i.e., those areas of a city in which SHOs can be implemented. Figure 74.14 shows those areas where SHO can be achieved using two BSs, and more rarely with three BSs. Observe that there are large areas in our example where SHOs cannot be implemented.

Let us now consider the high-rate service operating at 384 kb/s. Figure 74.15 displays the required MS transmit power. The processing gain for the high-rate services is much lower than for the speech service because the energy contained in each data bit is lower. This means that more signal energy must be

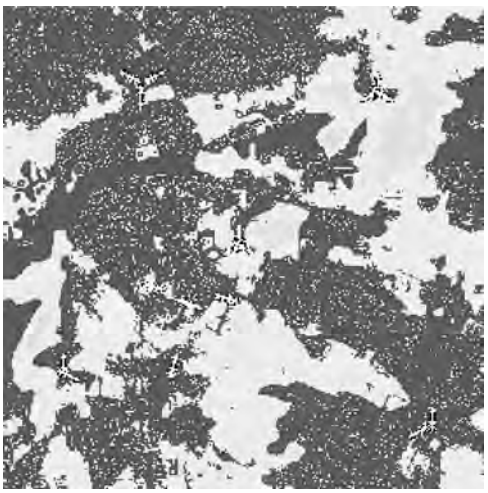


FIGURE 74.14 Soft handover regions for MSs using the speech service. The dark blue, cyan, and green areas represent regions in which an MS can be served by one, two or three BSs, respectively.

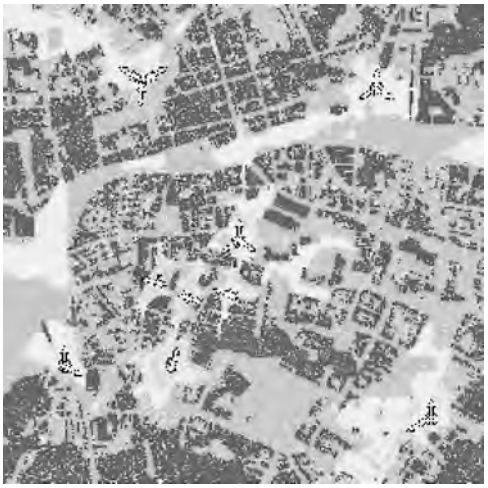


FIGURE 74.15 Required MS transmit power for the 384 kb/s high rate service. Original traffic levels shown in Fig. 74.9 apply.

provided to overcome the interference in the system, and we observe that the required MS transmit powers are generally higher than the corresponding values for the speech service displayed in Fig. 74.10.

74.7 Discussion

In the first generation (1G) FDMA networks, microcells were and remain rare, except in dense urban conurbations like Hong Kong. Microcells began to be more widely deployed in 2G networks but confined to cities. There are now small microcellular GSM BSs and they are to be found in city streets and within buildings. Sections 74.1 to 74.4 discussed microcells for highways, city streets and for buildings, while in Section 74.5 we commented on microcellular infrastructures. These FDMA and TDMA microcells are relatively straight forward to design, but in Section 74.6 we introduced microcells and oversailing minicells for cities having 3G CDMA networks. Designing these networks is difficult and we trust that the figures in Section 74.6 will give the reader an indication of the magnitude of the problem. These high user density CDMA networks providing multimedia service have not been fully deployed yet, so this brief chapter has looked back to the origins of microcells and moved forward to glimpse how they may be deployed in the future. We live in rapidly changing times.

References

1. R. Steele and V. K. Prabhu, "Mobile radio cellular structures for high user density and large data rates," *Proceedings of the IEE,* vol. Pt. F, no. 5, pp. 396–404, 1985.
2. S. T. S. Chia, R. Steele, E. Green, and A. Baran, "Propagation and bit error ratio measurements for a microcellular system," *JIERE, Supplement,* vol. 57, no. 6, pp. 5255–5266, 1987.
3. S. A. El-Dolil, W. C. Wong, and R. Steele, "Teletraffic performance of highway microcells with overlay macrocell," *IEEE Journal on Selected Areas of Communications,* vol. 7, no. 1, pp. 71–78, 1989.
4. R. Steele and M. Nofal, "Teletraffic performance of microcellular personal communication networks," *IEE Proceedings -I,* vol. 139, no. 4, pp. 448–461, 1992.
5. R. Steele, M. Nofal, and S. El-Dolil, "An adaptive algorithm for variable teletraffic demand in highway microcells," *IEE Electronic Letters,* vol. 26, no. 14, pp. 988–990, 1990.
6. E. Green, "Radio link design for microcellular systems," *British Telecom Technical Journal,* vol. 8, no. 1, pp. 85–96, 1990.
7. R. Steele and L. Hanzo, *Mobile Radio Communications,* 2nd edition, London, John Wiley, 1999.
8. R. Coombs and R. Steele, "Introducing microcells into macrocellular networks: a case study," *IEEE Transactions on Communications,* vol. 47, no. 4, pp. 568–576.
9. R. Steele, C.-C. Lee, and P. Gould, *GSM, cdmaOne and 3G Systems,* John Wiley, 2001.
10. R. Steele and J. E. B. Williams, "Third generation pcn and the intelligent multimode mobile portable," *IEE Electronics and Communications Engineering Journal,* vol. 5, no. 3, pp. 147–156, 1990.
11. J. M. Keenan and A. J. Motley, "Radio coverage in buildings," *British Telecom Technical Journal,* vol. 8, no. 1, pp. 19–24, 1990.
12. A. A. M. Saleh and R. A. Valenzula, "A statistical model for indoor multipath propagation," *IEEE Journal on Selected Areas of Communications,* vol. 5, no. 2, pp. 128–137, 1987.
13. C. Phongphanphanee, T. Mitchell, and R. Steele, "DECT coverage of indoor-outdoor environments," *IEEE Communications Letters,* vol. 3, no. 6, pp. 161–163, June 1999.
14. R. P. Merrett, A. J. Cooper, and I. C. A. Symington, "A cordless access system using radio-over-fiber techniques," in *IEEE 41st VTS Conference,* (Missouri, USA), pp. 921–924, May 1991.
15. S. Dehghan and R. Steele, "Small cell city," *IEEE Communications Magazine,* vol. 30, pp. 52–59, August 1997.

The IEEE Communications Magazine Special Issue on an update on personal communication, Vol. 30, No. 12, December 1992, provides a good introduction to microcells, particularly the paper by L. J. Greenstein et al.

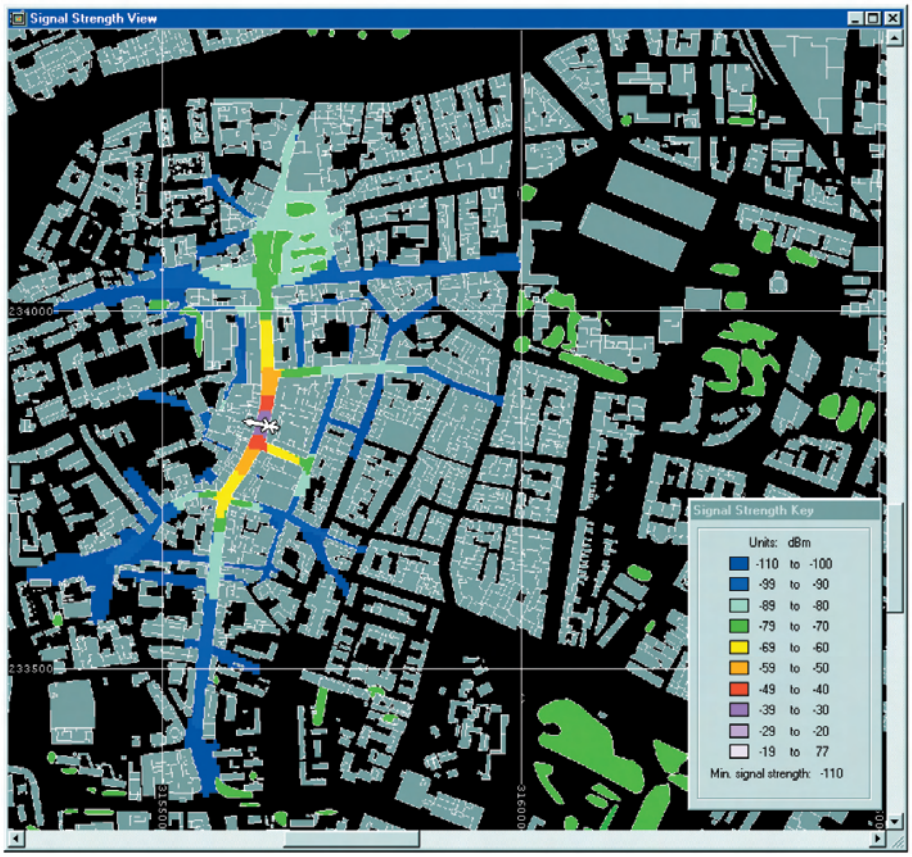


FIGURE 74.4 NP WorkPlace coverage plot of a city street microcell. The map gridsize is 500 meters.

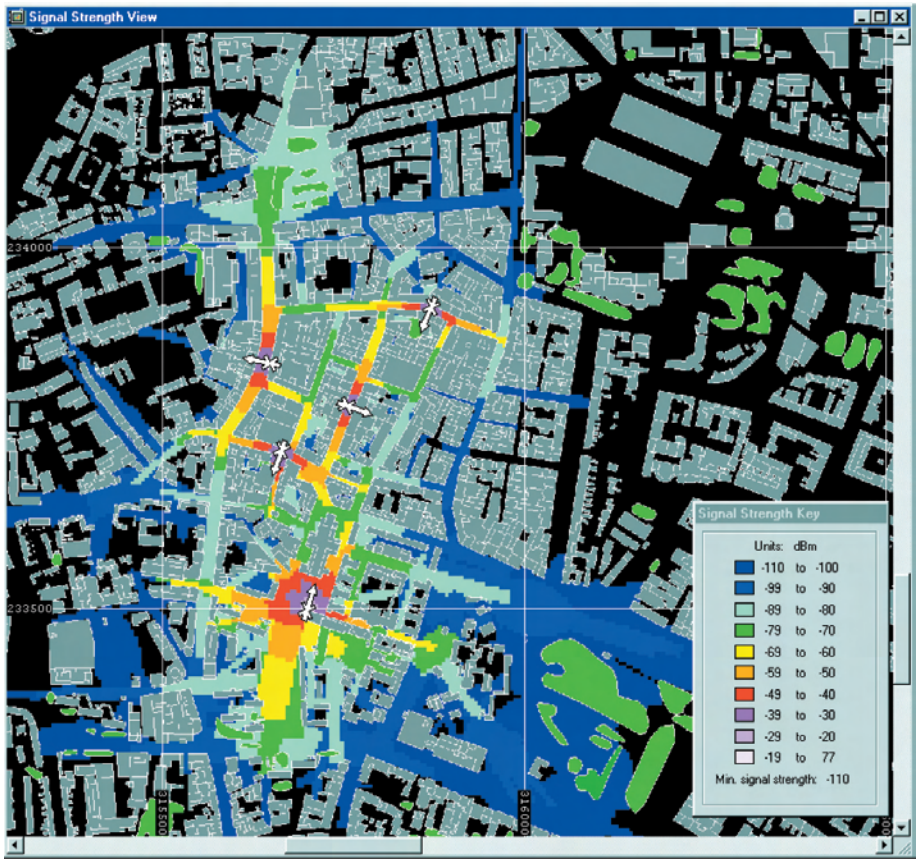


FIGURE 74.5 Cluster of microcells in a city center. The map gridsize is 500 meters.

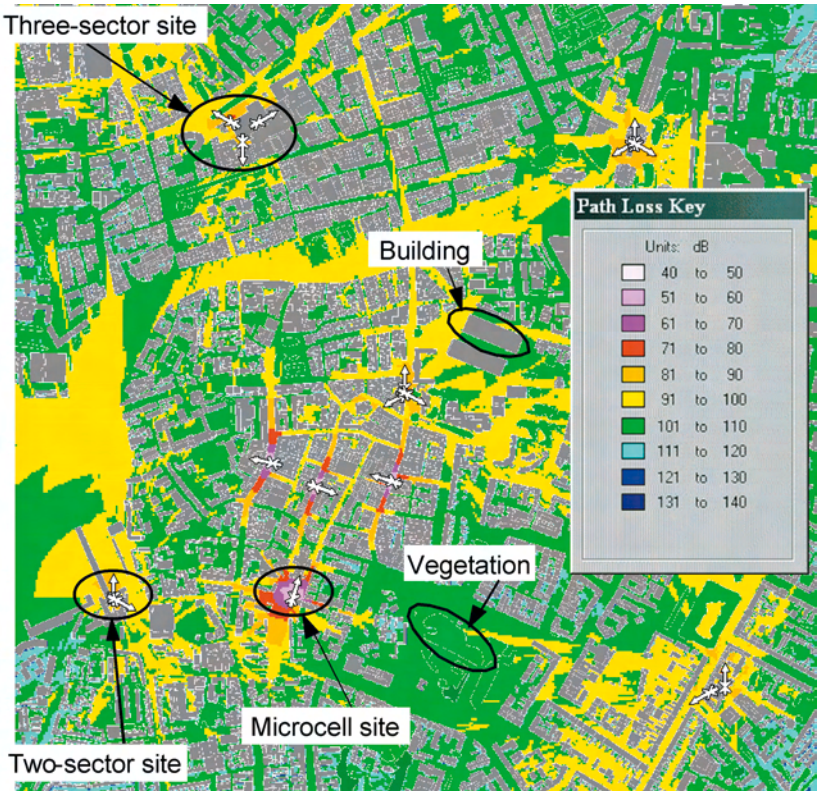


FIGURE 74.8 The minimum pathloss from all sites.

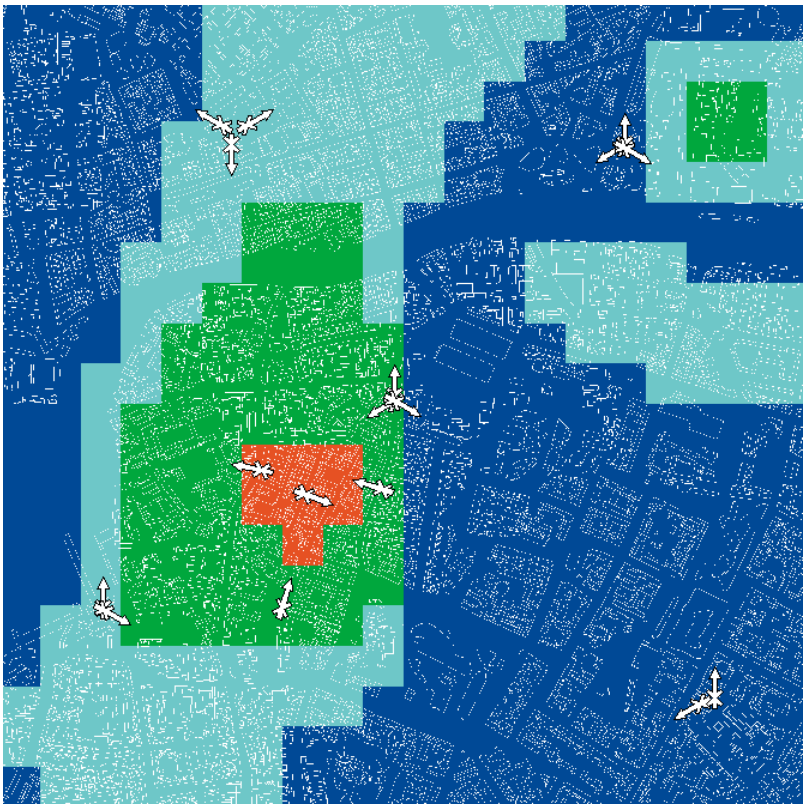


FIGURE 74.9 Variation of traffic density across the network. Red represents 128 $\mu E/m^2$, green represents 64 $\mu E/m^2$, light blue represents 32 $\mu E/m^2$, and dark blue represents 16 $\mu E/m^2$.

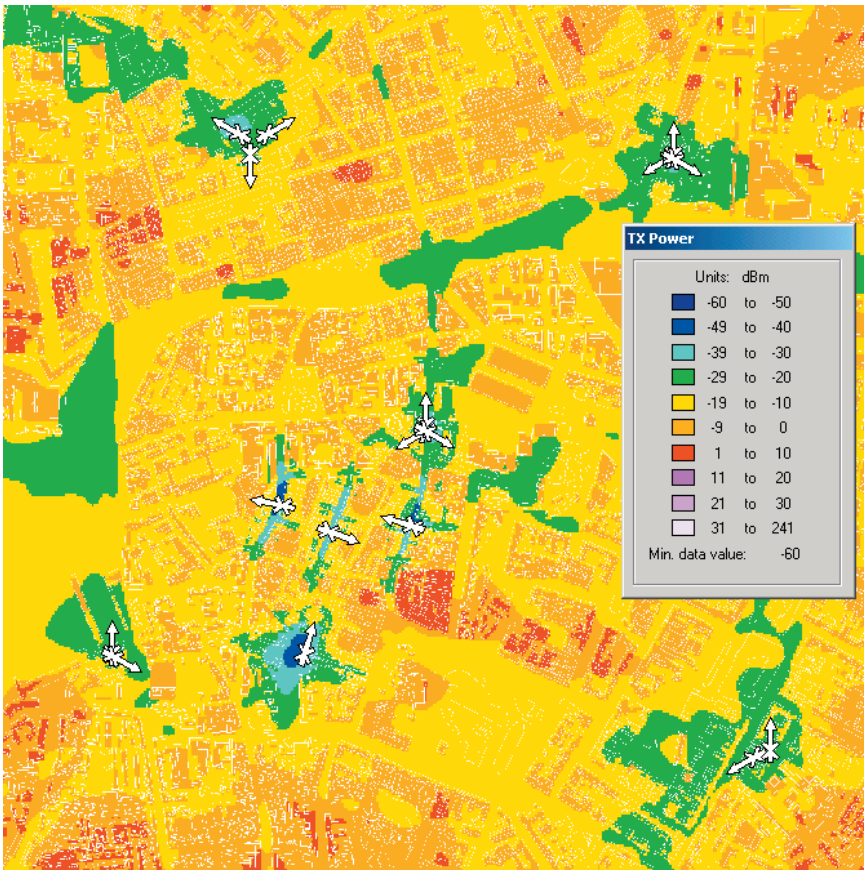


FIGURE 74.10 The required transmit power for the speech service.

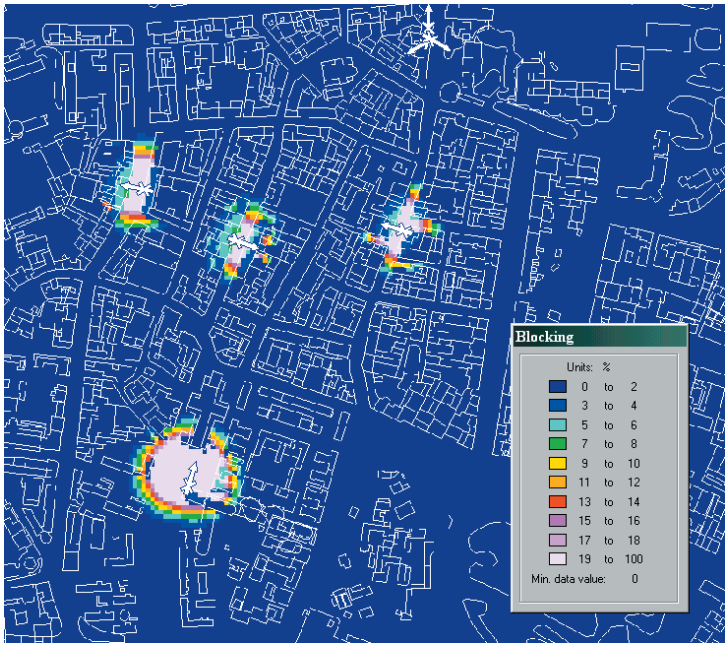


FIGURE 74.11 The probability of blocking for the speech service.

FIGURE 74.12 The required transmit power for MSs using the speech service, when the level of offered traffic has been increased by a factor of four.

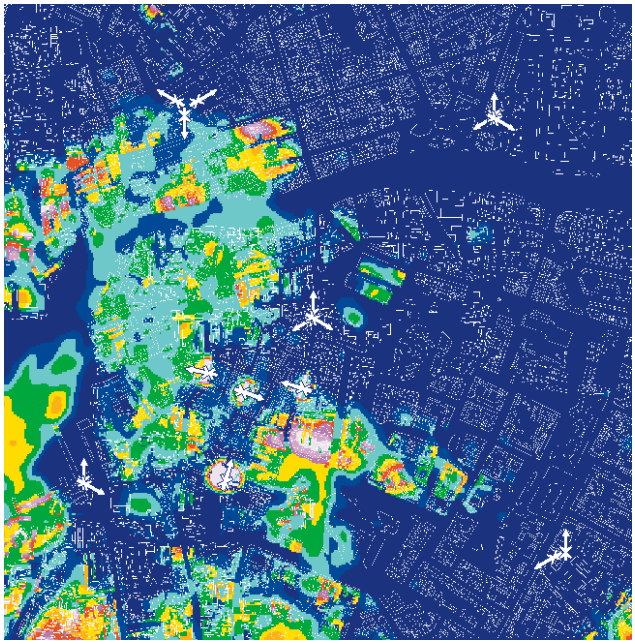


FIGURE 74.13 The probability of blocking for speech users when the offered traffic has been increased by a factor of four. For the key, see Figure 74.11.

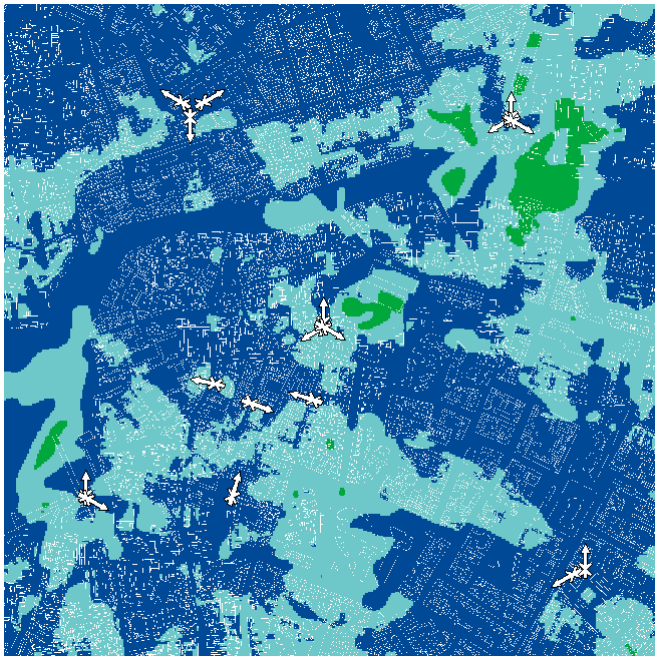


FIGURE 74.14 Soft handover regions for MSs using the speech service. The dark blue, cyan, and green areas represent regions in which an MS can be served by one, two, or three BSs, respectively.

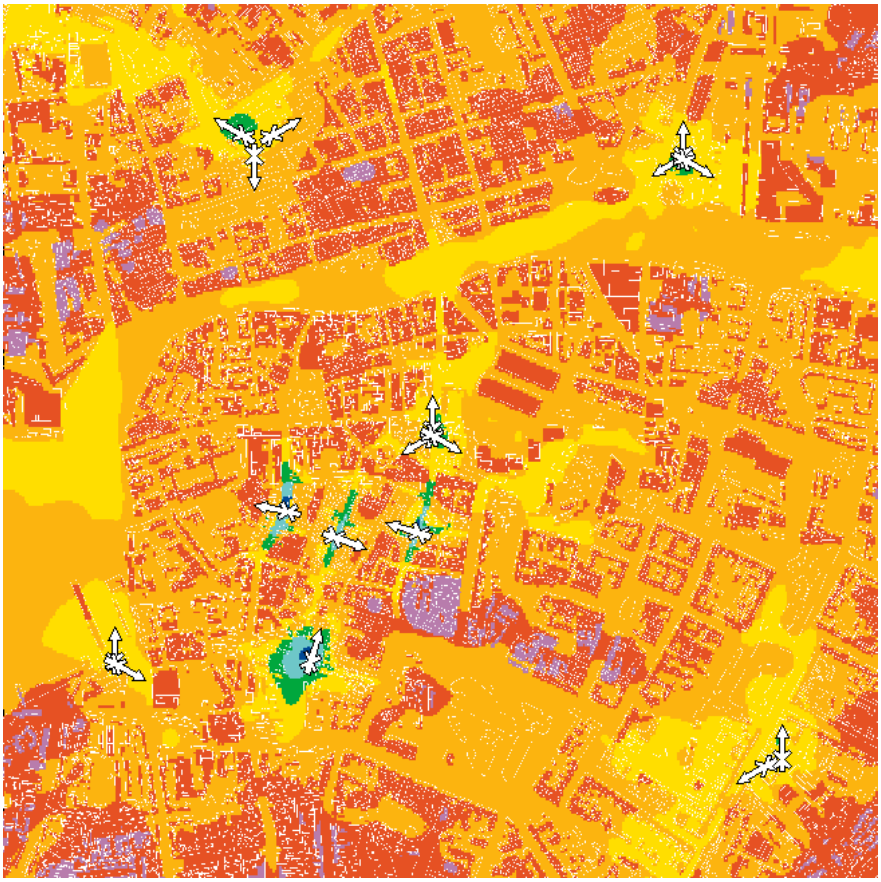


FIGURE 74.15 Required MS transmit power for the 384 kb/s high rate service. Original traffic levels shown in Figure 74.9 apply.

75

Microcellular Reuse Patterns

Michel Daoud Yacoub
FEEC/UNICAMP

Paula R. C. Gomez
FEEC/UNICAMP

This chapter addresses the basic principles underlying the construction and performance of reuse patterns applied to microcellular networks. These reuse patterns are based on a square grid and the performance of the system is assessed by means of the carrier-to-interference (C/I) ratio. Initially, the worst case condition is considered, in which the system is assumed to operate at full load and all mobiles are positioned for the highest interference situation. Following, in a more realistic condition, the mobiles are assumed to be randomly positioned within the network with the channel activity of each call connection varying in accordance with a given traffic intensity.

Amongst the possible reuse patterns (1, 2, 4, 5, 8, 9, 10, 13, 16... microcells per cluster), the performance analysis carried out here explores systems with clusters of 5, 8, 9, 10, and 13 microcells. General and closed-form formulas are presented for all possible reuse patterns. The chapter also addresses the pertinence of the use of simple but approximate formulas for the overall performance evaluation.

A substantial difference in performance for the uplink and downlink can be noticed, with the former performing less satisfactorily than the latter and with this difference diminishing as the size of the cluster increases. Some peculiarities of each reuse pattern, such as sudden degradation in the performance at certain points of the service area, are outlined. Both the 10- and the 9-microcell clusters exhibit an outstanding performance. On the other hand, if appropriately explored, a remarkable trade-off between capacity and interference may be encountered for the 5-microcell cluster.

75.1 Introduction

The expansion and the evolution of the wireless network shall be supported by an ample microcellular structure, not only to satisfy the high traffic demand in the dense urban regions but also to provide for the services requiring low mobility. The microcellular network concept is rather different from that of the macrocellular one widely employed in wireless networks. In particular, the macrocellular network makes use of a hexagonal cell array with reuse pattern being established with the supposition that reuse distances

are isotropic and that a cluster is comprised of a contiguous group of cells. In theory, the high-power sites combined with base station antennas being positioned well above the rooftops provide for a propagation symmetry, in which case, for system planning purposes, the hexagonal coverage grid has proven appropriate.

In microcellular systems, with low-power sites and antennas mounted at the street level (below the rooftops), the supposed propagation symmetry of the macrocellular network no longer applies and the hexagonal cell pattern does not make sense. The "microscopic" structure of the environment (e.g., street orientation, width of the streets, layout of the buildings, among others) constitutes a decisive element influencing the system performance.

The ubiquitous coverage of a service area based on a microcellular network will have to make use of an exceedingly greater number of base stations as compared with that of the macrocellular systems. Therefore, among the important factors to be accounted for in the microcellular system planning (cost of the site, size of the service area, etc.), the per-subscriber cost is determinant. This cost is intimately related to how efficiently the radio resources are reutilized in a given service area. The reuse efficiency depends on the interfering environment of the network and on the ability of the involved technology to cope with the interfering sources.

The studies of interference in macrocellular systems is greatly eased by the intrinsic symmetry of the problem. In the microcellular case, the inherent asymmetry due to the microscopic structures involved implies an additional complication. In such a case, the interference is dependent not only on the distance between transmitter and receiver but also, and mainly, on the line-of-sight (LOS) condition of the radio path. For instance, assuming base stations located at street intersections, mobiles on streets running radially from the base station may experience an interference pattern changing along the street as they depart from the vicinity of their serving base station, where the desired signal is strong and the relevant interfering signals are obstructed by buildings (non LOS—NLOS), and approach new intersections, where they may have an LOS condition not only to their serving base station but also to the interfering ones. The interfering situation will then follow a completely distinct pattern on the perpendicular streets. Again, the asymmetry of the problem is stressed by the traffic distribution, which is more likely to comply with an uneven configuration with the main streets accommodating more mobile users than the secondary ones.

75.2 Microcellular Reuse Patterns

A number of field measurements and investigations [1–6] show that an urban microcell service area can be reasonably well-approximated by a square diamond. Assuming a square cellular grid, as in Fig. 75.1, where R is the cell radius and $\sqrt{2}R$ [10] is the unit in the respective set of coordinates and the distance D between the centers of the cells is given by $D^2 = i^2 + j^2$, where $i = u_2 - u_1$ and $j = v_2 - v_1$ range over the integers. Considering that reuse distances are isotropic and that a cluster is a contiguous group of cells, the format of the clusters must be of a square type, whose area is given by D^2. Given that the area of the cell equals $2R^2$, the number N of cells per cluster is found to be $N = i^2 + j^2$, leading to clusters accommodating 1, 2, 4, 5, 8, 9, 10, 13, 16... cells.

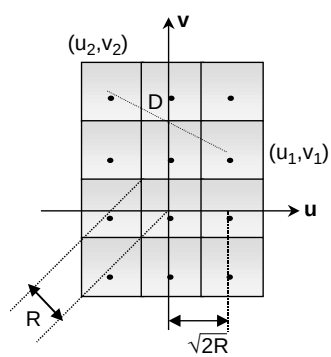


FIGURE 75.1 Square cell geometry.

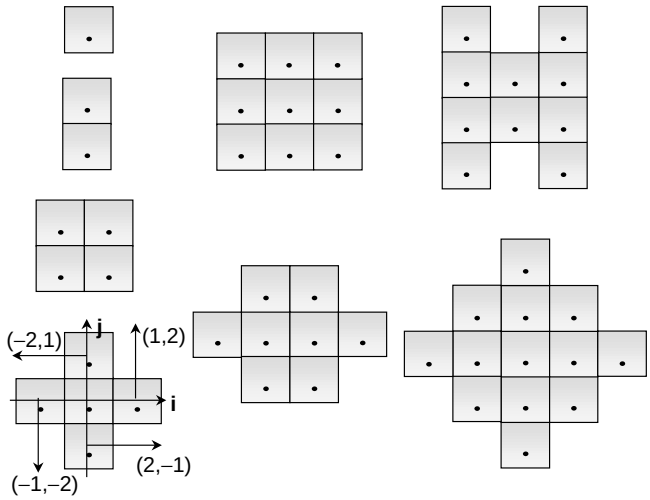


FIGURE 75.2 Some possible reuse patterns (N = 1, 2, 4, 5, 8, 9, 10, 13 microcells per cluster).

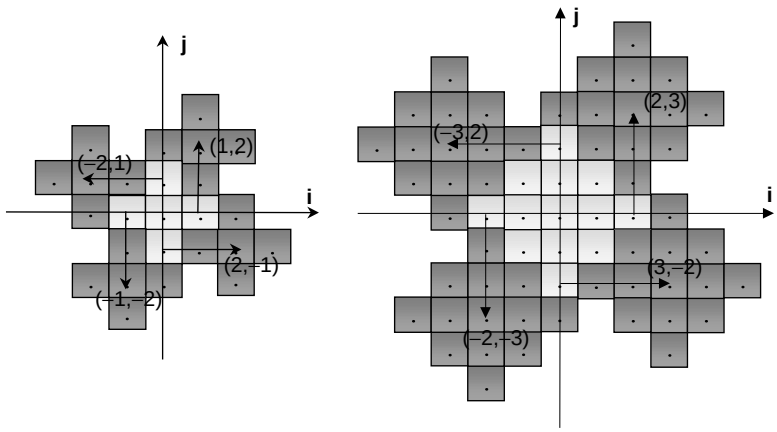


FIGURE 75.3 Partial tessellation for the reuse of 5 and 13.

The layout of the cells within the cluster is attained having as principal target *symmetry* and *compactness*. Figure 75.2 shows some of the suggested repeat patterns. The tessellation over the entire plane is then achieved by replicating the cluster in an isotropic manner. In other words, if the chosen reuse pattern is such that $i = p$ and $j = q$, then, for the reference cell located at the coordinates (0, 0), the four corresponding equidistant cocells are positioned at (p, q), (-q, p), (-p, -q), and (q, -p). Figure 75.2 illustrates this for the case $i = 1$ and $j = 2$ ($N = 5$). Figure 75.3 shows a partial tessellation for the reuses of 5 ($i = 1, j = 2$) and 13 ($i = 2, j = 3$).

In the performance analysis of the various reuse patterns, a parameter of interest is the distance between the interferers positioned at the co-cell of the Lth co-cell layer and at the target cell, the target cell taken as the reference cell. We define such a parameter as n_L and, for ease of manipulation, normalize it with respect to the cell radius, i.e., n_L is given in number of cell radii. We observe that this parameter is greatly dependent on the reuse pattern. It can be obtained by a simple visual inspection, but certainly for a very limited number of cell layers. For the general case, a more general formulation is required, and this is shown in the appendix.

75.3 Performance Analysis Model

The performance analysis to be carried out here considers a square cellular pattern with base stations positioned at every other intersection of the streets. This means that base stations are collinear and that each microcell covers a square area comprised of four 90∞-sectors, each sector corresponding to one-half of a block, with the streets running on the diagonals of this square. Figure 75.4 shows the cellular layout with respect to the streets. In Fig. 75.4, the horizontal and vertical lines represent the streets and the diagonal lines represent the microcells' borders. The central microcell is highlighted in the figure in order to evidence its location.

In order to give an insight into how the performance calculations are carried out, Figs. 75.5 and 75.6 illustrate the complete tessellation for clusters containing 5 (Fig. 75.5), 8, 9, 10, and 13 (Fig. 75.6) microcells, in which the highlighted cluster accommodates the target cell, and the other dark cells correspond to the

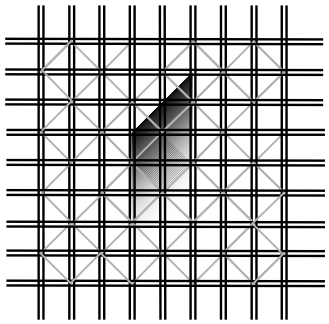


FIGURE 75.4 Cellular layout in an urban area.

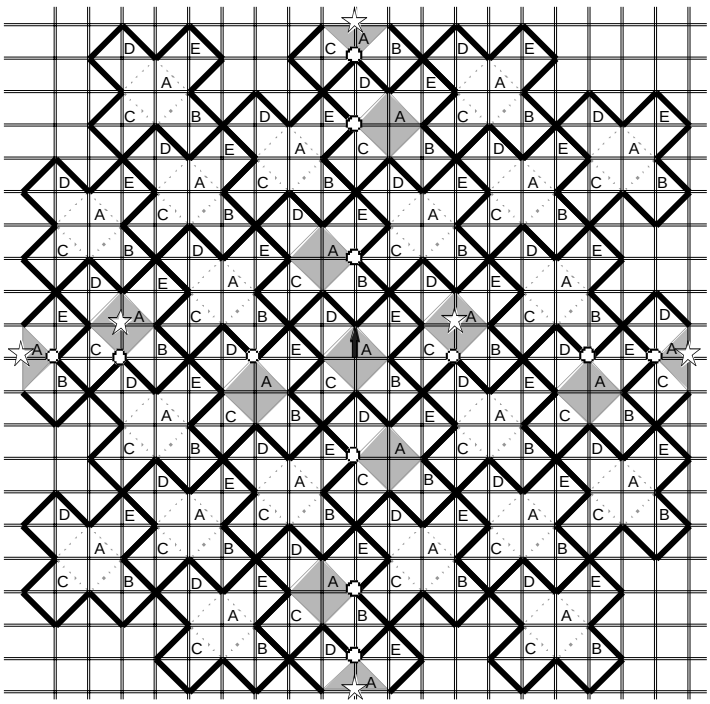


FIGURE 75.5 Five-microcell cluster tesselation—Prime Non-Collinear Group (see appendix).

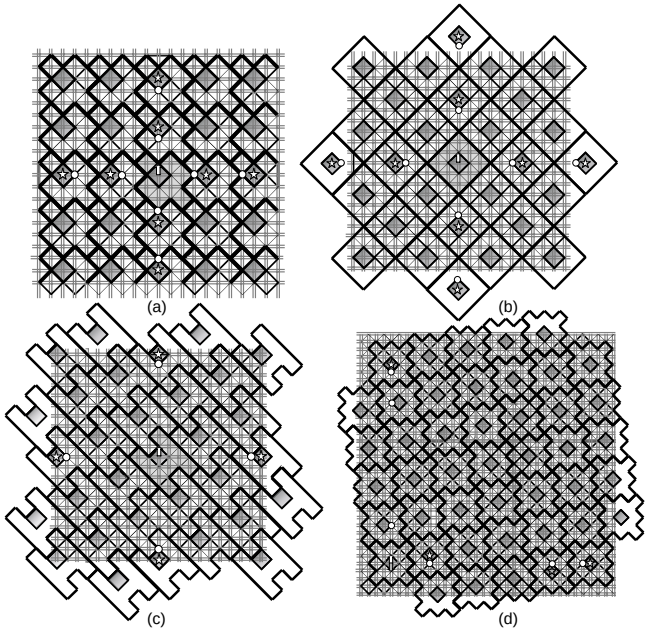


FIGURE 75.6 (a) Eight-microcell cluster tessellation—Collinear $i = j$ Group; (b) Nine-microcell cluster tessellation—Collinear $\sqrt{N} \in Z$ Group; (c) Ten-microcell cluster tessellation—Even Non-Collinear; (d) Thirteen-microcell cluster tessellation—Prime Non-Collinear Group (see appendix).

co-microcells that, at a certain time, may interfere with the mobile or base station of interest. Within a microcellular structure, distinct situations are found that affect in a different manner the performance of the downlink and of the uplink. In general, the set of microcells affecting the downlink constitutes a subset of those influencing the uplink. In Figs. 75.5 and 75.6, the stars indicate the sites contributing to the *C*/*I* performance of the downlink, whereas the circles point the worst case location of the mobile affecting the *C*/*I* performance of the uplink. The cluster attribute (collinear, non-collinear, etc.) indicated in the captions of these figures are defined in the appendix.

It is noteworthy that some of the patterns tessellate into staggered configurations with the closer interferers being either completely obstructed or obstructed for most of the time with an LOS interferer appearing many blocks away. It is also worth emphasizing that for clusters with a prime number of constituent cells, as is the case of the 5-cell cluster of Fig. 75.5, the base stations that interfere with the target mobile in the downlink change as the mobile moves along the street.

Concerning the radio coverage aspects, it is well known [1–6] that the LOS and the NLOS modes of propagation are very distinct from each other. The respective propagation losses at a distance d from the base station and for a wavelength λ, a transmitting antenna height h_t, and a receiving antenna height h_r are then given by

$$L_{LOS} = \frac{K_{LOS}}{d^2}\left[1 + \left(\frac{d}{d_B}\right)^2\right]^{-} \tag{75.1}$$

$$L_{NLOS} = \frac{K_{NLOS}}{d^{\alpha}} \tag{75.2}$$

where $K_{LOS} = (\lambda/4\pi)^2$, $d_B = 4h_t h_r/\lambda$ is the breakpoint distance, $K_{NLOS} = 0.16$ for 900 MHz and 0.0015 for 2 GHz, and $\alpha = 4.3$ for 900 MHz and $\alpha = 3.8$ for 2 GHz [4].

The following two sections of this chapter analyze the performance of a microcellular network system for the worst case condition and also for a more realistic condition. In the first case, the system is assumed to operate at full load and all mobiles are positioned for the highest interference situation. In the second case, the mobiles are assumed to be randomly positioned within the network with the channel activity of each call connection varying in accordance with a given traffic intensity. Because the contribution of the obstructed interferers to the overall C/I performance is negligible if compared to that of the LOS interferers, only the LOS condition of the interferers is used for the calculations. In such a case, the results for the worst case condition express the lower bound performance of the system whereas the realistic case condition express the upper bound performance.

75.4 Worst Case Condition

Carrier-to-Interference Ratio

Defining r as the serving base station-to-mobile station distance normalized with respect to the cell radius ($0 < r \leq 1$) and k as the ratio between the cell radius and the breakpoint distance ($k \geq 0$), it is possible to show that, for the interferers positioned at the worst case condition, the uplink C/I parameter is given by [7]

$$\frac{C}{I} = \frac{[1 + (rk)^2]^{-1}}{4r^2 \sum_{L=1}^{\infty} n_L^{-2} [1 + (n_L k)^2]^{-1}} \tag{75.3}$$

where $n_L = 1, 2, 3...$ is the distance between the interferers at the Lth layer and the target cell site, given in number of units of the cell radius. The parameter n_L is dependent on the reuse pattern, as shown in the appendix.

In the same way, a C/I can be found for the downlink. However, this ratio depends on the position of the target mobile with respect to its serving base station because, as the mobile moves from the cell center toward the next corner, it acquires a new LOS, and it may find new interfering base stations. Therefore, for the downlink and for the mobile at the vicinity of its serving base station, more specifically, at the intersection of the streets ($r \leq$ normalized distance from the site to the beginning of the street), the C/I ratio [7] is

$$\frac{C}{I} = \frac{[1 + (rk)^2]^{-1}}{r^2 \sum_{L=1}^{\infty} \{(n_L + r)^{-2} [1 + (n_L + r)^2 k^2]^{-1} + (n_L - r)^{-2} [1 + (n_L - r)^2 k^2]^{-1} + 2(n_L^2 + r^2)^{-1} [1 + (n_L^2 + r^2) k^2]^{-1}\}} \tag{75.4}$$

Away from the vicinity of the serving base station, which corresponds to most of the path, the C/I ratio [7] is

$$\frac{C}{I} = \frac{[1 + (rk)^2]^{-1}}{r^2 \sum_{L=1}^{\infty} \{(n_L + r)^{-2} [1 + (n_L + r)^2 k^2]^{-1} + (n_L - r)^{-2} [1 + (n_L - r)^2 k^2]^{-1}\}} \tag{75.5}$$

At the border of the cell (very end of the corner), new interferers may appear in an LOS condition. However, this is not the case for all reuse patterns. This phenomenon only happens for clusters with a prime number of cells, as, for example, in the case of 5-cell clusters as illustrated in Fig. 75.5. Hence, for clusters with a prime number of cells and the mobile away from its serving base station ($1 - r \leq$ normalized distance from the site

to the beginning of the street), the C/I ratio [7] is

$$\frac{C}{I} = \frac{[1+(rk)^2]^{-1}}{r^2[\sum_{L=1}^{\infty}\{(n_L+r)^{-2}[1+(n_L+r)^2k^2]^{-1}+(n_L-r)^{-2}[1+(n_L-r)^2k^2]^{-1}\}+\sum_{L'=1}^{\infty}\{(n_{L'}'^2+\bar{r}^2)^{-1}[1+(n_{L'}'^2+\bar{r}^2)(k^2)]^{-1}\}]} \tag{75.6}$$

where $\bar{r} = 1 - r$ and n_L' is defined in the appendix.

An Application Example

This section illustrates the C/I performance for clusters with 5, 8, 9, 10, and 13 microcells. The performance has been evaluated having the central microcell as the target cell and for the mobile subscriber departing from the center toward its edge, as shown by the arrow in the respective cell in Fig. 75.5. Figure 75.5 also shows, in gray, the co-microcells that, at a certain time, may interfere with the wanted mobile in an LOS condition.

For the numerical results, the calculations consider the radius of the microcell as 100 m, a street width of 15 m, the transmitter and receiver antennas heights, respectively, equal to 4 m and 1.5 m, an operation frequency of 890 MHz, which leads to $k = 1.405$, and a network consisting of an infinite number of cells (in practice, 600 layers of interfering cells).

Figures 75.7a and 75.7b show, respectively, the uplink and downlink performances for the cases of 5-, 8-, 9-, 10-, and 13-microcell clusters as a function of the normalized distance. In general, the bigger the cluster the better the carrier-to-interference ratio, as expected. However, the 5-microcell cluster has a notably outstanding behavior, with its C/I coinciding with that of the 8-microcell cluster for the uplink (lower curve in Fig. 75.7a) and with that of the 10-microcell cluster for the downlink for most of the extension of the path (curve below the upper one in Fig. 75.7b), with the separation of the curves in the latter occurring at the edge of the microcell, where two new interferers appear in an LOS condition.

There is a significant difference between the performance for the uplink and downlink, this difference becoming progressively smaller with the increase of the size of the cluster. This can be better observed in Fig. 75.8, where the 5- and 10-microcell clusters are compared.

It is interesting to examine the influence of the number of interfering layers in the C/I performance. For this purpose, we analyze the performance of a one-layer network ($L = 1$ in Eqs. (75.3) through (75.6)). Figures 75.9a and 75.9b show the performances for the uplink and downlink as a function of the normalized distance to the base station using both the exact (infinite number of layers) and the simplified (one layer) methods for clusters of 5 and 8 cells, respectively. The dotted lines correspond to the results for the infinite number of interferers case, and the solid lines represent the results for the simplified calculations. Analyzing the graphs and the numerical results, one observes that the difference between

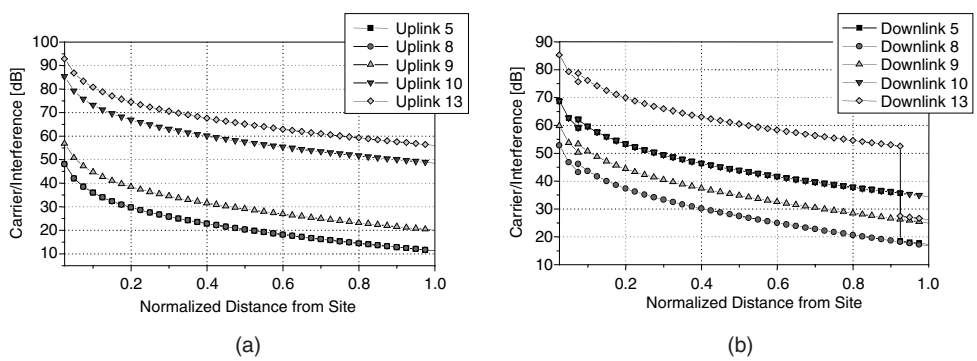


FIGURE 75.7 Carrier-to-interference ratio as a function of normalized distance: (a) uplink; (b) downlink.

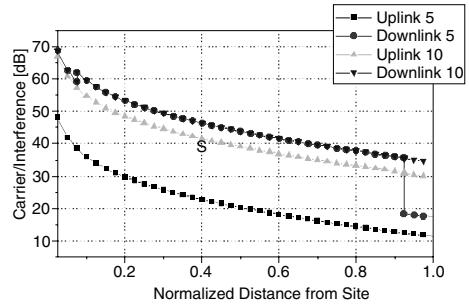


FIGURE 75.8 Carrier-to-interference ratio as a function of normalized distance for uplink and downlink compared.

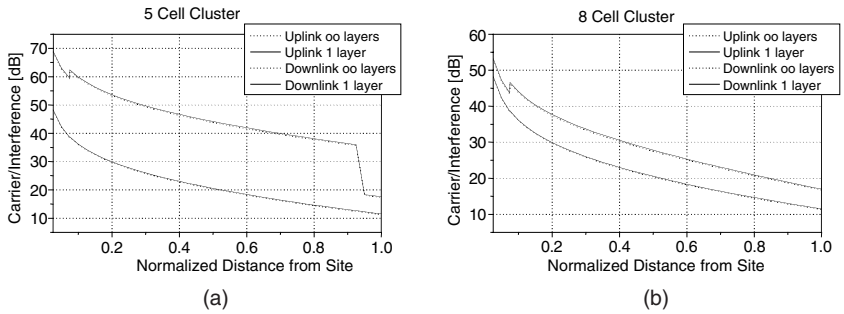


FIGURE 75.9 Performance considering an infinite number of interferers and a single layer of interferers for (a) 5-cell clusters and (b) 8-cell clusters.

the C/I ratio for an infinite-cell network and for a one-interfering-layer network is negligible. This conclusion also applies to the other patterns, with the biggest difference found in both methods for all reuse patterns being less than 0.35 dB.

Therefore, for the worst case condition, the C/I ratio can be estimated by considering only the interfering layer that is closest to the target cell. In such a case, Eqs. (75.3) through (75.6) reduce to very simple formulas. However, as will be discussed in the following section, for a more realistic condition more layers must be considered in the performance analyses.

75.5 A More Realistic Case Condition

Carrier-to-Interference Ratio

This section considers the interferer to be randomly positioned within the microcell network and the channel to be active according to a given traffic intensity. Assume a uniform probability distribution for the location of a mobile within a cell. Let i be the number of interferers and $i_{ACT\&LOS}$ $(0 \le i_{ACT\&LOS} \le i)$ a subset of the interferers that are active and in an LOS condition, and $i_{NACT\ or\ NLOS} = i - i_{ACT\&LOS}$ those that are not active or that are in an obstructed situation. Let p_{LOS} be the probability of a mobile being in an LOS condition, p_{ACT} the probability of a channel being active, and $p_{ACT\&LOS}$ the probability of having $i_{ACT\&LOS}$ interferers simultaneously active and in an LOS condition. Let $E[C/I|i_{ACT\&LOS}]$ be the carrier-to-interference ratio estimated in such a case. The overall mean carrier-to-interference ratio is, therefore [8],

$$E[C/I] = \sum_{i_{ACT\&LOS}=0}^{i} E[C/I\,|\,i_{ACT\&LOS}]\,p_{ACT\&LOS} \tag{75.7}$$

where

$$p_{ACT\&LOS} = \binom{i}{i_{ACT\&LOS}} (p_{LOS} \times p_{ACT})^{i_{ACT\&LOS}} (1 - p_{LOS} \times p_{ACT})^{i - i_{ACT\&LOS}} \tag{75.8}$$

or, equivalently,

$$p_{ACT\&LOS} = \binom{i^{(1)}}{i^{(1)}_{ACT\&LOS}} (p^{(1)}_{LOS} \times p_{ACT})^{i^{(1)}_{ACT\&LOS}} \cdot (1 - p^{(1)}_{LOS} \times p_{ACT})^{1 - i^{(1)}_{ACT\&LOS}}$$
$$\times \binom{i^{(2)}}{i^{(2)}_{ACT\&LOS}} (p^{(2)}_{LOS} \times p_{ACT})^{i^{(2)}_{ACT\&LOS}} \cdot (1 - p^{(2)}_{LOS} \times p_{ACT})^{i^{(2)}_{ACT\&LOS}} \tag{75.9}$$

$$i^{(1)} + i^{(2)} = i \tag{75.10}$$

$$i^{(1)}_{ACT\&LOS} + i^{(2)}_{ACT\&LOS} = i_{ACT\&LOS} \tag{75.11}$$

There is a unique value for p_{LOS} for clusters with a non-prime number of microcells, as can be observed from the cellular layout. Therefore, Eq. (75.8) is used to evaluate $p_{ACT\&LOS}$ in these cases. On the other hand, for clusters with a prime number of microcells, there is more than one possible value for p_{LOS}. From Fig. 75.5, one can observe that mobile interferers affecting the uplink (circles) from the first two layers are in an LOS condition only when positioned at the corner, whereas mobiles from the third layer are in an LOS condition at any point along the entire street. This is the case for all clusters with a prime number of cells. Therefore, calculations of the performance of the uplink for such cases make use of Eq. (75.9) for the value of $p_{ACT\&LOS}$.

An Application Example

The results to be shown here consider clusters of 5 and 8 cells for the parameters as given before. The probability p_{LOS} is different for each tessellation because for some of them there are co-microcells where the street offers an LOS condition at any of its points, whereas for others, the LOS condition is offered only at the corner of the street of the co-microcell. The calculations of the C/I ratio shown here assume each microcell to have 8 channels and a grade of service of 2% and they are performed for an "infinite" number of microcells. Certainly, performing the calculations outlined in Eq. (75.7) for an infinite number of interferers is impracticable. Limiting the number of interferers may compromise the results. For practical purposes, the interferers are then divided into two groups: those belonging to the layers closer to the target cell (inner layers) and those belonging to the other layers (outer layers). Again for practical purposes, only the former are given random properties, whereas the latter are positioned for the worst case condition. It has been found [9] that the minimum number of layers within the inner layers necessary to yield statistically reliable results is 7. As before, the calculations take into account 600 layers of interfering cells.

Figures 75.10a and 75.10b show the performances for the uplink and the downlink as a function of the normalized distance to the base station for clusters with 5 and 8 microcells, respectively, for the two limiting cases: for the inner layers containing one layer only (dotted lines) and for the inner layers containing seven layers (solid lines). Note that a remarkable difference between the two cases is found. This implies that, whereas the worst case condition can be easily modeled by means of a one-layer network, a more realistic condition can only be analyzed if a significantly greater number of layers is considered.

Figure 75.11 compares the lower bound performance and the upper bound performance for 5-cell clusters and 8-cell clusters. The plots show an outstanding increase in performance for both the uplink and downlink cases. It is noteworthy that this increase is enhanced for the case of 5-cell clusters, and this applies to all

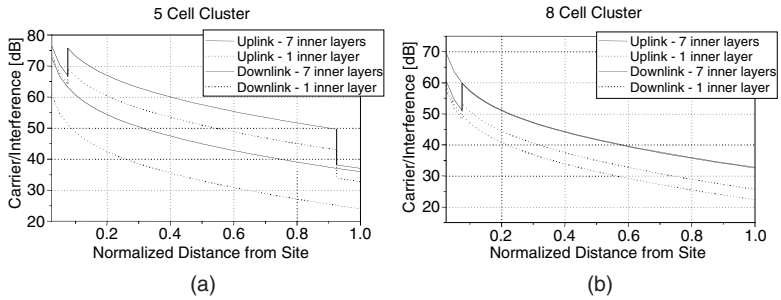


FIGURE 75.10 Upper bound performance as a function of normalized distance for the uplink and downlink for (a) 5-cell clusters; (b) 8-cell clusters.

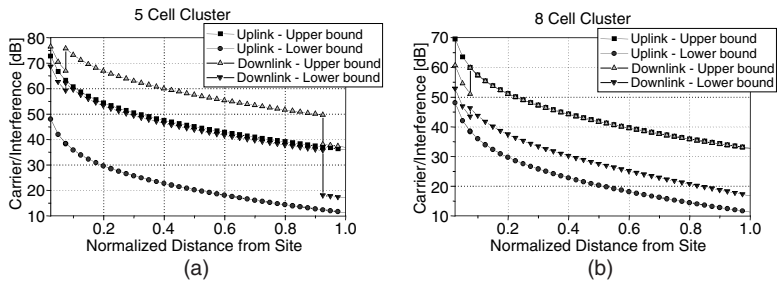


FIGURE 75.11 Carrier-to-interference ratio as a function of normalized distance for uplinks and downlinks for lower and upper bound performances compared: (a) 5-cell clusters; (b) 8-cell clusters.

clusters with prime number of cells. The reason for this is that, for clusters with prime number of cells, most of the co-microcells have only their corners in LOS, whereas the clusters with non-prime number of cells have the co-microcells collinear with the target cell, therefore, having a whole street in LOS. Special attention should be paid to the enhancement in efficiency at the corner opposite the serving base station for the cases of clusters with a prime number of cells. Those points showed a poor *C*/*I* ratio for the worst case, and now, considering a more realistic situation, they show a drastic improvement.

75.6 Conclusions

This chapter addressed the basic principles underlying the construction and performance of reuse patterns applied to microcellular networks. An initial assessment of these patterns, based on a square grid, was carried out by means of the carrier-to-interference ratio for the mobiles positioned at the worst case condition and the system operating at a full load. It is noteworthy that the worst case condition occurs with probability of $(p_{ACT} \times p_{LOS})^{4L}$, where L is the number of interfering layers, p_{ACT} is the probability of channel activity, and p_{LOS} is the probability of LOS condition. Roughly speaking, for the geometry explored here, the worst case condition occurs with probability $(3 \times 10^{-5})^{L} \times (p_{ACT})^{4L}$. Therefore, by taking into account the random distribution of the users and the channel occupancy, the mean *C*/*I* ratio is expected to improve drastically. This motivates the examination of a more realistic condition in which the mobiles are assumed to be randomly positioned within the network and the channel activity of each call connection varies in accordance with a given traffic intensity.

Among the possible reuse patterns (1, 2, 4, 5, 8, 9, 10, 13, 16... microcells per cluster), the performance analysis considered systems with clusters of 5, 8, 9, 10, and 13 microcells. General and closed-form formulas were presented for all possible reuse patterns for all cases.

When comparing the performance for the worst case condition and the more realistic condition, one can observe an outstanding increase in the carrier-to-interference ratio for the latter case. An analysis of

a microcellular system can be made by considering the results of the equations presented for the worst case condition as the lower bound performance and the results of the equations presented for the more realistic condition as the upper bound performance.

Finally, the chapter addressed the pertinence of the use of simple but approximate formulas for the overall performance evaluation. The graphs showed a remarkable similarity in the C/I ratio for the exact and the approximate methods when considering the worst case condition, from which it can be inferred that the first layer of interferers play the most influential role in the overall performance. On the other hand, if more realistic conditions are to be considered, then this influence just mentioned is not valid. The condition that yields statistically reliable results was discussed. It was found that convergence occurs for the number of interfering layers approaching seven. It could be observed that the more realistic conditions drastically improves the performance of the system as compared to the worst case condition. Both the 10- and the 9-microcell clusters exhibit an outstanding performance. On the other hand, if appropriately explored, a remarkable trade-off between capacity and interference may be encountered for the 5-microcell cluster.

Appendix

General Formulation for the Positioning of the Interferers

In the performance analysis of the various reuse patterns, a parameter of interest is the distance between the interferers at the Lth layer and the target cell. We define such a parameter as n_L and, for ease of manipulation, normalize it with respect to the cell radius, i.e., n_L is given in number of cell radii. We observe that this parameter is greatly dependent on the reuse pattern. It can be obtained by a simple visual inspection, but certainly for a very limited number of cell layers. For the general case, a more general formulation is required, and this is shown here.

The reuse patterns of the square grid cell configurations can be divided into four different groups according to the shape into which their cells tesselate [9]. Figures 75.5 and 75.6 illustrate samples of some groups. Given that the number of cells per cluster is defined by $N = i^2 + j^2$, where i and j are integers, the classification of reuse patterns for square cells is as follows:

- Collinear $\begin{cases} \bullet\ \sqrt{N} \in Z, \text{ i.e., } i = 0 \text{ or } j = 0 \quad \text{(square clusters)} \\ \qquad\qquad \text{or} \\ \bullet\ i = j, \text{ i.e., } N = x^2 \text{ where } x \in Z \text{ and } Z \text{ defines the integer set} \end{cases}$
- Non-Collinear $\begin{cases} \bullet\ N \text{ is prime} \\ \bullet\ N \text{ is non-prime} \end{cases}$.

Figures 75.6a and 75.6b show the tessellations of the first case (collinear). From the non-collinear tessellation group, there are three different subgroups: clusters with a prime number of cells, clusters with an even number of cells, and clusters with an odd and non-prime number of cells. The first of these subgroups is represented by a 5-cell cluster tessellation, as in Fig. 75.5, and by the 13-cell cluster tesselation, as in Fig. 75.6d, and is characterized by the appearance of new interferers as the mobile reaches the border of its serving base station. A member of the group of non-collinear tessellation with an even number of cells per cluster is illustrated in Fig. 75.6c.

The relations below show, for each of the groups defined above, the distance between the target cell and its interferers for the uplink and for the downlink, for the worst case condition. The minimum distance between collinear co-cells is represented by the variable d, given in number of units of the cell radius. In these relations, L is the number of the interfering layer and BS stands for base station.

1. ***Collinear***
   - $d = 2i$
   - *obs: when* $\sqrt{N} \in Z$, *if* $i = 0$, $d = 2j$
   - Uplink: $n_L = 2iL - 1$
   - Downlink: $\begin{cases} \textit{Near BS}: n_L = 2iL \\ \textit{Far from BS}: n_L = 2iL \end{cases}$
2. ***Even Non-Collinear***
   - $d = N$
   - Uplink: $n_L = N{\cdot}L - 1$
   - Downlink: $\begin{cases} \textit{Near BS}: n_L = N{\cdot}L \\ \textit{Far from BS}: n_L = N{\cdot}L \end{cases}$
3. ***Odd and Non-Prime Non-Collinear***
   - $\forall N = rN_R$
   - $n_L = r \cdot n_{L(R)}$

     where $r \in Z$, $N_R < N$, N_R = *number of cells per cluster of another reuse pattern R, and* $n_{L(R)}$ *is the distance between the interferers at the Lth layer and the target cell for reuse pattern R*
4. ***Prime Non-Collinear***
   - $d = 2N$
   - Uplink: $\begin{cases} n_{3(L-1)+1} = p + 2N(L-1) \\ n_{3(L-1)+2} = (2N-p) + 2N(L-1) \\ n_{3(L-1)+3} = 2NL - 1 \end{cases}$

     where p designates the prime number immediately prior to N
   - Downlink: $\begin{cases} \textit{Near BS}: n_L = 2NL \\ \textit{Far from BS}: n_L = 2NL \\ \textit{Near cell border}: \begin{cases} n_L = 2NL \\ n'_{2(L-1)+1} = p + 2N(L-1) \\ n'_{2(L-1)+2} = (2N-p) + 2N(L-1) \end{cases} \end{cases}$

References

1. J.H. Whitteker, Measurements of path loss at 910 MHz for proposed microcell urban mobile systems, *IEEE Trans. Veh. Technol.,* 37, 125, 1988.
2. A.J. Rustako, V. Erceg, R.S. Roman, T.M. Willis, and J. Ling, Measurements of microcellular propagation loss at 6 GHz and 2 GHz over nonline-of-sight paths in the city of Boston, in GLOBECOM'95 Conference, Singapore, 758.
3. V. Erceg, A.J. Rustako, Jr., and R.S. Roman, Diffraction around corners and its effects of the microcell coverage area in urban and suburban environments at 900 MHz, 2 GHz, and 6 GHz, *IEEE Trans. Veh. Technol.,* 43, 762, 1994.
4. M.V. Clark, V. Erceg, and L.J. Greenstein, Reuse efficiency in urban microcellular networks, *IEEE Trans. Veh. Technol.,* 46, 279, 1994.
5. A. Goldsmith and L.J. Greenstein, A measurement-based model for predicting coverage areas of urban microcells, *IEEE J. Select. Areas Commun.,* 11, 1013, 1993.
6. V. Erceg et al., Urban/suburban out-of-sight propagation modeling, *IEEE Commun. Mag.,* 30, 56, 1992.

7. M.D. Yacoub, A.F. Toledo, P.R.C. Gomez, P.J.E. Jeszensky, and E.B.R. Féris, Reuse pattern for microcellular networks, *Int. J. Wireless Inf. Networks,* 6(1), 1, 1999.
8. P.R.C. Gomez, M.D. Yacoub, and A.F. Toledo, Performance of a microcellular network in a more realistic condition, The 11th IEEE International Symposium on Personal Indoor and Mobile Radio Communications, 2000.
9. P.R.C. Gomez, M.D. Yacoub, and A.F. Toledo, Interference in wireless networks with a new reuse pattern, Brazilian Telecommunications Symposium (Simpósio Brasileiro de Telecomunicações), SBT99, Vitória, ES, Brazil, 1999 (in Portuguese).
10. M.D. Yacoub, *Wireless Technology,* CRC Press, Boca Raton, FL, 2001.

76

Fixed and Dynamic Channel Assignment

Bijan Jabbari
George Mason University

76.1 Introduction

One of the important aspect of frequency reuse-based cellular radio as compared to early land mobile telephone systems is the potential for dynamic allocation of channels to traffic demand. This fact had been recognized from the early days of research (e.g., see [18, Chapter 7], [7,8]) in this field. With the emergence of wireless personal communications and use of microcell with nonuniform traffic, radio resource assignment becomes essential to network operation and largely determines the available spectrum efficiency. The primary reason for this lies in the use of microcell in dense urban areas where distinct differences exist as compared to large cell systems due to radio propagation and fading effects that affect the interference conditions.

In this chapter, we will first review the channel reuse constraint and then describe methods to accomplish the assignment. Subsequently, we will consider variations of fixed channel assignment and discuss dynamic resource assignment. Finally, we will briefly discuss the traffic modeling aspect.

76.2 The Resource Assignment Problem

The resources in a wireless cellular network are derived either in frequency division multiple access (FDMA), time division multiple access (TDMA), or joint frequency–time (MC-TDMA) [1]. In these channel derivation techniques, the frequency reuse concept is used throughout the service areas comprised of cells and microcells. The same channel is used by distinct terminals in different cells, with the only constraint of meeting a given interference threshold. In spread spectrum multiple access (SSMA) such as the implemented code division multiple access (CDMA) system (IS-95) [24] each subscriber spreads its transmitted signal over the same frequency band by using a pseudorandom sequence simultaneously. As any active channel is influenced by the others, a new channel can be set up only if the overall interference is below a given threshold. Thus, the problem of resource assignment in CDMA relates to transmission power control in forward (base station to mobile terminal) and reverse (mobile terminal to base station) channels. Of course, the problem of power control applies to TDMA and FDMA as well, but not to the extent that it impacts the capacity of CDMA. Here we first focus on time- and

frequency-based access methods. We subsequently present some preliminary discussion on spread spectrum CDMA systems.

Fixed channel assignment (FCA) and dynamic channel assignment (DCA) techniques are the two extremes of allocating radio channels to mobile subscribers. For a specific grade of service and quality of transmission, the assignment scheme provides a tradeoff between spectrum utilization and implementation complexity. The performance parameters from a radio resource assignment point of view are interference constraints (quality of transmission link), probability of call blocking (grade of service), and the system capacity (spectrum utilization) described by busy hour erlang traffic that can be carried by the network. In a cellular system, however, there exist other functions, such as handoff and its execution or radio access control. These functions may be facilitated by the use of specific assignment schemes and, therefore, they should be considered in such a tradeoff [14].

The problem of channel assignment can be described as the following: Given a set of channels derived from the specified spectrum, assign the channels and their transmission power such that for every set of assigned channels to cell i, $(C/I)_i > (C/I)_0$. Here, $(C/I)_0$ represents the minimum allowed carrier to interference and $(C/I)_i$ represents carrier to interference at cell i.

76.3 Fixed Channel Assignment

In fixed channel assignment the interference constraints are ensured by a frequency plan independent of the number and location of active mobiles. Each cell is then assigned a fixed number of carriers, dependent on the traffic density and cell size. The corresponding frequency plan remains fixed on a long-term basis. In reconfigurable FCA (sometimes referred to as flexible FCA), however, it is possible to reconfigure the frequency plan periodically in response to near/medium term changes in predicted traffic demand.

In FCA, for a given set of communications system parameters, $(C/I)_0$ relates to a specific quality of transmission link (e.g., probability of bit error or voice quality). This parameter in turn relates to the number of channel sets [18] (or cluster size) given by $K = 1/3\ (D/R)^2$. Thus, the ratio D/R is determined by $(C/I)_0$. Here D is the cochannel reuse distance and R is the cell radius. For example, in the North American cellular system advanced mobile phone service (AMPS), $(C/I)_0 = 18$ dB, which results in $K = 7$ or $D = 4.6R$. Here, we have used a propagation attenuation proportional to the fourth power of the distance. The radius of the cell is determined mainly by the projected traffic density. In Fig. 76.1 a seven

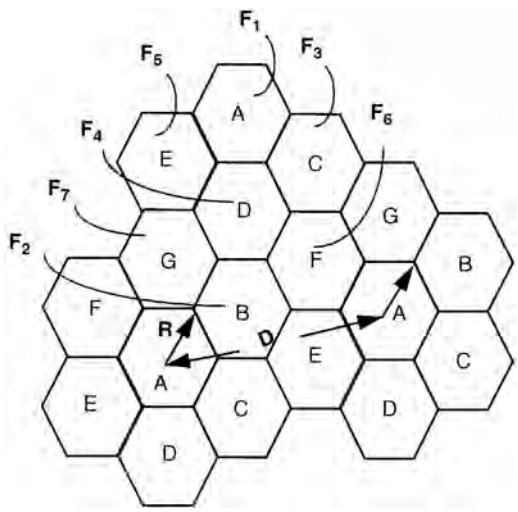


FIGURE 76.1 Fixed channel assignment.

cell cluster with frequency sets F1 through F7 (in cells designated A–G) has been illustrated. It is seen that the same set of frequencies is repeated two cells away.

The number of channels for each cell can be determined through the erlang-B formula (for example, see Cooper, 1990 [6]) by knowing the busy hour traffic and the desired probability of blocking (grade of service). Probability of blocking P_B is related to offered traffic A, and the number of channels per cell N by

$$P_B = \frac{A^N/N!}{\sum_{i=0}^{N} A^i/i!}$$

This applies to the case of blocked calls cleared. If calls are delayed, the grade of service becomes the probability of calls being delayed P_Q and is given by the erlang-C formula [6],

$$P_Q = \frac{\dfrac{A^N}{N!(1-A/N)}}{\sum_{i=0}^{N-1} A^i/i! + \dfrac{A^N}{N!(1-A/N)}}$$

FCA is used in almost all existing cellular mobile networks employing FDMA or TDMA. To illustrate resource assignment in FCA, we describe the FCA schemes used in global system for mobile communications (GSM or DCS) [17]. Here, the mobile terminal continuously monitors the received signal strength and quality of a broadcast channel along with the identity of the transmitting base station. This mechanism allows both the mobile and the network to keep track of the subscriber movements throughout the service area. When the mobile terminal tries and eventually succeeds in accessing the network (see Fig. 76.2), a two way control channel is assigned that allows for both authentication and ciphering mode establishment. On completion of these two phases, the setup phase initiates and a traffic channel is eventually assigned to the mobile terminal. Here, the terminal only knows the identity of the serving cell; the (control and traffic) radio channel assignment is a responsibility of the network and fulfills the original carrier to cell association, while simplifying the control functions during normal system operation.

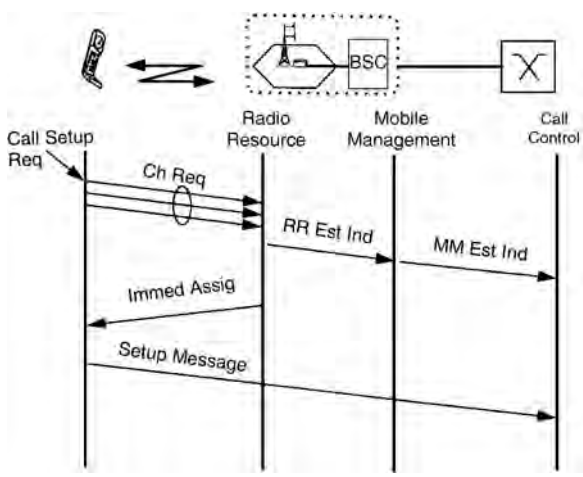


FIGURE 76.2 Radio resource assignment in GSM/DCS. *Source:* Jabbari, B. et al., Network issues for wireless personal communications. *IEEE Commun. Mag.*, 33(1), 1995.

76.4 Enhanced Fixed Channel Assignment

FCA has the advantage of having simple realization. Since the frequency sets are preassigned to cells based on long-term traffic demand, however, it cannot adapt to traffic variation across cells and, therefore, FCA will result in poor bandwidth utilization. To improve the utilization while maintaining the implementation simplicity, various strategies have been proposed as enhancements to FCA and deployed in existing networks. Two often used methods are *channel borrowing* and *directed retry*, which are described here.

In the channel borrowing strategy [11,12,18,29], channels that are not in use in their cells may be borrowed by adjacent cells with high offered traffic on a call-by-call basis. Borrowing of channels allows the arriving calls to be served in their own cell. This implies that there will be further restrictions in using the borrowed channels in other cells. Various forms of borrowing have been surveyed in Tekinay and Jabbari [23].

In directed retry [10,13,14], a call to or from a mobile subscriber may try other cells with channels with sufficient signal strength meeting the C/I constraint if there are no channels available in its own cell to be served. In some cases it may be necessary to direct some of the calls in progress in a given congested cell to adjacent lightly loaded cells (those calls that can be served by adjacent cells) in order to accommodate the new calls in that given cell. This is referred to as directed handoff [13,20]. The combination of these two capabilities provides a significant increase in bandwidth utilization.

76.5 Dynamic Channel Assignment

In dynamic channel assignment [2,3,4,7,8,9,12,21], the assignment of channels to cells occurs based on the traffic demand in the cells. In other words, channels are pooled together and assignments are made and modified in real time. Therefore, this assignment scheme has the potential to achieve a significantly improved bandwidth utilization when there are temporal or spatial traffic variations.

In DCA, the interference constraints are ensured by a real-time evaluation of the most suitable (less interfered) channels that can be activated in a given cell (reassignment may be needed). That is, the system behaves as if the frequency plan was dynamically changing to meet the actual radio link quality and traffic loads, realizing an implicit sharing of the frequency band under interference constraints.

The implementation of DCA generally is more complex due to the requirement for system-wide state information where the state refers to which channel in which cell is being used. Obviously, it is impractical to update the system state in a large cellular network at any time, especially those based on microcells, as the controller will be overloaded or call set delay will be unacceptable. Therefore, methods have been devised based on a limited state space centralized control [15] or based on a distributed control to perform the necessary updating. A method referred to as maximum packing, suggested in Everitt and Macfadyen [15], records only the number of channels in use. Distributed control schemes, however, where channel assignments are made at mobile stations or base stations may be attractive. These schemes are particularly suitable for a mobile controlled resource assignment where each mobile measures the actually perceived interference and decides to utilize a radio resource in a completely decentralized way.

Design of DCA algorithms is critical to achieve the potential advantages in efficiency and robustness to traffic heterogeneity throughout cells as compared to FCA and enhanced FCA. Poor DCA algorithms, however, might lead to an uncontrolled global situation, i.e., the locally selected channel might be very good for the specific mobile terminal and at the same time very poor for the interference level induced to other traffic sources.

Two classes of *regulated DCA* and *segregation DCA* are discussed here due to their importance. In a regulated DCA, appropriate thresholds tend to maintain the current channel, avoiding useless handoffs that can tie up channels; in a segregation DCA, channel acquisition obeys priorities assigned to each channel. In the latter, the channels successfully activated in a cell have a higher assignment probability than those found to have a high interference in the cell; of course, priorities change in time depending on changes in the system status. These types of assignment lead to a (frequency) plan still changing

dynamically but more slowly and, in general, only because of substantial load imbalance. In steady-state conditions, the plan either tends to be confirmed or fluctuates around a basic configuration, proven to be suboptimal in terms of bandwidth efficiency.

Both digital European cordless telecommunications (DECT) and cordless technology known as CT2 are employing the DCA technique [25], and the next generation cellular systems are foreseen to deploy it. In CT2 the handset and the base station jointly determine a free suitable channel to serve the call. In DECT, the mobile terminal (portable handset) not only recognizes the visited base station (radio fixed part) through a pilot signal but continuously scans all of the system channels and holds a list of the less interfered ones. The potentially available channels are ordered by the mobile terminal with respect to the measured radio parameters, namely, the radio signal strength indicator, measuring and combining cochannel, adjacent-channel, and intermodulation interference. When a connection has to be established, the best channel is used to communicate on the radio interface.

It is possible to have a hybrid of DCA and FCA in a cellular network in which a fraction of channels are fixed assigned and the remainder are allocated based on FCA. This scheme has less system implementation complexity than the DCA scheme but provides performance improvement (lower probability of call blocking) depending on the DCA–FCA channel partitioning [19].

In general, DCA schemes cannot be considered independently of the adopted power control mechanism because the transmitted power from both mobile terminal and base station substantially affects the interference within the network. For a detailed discussion of DCA and power control, the readers are referred to [5]. Despite the availability of several power control algorithms, much work is needed to identify realizable power and channel control algorithms that maximize bandwidth efficiency. Nevertheless, realization of power control mechanisms would require distribution of the system status information and information exchange between mobile terminal and network entities. This in turn will involve overhead and terminal power usage.

The performance of DCA depends on the algorithm implementing this capability [13,16]. In general, due to interactions between different cells the performance of the system will involve modeling the system as a whole, as opposed to in FCA where cells are treated independently. Therefore, mathematical modeling and performance evaluation of DCA becomes quite complex. Simplifying assumptions may, therefore, be necessary to obtain approximate results [15,22]. Simulation techniques have been widely used in evaluation of DCA performance. For a representative performance characteristics of DCA and a comparison to enhanced FCA schemes the readers are referred to [14].

76.6 CDMA Systems

Now we turn our attention to systems based on direct sequence spread spectrum CDMA. Examples include the second generation system based on IS-95 and the proposed third generation system based on Wideband CDMA (W-CDMA) [28]. Due to the simultaneous availability and reuse of the entire frequency spectrum in every cell, opportunity for dynamic channel assignment exists inherently. However, the problem of dynamic channel allocation turns into the problem of allocation of resources such as spreading codes, equipment limitations, intracell and intercell interference. Added to this is the practical implication of the soft capacity, which allows temporary degradation of the links. Furthermore, the use of voice activity compression technique allows increase in channel utilization while adding new performance measures such as speech freeze-out fraction.

The capacity of the CDMA system in terms of Erlang for both forward and reverse directions has been derived in [27]. The effect of imperfect power control, other cell interference has been taken into account for traffic sources with a given user activity. This capacity reflects the average traffic load the system can carry under a specified performance measure. This measure, although it is treated similar to the probability of blocking, in actuality is somewhat different from those commonly used in orthogonal time and frequency systems and indeed reflects the outage probability. The performance measure has been defined in [27] as the ratio of background noise density to total interference density, N_0/I_0, being greater than a specified value (for example, 0.1). For the simple equally loaded cells case where the limiting direction

is the reverse link, with a bandwidth of W, source data rate of R bits/second, required energy to interference ratio of E_b/I_0, and ratio of other cells to own cell (sector) interference of f, the condition to have an outage is given by [27]:

$$P_{\text{outage}} = \Pr\left[\sum_{i=1}^{K(1+f)} v_i (E_b/I_0)_i > (W/R)(1 - N_0/I_0)\right]$$

In the above formula, v_i represents the event the bi-state source is in the active state and K is the number of sources in a single cell. Note that $(E_b/I_0)_i$ is the energy per bit-to-interference ratio of source i as received at the observed base station. Additional scenarios and some approximations for special cases have been further carried out and discussed in [26]. Such analysis can be used to evaluate the system capacity for when traffic is nonuniformly distributed across the cells/sectors.

76.7 Conclusion

In this chapter we have classified and reviewed channel assignment techniques. We have emphasized the advantages of DCA schemes over FCA in terms of bandwidth utilization in a heterogeneous traffic environment at the cost of implementation complexity. The DCA schemes are expected to play an essential role in future cellular and microcellular networks.

CDMA inherently provides the DCA capability. However, the system capacity depends on factors including power control accuracy and other cell interference. We presented a brief discussion of capacity of CDMA.

References

1. Abramson, N., Multiple access techniques for wireless networks. *Proc. of IEEE,* 82(9), 1994.
2. Anderson, L.G., A simulation study of some dynamic channel assignment algorithms in a high capacity mobile telecommunications system. *IEEE Trans. on Comm.,* COM-21(11), 1973.
3. Beck, R. and Panzer, H., Strategies for handover and dynamic channel allocation in micro-cellular mobile radio systems. *Proceedings of the IEEE Vehicular Technology Conference,* 1989.
4. Chuang, J.C.-I., Performance issues and algorithms for dynamic channel assignment. *IEEE J. on Selected Areas in Comm.,* 11(6), 1993.
5. Chuang, J.C.-I., Sollenberger, N.R., and Cox, D.C., A pilot-based dynamic channel assignment schemes for wireless access TDMA/FDMA systems. *Internat. J. of Wireless Inform. Networks,* Jan. 1994.
6. Cooper, R.B., *Introduction to Queueing Theory,* 3rd ed., CEEPress Books, 1990.
7. Cox, D.C. and Reudnik, D.O., A comparison of some channel assignment strategies in large-scale mobile communications systems. *IEEE Trans. on Comm.,* COM-20(2), 1972.
8. Cox, D.C. and Reudnik, D.O., Increasing channel occupancy in large-scale mobile radio environments: dynamic channel reassignment. *IEEE Trans. on Comm.,* COM-21(11), 1973.
9. Dimitrijevic, D. and Vucetic, J.F., Design and performance analysis of algorithms for channel allocation in cellular networks. *IEEE Trans. on Vehicular Tech.,* 42(4), 1993.
10. Eklundh, B., Channel utilization and blocking probability in a cellular mobile telephone system with directed retry. *IEEE Trans. on Comm.,* COM 34(4), 1986.
11. Elnoubi, S.M., Singh, R., and Gupta, S.C., A new frequency channel assignment algorithm in high capacity mobile communications. *IEEE Trans. on Vehicular Techno.,* 31(3), 1982.
12. Engel, J.S. and Peritsky, M.M., Statistically-optimum dynamic server assignment in systems with interfering servers. *IEEE Trans. on Comm.,* COM-21(11), 1973.
13. Everitt, D., Traffic capacity of cellular mobile communications systems. *Computer Networks and ISDN Systems,* ITC Specialist Seminar, Sept. 25–29, 1989, 1990.

14. Everitt, D., Traffic engineering of the radio interface for cellular mobile networks. *Proc. of IEEE,* 82(9), 1994.
15. Everitt, D.E. and Macfadyen, N.W., Analysis of multicellular mobile radiotelephone systems with loss. *BT Tech. J.,* 2, 1983.
16. Everitt, D. and Manfield, D., Performance analysis of cellular mobile communication systems with dynamic channel assignment. *IEEE J. on Selected Areas in Comm.,* 7(8), 1989.
17. Jabbari, B., Colombo, G., Nakajima, A., and Kulkarni, J., Network issues for wireless personal communications. *IEEE Comm. Mag.,* 33(1), 1995.
18. Jakes, W.C. Ed. *Microwave Mobile Communications,* Wiley, New York, 1974, reissued by IEEE Press, 1994.
19. Kahwa, T.J. and Georganas, N.D., A hybrid channel assignment scheme in Large scale, cellular-structured mobile communications systems. *IEEE Trans. on Comm.,* COM-26(4), 1978.
20. Karlsson, J. and Eklundh, B., A cellular mobile telephone system with load sharing—an enhancement of directed retry. *IEEE Trans. on Comm.,* COM 37(5), 1989.
21. Panzer, H. and Beck, R., Adaptive resource allocation in metropolitan area cellular mobile radio systems. *Proceedings of the IEEE Vehicular Technology Conference,* 1990.
22. Prabhu, V. and Rappaport, S.S., Approximate analysis for dynamic channel assignment in large systems with cellular structure. *IEEE Trans. on Comm.,* COM-22(10), 1974.
23. Tekinay, S. and Jabbari, B., Handover and channel assignment in mobile cellular networks. *IEEE Comm. Mag.,* 29(11), 1991.
24. Telecommunications Industry Association. TIA Interim Standard IS-95, CDMA Specifications, 1993.
25. Tuttlebee, W.H.W., Cordless personal communications. *IEEE Comm. Mag.,* Dec. 1992.
26. Viterbi, A.J., *CDMA—Principles of Spread Spectrum Communication,* Addison Wesley, Reading, MA, 1995.
27. Viterbi, A.J. and Viterbi, A.J., Erlang capacity of a power controlled CDMA system. *IEEE JSAC,* 11(6), 892–900, 1993.
28. W-CDMA. Feature topic. *IEEE Communications Magazine,* Sept. 1998.
29. Zhang, M. and Yum, T.-S.P., Comparison of channel-assignment strategies in cellular mobile telephone systems. *IEEE Trans. on Vehicular Tech.,* 38(4), 1989.

77

Radiolocation Techniques

Gordon L. Stüber
Georgia Institute of Technology

James J. Caffery, Jr.
University of Cincinnati

77.1 Introduction

Several location technologies have been developed and commercially deployed for locating wireless radios including Decca, Loran, Omega, the Global Positioning System (GPS), and the Global Navigation Satellite System(GLONASS). GPS, originally developed for military use, is perhaps the most popular commercial location system today, providing location accuracies to within 100 m.

All of the above systems use a location technique known as radiolocation. A radiolocation system operates by measuring radio signals traveling between a mobile station (MS) and a set of fixed stations (FSs). The measurements are then used to determine the length and/or direction of the radio paths, and the MS's position is derived from known geometrical relationships. In general, measurements from $n + 1$ FSs are necessary to locate an MS in n dimensions. To achieve high accuracy in a radiolocation system, it is necessary that a line-of-sight (LoS) exist between the MS and FSs. Otherwise, large errors are likely to be incurred.

With the above mentioned radiolocation technologies, the MS formulates its own position by using signals received from the FSs. This form of location is often referred to as *self-positioning*. In these systems, a special receiver is placed in the MS to calculate the MS's position. Alternatively, the position of the MS could be calculated at a remote location by using signals received at the FSs. This form of radiolocation is known as *remote-positioning* and requires a transmitter for the MS.

Over the last several years, there has been increased interest in developing location services for wireless communications systems. An array of applications for such technology exists including location sensitive billing, fraud detection, cellular system design and resource management, fleet management, and Intelligent Transportation Services (ITS)[8]. The greatest driving force behind location system development in wireless systems has been the FCC's Emergency-911 (E-911) requirements, where a wireless E-911

caller must be located within an *rms* accuracy of 125 m in 67% of the cases [4]. Adding GPS to the handset is not a universal solution because of the large pool of existing handsets. Remote-positioning radiolocation is a natural choice since it requires no modification of existing MS handsets and most, if not all, of the complexity could be incorporated into the network side.

In this chapter, techniques for locating wireless users using measurements from radio signals are described. A few algorithms for location estimation are developed along with a discussion of measures of accuracy. The radiolocation methods that are appropriate for wireless location and the major sources of error in various mobile cellular networks are discussed.

77.2 Description of Radiolocation Methods

Radiolocation systems can be implemented that are based on angle-of-arrival (AoA), signal strength, time-of-arrival (ToA), time-difference-of-arrival (TDoA), or their combinations. These are briefly discussed below.

Angle of Arrival

AoA techniques estimate the MS location by first measuring the arrival angles of a signal from a MS at several FSs (Fig. 77.1). The AoA can be determined through the use of directive antennas or antenna arrays. Simple geometric relationships are then used to determine the location by finding the intersections of the lines-of-position (see Fig. 77.1). To generate a position fix, the AoA method requires that the signal transmitted from the MS be received by a minimum of two FSs.

Signal Strength

Radiolocation based on signal strength measurements uses a known mathematical model describing the **path loss** attenuation with distance. Since measurement of signal strength provides a distance estimate between a MS and FS, the MS must lie on a circle centered at the FS. Hence, for signal strength based radiolocation, the lines-of-position are defined by circles. By using measurements from multiple FSs, the location of the MS can be determined from the intersection of circles (Fig. 77.2).

A second method makes use of premeasured signal strength contours around each FS. Received signal strength measured at multiple FSs can be mapped to a location by overlaying the contours for each FS. This technique can be used to combat **shadowing**, as discussed in Section 77.6.

Time-Based Location

The final class of radiolocation techniques are those based on estimating the ToAs of a signal transmitted by the MS and received at multiple FSs or the TDoAs of a signal received at multiple pairs of FSs. In the ToA approach, the distance between a MS and FS is measured by finding the one-way propagation time

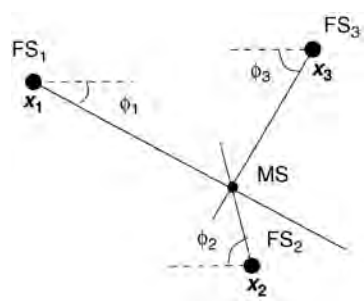


FIGURE 77.1 The measured angles, ϕ_i, determine the position of the MS for a given FS geometry.

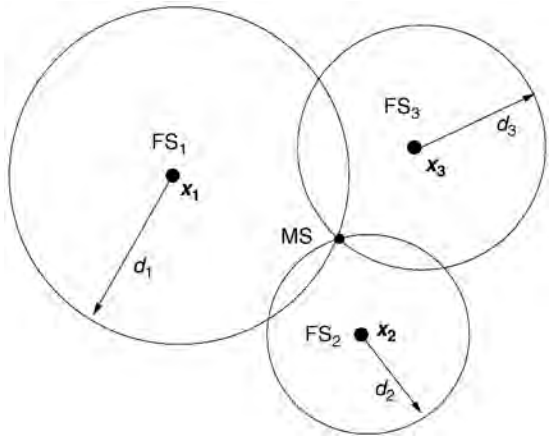


FIGURE 77.2 For signal strength and ToA based radiolocation systems the location of the MS is at the intersection of circles of radius d_i.

between the MS and FS. Geometrically, this provides a circle, centered at the FS, on which the MS must lie. Given a ToA at FS i, the equation for the circle is given by

$$\tau_i = D_i(\boldsymbol{x}_s)/c \tag{77.1}$$

where $D_i(\boldsymbol{x}_s) = \|\boldsymbol{x}_i - \boldsymbol{x}_s\|$, $\boldsymbol{x}_i$ is the position of ith FS, $\boldsymbol{x}_s$ is the position of MS and c is the speed of light. By using at least three base stations to resolve ambiguities, the MS's position is given by the intersection of circles (Fig. 77.2). Since the ToA and path loss based signal strength methods are based on distance measurements between the MS and FSs, they are often referred to as *ranging* systems.

In the TDoA approach, time differences of arrival are used. Hence, the time of signal transmission need not be known. Since the hyperbola is a curve of constant time *difference* of arrival for two FSs, a TDoA measurement defines a line-of-position as a hyperbola, with the foci located at one of the two FSs. For FSs i and j, the equation of the hyperbola, $\rho_{i,j}$, for a given TDoA is

$$\rho_{i,j} = \frac{D_i(\boldsymbol{x}_s) - D_j(\boldsymbol{x}_s)}{c} \tag{77.2}$$

The location of the MS is at the intersection of the hyperbolas (Fig. 77.3). In general, for N FSs receiving the signal from the MS, $N - 1$ non-redundant TDoA measurements can be made. Thus, a MS can be located in $N - 1$ dimensions.

77.3 Location Algorithms

When there are no measurement errors, the lines-of-position intersect at a point and a geometric solution can be obtained for the MS's location by finding the intersection of the lines of position. However, in practice, measurement errors occur and the lines of position do not intersect at a point. Consequently, other solution methods must be utilized. In the following, a solution approach is developed for two-dimensional location systems where the MS is located at $\boldsymbol{x}_s = [x_s, y_s]^T$ and the FSs are located at $\boldsymbol{x}_i = [x_i, y_i]^T$ for $i = 1, \ldots, N$.

Problem Formulation

In general, the $N \times 1$ vector of noisy measurements, $\boldsymbol{r}$, from a set of N FSs can be modeled by

$$\boldsymbol{r} = \boldsymbol{C}(\boldsymbol{x}_s) + \boldsymbol{n} \tag{77.3}$$

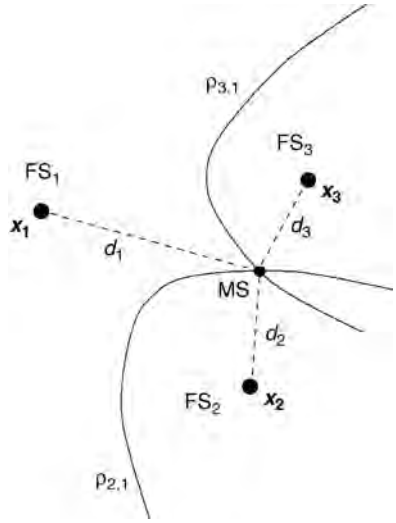


FIGURE 77.3 For TDoA based radiolocation the position of the MS is at the intersection of hyperbolas. The curves represent constant differences in distance to the MS with respect to the first FS, $\rho_{i,1} = d_i - d_1$.

where $\boldsymbol{n}$ is an $N \times 1$ measurement noise vector, generally assumed to have zero mean and an $N \times N$ covariance matrix $\boldsymbol{\Sigma}$. The system measurement model $\boldsymbol{C}(\boldsymbol{x}_s)$ depends on the location method used:

$$\boldsymbol{C}(\boldsymbol{x}_s) = \begin{cases} \boldsymbol{D}(\boldsymbol{x}_s) & \text{for ToA} \\ \mathcal{R}(\boldsymbol{x}_s) & \text{for TDoA} \\ \boldsymbol{\Phi}(\boldsymbol{x}_s) & \text{for AoA} \end{cases} \tag{77.4}$$

where

$$\boldsymbol{D}(\boldsymbol{x}_s) = [\tau_1, \tau_2, \ldots, \tau_N]^T \tag{77.5}$$

$$\mathcal{R}(\boldsymbol{x}_s) = [\rho_{2,1}, \rho_{3,1}, \ldots \rho_{N,1}]^T \tag{77.6}$$

$$\boldsymbol{\Phi}(\boldsymbol{x}_s) = [\phi_1, \phi_2, \ldots, \phi_N]^T \tag{77.7}$$

The terms τ_i and $\rho_{i,1}$ are the ToAs and TDoAs defined in Eqs. (77.1) and (77.2), respectively, where without loss of generality, the TDoAs are referenced to the first FS. If the time of transmission τ_s is needed to form the ToA estimates, it can be incorporated into $\boldsymbol{x}_s$ as a parameter to be estimated along with x_s and y_s. The unknown parameter vector can then be modified to $\boldsymbol{x}_s = [x_s, y_s, \tau_s]^T$, while the system measurement model becomes $\boldsymbol{C}(\boldsymbol{x}_s) = \boldsymbol{D}(x_s, y_s) + \tau_s \mathbf{1}$.

The AoAs are defined by

$$\phi_i = \tan^{-1}\left(\frac{y_i - y_s}{x_i - x_s}\right) \tag{77.8}$$

Although not explicitly shown in the above equations, τ_i, $\rho_{i,1}$ and ϕ_i are nonlinear functions of $\boldsymbol{x}_s$.

A well-known approach for determining an estimate from a noisy set of measurements is the method of **least squares** (LS) estimation. The weighted least squares (WLS) solution is formed as the vector $\hat{\boldsymbol{x}}_s$ that minimizes the cost function

$$\mathcal{E}(\hat{\boldsymbol{x}}_s) = [\boldsymbol{r} - \boldsymbol{C}(\hat{\boldsymbol{x}}_s)]^T \boldsymbol{W}[\boldsymbol{r} - \boldsymbol{C}(\hat{\boldsymbol{x}}_s)] \tag{77.9}$$

LS methods can achieve the **maximum likelihood** (ML) estimate when the measurement noise vector is Gaussian with $\mathrm{E}[\boldsymbol{n}] = 0$ and equal variances, i.e., $\boldsymbol{\Sigma} = \sigma_n^2 \boldsymbol{I}$. For unequal variances, WLS with $\boldsymbol{W} = \boldsymbol{\Sigma}^{-1}$ gives the ML estimate. In the following, $\boldsymbol{W} = \boldsymbol{I}$ will be assumed.

Location Solutions

As Eq. (77.4) indicates, $\boldsymbol{C}(\boldsymbol{x}_s)$ is a nonlinear function of the unknown parameter vector $\boldsymbol{x}_s$ so that the LS problem is a nonlinear one. One straightforward approach is to iteratively search for the minimum of the function using a **gradient descent method**. With this approach, an initial guess is made of the MS location and successive estimates are updated according to

$$\hat{\boldsymbol{x}}_s^{(k+1)} = \hat{\boldsymbol{x}}_s^{(k)} - \boldsymbol{\nu} \nabla \mathcal{E}\left(\hat{\boldsymbol{x}}_s^{(k)}\right) \tag{77.10}$$

where the matrix $\boldsymbol{\nu} = \mathrm{diag}(\nu_x, \nu_y)$ is the step size, $\hat{\boldsymbol{x}}_s^{(k)}$ is the estimate at time k, and $\nabla = \partial/\partial \boldsymbol{x}$ denotes the gradient vector with respect to the vector $\boldsymbol{x}$.

In order to mold the problem into a linear LS problem, the nonlinear function $\boldsymbol{C}(\boldsymbol{x}_s)$ can be linearized by using a Taylor series expansion about some reference point $\boldsymbol{x}_0$ so that

$$\boldsymbol{C}(\boldsymbol{x}_s) \approx \boldsymbol{C}(\boldsymbol{x}_0) + \boldsymbol{H}(\boldsymbol{x}_s - \boldsymbol{x}_0) \tag{77.11}$$

where $\boldsymbol{H}$ is the Jacobian matrix of $\boldsymbol{C}(\boldsymbol{x}_s)$. Then the LS solution can be formed as

$$\hat{\boldsymbol{x}}_s = \boldsymbol{x}_0 + (\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T[\boldsymbol{r} - \boldsymbol{C}(\boldsymbol{x}_0)] \tag{77.12}$$

This approach can be performed iteratively, with each successive estimate being closer to the final estimate. A key drawback to this approach is that an initial guess, $\boldsymbol{x}_0$, must be made of the MS's position.

The Taylor series approach introduces error when the linearized function $\boldsymbol{C}(\boldsymbol{x}_s)$ does not accurately approximate the nonlinear function. Other approaches have been developed for TDoA that avoid linearization by transforming the TDoA measurements into "pseudo-measurements." The pseudo-measurements are given by [5]

$$\boldsymbol{\varphi} = \boldsymbol{\Delta}\boldsymbol{x}_s + D_1(\boldsymbol{x}_s)\boldsymbol{r} \tag{77.13}$$

where

$$\boldsymbol{\Delta} = \begin{bmatrix} (\boldsymbol{x}_2 - \boldsymbol{x}_1)^T \\ \vdots \\ (\boldsymbol{x}_N - \boldsymbol{x}_1)^T \end{bmatrix} \qquad \boldsymbol{\varphi} = \frac{1}{2}\begin{bmatrix} \|\boldsymbol{x}_2\|^2 - \|\boldsymbol{x}_1\|^2 - \rho_{2,1}^2 \\ \vdots \\ \|\boldsymbol{x}_N\|^2 - \|\boldsymbol{x}_1\|^2 - \rho_{N,1}^2 \end{bmatrix} \tag{77.14}$$

The term $D_1(\boldsymbol{x}_s)$ is nonlinear in the unknown vector $\boldsymbol{x}_s$ and can be removed by using a projection matrix that has $\boldsymbol{r}$ in its null space. A suggested projection is $\boldsymbol{P} = (\boldsymbol{I} - \boldsymbol{Z})[\mathrm{diag}(\boldsymbol{r})]^{-1}$ where $\boldsymbol{Z}$ is a circular shift

matrix [5]. Projecting Eq. (77.13) with $\boldsymbol{P}$, the following linear equation results:

$$\boldsymbol{P\varphi} = \boldsymbol{P\Delta x}_s \tag{77.15}$$

which leads to the following linear LS solution for the location of the MS

$$\hat{\boldsymbol{x}}_s = (\boldsymbol{\Delta}^T\boldsymbol{P}^T\boldsymbol{P\Delta})^{-1}\boldsymbol{\Delta}^T\boldsymbol{P}^T\boldsymbol{P\varphi} \tag{77.16}$$

77.4 Measures of Location Accuracy

To evaluate the performance of a location method, several benchmarks have been proposed. A common measure of accuracy is the comparison of the mean-squared-error (MSE) of the location estimate with the Cramér-Rao lower bound (CRLB) [10]. The concepts of circular error probability (CEP) [9] and geometric dilution of precision (GDOP) [6] have also been used as accuracy measures.

Cramér-Rao Lower Bound

For location in M dimensions, the MSE of the position estimate is given by

$$\text{MSE} = \sqrt{\text{E}[(\boldsymbol{x}_s - \hat{\boldsymbol{x}}_s)^T(\boldsymbol{x}_s - \hat{\boldsymbol{x}}_s)]} \tag{77.17}$$

where E[·] denotes expectation. The calculated MSE is often compared to the theoretical minimum MSE given by the CRLB which sets a lower bound on the variance of any unbiased estimator. The CRLB is the inverse of the information matrix $\boldsymbol{J}$ defined as [10]

$$\boldsymbol{J} = \text{E}\left[\left(\frac{\partial p(\boldsymbol{r}|\boldsymbol{x})}{\partial \boldsymbol{x}}\right)\left(\frac{\partial p(\boldsymbol{r}|\boldsymbol{x})}{\partial \boldsymbol{x}}\right)^T\right]\Bigg|_{\boldsymbol{x}=\boldsymbol{x}_s} \tag{77.18}$$

where $\boldsymbol{r}$ is the vector of TDoA, ToA, or AoA estimates and $p(\boldsymbol{r}|\boldsymbol{x})$ is the probability density function of $\boldsymbol{r}$ conditioned on the parameter vector $\boldsymbol{x}$. Assuming Gaussian measurement noise, $p(\boldsymbol{r}|\boldsymbol{x})$ is Gaussian with mean $\boldsymbol{r}_0$ and covariance matrix $\boldsymbol{Q}$, and the CRLB reduces to

$$\text{CRLB} = \boldsymbol{J}^{-1} = c^2\left(\frac{\partial \boldsymbol{r}_o^T}{\partial \boldsymbol{x}}\boldsymbol{Q}^{-1}\frac{\partial \boldsymbol{r}_{0)}}{\partial \boldsymbol{x}^T}\right)^{-1}\Bigg|_{\boldsymbol{x}=\boldsymbol{x}_s} \tag{77.19}$$

Circular Error Probability

A simple measure of accuracy is the CEP which is defined as the radius of the circle that has its center at the mean and contains half the realizations of a random vector. The CEP is a measure of the uncertainty in the location estimator $\hat{\boldsymbol{x}}_s$ relative to its mean E[$\hat{\boldsymbol{x}}_s$]. If the location estimator is unbiased, the CEP is a measure of the estimator uncertainty relative to the true MS position. If the magnitude of the bias vector is bounded by B, then with a probability of one-half, a particular estimate is within a distance of B + CEP from the true position.

Because it is difficult to derive an exact expression for the CEP, an approximation that is accurate to within 10% is often used. The approximation for CEP is given as [9]

$$\text{CEP} \approx 0.75\sqrt{\text{E}[(\hat{\boldsymbol{x}}_s - \hat{\boldsymbol{\mu}})^T(\hat{\boldsymbol{x}}_s - \hat{\boldsymbol{\mu}})]} \tag{77.20}$$

$$= 0.75\sqrt{\sum_{i=1}^{M}\sigma^2_{\hat{x}_{s,i}}} \tag{77.21}$$

where $\hat{\boldsymbol{\mu}} = \mathrm{E}[\hat{\boldsymbol{x}}_s]$ is the mean location estimate and $\sigma^2_{\hat{x}_{s,i}}$ is the variance of the ith estimated coordinate, $i = 1,\ldots,M$.

Geometric Dilution of Precision

The GDOP provides a measure of the effect of the geometric configuration of the FSs on the location estimate. It is defined as the ratio of the *rms* position error to the *rms* ranging error [9,6]. Hence, for an unbiased estimator, the GDOP is given by

$$\mathrm{GDOP} = \frac{\sqrt{\mathrm{E}[(\hat{\boldsymbol{x}}_s - \hat{\boldsymbol{\mu}})^T(\hat{\boldsymbol{x}}_s - \hat{\boldsymbol{\mu}})]}}{\sigma_r} \tag{77.22}$$

where σ_r denotes the fundamental ranging error for ToA and TDoA systems. For AoA, σ_n^2 is the average variance of the distance between each FS and a reference point near the true position of the MS.

The GDOP is an indicator of the extent to which the fundamental ranging error is magnified by the geometric relation between the MS and FSs. Furthermore, comparing (77.21) and (77.22), we find that the CEP and GDOP are related by

$$\mathrm{CEP} \approx (0.75\sigma_r)\,\mathrm{GDOP} \tag{77.23}$$

The GDOP serves as a useful criterion for selecting the set of FSs from a large set to produce the minimum location error. In addition, it may aid cell site planning for cellular networks which plan to provide location services to their users.

77.5 Location in Cellular Systems

The location requirements set forth by the FCC [4] must be met not only by the new digital cellular systems, but the older analog system as well. In cellular networks, the BSs serve the role of the FSs in the algorithms of Section 77.3. With several different wireless systems on the market (**AMPS**, IS-54/136 **TDMA**, **GSM**, IS-95 **CDMA**), different methods may be necessary to implement location services in each of those systems. The signal strength method is often not implemented for cellular systems because of the large variability of received signal strength resulting from shadowing and **multipath fading** (see Section 77.6). AoA requires the placement of antenna arrays at the BSs which may be extremely costly. The AoA measurements can be obtained from array signal processing and is not dependent on the type of cellular system deployed. Unlike AoA, the ToA and TDoA methods require that timing information be obtained from the signals transmitted by a MS which may be implemented in different ways for each cellular system. The time-based methods may also require strict synchronization of the BSs, especially the TDoA approach. The remainder of this section discusses implementation strategies for the ToA and TDoA location methods in current and future generation cellular systems.

The most straightforward approach for obtaining timing information for ToA or TDoA location is the use of signal correlation methods. Specifically, maximizing cross-correlations between the signals received at pairs of BSs will provide an estimate of the TDoAs for each pair of BSs. Of course, this approach requires that the BSs be synchronized. These techniques are necessary for implementing a location system in AMPS since no system message parameters provide useful radiolocation information.

For CDMA, different methods can be used for the uplink (MS to BS) and downlink (BS to MS). On the uplink, the timing information for ToA or TDoA can be obtained using correlation techniques. Since the BSs in IS-95 are synchronized to a GPS time reference, the time of detection of the signal from the MS can serve as a ToA time stamp. Similarly, segments of the detected signal can be sent to a central processing office for cross-correlation in order to determine the set TDoAs for the BSs. The signals for the ToA/TDoA measurements can come from the reverse traffic channel or the access channel. The reverse traffic channel could be used for E-911 calls, for example, since a voice call must be initially made.

For other location applications, the location may be desired when the MS is not actively transmitting. In these cases, the MS could be prompted to transmit messages on the access channel in response to commands from its serving BS on the paging channels. Unfortunately, it may be impossible to detect the MS transmissions at other BSs due to the **near–far effect** (see Section 77.6), although this problem can be alleviated by having the MS power-up to its maximum power for a short time. However, the use of the power up function must be limited to emergencies (such as E-911) in order to avoid excessive interference to other users.

An alternative for location in CDMA is to utilize pilot monitoring in the MS on the downlink. To assist in the handoff process, the MS monitors the strongest pilots from the surrounding BSs. The serving BS can send a pilot measurement request order (PMRO) causing the BS to respond with a message which includes the magnitudes of the pilots in the candidate set as well as the code phase of each pilot relative to its serving BS [2]. Hence, it is possible to construct TDoA estimates from these system messages. The accuracy of the TDoA estimates is dependent on the resolution of the code phase and the synchronization of the BSs. Fortunately, for IS-95, the BSs are synchronized to a GPS time reference. However, the code phase resolution is limited to a chip time, T_c, which implies a TDoA resolution of approximately 244 m. Finally, the **soft handoff**, during which the MS communicates with nearby BSs during a handoff, can be used for location in CDMA systems as long as at least three BSs are in a soft handoff with the MS.

The TDMA-based systems also provide timing information in their system messages that can be used for ToA or TDoA location. The time alignment parameter in IS-54/136 and timing advance in GSM (both abbreviated TA) are used by each of those networks to ensure that the transmissions of MSs arrive at their serving BSs in the appropriate time slots. Each BS sends the MSs a TA value which is the amount the MS must advance or retard the timing of its transmissions. Additionally, the TA serves as a measure of the propagation time between the MS and BS. By artificially forcing the MS to handoff to two or more BSs, the location of the MS could be found using the ToA method. A primary consideration is the accuracy of the TA. For IS-54, the timing of MS transmissions are advanced or retarded in units of $T_b/2$, where $T_b = 20.6\ \mu$s is the bit duration [1]. Hence, the TAs are accurate to $T_b/4$, or 1543 m. For GSM, the TA messages are reported in units of bits, with $T_b = 3.7\ \mu$s, which gives a TA resolution of $T_b/2$, or 554 m, in GSM [3].

An alternative for GSM is to use the observed time difference (OTD) measurements which are made at the MS without forcing additional handoffs. The OTDs are used to facilitate handoffs by estimating the amount the timing of the MS would have to be advanced or retarded if it were to be handed over to another BS. With a synchronized network, the OTDs could be used to implement a TDoA location system. Unfortunately, the GSM standard does not require that the network be synchronized. Additionally, the OTD measurements are made to the same accuracy of the TA measurements, 554 m.

Because of the high chip rate and good correlation properties of the spreading code sequences used in CDMA systems, these systems have greater potential than the other systems for accurate location estimates. It is apparent that the resolution of the timing parameters in the system messages needs to be improved in order to provide more accurate estimates of location.

77.6 Sources of Location Error

In all cellular systems, several factors can introduce error in the location estimates. Sources of error that are common to all cellular systems include multipath propagation, non-line-of-sight (NLoS) propagation and multiple-access interference (MAI). However, MAI poses a more significant problem in CDMA systems because of **power control** and the near–far effect. These effects are described below.

Multipath

Multipath propagation can introduce error in signal strength, AoA, ToA, and TDoA measurements. For signal-strength-based location systems, multipath fading and shadowing cause variations in the signal strength that can be as great as 30–40 dB over distances in the order of a half wavelength. Signal strength

averaging can help, but low mobility MSs may not be able to average out the effects of multipath fading and there will still be the variability due to shadowing. The errors due to shadowing can be combated by using premeasured signal strength contours that are centered at the BSs. However, this approach assumes a constant physical topography since shadows will change with the tree foliage, construction/destruction of structures, etc.

For AoA-based systems, scattering near and around the MS and BS will affect the measured AoA. Multipath will interfere with the angle measurement even when a LoS component is present. For macrocells, scattering objects are primarily within a small distance of the MS and the BSs are usually elevated well above the local terrain. Consequently, the signals arrive with a relatively narrow AoA spread at the BSs. For microcells, the BSs may be placed below roof top level. Consequently, the BSs will often be surrounded by local scatterers such that the signals arrive at the BSs with a large AoA spread. Thus, the AoA method may be impractical for microcells.

In time-based location systems, the ToA or TDoA estimates can be in error even when there is a LoS path between the MS and BS. Conventional delay estimators, which are usually based on correlation techniques, are influenced by the presence of multipath fading. The result is a shift in the peak of the correlation away from the true value. Conventional delay estimators will detect a delay in the vicinity of these later arriving rays.

NLoS Propagation

With NLoS propagation, the signal transmitted from the MS (or BS) is reflected or diffracted and takes a path that is longer than the direct path or received at a different angle (Fig. 77.4). Obviously, the effect on an AoA system can be disastrous if the received AoAs are in a much different direction than the true AoAs. For the time-based systems, the measured distances can be considerably greater than true distances. For instance, for ToA location in the GSM system, the typical ranging error introduced by NLoS propagation can average 400–700 meters [7]. NLoS propagation will corrupt the ToA or TDoA measurements even when high resolution timing techniques are employed and even if there is no multipath fading.

Multiple-Access Interference

All cellular systems suffer from cochannel interference. The transmissions of other users interfere with the signal of the desired user reducing the accuracy with which location measurements can be made. The problem is most evident in CDMA systems where the users share the same frequency band. As a result, signals from higher-powered users may mask the signals of the lower-powered users, a phenomenon known as the near–far effect. To combat the near–far effect, power control is used. However, for a

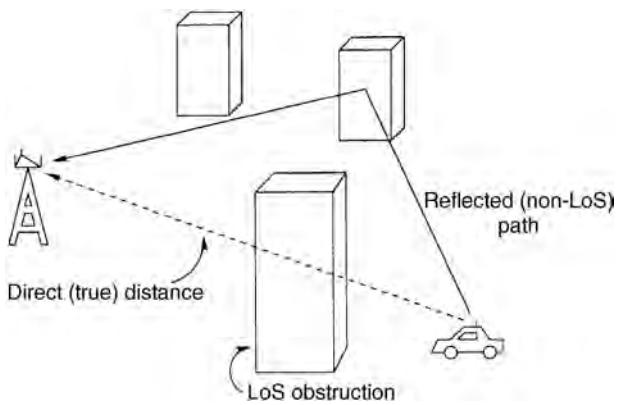


FIGURE 77.4 Propagation in a NLoS environment where signals are received from a reflection rather than a direct (LoS) path.

location system where multiple BSs must receive the transmission from the MS, the near–far problem still exists because the MS is not power-controlled to the other BSs. Consequently, the signal from the MS may not be detected at enough BSs to form a location estimate. As mentioned in Section 77.5, it may be possible, for instance, in E-911 situations, for the MS to power up to maximum level and, therefore, mitigate the near–far effect. A further possibility is to take advantage of soft handoffs. However, the MS must be in position for a three-way soft handoff to be located.

77.7 Summary and Conclusions

This chapter has provided a brief introduction to radiolocation techniques for wireless systems. Several algorithms were developed for locating a MS using AoA, ToA, and TDoA, and measures of locator accuracy were described. For location services in mobile cellular networks, many possibilities exist. However, it is apparent that none has the current capability of providing high accuracy location estimates to meet the FCC requirements.

Defining Terms

AMPS (advanced mobile phone service): Analog cellular system in North America.

CDMA (code-division multiple-access): A technique for spread-spectrum multiple-access digital communications that separates users through the use of unique code sequences.

Gradient descent method: A minimization technique which searches for the minimum of an error surface by taking steps along the direction of greatest slope.

GSM (global system for mobile communications): Pan-European digital cellular standard.

Least squares estimation: A method whose estimate is chosen as the value that minimizes the sum of squares of the measured error.

Maximum likelihood estimation: A method whose estimate is chosen as the parameter value from which the observed data was most likely to come.

Multipath fading: Rapid fluctuation of the complex envelope of the received signal caused by reception of multiple copies of the transmitted signal, each with different amplitude, phase, and delay.

Near–far effect: A phenomenon that arises from unequal received power levels from the MSs. Stronger signals mask the weaker signals.

Path loss: Description of the attenuation of signal power with distance from a transmitter.

Power control: System for controlling the transmission power of the MS. Used to reduce cochannel interference and mitigate the near–far effect on the uplink.

Shadowing: Slow variation in the mean envelope over a distance corresponding to tens of wavelengths.

Soft handoff: Reception and transmission of radio signals between an MS and two or more BSs to achieve a macrodiversity gain.

TDMA (time-division multiple access): A form of multiple access giving each user a different time slot for transmission and reception of signals.

References

1. EIA/TIA Interim Standard, IS-54., Cellular System Dual Mode Mobile Station–Land Station Compatibility Specifications, May 1990.
2. EIA/TIA Interim Standard IS-95., Mobile Station–Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular System, May 1992.
3. ETSI-SMG., GSM 05.10 Technical Specification, Version 4.0.1, Apr. 1992.
4. FCC Docket No.96-254., Report and Order and Further Notice of Proposed Rulemaking in the Matter of Revision of the Commission's Rules to Ensure Compatibility with Enhanced 911 Emergency Calling Systems, Jun. 12, 1996.

5. Friedlander, B., A passive location algorithm and its accuracy analysis, *IEEE Journal of Oceanic Engineering,* 234–244, Jan. 1987.
6. Massatt, P. and Rudnick, K., Geometric formulas for dilution of precision calculations, *Journal of the Institute of Navigation,* 37, 379–391, Winter, 1991.
7. Silventoinen, M. and Rantalainen, T., Mobile station emergency locating in GSM, in *Proceedings of the IEEE International Conference on Personal Wireless Communications,* 232–238, 1996.
8. Stilp, L., Carrier and end-user applications for wireless location systems. *Proc. SPIE,* 119–126, 1996.
9. Torrieri, D.J., Statistical theory of passive location systems, *IEEE Trans. on Aerospace and Electronic Systems,* AES-20, 183–197, Mar. 1984.
10. Van Trees, H.L., *Detection, Estimation and Modulation Theory, Part I,* John Wiley & Sons, New York, Ch. 2, 1968.

Further Information

More information on radiolocation can be found in the April, 1998, issue of *IEEE Communications Magazine,* which provided several articles regarding location service issues for wireless communications networks.

Articles discussing location techniques and algorithms can be found in many IEEE journals including *Transactions on Vehicular Technology*, *Transactions on Aerospace and Electronic Systems*, and the *Journal of Oceanic Engineering.*

Proceedings of various IEEE conferences such as the *Vehicular Technology Conference* document some of the latest developments for location in cellular networks.

78

Power Control

Roman Pichna
Nokia

Qiang Wang
University of Victoria

78.1 Introduction

The growing demand for mobile communications is pushing the technological barriers of wireless communications. The available spectrum is becoming crowded and the old analog **FDMA** (frequency division multiple access) cellular systems no longer meet the growing demand for new services, higher quality, and spectral efficiency. A second generation of digital cellular mobile communication systems are being deployed all around the world. The second generation systems are represented by three major standards: the **GSM**, IS-136, and IS-95. The first two are **TDMA**-based digital cellular systems and offer a significant increase in spectral efficiency and quality of service as compared to the first generation systems, e.g., **AMPS**, **NMT**, and **TACS**. IS-95 is based on **DS/CDMA** technology. The standardization of the third generation systems, IMT-2000 (formerly known as FPLMTS), is being pursued at ITU. Similar efforts are being conducted at regional standardization bodies.

The channel capacity of any cellular system is significantly influenced by the cochannel interference. To minimize the cochannel interference, several techniques are proposed: frequency reuse patterns, which ensure that the same frequencies are not used in adjacent cells; efficient **power control**, which minimizes the transmitted power; cochannel interference cancellation techniques; and orthogonal signalling (time, frequency, or code). All of these are being intensively researched, and some have already been implemented.

This chapter provides a short overview of power control. Since power control is a very broad topic, it is not possible to exhaustively cover all facets associated with power control. The interested reader can find additional information in the recommended reading that is appended at the end of this chapter.

The following section (Section 78.2) provides a brief introduction into cellular networks and demonstrates the necessity of power control. The various types of power control are presented in this section. The next section (Section 78.3) illustrates some applications of power control employed in various systems such as analog AMPS, GSM, DS/CDMA cellular standard IS-95, and digital cordless telephone standard **CT2**. A glossary of definitions is provided at the end of the chapter.

78.2 Cellular Systems and Power Control

In cellular communication systems, the service area is divided into cells, each covered by a single base station. If, in the **forward link** (base station to mobile), all users served by all base stations share the same frequency, each communication between a base station and a particular user would also reach all other users in the form of cochannel interference. However, the greater the distance between the mobile

and the interfering transmitter, the weaker the interference becomes due to the propagation loss. To ensure a good quality of service throughout the cell, the received signal in the fringe area of the cell must be strong. Once the signal has crossed the boundary of a cell, however, it becomes interference and is required to be as weak as possible. Since this is difficult, the channel frequency is usually not reused in adjacent cells in most of the cellular systems. If the frequency is reused, the cochannel interference impairs the signal reception in the adjacent cell, and the quality of service severely degrades unless other measures are taken to mitigate the interference. Therefore, a typical reuse pattern reuses the frequency in every seventh cell (frequency reuse factor = 1/7). The only exception is for CDMA-based systems where the users are separated by codes, and the allocated frequency may be shared by all users in all cells.

Even if the frequency is reused in every seventh cell, there is still some cochannel interference arriving at the receiver. It is, therefore, very important to maintain a minimal transmitted level at the base station to keep the cochannel interference low, frequency reuse factor high, and therefore the capacity of the system and quality of service high.

The same principle applies in the **reverse link** (mobile to base station)—the power control maintains the minimum necessary transmitted power for reliable communication. Several additional benefits can be gained from this strategy. The lower transmitted power conserves the battery energy allowing the mobile terminal (the portable) to be lighter and stay on the air longer. Furthermore, recent concerns about health hazards caused by the portable's electromagnetic emissions are also alleviated.

In the reverse link, the power control also serves to alleviate the near–far effect. If all mobiles transmitted at the same power level, the signal from a near mobile would be received as the strongest. The difference between the received signal strength from the nearest and the farthest mobile can be in the range of 100 dB, which would cause saturation of the weaker signals' receivers or an excessive amount of adjacent channel interference. To avoid this, the transmitted power at the mobile must be adjusted inversely proportional to the effective distance from the base station. The term effective distance is used since a closely located user in a propagation shadow or in a deep fade may have a weaker signal than a more distant user having excellent propagation conditions.

In a DS/CDMA system, power control is a vital necessity for system operation. The capacity of a DS/CDMA cellular system is interference limited since the channels are separated neither in frequency nor in time, and the cochannel interference is inherently strong. A single user exceeding the limit on transmitted power could inhibit the communication of all other users.

The power control systems have to compensate not only for signal strength variations due to the varying distance between base station and mobile but must also attempt to compensate for signal strength fluctuations typical of a wireless channel. These fluctuations are due to the changing propagation environment between the base station and the user as the user moves across the cell or as some elements in the cell move. There are two main groups of channel fluctuations: slow (i.e., **shadowing**) and fast **fading**.

As the user moves away from the base station, the received signal becomes weaker because of the growing propagation attenuation with the distance. As the mobile moves in uneven terrain, it often travels into a propagation shadow behind a building or a hill or other obstacle much larger than the wavelength of the frequency of the wireless channel. This phenomenon is called shadowing. Shadowing in a land-mobile channel is usually described as a stochastic process having log-normal distributed amplitude. For other types of channels other distributions are used, e.g., Nakagami.

Electromagnetic waves transmitted from the transmitter may follow multiple paths on the way from the transmitter to the receiver. The different paths have different delays and interfere at the antenna of the receiver. If two paths have the same propagation attenuation and their delay differs in an odd number of half-wavelengths (half-periods), the two waves may cancel each other at the antenna completely. If the delay is an even multiple of the half-wavelengths (half-periods), the two waves may constructively add, resulting in a signal of double amplitude. In all other cases (nonequal gains, delays not a multiple of half-wavelength), the resultant signal at the antenna of the receiver is between the two mentioned limiting cases. This fluctuation of the channel gain is called fading. Since the scattering and reflecting

surfaces in the service area are randomly distributed (buildings, trees, furniture, walls, etc.), the amplitude of the resulting signal is also a random variable. The amplitude of fading is usually described by a Rayleigh, Rice, or Nakagami distributed random variable.

Since the mobile terminal may move at the velocity of a moving car or even of a fast train, the rate of channel fluctuations may be quite high and the power control has to react very quickly in order to compensate for it.

The performance of the **reverse link** of DS/CDMA systems is most affected by the near–far effect and, therefore, very sophisticated power control systems in the reverse link that attempt to alleviate the effects of channel fluctuations must be used. Together with other techniques, such as micro- and macrodiversity, interleaving and coding, interference cancellation, multiuser detection, and adaptive antennae, the DS/CDMA cellular system is able to cope with the wireless channel extremely well.

The effective use of the **power control** in DS/CDMA cellular system enables the frequency to be reused in every cell, which in turn enables features such as the soft hand-off and base station diversity. All together, these help enhance the capacity of the system.

In the **forward link** of a DS/CDMA system, power control may also be used. It may vary the transmitted power to the mobile, but the dynamic range is smaller due to the shared spectrum and, thus, shared interference.

We can distinguish between two kinds of power control, the open-loop power control and the closed-loop power control. The open-loop power control estimates the channel and adjusts the transmitted power accordingly but does not attempt to obtain feedback information on its effectiveness. Obviously, the open-loop power control is not very accurate, but since it does not have to wait for the feedback information it may be relatively fast. This can be advantageous in the case of a sudden channel fluctuation, such as a mobile driving from behind a big building or in case it should provide only the initial or rough transmitted power setting.

The principle operation of open-loop power control is shown in Fig. 78.1. The open-loop power control must base its action on the estimation of the channel state. In the reverse link it estimates the channel by measuring the received power level of the pilot from the base station in the forward link and

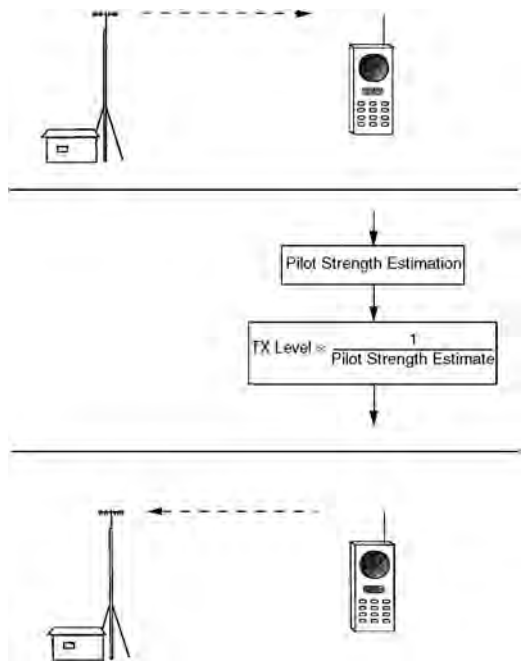


FIGURE 78.1 Reverse link open-loop power control.

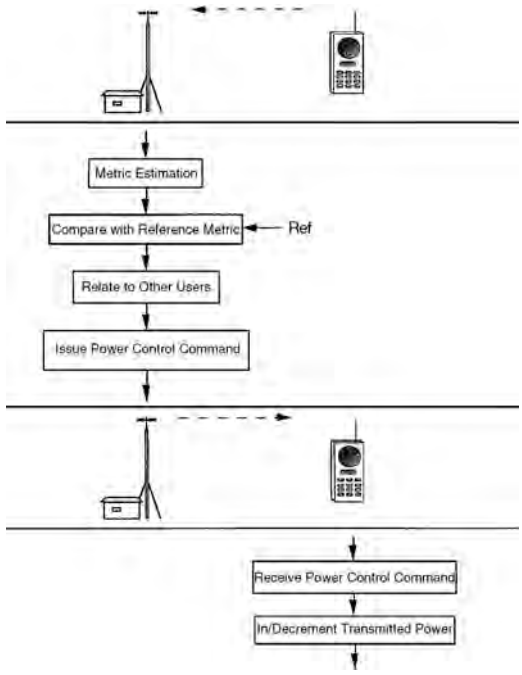


FIGURE 78.2a Reverse link closed-loop power control.

sets the transmitted power level inversely proportional to it. Estimating the power of pilot is, in general, more reliable than estimating the power of the voice (or data) channel since the pilot is usually transmitted at higher power levels. Using the estimated value for setting the transmitted power ensures that the average power level received from the mobile at the base station remains constant irrespective of the channel variations. However, this approach assumes that the forward and the reverse link signal strengths are closely correlated. Although forward and reverse link may not share the same frequency and, therefore, the fading is significantly different, the long-term channel fluctuations due to shadowing and propagation loss are basically the same.

The closed-loop power control system (Fig. 78.2a) may base its decision on an actual communication link performance metric, e.g., received signal power level, received signal-to-noise ratio, received bit-error rate, or received frame-error rate, or a combination of them. In the case of the reverse link power control, this metric may be forwarded to the mobile as a base for an autonomous power control decision, or the metric may be evaluated at the base station and only a power control adjustment command is transmitted to the mobile. If the reverse link power control decision is made at the base station, it may be based on the additional knowledge of the particular mobile's performance and/or a group of mobiles' performance (such as mobiles in a sector, cell, or even in a cluster of cells). If the power control decision for a particular mobile is made at the base station or at the switching office for all mobiles and is based on the knowledge of all other mobile's performance, it is called a centralized power control system. A centralized power control system may be more accurate than a distributed power control system, but it is much more complex in design, more costly, and technologically challenging.

In principle, the same categorization may be used for the power control in the forward link (Fig. 78.2b) except that in the reverse link pilots from the mobiles are usually unavailable and only closed-loop power control is applied.

A special case for the design of power control are **TDD**-based systems [3]. In TDD systems, the forward and reverse link are highly correlated, and therefore a very good estimate of the channel gain in the

TABLE 78.1 Capacity Reduction versus Power Control Error

	$\sigma_E = 0.5$ dB, %	$\sigma_E = 1$ dB, %	$\sigma_E = 2$ dB, %	$\sigma_E = 3$ dB, %
Forward link	10	29	64	83
Reverse link	10	31	61	81

Source: Kudoh, E., On the capacity of DS/CDMA cellular mobile radios under imperfect transmitter power control. *IEICE Trans. Commun.*, E76-B, 886–893, Apr. 1993.

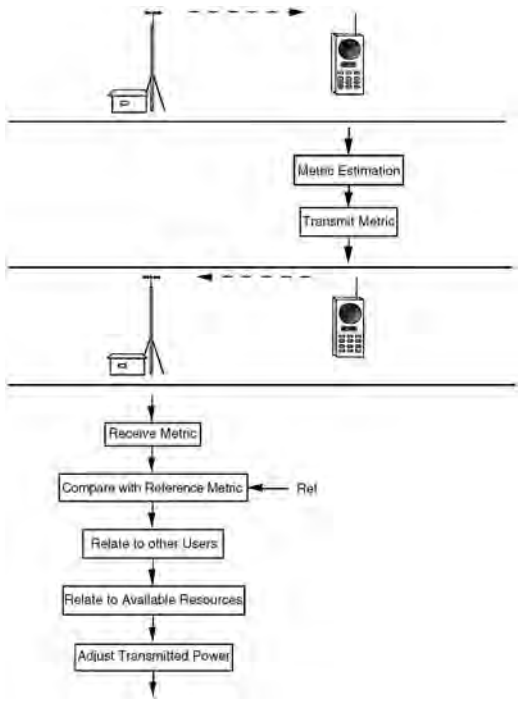


FIGURE 78.2b Forward link closed-loop power control.

forward link can be obtained from the estimate of the reverse link gain and vice versa. An open-loop power control then performs with the precision of a closed-loop power control but much faster since no feedback information has to be transmitted.

In the ideal case, power control compensates for the propagation loss, shadowing, and fast fading. However, there are many effects that prevent the power control from becoming ideal. Fast fading rate, finite delays of the power control system, nonideal channel estimation, error in the power control command transmission, limited dynamic range, etc., all contribute to degrading the performance of the power control system. It is very important to examine the performance of power control under nonideal conditions since the research done has shown that the power control system is quite sensitive to some of these conditions [11]. Kudoh [5] simulated a nonideal closed-loop power control system. Errors in the system were represented by a log-normal distributed control error with standard deviation σ_E (dB). Some results on capacity reduction are presented in Table 78.1.

The authors have also studied the effects of Doppler and delay and feedback errors in power control loop on power control [8].

78.3 Power Control Examples

In the following section, several applications of power control of analog and digital cellular systems are presented.

In the analog networks we may see power control implemented in both the reverse link and forward link [6]. Power control in the reverse link:

- reduces the chance of receiver saturation by a closely located mobile
- reduces the cochannel interference and thus increases the frequency-reuse factor and capacity, and
- reduces the average transmitted power at the mobile thus conserving battery energy at the mobile.

The power control in the forward link:

- reduces cochannel interference and thus increases the frequency reuse factor and capacity and
- reduces adjacent-channel interference and improves the quality of service.

One example of a power control system shown by Lee [7] was of an air-to-ground communication system. The relevant airspace is divided into six zones based on the aircraft altitude. The transmitted power at the aircraft is then varied in six steps based on the zone in which the aircraft is located. The power control system exhibits a total of approximately 28 dB of dynamic range. This reduces the cochannel interference and, due to the excellent propagation conditions in the free air, has a significant effect on the capacity of the system.

Another example of a power control system in an analog wireless network is in the analog part of the TIA standard IS-95 [10]. IS-95 standardizes a dual-mode FDMA/CDMA cellular system compatible with the present day AMPS analog FDMA cellular system.

The analog part of IS-95 divides the mobiles into three classes according to nominal **ERP** (effective radiated power with respect to half-wave dipole) at the mobile. For each class, the standard specifies eight power levels. Based on the propagation conditions, the mobile station may receive a power control command that specifies at what power level the mobile should transmit. The maximum change is 4 dB per step. (See Table 78.2.)

IS-95 supports further discontinuous transmission. This feature allows the mobile to vary its transmitted power between two states: low and high. These two states must be at least 8 dB apart.

As for the power control in a digital wireless system, three examples will be shown: GSM [1], CT2/CT2PLUS standard [2] for digital cordless telephones of second generation, and the IS-95 standard for digital cellular DS/CDMA system [10].

TABLE 78.2 Nominal ERP of the Mobile

	Nominal ERP (dBW) of Mobile		
Power Level	I	II	III
0	6	2	−2
1	2	2	−2
2	−2	−2	−2
3	−6	−6	−6
4	−10	−10	−10
5	−14	−14	−14
6	−18	−18	−18
7	−22	−22	−22

Source: Telecommunications Industry Association/Electronic Industries Association. Mobile Station-Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular System, *TIA/EIA/IS-95 Interim Standard,* Jul. 1993.

TABLE 78.3 GSM Transmitter Classes

Power Class	Base Station Power (W)	Mobile Station Power (W)
1	320	20
2	160	8
3	80	5
4	40	2
5	20	0.8
6	10	
7	5	
8	2.5	

Source: Balston, D.M. and Macario, R.C.V., *Cellular Radio Systems*, Artech House, Norwood, MA, 1993.

GSM is a Pan-European digital cellular system that was introduced in many countries during the 1992–1993 period. GSM is a digital TDMA system with a frequency hopping feature. The power control in GSM ensures that the mobile station uses only the minimum power level necessary for reliable communication with the base station. GSM defines eight classes of base stations and five classes of mobiles according to their power output, as shown in Table 78.3.

The transmitted power at the base station is controlled, nominally in 2-dB steps. The adjustment of the transmitted power reduces the intercell interference and, thus, increases the frequency reuse factor and capacity. The transmitted power at the base station may be decremented to a minimum of 13 dBm.

The power control of the mobile station is a closed-loop system controlled from the base station. The power control at the mobile sets the transmitted power to one of 15 transmission power levels spaced by 2 dB. Any change can be made only in steps of 2 dB during each time slot. Another task for the power control in GSM is to control graceful ramp-on and ramp-off of the TDMA bursts since too steep slopes would cause spurious frequency emissions.

The dynamic range of the received signal at the base station may be up to 116 dB [1] and, thus, the near–far problem may also by experienced, especially if the problem occurs in adjacent time slots. In addition to power control, a careful assignment of adjacent slots can also alleviate the near–far effect.

The CT2PLUS standard [2] is a Canadian enhancement of the **ETSI** CT2 standard. Both these standards allow power control in the forward and in the reverse link. Due to the expected small cell radius and relatively slow signal level fluctuation rate, given by the fact that the user of the portable is a pedestrian, the power control specifications are relatively simple. The transmission at the portable can have two levels: *normal* (full) and *low*. The low–normal difference is up to 20 dB.

The IS-95 standard represents a second generation digital wireless cellular system using direct-sequence code division multiple access (DS/CDMA). Since in a DS/CDMA system all users have the same frequency allocation, the cochannel interference is crucial for the performance of the system [4]. The near–far effect may cause the received signal level to change up to 100 dB [12]. This considerable dynamic range is disastrous for a DS/CDMA where the channels are separated by a finite correlation between spreading sequences. This is further aggravated by the shadowing and the fading. The fading may have a relatively high rate since the mobile terminal is expected to move at the speed of a car. Therefore, the power control system must be very sophisticated. Power control is employed in both the reverse link and in the forward link.

The reverse link power control serves to do the following:

- Equalizes the received power level from all mobiles at the base station. This function is vital for system operation. The better the power control performs, the more it reduces the cochannel interference and, thus, increases the capacity. The power control compensates for the near–far effect, shadowing, and partially for slow fading.
- Minimizes the necessary transmission power level to achieve good quality of service. This reduces the cochannel interference, which increases the system capacity and alleviates health concerns.

In addition, it saves the battery power. Viterbi [12] has shown up to 20–30 dB average power reduction compared to the AMPS mobile user as measured in field trials.

The forward link power control serves to:

- Equalize the system performance over the service area (good quality signal coverage of the worst-case areas),
- Provide load shedding between unequally loaded cells in the service areas (e.g., along a busy highway) by controlling the intercell interference to the heavy loaded cells, and
- Minimize the necessary transmission power level to achieve good quality of service. This reduces the cochannel interference in other cells, which increases the system capacity and alleviates health concerns in the area around the base station.

The reverse link power control system is composed of two subsystems: the closed-loop and the open-loop. The system operates as follows. Prior to the application to access, closed-loop power control is inactive. The mobile estimates the mean received power of the received pilot from the base station and the open-loop power control estimates the mean output power at the access channel [10]. The system then sets the closed-loop probing and estimates the mean output power,

$$\begin{aligned}\text{mean output power (dBm)} = &-\text{mean input power (dBm)} - 73 + \text{NOM\_PWR (dB)}\\ &+ \text{INIT\_PWR (dB)}\end{aligned} \tag{78.1}$$

where NOM_PWR and INIT_PWR are parameters obtained by the mobile prior to transmission. Subsequent probes are sent at increased power levels in steps until a response is obtained. The initial transmission on the reverse traffic channel is estimated as

$$\begin{aligned}\text{mean output power (dBm)} = &-\text{mean input power (dBm)} - 73\\ &+ \text{NOM\_PWR (dB)} + \text{INIT\_PWR (dB)}\\ &+ \text{the sum of all access probe corrections (dB)}\end{aligned} \tag{78.2}$$

Once the first closed-loop power control bit is received the mean output power is estimated as

$$\begin{aligned}\text{mean output power (dBm)} = &-\text{mean input power (dBm)} - 73 + \text{NOM\_PWR (dB)}\\ &+ \text{INIT\_PWR (dB)} + \text{the sum of all access probe corrections (dB)}\\ &+ \text{the sum of all closed-loop power control corrections (dB)}\end{aligned} \tag{78.3}$$

The ranges of the parameters NOM_PWR and INIT_PWR are shown in Table 78.4.

The closed-loop power control command arrives at the mobile every 1.25 ms (i.e., 800 b/s). Therefore, the base station estimates the received power level for approximately 1.25 ms. A closed-loop power control command can have only two values: 0 to increase the power level and 1 to decrease the power level. The mobile must respond to the power control command by setting the required transmitted power level

TABLE 78.4 NOM_PWR and INIT_PWR Parameters

	Nominal Value, dB	Range, dB
NOM_PWR	0	−8–7
INIT_PWR	0	−16–15

Source: Telecommunications Industry Association/Electronic Industries Association. Mobile Station-Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular System. *TIA/EIA/IS-95 Interim Standard,* Jul. 1993.

within 500 μs. The total range of the closed-loop power control system is ±24 dB. The total supported range of power control (closed-loop and open-loop) must be at least ±32 dB.

The behavior of the closed-loop power control system while the mobile receives base station diversity transmissions is straightforward. If all diversity transmitting base stations request the mobile to increase the transmitted power (all power control commands are 0), the mobile increases the power level. If at least one base station requests the mobile to decrease its power, the mobile decreases its power level.

The system also offers a feature of gated transmitted power for variable rate transmission mode. The gate-off state reduces the output power by at least 20 dB within 6 μs. This reduces the interference to the other users at the expense of transmitted bit rate. This feature may be used together with variable rate voice encoder or voice activated keying of the transmission.

The forward link power control works as follows. The mobile monitors the errors in the frames arriving from the base station. It reports the frame-error rate to the base station periodically. (Another mode of operation may report the error rate only if the error rate exceeds a preset threshold.) The base station evaluates the received frame-error rate reports and slightly adjusts its transmitting power. In this way, the base station may equalize the performance of the forward links in the cell or sector.

A system conforming with the standard has been field tested, and the results show that the power control is able to combat the channel fluctuation (together with other techniques such as RAKE reception) and achieve the bit energy to interference power density (E_b/I_0) necessary for a reliable service [12].

Power control together with soft handoff determines the feasibility of the DS/CDMA cellular system and is crucial to its performance. QUALCOMM, Inc. has shown on field trials that their system conforms with the theoretical predictions and surpasses the capacity of other currently proposed cellular systems [12].

78.4 Summary

We have shown the basic principles of power control in wireless cellular networks and have presented some examples of power control systems employed in some networks.

In a wireless channel the channel transmission or channel gain is a random variable. If all transmitters in the system transmitted at equal and constant power levels, the received powers would be random.

In the reverse link (mobile to base station) each user has its own wireless channel, generally uncorrelated with all other users. The received signals at the base station are independent and random. Furthermore, since the users are randomly distributed over the cell, the distance between the mobiles and the base station may vary and so does the propagation loss. The differences between the strongest and the weakest received signal level may approach the order of 100 dB. This power level difference may cause saturation of the receivers at the base station even if they are allocated a different frequency or time slot. This phenomenon is called the near–far effect.

The near–far effect is especially detrimental for a DS/CDMA system where the frequency band is shared by all users and, for any given user, all other users' transmissions form the cochannel interference. Therefore, for the DS/CDMA system it is vitally important to efficiently mitigate the near–far effect.

The most natural way to mitigate the near–far effect is to power control the transmission in such a way that the transmitted power counterfollows the channel fluctuations and compensates for them. Then the received signal at the base station arrives at a constant amplitude.

The use of power control is not limited to the reverse link, but is also employed in the forward link. The controlled transmission maintaining the transmitted level at the minimum acceptable level reduces the cochannel interference, which translates into an increased capacity of the system.

Since the DS/CDMA systems are most vulnerable to the near–far effect they have a very sophisticated power control system. In giving examples, we have concentrated on the DS/CDMA cellular system. We have also shown the power control used in other systems such as GSM, AMPS, and CT2.

Although there are more techniques available for mitigation of the near–far effect, power control is the most efficacious. As such, power control forms the core in the effort in combatting the near–far effect and channel fluctuations in general [12].

Defining Terms

AMPS: Advanced Mobile Phone Service. Analog cellular system in North America.
CT2: Cordless Telephone, Second Generation. A digital FDMA/TDD system.
DS/CDMA: Direct Sequence Code Division Multiple Access.
DOC: Department of Communications.
ERP: Effective Radiated Power.
ETSI: European Telecommunications Standard Institute.
Fading: Fast varying fluctuations of the wireless channel mainly due to the interference of time-delayed multipaths.
FDMA: Frequency Division Multiple Access.
Forward link: Link from the base (fixed) station to the mobile (user, portable).
GSM: Groupe Spéciale Mobile, recently referred to as the **Global System for Mobility**. An ETSI standard for digital cellular and microcellular systems.
NMT: Nordic Mobile Telephone. A cellular telephony standard used mainly in Northern Europe.
Power control: Control system for controlling the transmission power. Used to reduce the cochannel interference and mitigate the near–far effect in the reverse link.
Reverse link: Link from the mobile (user, portable) to the base (fixed) station.
Shadowing: Slowly varying fluctuations of the wireless channel due mainly to the shades in propagation of electromagnetic waves. Often described by log-normal probability density function.
TACS: Total Access Communication System. An analogue cellular system used mainly in UK.
TDD: Time Division Duplex.
TDMA: Time Division Multiple Access.

References

1. Balston, D.M. and Macario, R.C.V., *Cellular Radio Systems,* Artech House, Norwood, MA, 1993.
2. Department of Communications. ETI Interim Standard # I-ETS 300 131, Annex 1, Issue 2, Attachment 1. In *CT2PLUS Class 2: Specification for the Canadian Common Air Interface for Digital Cordless Telephony, Including Public Access Services, RS-130.* Communications Canada, Ottawa, ON, 1993.
3. Esmailzadeh, R., Nakagawa, M., and Sourour, E.A., Time-division duplex CDMA communications, *IEEE Personal Comm.,* 4(3), 51–56, Apr. 1997.
4. Gilhousen, K.S., Jacobs, I.S., Padovani, R, Viterbi, A.J., Weaver, L.A., and Wheatley III, C.E., On the capacity of cellular CDMA system. *IEEE Trans. Veh. Tech.,* 40, 303–312, May 1991.
5. Kudoh, E., On the capacity of DS/CDMA cellular mobile radios under imperfect transmitter power control. *IEICE Trans. Commun.,* E76-B, 886–893, Apr. 1993.
6. Lee, W.C.L., *Mobile Cellular Telecommunications Systems,* McGraw-Hill, New York, 1989.
7. Lee, W.C.L., *Mobile Communications Design Fundamentals,* 2nd ed., John Wiley & Sons, New York, 1993.
8. Pichna, R., Kerr, R., Wang, Q., Bhargava, V.K., and Blake, I.F., *CDMA Cellular Network Analysis Software.* Final Rep. Ref. No. 36-001-2-3560/01-ST, prepared for Department of Communications, Communications Research Centre, Ottawa, ON, Mar. 1993.
9. Simon, M.K., Omura, J.K., Scholtz, R.A., and Levitt, B.K., *Spread Spectrum Communication Handbook,* McGraw-Hill, New York, 1994.
10. Telecommunications Industry Association/Electronic Industries Association. Mobile Station-Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular System. *TIA/EIA/IS-95 Interim Standard,* Jul. 1993.
11. Viterbi, A.J. and Zehavi, E., Performance of power-controlled wideband terrestrial digital communication. *IEEE Trans. Comm.,* 41, 559–569, Apr. 1993.
12. Viterbi, A.J., The orthogonal-random waveform dichotomy for digital mobile personal communication. *IEEE Personal Comm.,* 1(1st qtr.), 18–24, 1994.

Further Information

For general information see the following overview books:

1. Balston, D.M. and Macario, R.C.V., *Cellular Radio Systems,* Artech House, Norwood, MA, 1993.
2. Simon, M.K., Omura, J.K., Scholtz, R.A., and Levitt, B.K., *Spread Spectrum Communication Handbook,* McGraw-Hill, New York, 1994.

For more details on power control in DS/CDMA systems consult the following:

3. Gilhousen, K.S., Jacobs, I.S., Padovani, R., Viterbi, A.J., Weaver, L.A., and Wheatley C.E., III., On the capacity of cellular CDMA system. *IEEE Trans. Veh. Tech.,* 40, 303–312, May 1991.
   or:
4. Viterbi, A.J. and Zehavi, E., Performance of power-controlled wideband terrestrial digital communication. *IEEE Trans. Comm.,* 41, 559–569, Apr. 1993.

Readers deeply interested in power control are recommended the *IEEE Transactions on Communications, IEEE Transactions on Vehicular Technology,* and relevant issues of *IEEE Journal on Selected Areas in Communications.*

79

Enhancements in Second Generation Systems

Marc Delprat
Alcatel Mobile Network Division

Vinod Kumar
Alcatel Research & Innovation

79.1 Introduction

Present digital **cellular** and **cordless** systems were optimized for voice services. At the very best, they can provide low and medium bit rate information services. The development of enhanced versions of these radio interfaces is motivated by the following:

- Provision of additional services, including high bit rate circuit-switched and packet-switched services that meet the short-term needs of mobile multimedia services.
- Improvements in the radio coverage of existing cellular systems. An extension to allow cordless coverage in the home is also being considered.
- Even more efficient utilization of the available frequency spectrum, which is a valuable but limited resource.

This chapter deals mainly with the air interface of second generation systems. The subject of enhancement in system performance is addressed from two directions. On the one hand, system features like adaptive multirate coders, packet transmission, which have been explicitly included in the standard, are discussed in detail. On the other hand, methods of equipment design or network design possible with the new or already existing features of the air interface are presented.

79.2 Overview of Second Generation Systems

Initially the need of a pan-European system to replace a large variety of disparate analog cellular systems was the major motivating factor behind the creation of the Global System for Mobile communications (GSM). In North America and Japan, where unique analog systems existed, the need to standardize respectively IS-54, IS-95, and Personal Digital Cellular (PDC) for digital cellular applications arose from the lack of spectrum to serve the high traffic density areas [3]. Additionally, some of the second generation systems like Digital European Cordless Telecommunications (DECT) and Personal Handy Phone Systems (PHS) are the result of a need to offer wireless services in residential and office environments with low cost subscriber equipment [16].

The physical layer characteristics of all these systems offer robust radio links paired with good spectral efficiency. The network related functionalities have been designed to offer secure communication to authenticated users even when roaming between various networks based on the same system. Table 79.1 provides the essential characteristics of the second generation systems as initially designed in the late eighties and early nineties. Since then several additions have been made to those standards. Enhancements of air interface as well as network subsystem functionalities have been incorporated.

79.3 Capacity Enhancement

The **capacity** of a mobile network can be defined as the Erlangs throughput by a cell, a cluster of cells, or by a portion of a network. For a given radio interface, the achievable capacity is a function of the robustness of the physical layer, the effectiveness of the medium access control (MAC) layer and the multiple access technique. Moreover, it is strongly dependent on the radio spectrum available for network planning. If we define:

- BW Available radio spectrum
- Wc Spectrum occupied by a single radio carrier
- Nc Number of circuits handled by a single carrier (e.g., number of time slots per frame in a TDMA system)
- Cs Permissible cluster size which guarantees a good quality for a vast majority of active calls (like 90 to 95%)

The number of circuits per cell is given by $(BW/Wc) \times (Nc/Cs)$.

The Erlang capacity can be derived from this expression after consideration of the signalling overhead required by the pilot channel, signalling and traffic channel overheads for handover, and the specified call blocking rate.

This definition of capacity is applicable in case of TDMA/FDMA air interfaces and when the networks are designed using "deterministic frequency allocation." A slightly different approach to the problem is necessary in systems like DECT which use dynamic channel selection and for DS-CDMA systems. IS-95 networks pretend to use a cluster size (Cs) of one. However, the number of good quality circuits in a cell is a function of the available noninterfered spreading codes.

Capacity enhancement can be obtained through:

- An increase in number of radio carriers and/or the number of traffic channels (e.g., voice circuits)
- Improved interference management techniques
- Novel cellular configurations

Capacity Enhancement Through Increase in Number of Carriers and/or Voice Circuits

This family of relatively simple solutions can be subdivided in three categories related to the availability of spectrum in the same band, or in a different band where the same radio interface can be utilized or the situation in which the number of voice circuits can be increased by the introduction of multiple speech codecs.

TABLE 79.1 Air Interface Characteristics of Second Generation Systems

	Cellular				Cordless	
Standard	GSM	IS-54	IS-95	PDC	DECT	PHS
Frequency band (MHz)	Europe	USA	USA	Japan	Europe	Japan
Uplink	890–915 (1710–1785)	824–849	824–849	940–956 (1429–1441)	1880–1900	1895–1907
Downlink	935–960 (1805–1880)	869–894	869–894	810–826 (1477–1489)		
Duplex Spacing (MHz)	45 (95)	45	45	130 (48)	—	—
Carrier spacing (kHz)	200	30	1250	25	1728	300
Number of radio channels in the frequency band	124 (374)	832	20	640 (480)	10	77
Multiple access	TDMA	TDMA	CDMA	TDMA	TDMA	TDMA
Duplex mode	FDD	FDD	FDD	FDD	TDD	TDD
Number of channels per carrier	8 (half rate: 16)	3 (half rate: 6)	128	3 (half rate: 6)	12	4
Modulation	GMSK	II/4 DQPSK	QPSK BPSK	II/4 DQPSK	GFSK	II/4 DQPSK
Carrier bit rate (kb/s)	270.8	48.6	1288	42	1152	384
Speech coder (full rate)	RPE-LTP	VSELP	QCELP	VSELP	ADPCM	ADPCM
Net bit rate (kb/s)	13	7.95	(var.rate: 8, 4, 2, 0.8)	6.7	32	32
Channel coder for speech channels	1/2 rate convolutional + CRC	1/2 rate convolutional + CRC	1/2 (downlink), 1/3 (uplink) convolutional + CRC	1/2 rate convolutional + CRC	no	no
Gross bit rate speech + channel coding (kb/s)	22.8	13	var. rate 19.2, 9.6, 4.8, 2.4	11.2	—	—
Frame size (ms)	4.6	40	20	20	10	5
	Peak Aver.	Pak Aver.				
	8 1	9 3		Peak Aver.	Peak Aver.	Peak Aver.
MS transmission power (W)	2 0.25 (1) (0.125)	4.8 1.6 1.8 0.6	0.6	2 0.66	0.25 0.01	0.08 0.01
Power control						
MS	Y	Y	Y	Y	N	Y
BS	Y	Y	Y	Y	N	Y
Operational C/I(dB)	9	16	6	17	21	26
Equalizer	Needed	Needed	Rake receiver	Option	Option	No
Handover	Y	Y	Soft handoff	Y	Y	Y

TABLE 79.2 Applicability Matrix of Methods for Capacity Enhancements Based on Increased Spectrum Availability

Second Generation System	Method for Capacity Enhancement: Additional Spectrum in Same Band	Additional Spectrum in Another Band	Multiple Speech Codecs
GSM Family	Available	Under trial in real networks	Under trial in real networks
IS-54/IS-136	Available	Applicable	Applicable
PDC	Available	Applicable	Applicable
IS-95 Family using DS-CDMA	Some of these solutions are applicable. However, their implementation shall adversely effect the functioning of soft handover which is an essential feature of DS-CDMA networks.		
DECT	Available	Under Consideration	Under Consideration
PHS	Available	Not Applicable	Applicable

Increased Spectrum Availability

In such a fortunate situation, extra carriers can be added to the already installed base stations (BS). The existing Cs is maintained and the increase in capacity is according to the added carriers/circuits. An additional gain in Erlang capacity is available due to increased trunking efficiency.

Supplementary Spectrum Availability (in Different Frequency Bands)

Most of the second generation systems were implemented in the 800 and 900 MHz frequency bands. Their application has been extended to the 1800 and 1900 MHz bands too (e.g., GSM 1800 and PCS1900). Carriers from both the bands can be used at the same BSs—either according to a common or independent frequency reuse scheme. Increase in traffic throughput can be maximized by using dual band mobile stations (MSs) and by providing adequate handover mechanisms between the carriers of two bands. Also, due to difference in propagation for carriers from two widely separated bands, certain adjustments related to coverage planning might be required when the carriers are co-sited.

Multiple Speech Codecs for Increased Number of Voice Circuits

Originally, only full rate (FR) speech codecs were used in second generation systems. Hence, a one-to-one correspondence between the physical channels on the air interface and the available voice circuits was established. If the number of installed carriers in a cell is kept unchanged, the introduction of half rate (HR) codecs will double the number of voice circuits in the cell and a more than two-fold increase in Erlang capacity in the cell will be achievable. A similar possibility is offered by the adaptive multirate codec (AMR). The output bit rates of speech codec and channel codec can be adapted in order to minimize the carrier occupancy time necessary to offer a predetermined call quality. The statistical multiplexing gain thus obtained can be exploited to enhance the traffic throughput.

Such capacity enhancement methods can be implemented only if corresponding MSs or MSs with multiple codec capabilities are commercialized. Moreover, every cell site in a network will have to maintain a certain number of FR voice circuits necessary to offer downwards compatibility. Table 79.2 provides an applicability matrix related to the above mentioned capacity enhancement methods.

Improved Interference Management Techniques

Initially, the design of a cellular network is based on some simplifying assumptions like uniformly distributed subscriber density and traffic patterns or homogeneous propagation conditions. Usually, such design results in a worst case value of Cs. The situation can be improved through better management of cochannel interference by implementing one or a combination of the following:

- Slow frequency hopping (SFH)
- Voice activity detection (VAD) and discontinuous transmission (DTx)

- Transmit power control (PC)
- Antenna beamforming

Slow Frequency Hopping

Every call is "spread" over all the carrier frequencies available in the cell. To avoid intracell interference, orthogonal (random or pseudo random) frequency hopping laws are used for different calls. Also, the worst-case intercell interference situation can last only one hop and an averaging out of cochannel interference in the network occurs. Statistically, the distribution of carrier-to-cochannel interference ratio in the network is more compact with frequency hopping than without it, and this can be exploited to reduce the cluster size Cs and increase capacity.

Voice Activity Detection (VAD) and Discontinuous Transmission (DTx)

Collection of statistics related to telephone conversations has demonstrated that the duty cycle of voice sources is around 40%. With DTx, the radio signal is transmitted according to the activity of voice source. An overall reduction in the averaged cochannel interference experienced by the calls can thus be observed. This offers the possibility of implementing smaller Cs. Also, saving energy with DTx in the up-link (UL) results in a prolonged autonomy for the MS.

Transmit Power Control

Usually, the full available transmit power is necessary for the initial access and for a short duration after call establishment. For the rest of the call, both DL and UL transmit powers can be reduced to a level necessary to maintain a good link quality. The overall improvement in radio interference in the network thus obtained is helpful for the reduction of cluster size. Like VAD/DTx, the DL transmit power control is not permitted for pilot carriers if the mobile assisted handover is implemented in the system.

Antenna Beamforming

Interference related to every call can be individually managed by "dynamic cell sectorization." Actually, the base station receiver captures the up-link signal in a narrow antenna beam dynamically "placed" around the MS. Similarly, the down-link signal is transmitted in a beam focused towards the MS. This sort of spatial filtering of cochannel interference is useful for implementing very compact frequency reuse schemes. Generally, both the down-link and up-link beamforming capabilities are placed at the BS where an antenna array and signal processing algorithms related to direction finding, signal source separation, and beam synthesis have to be implemented. Table 79.3 provides an applicability matrix for the second generation systems.

Capacity enhancements of 200% or more have been reported [10,13] through the implementation of combinations of SFH, VAD/DTx, and PC in GSM networks. With SFH and antenna beam forming, Cs of three is achievable for GSM networks [1]. Since VAD/DTx and PC cannot be applied to the pilot

TABLE 79.3 Applicability Matrix of Methods for Capacity Enhancements Based on Improved Interference Management

	Method for Capacity Enhancement			
Second Generation System	Slow Frequency Hopping	VAD/DTx	Power Control	Antenna Beam Forming
GSM Family	AAP	AAP	AAP	APP
IS-95 (DS-CDMA)	Not Applicable	Essential requirements for satisfactory system operation and not capacity enhancement features		APP
IS-54/IS-136	ANP	ANP	AAP	APP
PDC	ANP	ANP	AAP	APP
DECT	ANP	ANP	ANP	APP
PHS	ANP	ANP	ANP	APP

Note: AAP Applicable and already provided by the standard.
ANP Applicable but not explicitly provided by the standard.
APP Applicable depending on BS equipment design.

carriers, most of the operational networks deploy a dual cluster scheme where Cs for pilot carriers is slightly higher than Cs for traffic carriers. Moreover, some other issues related to pilot channels/carriers have somewhat impeded the introduction of antenna beamforming in GSM or IS-95 (DS-CDMA) networks for cellular applications. However, in wireless local loop applications, substantial capacity gains have been reported for IS-95 [12].

Novel Cellular Configurations

The traffic capacity throughput by a regular grid of homogeneous cells can be increased by cell splitting. Theoretically speaking, if the cell size is divided by two by adding base stations in the middle of existing ones and the current frequency reuse scheme is maintained the achievable traffic capacity is multiplied by four. However, a reduction in cell size beyond a limit leads to excessive overlap between cells due to increased difficulty of coverage prediction. Moreover, since the dwelling time of MS in a cell is reduced, the average number of handovers per call increases. Conventional cellular organizations can no longer meet the requirements for good quality of service due to the failure of handover mechanisms. For high capacity coverage, the following novel cellular configurations have been suggested/implemented:

- Microcells and the associated hierarchical network organization
- Concentric cells
- Frequency reuse of one through generalized slow frequency hopping

Microcells and Associated Hierarchical Network Organization

Hot spot coverage in dense urban areas is realized in the form of isolated islands of microcells. This is implemented through micro-BS antennas placed below the roof tops of the surrounding buildings. Each antenna radiates very low power. Each island is covered by an umbrella cell which is a part of a continuous macrocellular network over a wider area. Traffic throughput is optimized by intelligent spectrum management performed either off-line or on-line. A set of carriers is assigned to the macrocellular layer organized in a conventional manner. The remaining carriers are repeatedly used in the islands of microcells. Cell selection parameters in the MS and the call admission control algorithms (e.g., Forced Directed Retry) in the base station controllers are adjusted such that a maximum of traffic in the hot spot is taken by the microcells. The umbrella cells are dimensioned to handle the spill-over traffic. This can be the fast moving MS which could generate too many handovers if kept with the microcells. An MS which experiences a sudden degradation of link budget with respect to its serving microcell and/or in the absence of a good target microcell can be temporarily handed over to the umbrella cell.

Despite its difficulty related to spectrum management, this technique has proved to be quite popular for the densification of parts of existing networks based on second generation systems using TDMA/FDMA. [6,11] provide analysis and guidelines for spectrum management for optimized efficiency in hierarchical networks.

Concentric Cells

A concentric cell coverage is implemented by splitting the available traffic carriers in two groups. One group transmits at full power required to cover the complete cell and the other group transmits at a lower level thus providing the coverage of an inner zone concentric with the original cell. The pilot carrier is transmitted at full power. The localized transmission in the inner zone creates a lower level of interference for other cells. Hence, a smaller cluster size can be used for the frequencies of the inner zone leading to traffic capacity enhancement. Call admission control is designed to keep the MSs near the BS on the carriers of inner zone. Simple intracell handover mechanisms are used to ensure call continuity for the MSs moving across the boundary of the inner zone. Analysis and simulations have shown that optimized capacity enhancement is achieved if the inner zone is limited to 40% of the total cell area. Field trials of concentric cell networks have demonstrated 35% higher spectrum efficiency as compared to single cell networks.

TABLE 79.4 Applicability Matrix of Methods for Capacity Enhancements Based on Novel Cellular Configurations

Second Generation System	Method for Capacity Enhancement		
	Micro-cells and Hierarchical N/W	Concentric Cells	Reuse of One with GSFH
GSM Family	AAP	AAP	AAP
IS-54/IS-136	AAP	AAP	ANP
PDC	AAP	AAP	ANP
IS-95 Family using DS-CDMA	Not applicable due to the resulting near–far problem and the power control complexity		Inherent due to DS-CDMA
DECT	AAP	AAP	ANP. Reuse of one possible with DCS
PHS	AAP	AAP	ANP

Note: AAP Applicable and already provided by the standard.
ANP Applicable but not explicitly provided by the standard.
APP Applicable depending on BS equipment design.

Reuse of One Through Generalized Frequency Hopping and Fractional Loading

Micro- or picocellular networks with TDMA/FDMA systems can be deployed with frequency "reuse of one." Every base station has the capability of using all the available carriers by slow frequency hopping. The allocation of frequency hopping laws for MSs in clusters of adjacent cells is managed by a centralized control.

During the steady state of operation in a loaded network, only a fraction of the total available bandwidth is in active use in every cell (fractional loading). The level of interference for active calls, for unused circuits and the availability of noninterfered frequency hopping laws is constantly monitored by the network. New calls in a cell are accepted according to the availability of interference-free circuits. In case of unevenly distributed traffic load, an average circuit occupancy in the network is maintained at a level necessary to keep the interference level for active calls below a predetermined threshold. Extreme situations where the same circuit and/or same frequency hopping law is used in two adjacent cells are remedied by intracell handover of one of the two calls. Very high capacity enhancement has been demonstrated in operational GSM networks by using this technique.

Table 79.4 provides the applicability matrix for novel cellular configurations.

79.4 Quality Enhancement

Quality Aspects and Definitions

The quality of service in a telecommunications network can simply be defined as the average performance perceived by the end user in setting up and maintaining a communication. However, its assessment is complex since it is influenced by many parameters, especially in digital wireless systems. In these systems the quality is primarily based on the end-to-end bit error rate and on the continuity of radio links between the two ends. Interference-free radio coverage with sufficient desired signal strength needs to be provided to achieve the above. Moreover, communication continuity has to be ensured between coverage areas with high traffic density (microcells) and low/medium traffic density (macrocells). Three main quality aspects will be distinguished in the following, namely the call handling quality, the communication quality, and the coverage quality.

For speech transmission, the communication quality strongly depends on the intrinsic performance of the speech coder, and its evaluation normally requires intensive listening tests. When it is comparable to the quality achieved on modern wire-line telephone networks, it is called "toll quality." But the **speech quality** is also influenced by other parameters linked to the communication characteristics like radio channel impairments (bit error rate), transmission delay, echo, background noise and tandeming (i.e., when several coding/decoding operations exist in the link).

For data transmission, the communication quality can be more easily quantified based on bit error rate and transmission delay. In synchronous circuit mode, delay is fixed and bit error rate depends on radio channel quality. In packet mode, bit error rate can be kept low thanks to retransmission mechanisms, but average delay increases and throughput decreases as the radio channel degrades.

The **coverage quality** is the percentage of the served area where a communication can be established. It is determined by the acceptable path loss of the radio link and by the propagation characteristics in the area. The radio link budget generally includes some margin depending on the type of terrain (for shadowing effects) and on operator's requirements (for indoor penetration). A coverage quality of 90% is a typical value for cellular networks.

The call handling quality mainly depends on the capacity of the mobile network. When the user is under coverage of the network, the call set-up performance is measured by the blocking rate which depends on the network load. Since in cellular networks mobility of the users is high, capacity requirements are less predictible than in fixed networks which results in a higher blocking rate (typically 1%). Requirements are even more stringent in cordless systems.

A last call-handling quality attribute specific to mobile networks is the success rate in maintaining the communication for mobile users. The handover procedure triggered in cellular networks when the user moves from one cell to another implies the establishment of a new link with some risk to lose the call. The performance is here given by the handover success rate. It must be noted that even if successful, a handover generally results in a short transmission break which degrades the communication quality.

In fact, capacity and quality are dual parameters and a compromise is needed when designing a mobile network. The call handling quality is strongly linked to the correct dimensioning of the network capacity. On the other hand, offering a high capacity in a mobile network implies an intensive use of the available radio spectrum and hence a high average interference level, which may in turn degrade the communication quality. In the following, some major enhancements in speech quality and coverage quality standardized or implemented in second generation systems are reviewed.

Speech Quality Enhancements

In digital cellular systems the first generation of speech coders (full rate) were standardized in the late 1980s to provide good communication quality at medium bit rate (6.7 to 13 kb/s). Operation at a fixed bit rate matched to the intended communication channel was the most important requirement in creating these standards. The main characteristics of the full rate speech coders standardized for second generation cellular systems are listed in Table 79.1 together with other air interface parameters.

With the recent advances in speech coding techniques, low delay toll quality at 8 kb/s and near-toll quality below 6 kb/s are now available [2]. This has enabled the standardization of half rate codecs in cellular systems, though with an increased complexity of implementation. In GSM/DCS, IS-54, and PDC this evolution was anticipated at an early stage, so the need for a half rate speech channel was taken into account in the design of the TDMA frame structure. Half rate codecs provide a doubling of the systems capacity (in terms of number of voice circuits per cell site) while maintaining a speech quality comparable to that available from related full rate codecs. A 5.6 kb/s VSELP coder for GSM and a 3.45 kb/s PSI-CELP coder for PDC were standardized in 1993. However, they have not been widely introduced up to now because of their slightly lower quality in some conditions (e.g., background noise, tandeming) compared to their full rate counterpart.

On the contrary, operators have pushed toward speech quality enhancements, and a new generation of speech coders has emerged. In 1996 ETSI standardized the GSM EFR (enhanced full rate) coder, adopting without any competitive selection process the US1 coder defined for PCS1900 in the U.S. The EFR coder has the remarkable feature of keeping the same channel coding as for the full rate channel, hence simplifying its implementation in the infrastructure. The ACELP coder (using an algebraic codebook with some benefits in computational efficiency like in the the 8 kb/s G.729 ITU standard) has a net bit rate of 12.2 kb/s which leaves room for additional protection (CRC).

Similarly the IS-641 coder has been standardized for IS-54 cellular systems. Its structure is very similar to that of G.729 but with a frame size of 20 ms and a bit rate of 7.4 kb/s. Also a 13 kb/s QCELP coder has been standardized for IS-95. These new coders provide near toll quality in ideal transmission conditions.

Concerning cordless systems, both DECT and PHS use 32 kb/s ADPCM and, therefore, provide toll quality in normal conditions thanks to their relatively high operational C/I. For these systems the concern is rather to introduce lower rate coders with equivalent quality in some kind of half rate mode, allowing for capacity increase, but such coders have not been standardized yet.

With the emergence of variable bit rate (VBR) techniques, some further developments in speech coding are now taking place to satisfy the requirements of cellular operators for toll quality speech with better robustness to radio channel impairments, combined with the capacity increase achievable with half rate operation. One VBR approach is to adapt the bit rate according to the source requirements, taking advantage of silence and stationary segments in the speech signal. Another VBR approach is to adapt the bit rate according to the radio channel quality, either by varying the bit rate allocation between speech and channel coding within a constant gross bit rate (for better quality) or by reducing the gross bit rate in good transmission conditions (for higher capacity). Adaptation to the source characteristics is fast (on a frame-by-frame basis) and adaptation to the channel quality is slower (maximum, a few times per second). In all cases, the underlying idea is that current coders are dimensioned for worst-case operation so that the system quality and capacity can be increased by exploiting the large variations over time of the bit rate requirements (for a given quality).

VBR capabilities with source driven adaptation were introduced since the beginning in IS-95. In CDMA systems the resulting average bit rate reduction directly translates into increased capacity. The initial IS-96B QCELP coder supports four bit rates (8, 4, 2, and 0.8 kb/s) with a scalable CELP architecture, and bit rate adjustment is performed based on adaptive energy thresholds. In practice the two extreme rates are used most frequently. The achievable speech quality was estimated to be lower than that of the IS-54 VSELP coder. Subsequently an enhanced variable rate coder (EVR) was standardized as IS-127. It is based on the relaxed CELP (RCELP) coding technique and supports three rates (8.5, 4, and 0.8 kbit/s).

In Europe, ETSI has launched in 1997 the standardization of an Adaptive Multirate coder (AMR) for GSM systems. Its output bit rate is continuously adapted to radio channel conditions and traffic load. The objective is to find the best compromise between speech quality and capacity by selecting an optimum combination of channel mode and codec mode. The AMR codec can operate in two channel modes, full rate and half rate. For each channel mode there are several possible codec modes (typically three) with different bit rate allocations between speech and channel coding. An adaptation algorithm tracks the variations in speech quality using specific metrics and decides upon the changes in codec mode (up to several times per second). Changes in codec mode is detected by the receiver either via in-band signalling or by automatic mode identification. As shown in Fig. 79.1, the multiplicity of codec modes gives significant performance improvement over any of the corresponding fixed rate codecs.

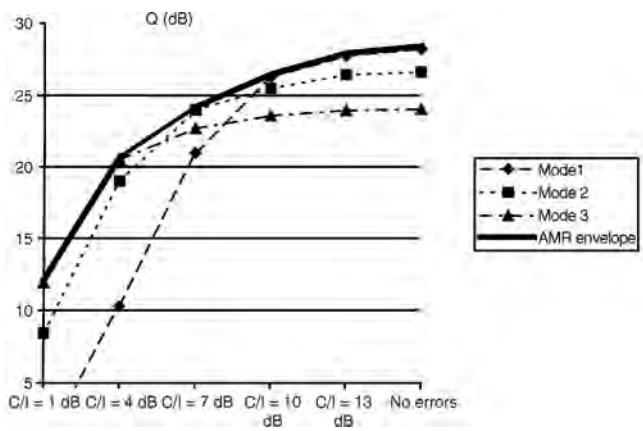


FIGURE 79.1 AMR codec quality as a function of C/I. (Typical results derived from the first AMR selection tests. Quality is given on the equivalent MNRU scale in dB.)

AMR codec will also allow handovers between half rate and full rate channels using intracell handover mechanisms. The AMR codec was selected by October 1998 and the full standard was available in mid-1999. A potential extension is wideband coding, which could be added later to AMR as an option. The wideband option would extend the audio bandwidth from the current 300–3400 Hz to 50–5000 Hz or even 50–7000 Hz.

Speech quality enhancements in cellular networks do not concern only the speech coder itself. First, it is well known that tandeming (i.e., several cascaded coding/decoding operations) can be a source of significant degradation of speech quality. In the case of mobile-to-mobile calls there is no need to decode the speech signal in the network (assuming both mobiles use the same speech codec). Therefore, a tandem-free operation (TFO) will soon be introduced for GSM using in-bandsignalling between peer transcoders. The benefit of TFO, however, will largely depend on the percentage of intra-GSM mobile-to-mobile calls.

In CDMA systems like IS-95, the soft handoff feature enables a smooth transition from one cell to another. On the contrary in TDMA systems, handovers produce short breaks in the communication which locally degrade the speech quality in spite of the speech extrapolation mechanism used in the decoder. In GSM the duration of the transmission break can be as long as 160 ms due to the necessary time alignment of the mobile in the new cell. The speech interruption can be reduced by 40 ms in a synchronized network, but in practice local synchronization is only offered for co-sited cells. The same performance can be more easily achieved with the presynchronized handover where the mobile receives an indication of the distance to the base station together with the handover command. Some improvement can also be obtained on the uplink by switching at the right point in time and on the downlink by broadcasting the speech information to both the old and the new cell. As a result of all these improvements, the interruption time can be reduced down to 60 ms.

Some advances have also been made concerning the robustness to radio channel impairments. Unequal error protection (UEP) is now used in most codecs designed for cellular systems and it is often based on rate-compatible punctured convolutional codes. UEP enables the adjustment of the protection rate as a function of the sensitivity of bits output by the speech coder. Besides, most linear predictive coders use a speech extrapolation procedure, replacing potentially corrupted information in the current frame with more reliable information from a previously received frame. More sophisticated error concealment techniques, using both reliability information at the output of the channel decoder and a statistical source model (*a priori* knowledge), have been reported to provide up to 3 dB improvement in E_b/N_0 under adverse channel conditions [5].

Robustness to background noise is another topic of interest. Speech quality improvements in medium- to low-bit rate coders have been obtained with coding algorithms optimized for speech signals. Such algorithms may produce poor results in the presence of background noise. Therefore, optional noise cancellation processing has been introduced in recent standards. This can be performed either in the time domain (Kalman filtering) as in JDC half rate or in the frequency domain (spectral subtraction) as in IS-127 EVR.

Coverage Quality Enhancements

Various second generation systems provide the possibility of implementing mechanisms like slow frequency hopping (as in GSM), antenna diversity (also called microdiversity), macrodiversity or multisite transmission, and dynamic channel selection (as in DECT). Such mechanisms are useful to alleviate the effects of radio transmission phenomena like shadowing, fading, and cochannel interference. They are, therefore, particularly interesting to enhance the coverage quality in interference-limited or strong multipath environments (e.g., urban areas). Optional equalizers have been defined for cordless systems (DECT and PHS) with the introduction of a suitable training sequence ("prolonged preamble" in DECT), allowing for channel estimation in the presence of longer delay spread and thus providing increased coverage and/or quality.

Operators also have to face network planning issues linked to natural or artificial obstacles. Various solutions have been designed for the cases where the coverage cannot be efficiently ensured with regular base stations. Radiating cables (leaky feeders) are typically used for the coverage of tunnels. Radio repeaters have been standardized for GSM and for DECT (Wireless Relay Station, WRS) and are useful to fill coverage holes and to provide outdoor-to-indoor or outdoor-to-underground coverage extensions. Microcells (and microbase stations) may also be used as "gap fillers."

The use of smart antennas at cell sites helps to extend the cell range. In rural environments where traffic is low, the required number of cell sites is minimized by using antenna arrays and signal processing algorithms to ensure high sensitivity reception at the base station. Algorithms that implement either n-fold receive diversity or mobile station direction finding followed by beamforming for mobile station tracking have been shown to perform well in such radio channels. Adaptive beamforming techniques with an M-element antenna array generally provide a directivity gain of M, e.g., 9 dB with $M = 8$ (plus some diversity gain depending on channel type). It results in a significant increase of the uplink range (typically by a factor >1.5 with $M = 8$), but an increased range in the downlink is also needed to get an effective reduction of the number of cell sites. An increased downlink range can be achieved using adaptive beamforming (but with a much higher complexity compared to the uplink-only implementation), a multibeam antenna (i.e., a phased array doing fixed beamforming), or an increased transmit power of the base station. However, the success of smart antenna techniques for range extension applications in second generation systems has been slowed down by their complexity of implementation and by operational constraints (multiple feeders, large antenna panels).

79.5 High Bit Rate Data Transmission

Circuit Mode Techniques

All second generation wireless systems support circuit mode data services with basic rates typically ranging from 9.6 kb/s (in cellular systems) to 32 kb/s (in cordless systems) for a single physical radio resource. With the growing needs for higher rates, new services have been developed based on multiple allocation or grouping of physical resource.

In GSM, HSCSD (High Speed Circuit Switched Data) enables multiple Full Rate Traffic Channels (TCH/F) to be allocated to a call so that a mobile subscriber can use n times the transmission capacity of a single TCH/F channel (Fig. 79.2). The n full rate channels over which the user data stream is split are handled completely independently in the physical layer and for layer 1 error control. The HSCSD channel resulting from the logical combination of n TCH/F channels is controlled as a single radio link during cellular operations such as handover. At the A interface, calls will be limited to a single 64 kb/s circuit. Thus HSCSD will support transparent (up to 64 kb/s) and nontransparent modes (up to $4 \times 9.6 = 38.4$ kb/s and, later, $4 \times 14.4 = 57.6$ kb/s). The initial allocation can be changed during a call if

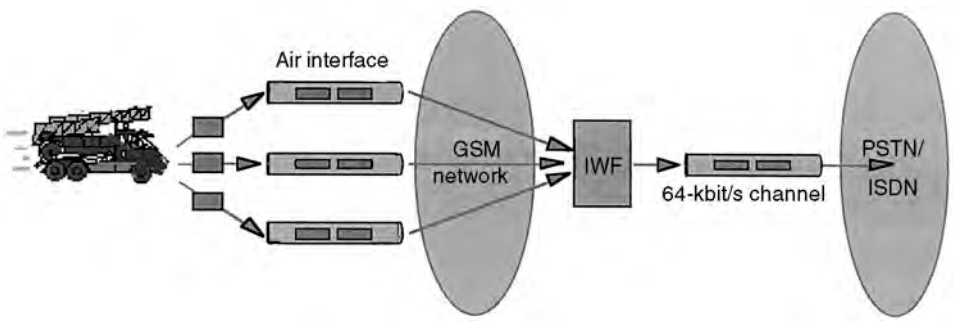


FIGURE 79.2 Simplified GSM network configuration for HSCSD.

required by the user and authorized by the network. Initially the network allocates an appropriate HSCSD connection according to the requested user bit rate over the air interface. Both symmetric and asymmetric configurations for bidirectional HSCSD operation are authorized. The required TCH/F channels are allocated over consecutive or nonconsecutive timeslots.

Similar multislot schemes are envisaged or standardized for other TDMA systems. In IS-54 and PDC, where radio channels are relatively narrowband, no more than three time slots can be used per carrier and the achievable data rate is therefore limited to, say, 32 kb/s. On the contrary, in DECT up to 12 time slots can be used at 32 kb/s each, yielding a maximum data rate of 384 kb/s. Moreover, the TDD access mode of DECT allows asymmetric time slot allocation between uplink and downlink, thus enabling even higher data rates in one direction.

Packet Mode Techniques

There is a growing interest for packet data services in second generation wireless systems to support data applications with intermittent and bursty transmission requirements like the Internet, with a better usage of available radio resources, thanks to the multiplexing of data from several mobile users on the same physical channel. Cellular Digital Packet Data (CDPD) has been defined in the U.S. as a radio access overlay for AMPS or D-AMPS (IS-54) systems, allowing packet data transmission on available radio channels. However, CDPD is optimized for short data transmission and the bit rate is limited to 19.2 kb/s. A CDMA packet data standard has also been defined (IS-657) which supports CDPD and Internet protocols with a similar bit rate limitation but allowing use of the same backhaul as for voice traffic.

In Europe, ETSI has almost completed the standardization of GPRS (General Packet Radio Service) for GSM. A GPRS subscriber will be able to send and receive in an end-to-end packet transfer mode. Both point-to-point and point-to-multipoint modes are defined. A GPRS network coexists with a GSM PLMN as an autonomous network. In fact, the Serving GPRS Support Node (SGSN) interfaces with the GSM Base Station Controller (BSC), an MSC and a Gateway GPRS Service Node (GGSN). In turn, the GGSN interfaces with the GGSNs of other GPRS networks and with public Packet Data Networks (PDN). Typically, GPRS traffic can be set up through the common control channels of GSM, which are accessed in slotted ALOHA mode. The layer 2 protocol data units, which are about 2 kbytes in length, are segmented and transmitted over the air interface using one of the four possible channel coding schemes. The system is highly scaleable as it allows from one mobile using 8 radio time slots up to 16 mobiles per time slot, with separate allocation in up- and downlink. The resulting peak data rate per user ranges from 9 kb/s up to 170 kb/s. Time slot concatenation and variable channel coding to maximize the user information bit rate are envisaged for future implementations. This is indicated by the mobile station, which provides information concerning the desire to initiate in-call modifications and the channel coding schemes that can be used during the call set up phase. It is expected that use of the GPRS service will initially be limited and traffic growth will depend on the introduction of GPRS capable subscriber terminals. Easy scalability of the GPRS backbone (e.g., by introducing parallel GGSNs) is an essential feature of the system architecture (Fig. 79.3).

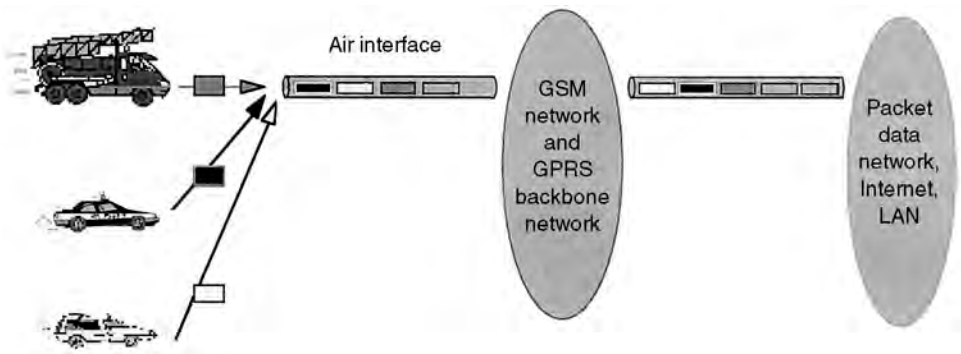


FIGURE 79.3 Simplified view of the GPRS architecture.

New Modulation Schemes

New modulation schemes are being studied as an option in several second generation wireless standards. The aim is to offer higher rate data services equivalent or close to the 2 Mb/s objective of the forthcoming third generation standards. Multilevel modulations (i.e., several bits per modulated symbol) represent a straightforward means to increase the carrier bit rate. However, it represents a significant change in the air interface characteristics, and the increased bit rate is achieved at the expense of a higher operational signal-to-noise plus interference ratio, which is not compatible with large cell dimensions. Therefore, the new high bit rate data services are mainly targetting urban areas, and the effective bit rate allocated to data users will depend on the system load.

Such a new air interface option is being standardized for GSM under the name of EDGE (Enhanced Data rates for GSM Evolution). The selected modulation scheme is 8-PSK, suitable coding schemes are under study, whereas the other air interface parameters (carrier spacing, TDMA frame structure,…) are kept unchanged. Reusing HSCSD (for circuit data) and GPRS (for packet data) protocols and service capabilities, EDGE will provide similar ECSD and EGPRS services but with a three-fold increase of the user bit rate. The higher level modulation requires better radio link performances, typically a loss of 3 to 4 dB in sensitivity and a C/I increased by 6 to 7 dB. Operation will also be restricted to environments with limited time dispersion and limited mobile speed. Nevertheless, EGPRS will roughly double the mean throughput compared to GPRS (for the same average transmitted power). EDGE will also increase the maximum achievable data rate in a GSM system to 553.6 kb/s in multislot (unprotected) operation. Six different protection schemes are foreseen in EGPRS using convolutional coding with a rate ranging from 1/3 to 1 and corresponding to user rates between 22.8 and 69.2 kb/s per time slot. This is in addition to the four coding schemes already defined for GPRS. An intelligent link adaptation algorithm will dynamically select the most appropriate modulation and coding schemes, i.e., those yielding the highest throughput for a given channel quality. The first phase of EDGE standardization was completed by end 1999. It should be noted that a similar EDGE option is being studied for IS-54/IS-136 (and their PCS derivatives). Initially, the 30 kHz channel spacing will be maintained and then extension to a 200 kHz channel will be provided in order to offer a convergence with its GSM counterpart.

A higher bit rate option is also under standardization for DECT. Here it is seen as an essential requirement to maintain backward compatibility with existing equipment so the new multilevel modulation will only affect the payload part of the bursts, keeping the control and signalling parts unchanged. This ensures that equipment with basic modulation and equipment with a higher rate option can efficiently share a common base station infrastructure. Only 4-level and 8-level modulations are considered and the symbol length, carrier spacing, and slot structure remain unchanged. The requirements on transmitter modulation accuracy need to be more stringent for 4- and 8-level modulation than for the current 2-level scheme. An increased accuracy can provide for coherent demodulation, whereby some (or most) of the sensitivity and C/I loss when using the multilevel mode can be regained. In combination with other new air interface features like forward error correction and double slots (with reduced overhead), the new modulation scheme will provide a wide range of data rates up to 2 Mb/s. For instance using (II /4-DQPSK modulation (a possible/suitable choice), an unprotected connection with two double slots in each direction gives a data rate of 384 kb/s. Asymmetric connections with a maximum of 11 double slots in one direction will also be supported.

79.6 Conclusion

Since their introduction in the early 1990s, most of the second generation systems have been enjoying exponential growth. With more than 100 million subscribers acquired worldwide in less than ten years of lifetime, the systems based on the GSM family of standards have demonstrated the most spectacular development. Despite a more regional implementation of other second generation systems, each one of those can boast a multimillion subscriber base in mobile or fixed wireless networks.

A variety of service requirements of third generation mobile communication systems are being already met by the upcoming enhancements of second generation systems. Two important trends are reflected by this:

- The introduction of third generation systems like Universal Mobile Telecommunication System (UMTS) or International Mobile Telecommunication-2000 (IMT-2000) might be delayed to a point in time where the evolutionary capabilities of second generation systems have been exhausted.
- The deployment of networks based on third generation systems will be progressive. Any new radio interface will be imposed worldwide if and only if it provides substantial advantages as compared to the present systems. Another essential requirement is the capability of downward compatibility to second generation systems.

Defining Terms

Capacity: In a mobile network it can be defined as the Erlangs throughput by a cell, a cluster of cells, or by a portion of a network. For a given radio interface, the achievable capacity is a function of the robustness of the physical layer, the effectiveness of the medium access control (MAC) layer and the multiple access technique. Moreover, it is strongly dependent on the radio spectrum available for network planning.

Cellular: Refers to public land mobile radio networks for generally wide area (e.g., national) coverage, to be used with medium- or high-power vehicular mobiles or portable stations and for providing mobile access to the Public Switched Telephone Network (PSTN). The network implementation exhibits a cellular architecture which enables frequency reuse in nonadjacent cells.

Cordless: These are systems to be used with simple low power portable stations operating within a short range of a base station and providing access to fixed public or private networks. There are three main applications, namely, residential (at home, for Plain Old Telephone Service, POTS), public-access (in public places and crowded areas, also called Telepoint), and Wireless Private Automatic Branch eXchange (WPABX, providing cordless access in the office environment), plus emerging applications like radio access for local loop.

Coverage quality: It is the percentage of the served area where a communication can be established. It is determined by the acceptable path loss of the radio link and by the propagation characteristics in the area. The radio link budget generally includes some margin depending on the type of terrain (for shadowing effects) and on operator's requirements (for indoor penetration). A coverage quality of 90% is a typical value for cellular networks.

Speech quality: It strongly depends on the intrinsic performance of the speech coder and its evaluation normally requires intensive listening tests. When it is comparable to the quality achieved on modern wire-line telephone networks, it is called "toll quality." In wireless systems it is also influenced by other parameters linked to the communication characteristics like radio channel impairments (bit error rate), transmission delay, echo, background noise, and tandeming (i.e., when several coding/decoding operations are involved in the link).

References

1. Anderson, S., Antenna Arrays in Mobile Communication Systems, *Proc. Second Workshop on Smart Antennas in Wireless Mobile Communications,* Stanford University, Jul. 1995.
2. Budagavi, M. and Gibson, J.D., Speech coding in mobile radio communications, *Proceedings of the IEEE,* 86(7), 1402–1412, Jul. 1998.
3. Cox, D.C., Wireless network access for personal communications, *IEEE Communications Magazine,* 96–115, Dec. 1992.
4. DECT, *Digital European Cordless Telecommunications Common Interface,* ETS-300–175, ETSI, 1992.
5. Fingscheidt, T. and Vary, P., Robust Speech Decoding: A Universal Approach to Bit Error Concealment, *Proc. IEEE ICASSP,* 1667–1670, Apr. 1997.

6. Ganz, A., et al., On optimal design of multitier wireless cellular systems, *IEEE Communications Magazine,* 88–93, Feb. 1997.
7. GSM, *GSM Recommendations Series 01-12,* ETSI, 1990.
8. IS-54, *Cellular System, Dual-Mode Mobile Station-Base Station Compatibility Standard,* EIA/TIA Interim Standard, 1991.
9. IS-95, *Mobile Station-Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular System,* EIA/TIA Interim Standard, 1993.
10. Kuhn, A., et al., Validation of the Feature Frequency Hopping in a Live GSM Network, *Proc. 46th IEEE Vehic. Tech. Conf.,* 321–325, Apr. 1996.
11. Lagrange, X., Multitier cell design, *IEEE Communications Magazine,* 60–64, Aug. 1997.
12. Lee, D. and Xu, C., The effect of narrowbeam antenna and multiple tiers on system capacity in CDMA wireless local loop, *IEEE Communications Magazine,* 110–114, Sep. 1997.
13. Olofsson, H., et al., Interference Diversity as Means for Increased Capacity in GSM, *Proc. EPMCC'95,* 97–102, Nov. 1995.
14. PDC, *Personal Digital Cellular System Common Air Interface,* RCR-STD27B, 1991.
15. PHS, *Personal Handy Phone System: Second Generation Cordless Telephone System Standard,* RCR-STD28, 1993.
16. Tuttlebee, W.H.W., Cordless personal communications, *IEEE Communications Magazine,* 42–53, Dec. 1992.

Further Information

European standards (GSM, CT2, DECT, TETRA) are published by ETSI Secretariat, 06921 Sophia Antipolis Cedex, France.

U.S. standards (IS-54, IS-95, APCO) are published by Electronic Industries Association, Engineering Department, 2001 Eye Street, N.W., Washington D.C. 20006, U.S.A.

Japanese standards (PDC, PHS) are published by RCR (Research and Development Center for Radio Systems), 1-5-16, Toranomon, Minato-ku, Tokyo 105, Japan.

80

The Pan-European Cellular System

Lajos Hanzo
University of Southampton

80.1 Introduction

Following the standardization and launch of the Pan-European digital mobile cellular radio system known as GSM, it is of practical merit to provide a rudimentary introduction to the system's main features for the communications practitioner. Since GSM operating licenses have been allocated to 126 service providers in 75 countries, it is justifiable that the GSM system is often referred to as the Global System of Mobile communications.

The GSM specifications were released as 13 sets of recommendations [1], which are summarized in Table 80.1, covering various aspects of the system [3].

After a brief system overview in Section 80.2 and the introduction of physical and logical channels in Section 80.3 we embark upon describing aspects of mapping logical channels onto physical resources for speech and control channels in Sections 80.4 and 80.5, respectively. These details can be found in recommendations R.05.02 and R.05.03. These recommendations and all subsequently enumerated ones are to be found in [1]. Synchronization issues are considered in Section 80.6. Modulation (R.05.04), transmission via the standardized wideband GSM channel models (R.05.05), as well as adaptive radio link control (R.05.06 and R.05.08), discontinuous transmission (**DTX**) (R.06.31), and voice activity detection (**VAD**) (R.06.32) are highlighted in Sections 80.7–80.10, whereas a summary of the fundamental GSM features is offered in Section 80.11.

80.2 Overview

The system elements of a GSM public land mobile network (**PLMN**) are portrayed in Fig. 80.1, where their interconnections via the standardized interfaces A and Um are indicated as well. The mobile station (**MS**) communicates with the serving and adjacent base stations (**BS**) via the radio interface Um, whereas the BSs are connected to the mobile switching center (**MSC**) through the network interface A. As seen

TABLE 80.1 GSM Recommendations [R.01.01]

R.00	*Preamble* to the GSM recommendations
R.01	*General structure* of the recommendations, description of a GSM network, associated recommendations, vocabulary, etc.
R.02	*Service aspects:* bearer-, tele- and supplementary services, use of services, types and features of mobile stations (MS), licensing and subscription, as well as transferred and international accounting, etc.
R.03	*Network aspects,* including network functions and architecture, call routing to the MS, technical performance, availability and reliability objectives, handover and location registration procedures, as well as discontinuous reception and cryptological algorithms, etc.
R.04	*Mobile/base station (BS) interface and protocols,* including specifications for layer 1 and 3 aspects of the open systems interconnection (OSI) seven-layer structure.
R.05	*Physical layer on the radio path,* incorporating issues of multiplexing and multiple access, channel coding and modulation, transmission and reception, power control, frequency allocation and synchronization aspects, etc.
R.06	*Speech coding specifications,* such as functional, computational and verification procedures for the speech codec and its associated voice activity detector (VAD) and other optional features.
R.07	*Terminal adaptors for MSs,* including circuit and packet mode as well as voiceband data services.
R.08	*Base station and mobile switching center* (MSC) *interface,* and transcoder functions.
R.09	*Network interworking* with the public switched telephone network (PSTN), integrated services digital network (ISDN) and, packet data networks.
R.10	*Service interworking, short message service.*
R.11	*Equipment specification and type approval specification* as regards to MSs, BSs, MSCs, home (HLR) and visited location register (VLR), as well as system simulator.
R.12	*Operation and maintenance,* including subscriber, routing tariff and traffic administration, as well as BS, MSC, HLR and VLR maintenance issues.

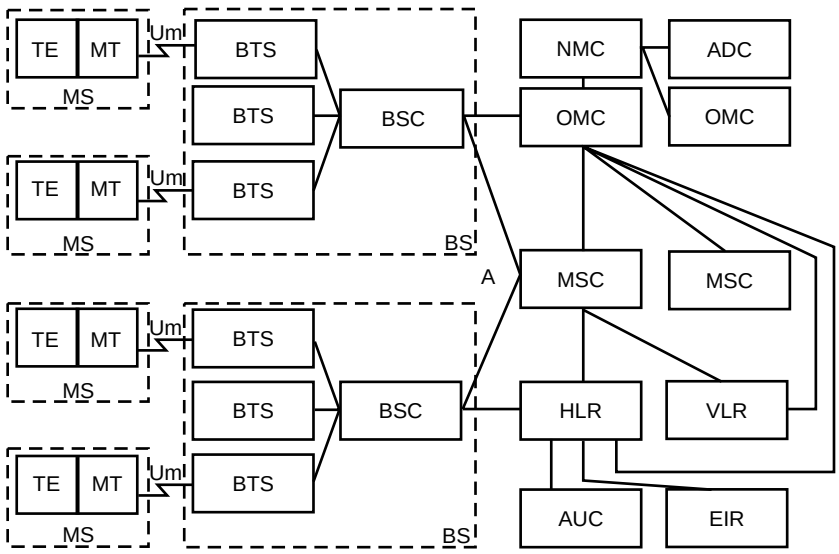


FIGURE 80.1 Simplified structure of GSM PLMN, ©ETT [4].

in Fig. 80.1, the MS includes a mobile termination (MT) and a terminal equipment (TE). The TE may be constituted, for example, by a telephone set and fax machine. The MT performs functions needed to support the physical channel between the MS and the base station, such as radio transmissions, radio channel management, channel coding/decoding, speech encoding/decoding, and so forth.

The BS is divided functionally into a number of base transceiver stations (BTS) and a base station controller (BSC). The BS is responsible for channel allocation (R.05.09), link quality and power budget control (R.05.06 and R.05.08), signalling and broadcast traffic control, frequency hopping (**FH**) (R.05.02), handover (**HO**) initiation (R.03.09 and R.05.08), etc. The MSC represents the gateway to other networks, such as the public switched telephone network (**PSTN**), integrated services digital network (**ISDN**) and packet data networks using the interworking functions standardized in recommendation R.09. The MSC's further functions include paging, MS location updating (R.03.12), HO control (R.03.09), etc. The MS's mobility management is assisted by the home location register (**HLR**) (R.03.12), storing part of the MS's location information and routing incoming calls to the visitor location register (**VLR**) (R.03.12) in charge of the area, where the paged MS roams. Location update is asked for by the MS, whenever it detects from the received and decoded broadcast control channel (**BCCH**) messages that it entered a new location area. The HLR contains, amongst a number of other parameters, the international mobile subscriber identity (**IMSI**), which is used for the authentication (R.03.20) of the subscriber by his authentication center (**AUC**). This enables the system to confirm that the subscriber is allowed to access it. Every subscriber belongs to a home network and the specific services that the subscriber is allowed to use are entered into his HLR. The equipment identity register (**EIR**) allows for stolen, fraudulent, or faulty mobile stations to be identified by the network operators. The VLR is the functional unit that attends to a MS operating outside the area of its HLR. The visiting MS is automatically registered at the nearest MSC, and the VLR is informed of the MSs arrival. A roaming number is then assigned to the MS, and this enables calls to be routed to it. The operations and maintenance center (**OMC**), network management center (**NMC**) and administration center (**ADC**) are the functional entities through which the system is monitored, controlled, maintained and managed (R.12).

The MS initiates a call by searching for a BS with a sufficiently high received signal level on the BCCH carrier; it will await and recognize a frequency correction burst and synchronize to it (R.05.08). Now the BS allocates a bidirectional signalling channel and also sets up a link with the MSC via the network. How the control frame structure assists in this process will be highlighted in Section 80.5. The MSC uses the IMSI received from the MS to interrogate its HLR and sends the data obtained to the serving VLR. After authentication (R.03.20) the MS provides the destination number, the BS allocates a traffic channel, and the MSC routes the call to its destination. If the MS moves to another cell, it is reassigned to another BS, and a handover occurs. If both BSs in the handover process are controlled by the same BSC, the handover takes place under the control of the BSC, otherwise it is performed by the MSC. In case of incoming calls the MS must be paged by the BSC. A paging signal is transmitted on a paging channel (**PCH**) monitored continuously by all MSs, and which covers the location area in which the MS roams. In response to the paging signal, the MS performs an access procedure identical to that employed when the MS initiates a call.

80.3 Logical and Physical Channels

The GSM logical traffic and control channels are standardized in recommendation R.05.02, whereas their mapping onto physical channels is the subject of recommendations R.05.02 and R.05.03. The GSM system's prime objective is to transmit the logical traffic channel's (**TCH**) speech or data information. Their transmission via the network requires a variety of logical control channels. The set of logical traffic and control channels defined in the GSM system is summarized in Table 80.2. There are two general forms of speech and data traffic channels: the full-rate traffic channels (**TCH/F**), which carry information at a gross rate of 22.8 kb/s, and the half-rate traffic channels (**TCH/H**), which communicate at a gross rate of 11.4 kb/s. A physical channel carries either a full-rate traffic channel, or two half-rate traffic channels. In the former, the traffic channel occupies one timeslot, whereas in the latter the two half-rate traffic channels are mapped onto the same timeslot, but in alternate frames.

For a summary of the logical control channels carrying signalling or synchronisation data, see Table 80.2. There are four categories of logical control channels, known as the BCCH, the common control channel (**CCCH**), the stand-alone dedicated control channel (**SDCCH**), and the associated

TABLE 80.2 GSM Logical Channels, ©ETT [4]

Logical Channels					
Duplex BS ´ MS Traffic Channels: TCH		Control Channels: CCH			
FEC-coded Speech	FEC-coded Data	Broadcast CCH BCCH BS → MS	Common CCH CCCH	Stand-alone Dedicated CCH SDCCH BS ´ MS	Associated CCH ACCH BS ´ MS
TCH/F 22.8 kb/s	TCH/F9.6 TCH/F4.8 TCH/F2.4 22.8 kb/s	Freq. Corr. Ch: FCCH	Paging Ch: PCH BS → MS	SDCCH/4	Fast ACCH: FACCH/F FACCH/H
TCH/H 11.4 kb/s	TCH/H4.8 TCH/H2.4 11.4 kb/s	Synchron. Ch: SCH	Random Access Ch: RACH MS → BS	SDCCH/8	Slow ACCH: SACCH/TF SACCH/TH SACCH/C4 SACCH/C8
		General Inf.	Access Grant Ch: AGCH BS → MS		

control channel (**ACCH**). The purpose and way of deployment of the logical traffic and control channels will be explained by highlighting how they are mapped onto physical channels in assisting high-integrity communications.

A physical channel in a time division multiple access (**TDMA**) system is defined as a timeslot with a timeslot number (TN) in a sequence of TDMA frames. The GSM system, however, deploys TDMA combined with frequency hopping (FH) and, hence, the physical channel is partitioned in both time and frequency. Frequency hopping (R.05.02) combined with interleaving is known to be very efficient in combatting channel fading, and it results in near-Gaussian performance even over hostile Rayleigh-fading channels. The principle of FH is that each TDMA burst is transmitted via a different RF channel (**RFCH**). If the present TDMA burst happened to be in a deep fade, then the next burst most probably will not be. Consequently, the physical channel is defined as a sequence of radio frequency channels and timeslots. Each carrier frequency supports eight physical channels mapped onto eight timeslots within a TDMA frame. A given physical channel always uses the same TN in every TDMA frame. Therefore, a timeslot sequence is defined by a TN and a TDMA frame number FN sequence.

80.4 Speech and Data Transmission

The speech coding standard is recommendation R.06.10, whereas issues of mapping the logical speech traffic channel's information onto the physical channel constituted by a timeslot of a certain carrier are specified in recommendation R.05.02. Since the error correction coding represents part of this mapping process, recommendation R.05.03 is also relevant to these discussions. The example of the full-rate speech traffic channel (**TCH/FS**) is used here to highlight how this logical channel is mapped onto the physical channel constituted by a so-called normal burst (**NB**) of the TDMA frame structure. This mapping is explained by referring to Figs. 80.2 and 80.3. Then this example will be extended to other physical bursts such as the frequency correction (**FCB**), synchronization (**SB**), access (**AB**), and dummy burst (**DB**) carrying logical control channels, as well as to their TDMA frame structures, as seen in Figs. 80.2 and 80.6.

The regular pulse excited (**RPE**) speech encoder is fully characterized in the following references: [3,5,7]. Because of its complexity, its description is beyond the scope of this chapter. Suffice to say that, as it can be seen in Fig. 80.3, it delivers 260 b/20 ms at a bit rate of 13 kb/s, which are divided into three significance classes: class 1a (50 b), class 1b (132 b) and class 2 (78 b). The class-1a bits are encoded by a systematic (53, 50) cyclic error detection code by adding three parity bits. Then the bits are reordered and four zero tailing

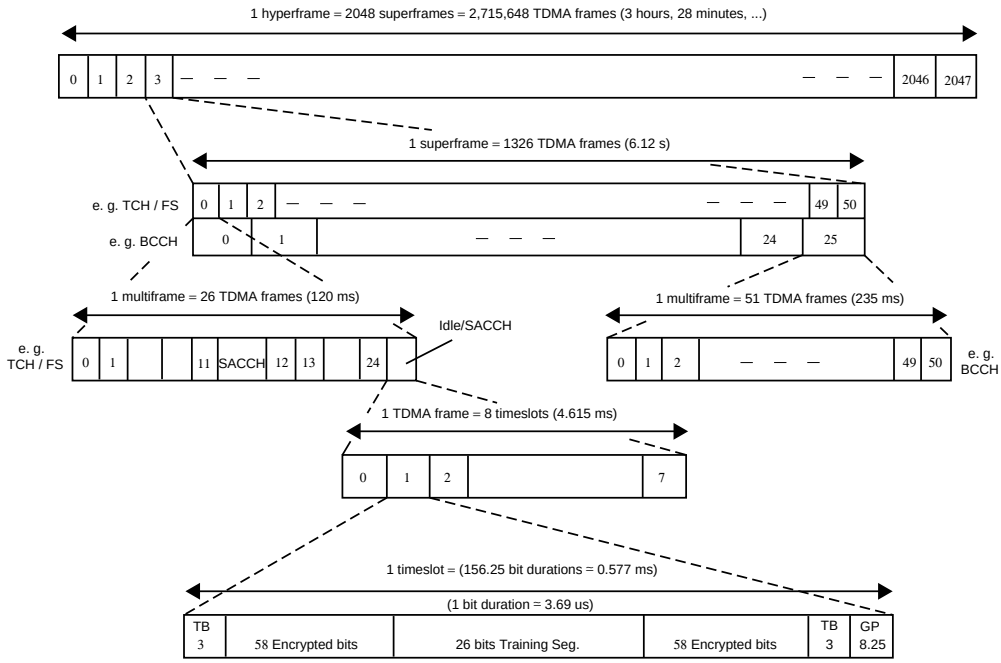


FIGURE 80.2 The GSM TDMA frame structure, ©ETT [4].

bits are added to periodically reset the memory of the subsequent half-rate, constraint length five convolutional codec (**CC**) CC(2, 1, 5), as portrayed in Fig. 80.3. Now the unprotected 78 class-2 bits are concatenated to yield a block of 456 b/20 ms, which implies an encoded bit rate of 22.8 kb/s. This frame is partitioned into eight 57-b subblocks that are block diagonally interleaved before undergoing intraburst interleaving. At this stage each 57-b subblock is combined with a similar subblock of the previous 456-b frame to construct a 116-b burst, where the flag bits *hl* and *hu* are included to classify whether the current burst is really a TCH/FS burst or it has been stolen by an urgent fast associated control channel (**FACCH**) message. Now the bits are encrypted and positioned in an NB, as depicted at the bottom of Fig. 80.2, where three tailing bits (**TB**) are added at both ends of the burst to reset the memory of the Viterbi channel equalizer (**VE**), which is responsible for removing both the channel-induced and the intentional controlled intersymbol interference [6].

The 8.25-b interval duration guard period (GP) at the bottom of Fig. 80.2 is provided to prevent burst overlapping due to propagation delay fluctuations. Finally, a 26-b equalizer training segment is included in the center of the normal traffic burst. This segment is constructed by a 16-b Viterbi channel equalizer training pattern surrounded by five quasiperiodically repeated bits on both sides. Since the MS has to be informed about which BS it communicates with, for neighboring BSs one of eight different training patterns is used, associated with the so-called BS color codes, which assist in identifying the BSs.

This 156.25-b duration TCH/FS NB constitutes the basic timeslot of the TDMA frame structure, which is input to the Gaussian minimum shift keying (**GMSK**) modulator to be highlighted in Section 80.7, at a bit rate of approximately 271 kb/s. Since the bit interval is 1/(271 kb/s) = 3.69 μs, the timeslot duration is $156.25 \cdot 3.69 \approx 0.577$ ms. Eight such normal bursts of eight appropriately staggered TDMA users are multiplexed onto one radio frequency (**RF**) carrier giving a TDMA frame of $8 \cdot 0.577 \approx 4.615$-ms duration, as shown in Fig. 80.2. The physical channel as characterized earlier provides a physical timeslot with a throughput of 114 b/4.615 ms = 24.7 kb/s, which is sufficiently high to transmit the 22.8 kb/s TCH/FS information. It even has a reserved capacity of 24.7 - 22.8 = 1.9 kb/s, which can be exploited to transmit slow control information associated with this specific traffic channel, i.e., to construct a so-called slow associated control channel (**SACCH**), constituted by the SACCH TDMA frames, interspersed with traffic frames at multiframe level of the hierarchy, as seen in Fig. 80.2.

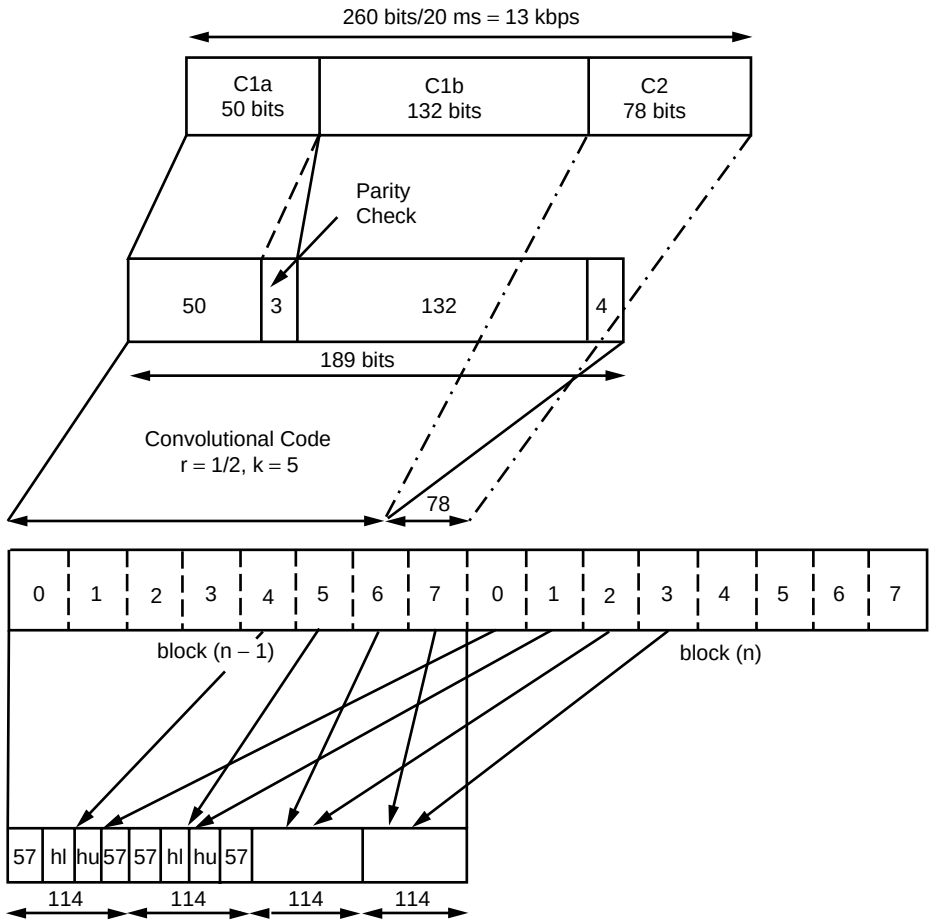


FIGURE 80.3 Mapping the TCH/FS logical channel onto a physical channel, ©ETT [4].

Mapping logical data traffic channels onto a physical channel is essentially carried out by the channel codecs [8], as specified in recommendation R.05.03. The full- and half-rate data traffic channels standardized in the GSM system are: **TCH/F9.6, TCH/F4.8, TCH/F2.4,** as well as **TCH/H4.8, TCH/H2.4,** as was shown earlier in Table 80.2. Note that the numbers in these acronyms represent the data transmission rate in kilobits per second. Without considering the details of these mapping processes we now focus our attention on control signal transmission issues.

80.5 Transmission of Control Signals

The exact derivation, forward error correcting (**FEC**) coding and mapping of logical control channel information is beyond the scope of this chapter, and the interested reader is referred to ETSI, 1988 (R.05.02 and R.05.03) and Hanzo and Stefanov [4] for a detailed discussion. As an example, the mapping of the 184-b SACCH, FACCH, BCCH, SDCCH, PCH, and access grant control channel (**AGCH**) messages onto a 456-b block, i.e., onto four 114-b bursts is demonstrated in Fig. 80.4. A double-layer concatenated FIRE-code/convolutional code scheme generates 456 bits, using an overall coding rate of $R = 184/456$, which gives a stronger protection for control channels than the error protection of traffic channels.

Returning to Fig. 80.2 we will now show how the SACCH is accommodated by the TDMA frame structure. The TCH/FS TDMA frames of the eight users are multiplexed into multiframes of 24 TDMA

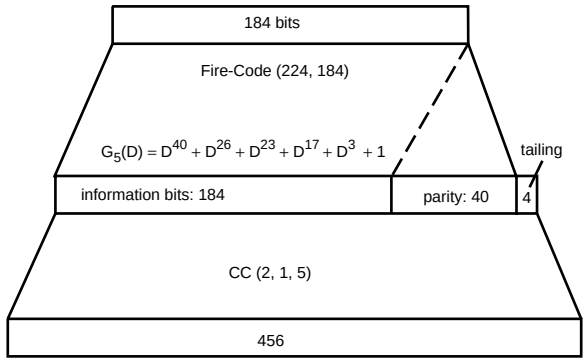


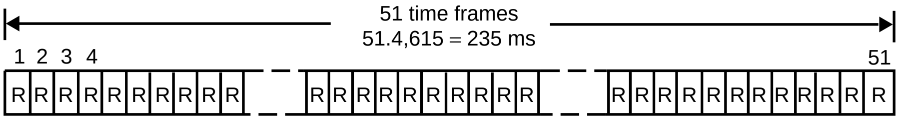
FIGURE 80.4 FEC in SACCH, FACCH, BCCH, SDCCH, PCH, and AGCH, ©ETT [4].

frames, but the 13th frame will carry a SACCH message, rather than the 13th TCH/FS frame, whereas the 26th frame will be an idle or dummy frame, as seen at the left-hand side of Fig. 80.2 at the multiframe level of the traffic channel hierarchy. The general control channel frame structure shown at the right of Fig. 80.2 is discussed later. This way 24-TCH/FS frames are sent in a 26-frame multiframe during 26 · 4.615 = 120 ms. This reduces the traffic throughput to (24/26) · 24.7 = 22.8 kb/s required by TCH/FS, allocates (1/26) · 24.7 = 950 b/s to the SACCH and wastes 950 b/s in the idle frame. Observe that the SACCH frame has eight timeslots to transmit the eight 950-b/s SACCHs of the eight users on the same carrier. The 950-b/s idle capacity will be used in case of half-rate channels, where 16 users will be multiplexed onto alternate frames of the TDMA structure to increase system capacity. Then 16, 11.4-kb/s encoded half-rate speech TCHs will be transmitted in a 120-ms multiframe, where also 16 SACCHs are available.

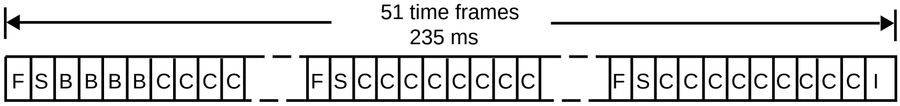
The FACCH messages are transmitted via the physical channels provided by bits stolen from their own host traffic channels. The construction of the FACCH bursts from 184 control bits is identical to that of the SACCH, as also shown in Fig. 80.4 but its 456-b frame is mapped onto eight consecutive 114-b TDMA traffic bursts, exactly as specified for TCH/FS. This is carried out by stealing the even bits of the first four and the odd bits of the last four bursts, which is signalled by setting $hu = 1$, $hl = 0$ and $hu = 0$, $hl = 1$ in the first and last bursts, respectively. The unprotected FACCH information rate is 184 b/20 ms = 9.2 kb/s, which is transmitted after concatenated error protection at a rate of 22.8 kb/s. The repetition delay is 20 ms, and the interleaving delay is 8 · 4.615 = 37 ms, resulting in a total of 57-ms delay.

In Fig. 80.2 at the next hierarchical level, 51-TCH/FS multiframes are multiplexed into one superframe lasting 51 · 120 ms = 6.12 s, which contains 26 · 51 = 1326 TDMA frames. In the case of 1326 TDMA frames, however, the frame number would be limited to $0 \leq \boldsymbol{FN} \leq 1326$ and the encryption rule relying on such a limited range of *FN* values would not be sufficiently secure. Then 2048 superframes were amalgamated to form a hyperframe of 1326 · 2048 = 2,715,648 TDMA frames lasting 2048 · 6.12 s ≈ 3 h 28 min, allowing a sufficiently high *FN* value to be used in the encryption algorithm. The uplink and downlink traffic-frame structures are identical with a shift of three timeslots between them, which relieves the MS from having to transmit and receive simultaneously, preventing high-level transmitted power leakage back to the sensitive receiver. The received power of adjacent BSs can be monitored during unallocated timeslots.

In contrast to duplex traffic and associated control channels, the simplex BCCH and CCCH logical channels of all MSs roaming in a specific cell share the physical channel provided by timeslot zero of the so-called BCCH carriers available in the cell. Furthermore, as demonstrated by the right-hand side section of Fig. 80.2, 51 BCCH and CCCH TDMA frames are mapped onto a 51 · 4.615 = 235-ms duration multiframe, rather than on a 26-frame, 120-ms duration multiframe. In order to compensate for the extended multiframe length of 235 ms, 26 multiframes constitute a 1326-frame superframe of 6.12-s duration. Note, in Fig. 80.5, that the allocation of the uplink and downlink frames is different, since these control channels exist only in one direction.



(a) Uplink Direction



(b) Downlink Direction

R: Random Access Channel
F : Frequency Correction Channel
S: Synchronisation Channel
B: Broadcast Control Channel
C: Access Grant/Paging Channel
I : Idle Frame

FIGURE 80.5 The control multiframe, ©ETT [4].

Specifically, the random access channel (**RACH**) is only used by the MSs in the uplink direction if they request, for example, a bidirectional SDCCH to be mapped onto an RF channel to register with the network and set up a call. The uplink RACH has a low capacity, carrying messages of 8-b/235-ms multiframe, which is equivalent to an unprotected control information rate of 34 b/s. These messages are concatenated FEC coded to a rate of 36 b/235 ms = 153 b/s. They are not transmitted by the NB derived for TCH/FS, SACCH, or FACCH logical channels, but by the AB, depicted in Fig. 80.6 in comparison to a NB and other types of bursts to be described later. The FEC coded, encrypted 36-b AB messages of Fig. 80.6 contain among other parameters, the encoded 6-b BS identifier code (**BSIC**) constituted by the 3-b PLMN color code and 3-b BS color code for unique BS identification. These 36 b are positioned after the 41-b synchronization sequence, which has a high wordlength in order to ensure reliable access burst recognition and a low probability of being emulated by interfering stray data. These messages have no interleaving delay, while they are transmitted with a repetition delay of one control multiframe length, i.e., 235 ms.

Adaptive time frame alignment is a technique designed to equalize propagation delay differences between MSs at different distances. The GSM system is designed to allow for cell sizes up to 35 km radius. The time a radio signal takes to travel the 70 km from the base station to the mobile station and back again is 233.3 μs. As signals from all the mobiles in the cell must reach the base station without overlapping each other, a long guard period of 68.25 b (252 μs) is provided in the access burst, which exceeds the maximum possible propagation delay of 233.3 μs. This long guard period in the access burst is needed when the mobile station attempts its first access to the base station or after a handover has occurred. When the base station detects a 41-b random access synchronization sequence with a long guard period, it measures the received signal delay relative to the expected signal from a mobile station of zero range. This delay, called the timing advance, is signalled using a 6-b number to the mobile station, which advances its timebase over the range of 0–63 b, i.e., in units of 3.69 μs. By this process the TDMA bursts arrive at the BS in their correct timeslots and do not overlap with adjacent ones. This process allows the

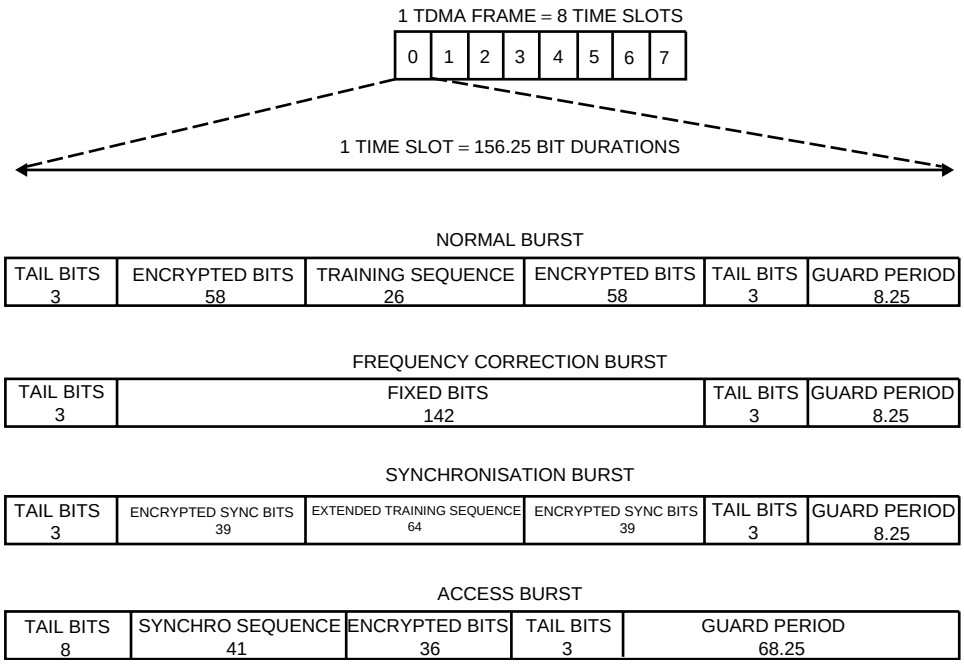


FIGURE 80.6 GSM burst structures, ©ETT [4].

guard period in all other bursts to be reduced to $8.25 \cdot 3.69\ \mu s \approx 30.46\ \mu s$ (8.25 b) only. During normal operation, the BS continuously monitors the signal delay from the MS and, if necessary, it will instruct the MS to update its time advance parameter. In very large traffic cells there is an option to actively utilize every second timeslot only to cope with higher propagation delays, which is spectrally inefficient, but in these large, low-traffic rural cells it is admissible.

As demonstrated by Fig. 80.2, the downlink multiframe transmitted by the BS is shared amongst a number of BCCH and CCCH logical channels. In particular, the last frame is an idle frame (I), whereas the remaining 50 frames are divided in five blocks of ten frames, where each block starts with a frequency correction channel (**FCCH**) followed by a synchronization channel (**SCH**). In the first block of ten frames the FCCH and SCH frames are followed by four BCCH frames and by either four AGCH or four PCH. In the remaining four blocks of ten frames, the last eight frames are devoted to either PCHs or AGCHs, which are mutually exclusive for a specific MS being either paged or granted a control channel.

The FCCH, SCH, and RACH require special transmission bursts, tailored to their missions, as depicted in Fig. 80.6. The FCCH uses frequency correction bursts (FCB) hosting a specific 142-b pattern. In partial response GMSK it is possible to design a modulating data sequence, which results in a near-sinusoidal modulated signal imitating an unmodulated carrier exhibiting a fixed frequency offset from the RF carrier utilized. The synchronization channel transmits SB hosting a $16 \cdot 4 = 64$-b extended sequence exhibiting a high-correlation peak in order to allow frame alignment with a quarter-bit accuracy. Furthermore, the SB contains $2 \cdot 39 = 78$ encrypted FEC-coded synchronization bits, hosting the BS and PLMN color codes, each representing one of eight legitimate identifiers. Lastly, the AB contain an extended 41-b synchronization sequence, and they are invoked to facilitate initial access to the system. Their long guard space of 68.25-b duration prevents frame overlap, before the MS's distance, i.e., the propagation delay becomes known to the BS and could be compensated for by adjusting the MS's timing advance.

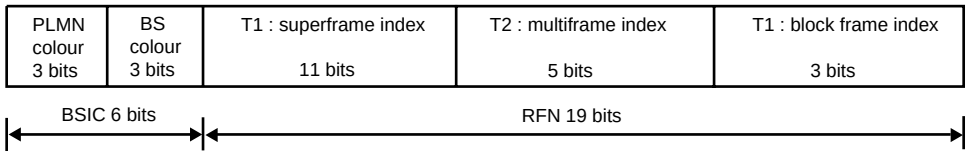


FIGURE 80.7 Synchronization channel (SCH) message format, ©ETT [4].

80.6 Synchronization Issues

Although some synchronization issues are standardized in recommendations R.05.02 and R.05.03, the GSM recommendations do not specify the exact BS-MS synchronization algorithms to be used, these are left to the equipment manufacturers. A unique set of timebase counters, however, is defined in order to ensure perfect BS-MS synchronism. The BS sends FCB and SB on specific timeslots of the BCCH carrier to the MS to ensure that the MS's frequency standard is perfectly aligned with that of the BS, as well as to inform the MS about the required initial state of its internal counters. The MS transmits its uniquely numbered traffic and control bursts staggered by three timeslots with respect to those of the BS to prevent simultaneous MS transmission and reception, and also takes into account the required timing advance (**TA**) to cater for different BS-MS-BS round-trip delays.

The timebase counters used to uniquely describe the internal timing states of BSs and MSs are the quarter-bit number (***QN*** = 0–624) counting the quarter-bit intervals in bursts, bit number (***BN*** = 0–156), timeslot number (***TN*** = 0–7) and TDMA Frame Number (FN = 0–26 · 51 · 2048), given in the order of increasing interval duration. The MS sets up its timebase counters after receiving a SB by determining QN from the 64-b extended training sequence in the center of the SB, setting $TN = 0$ and decoding the 78-encrypted, protected bits carrying the 25-SCH control bits.

The SCH carries frame synchronization information as well as BS identification information to the MS, as seen in Fig. 80.7, and it is provided solely to support the operation of the radio subsystem. The first 6 b of the 25-b segment consist of three PLMN color code bits and three BS color code bits supplying a unique BS identifier code (BSIC) to inform the MS which BS it is communicating with. The second 19-bit segment is the so-called reduced TDMA frame number (**RFN**) derived from the full TDMA frame number FN, constrained to the range of [0 - (26 · 51 · 2048) -1] = (0–2, 715, 647) in terms of three subsegments $T1$, $T2$, and $T3$. These subsegments are computed as follows: $T1$(11 b) = [FN div (26 · 51)], $T2$(5 b) = (FN mod 26) and $T3'$(3 b) = [($T3$ - 1) div 10], where $T3$ = (FN mod 5), whereas div and mod represent the integer division and modulo operations, respectively. Explicitly, in Fig. 80.7 $T1$ determines the superframe index in a hyperframe, $T2$ the multiframe index in a superframe, $T3$ the frame index in a multiframe, whereas $T3'$ is the so-called signalling block index [1–5] of a frame in a specific 51-frame control multiframe, and their roles are best understood by referring to Fig. 80.2. Once the MS has received the SB, it readily computes the FN required in various control algorithms, such as encryption, handover, etc., as

$$N = 51[(T3 - T2) \bmod 26] + T3 + 51 \cdot 26 \cdot T1, \quad \text{where } T3 = 10 \cdot T3' + 1$$

80.7 Gaussian Minimum Shift Keying Modulation

The GSM system uses constant envelope partial response GMSK modulation [6] specified in recommendation R.05.04. Constant envelope, continuous-phase modulation schemes are robust against signal fading as well as interference and have good spectral efficiency. The slower and smoother are the phase changes, the better is the spectral efficiency, since the signal is allowed to change less abruptly, requiring lower frequency components. The effect of an input bit, however, is spread over several bit periods, leading to a so-called partial response system, which requires a channel equalizer in order to remove this controlled, intentional intersymbol interference (ISI) even in the absence of uncontrolled channel dispersion.

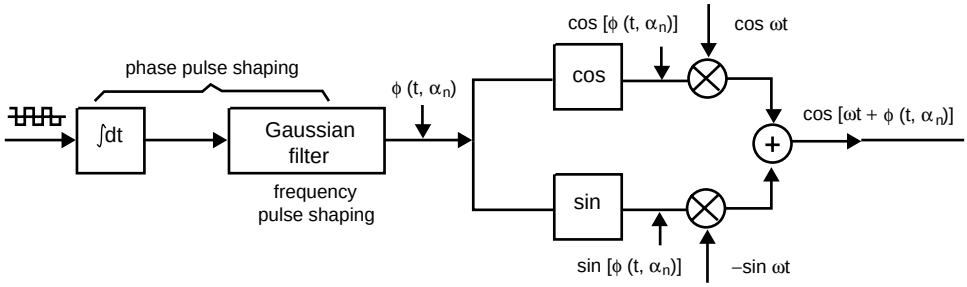


FIGURE 80.8 GMSK modulator schematic diagram, ©ETT [4].

The widely employed partial response GMSK scheme is derived from the full response minimum shift keying (MSK) scheme. In MSK the phase changes between adjacent bit periods are piecewise linear, which results in discontinuous-phase derivative, i.e., instantaneous frequency at the signalling instants, and hence widens the spectrum. Smoothing these phase changes, however, by a filter having a Gaussian impulse response [6], which is known to have the lowest possible bandwidth, this problem is circumvented using the schematic of Fig. 80.8, where the GMSK signal is generated by modulating and adding two quadrature carriers. The key parameter of GMSK in controlling both bandwidth and interference resistance is the 3-dB down filter-bandwidth × bit interval product ($B \cdot T$), referred to as normalized bandwidth. It was found that as the $B \cdot T$ product is increased from 0.2 to 0.5, the interference resistance is improved by approximately 2 dB at the cost of increased bandwidth occupancy, and best compromise was achieved for $B \cdot T = 0.3$. This corresponds to spreading the effect of 1 b over approximately 3-b intervals. The spectral efficiency gain due to higher interference tolerance and, hence, more dense frequency reuse was found to be more significant than the spectral loss caused by wider GMSK spectral lobes.

The channel separation at the TDMA burst rate of 271 kb/s is 200 kHz, and the modulated spectrum must be 40 dB down at both adjacent carrier frequencies. When TDMA bursts are transmitted in an on-off keyed mode, further spectral spillage arises, which is mitigated by a smooth power ramp up and down envelope at the leading and trailing edges of the transmission bursts, attenuating the signal by 70 dB during a 28- and 18-μs interval, respectively.

80.8 Wideband Channel Models

The set of 6-tap GSM impulse responses [2] specified in recommendation R.05.05 is depicted in Fig. 80.9, where the individual propagation paths are independent Rayleigh fading paths, weighted by the appropriate coefficients h_i corresponding to their relative powers portrayed in the figure. In simple terms the wideband channel's impulse response is measured by transmitting an impulse and detecting the received echoes at the channel's output in every D-spaced so-called delay bin. In some bins no delayed and attenuated multipath component is received, whereas in others significant energy is detected, depending on the typical reflecting objects and their distance from the receiver. The path delay can be easily related to the distance of the reflecting objects, since radio waves are travelling at the speed of light. For example, at a speed of 300,000 km/s, a reflecting object situated at a distance of 0.15 km yields a multipath component at a round-trip delay of 1 μs.

The typical urban (**TU**) impulse response spreads over a delay interval of 5 μs, which is almost two 3.69-μs bit-intervals duration and, therefore, results in serious ISI. In simple terms, it can be treated as a two-path model, where the reflected path has a length of 0.75 km, corresponding to a reflector located at a distance of about 375 m. The hilly terrain (HT) model has a sharply decaying short-delay section due to local reflections and a long-delay path around 15 μs due to distant reflections. Therefore, in

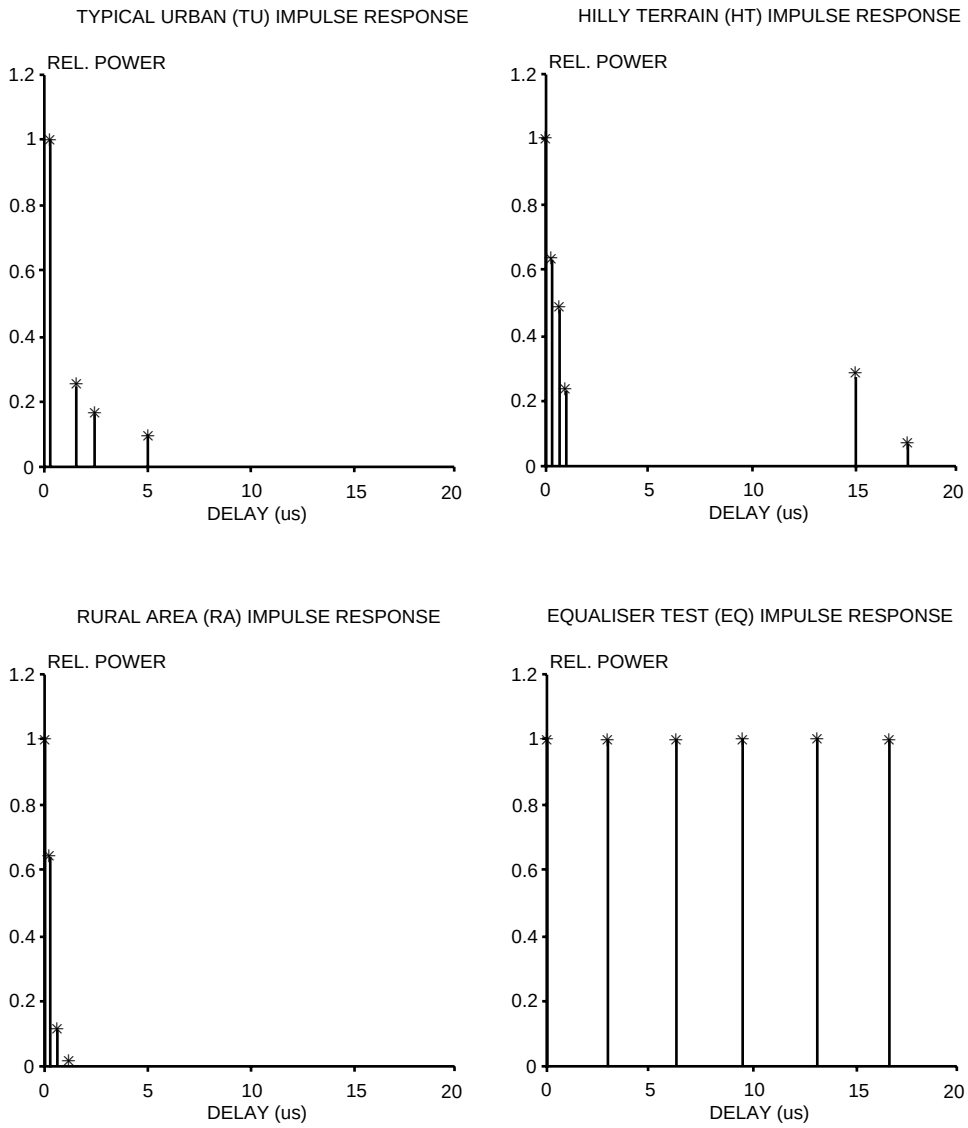


FIGURE 80.9 Typical GSM channel impulse responses, ©ETT [4].

practical terms it can be considered a two- or three-path model having reflections from a distance of about 2 km. The rural area (**RA**) response seems the least hostile amongst all standardized responses, decaying rapidly inside 1-b interval and, therefore, is expected to be easily combated by the channel equalizer. Although the type of the qualizer is not standardized, partial response systems typically use VEs. Since the RA channel efffectively behaves as a singlepath nondispersive channel, it would not require an equalizer. The fourth standardized impulse response is artificially contrived in order to test the equalizer's performance and is constituted by six equidistant unit-amplitude impulses representing six equal-powered independent Rayleigh-fading paths with a delay spread over 16 μs. With these impulse responses in mind, the required channel is simulated by summing the appropriately delayed and weighted received signal components. In all but one case the individual components are assumed to have Rayleigh amplitude distribution, whereas in the RA model the main tap at zero delay is supposed to have a Rician distribution with the presence of a dominant line-of-sight path.

80.9 Adaptive Link Control

The adaptive link control algorithm portrayed in Fig. 80.10 and specified in recommendation R.05.08 allows for the MS to favor that specific traffic cell which provides the highest probability of reliable communications associated with the lowest possible path loss. It also decreases interference with other cochannel users and, through dense frequency reuse, improves spectral efficiency, whilst maintaining an

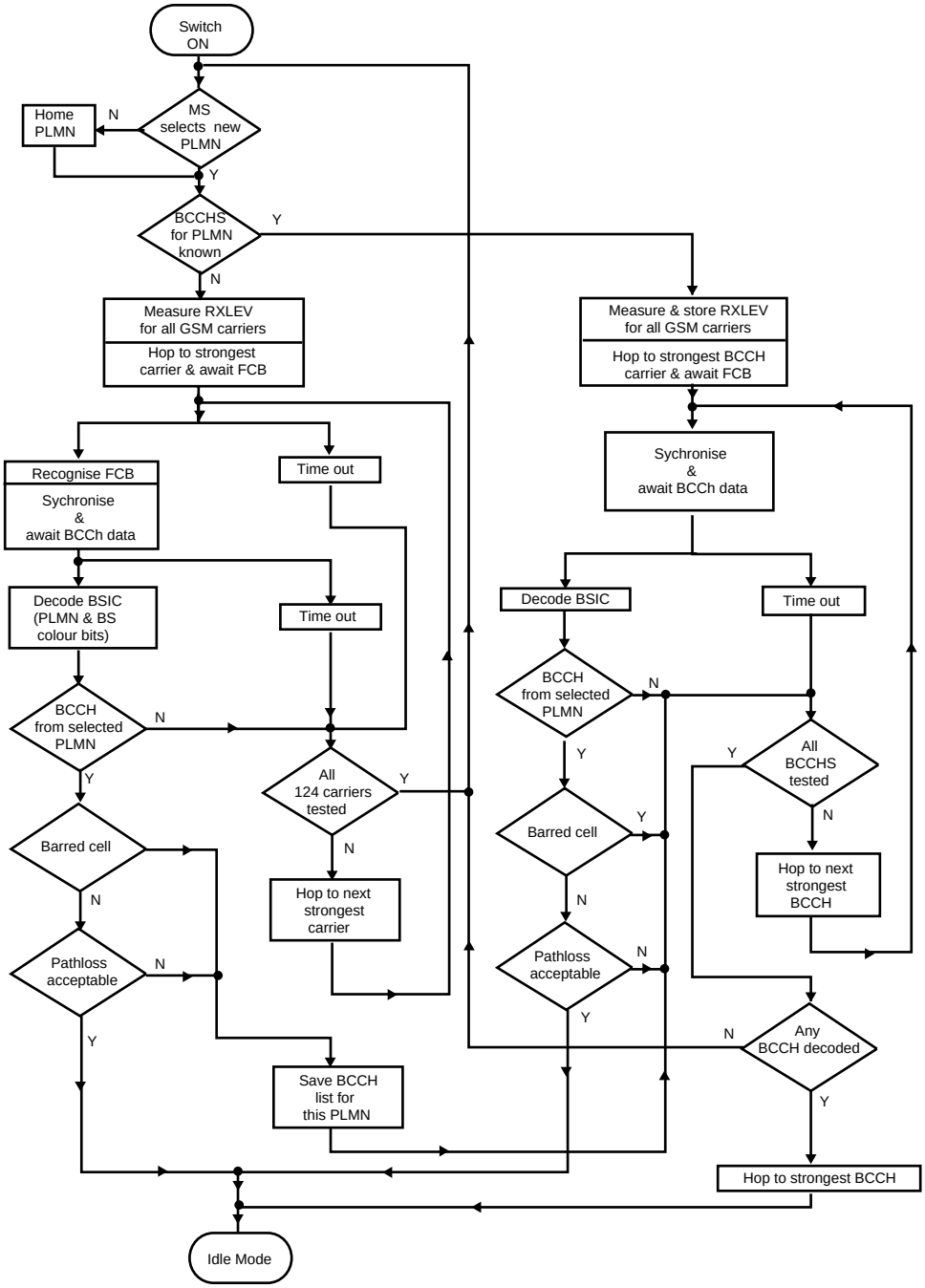


FIGURE 80.10 Initial cell selection by the MS, ©ETT [4].

adequate communications quality, and facilitates a reduction in power consumption, which is particularly important in hand-held MSs. The handover process maintains a call in progress as the MS moves between cells, or when there is an unacceptable transmission quality degradation caused by interference, in which case an intracell handover to another carrier in the same cell is performed. A radio-link failure occurs when a call with an unacceptable voice or data quality cannot be improved either by RF power control or by handover. The reasons for the link failure may be loss of radio coverage or very high-interference levels. The link control procedures rely on measurements of the received RF signal strength (**RXLEV**), the received signal quality (**RXQUAL**), and the absolute distance between base and mobile stations (DISTANCE).

RXLEV is evaluated by measuring the received level of the BCCH carrier which is continuously transmitted by the BS on all time slots of the B frames in Fig. 80.5 and without variations of the RF level. A MS measures the received signal level from the serving cell and from the BSs in all adjacent cells by tuning and listening to their BCCH carriers. The root mean squared level of the received signal is measured over a dynamic range from -103 to -41 dBm for intervals of one SACCH multiframe (480 ms). The received signal level is averaged over at least 32 SACCH frames (≈15 s) and mapped to give RXLEV values between 0 and 63 to cover the range from -103 to -41 dBm in steps of 1 dB. The RXLEV parameters are then coded into 6-b words for transmission to the serving BS via the SACCH.

RXQUAL is estimated by measuring the bit error ratio (**BER**) before channel decoding, using the Viterbi channel equalizer's metrics [6] and/or those of the Viterbi convolutional decoder [8]. Eight values of RXQUAL span the logarithmically scaled BER range of 0.2–12.8% before channel decoding.

The absolute DISTANCE between base and mobile stations is measured using the timing advance parameter. The timing advance is coded as a 6-b number corresponding to a propagation delay from 0 to $63 \cdot 3.69\ \mu s = 232.6\ \mu s$, characteristic of a cell radius of 35 km.

While roaming, the MS needs to identify which potential target BS it is measuring, and the BCCH carrier frequency may not be sufficient for this purpose, since in small cluster sizes the same BCCH frequency may be used in more than one surrounding cell. To avoid ambiguity a 6-b BSIC is transmitted on each BCCH carrier in the SB of Fig. 80.6. Two other parameters transmitted in the BCCH data provide additional information about the BS. The binary flag called **PLMN_PERMITTED** indicates whether the measured BCCH carrier belongs to a PLMN that the MS is permitted to access. The second Boolean flag, **CELL_BAR_ACCESS**, indicates whether the cell is barred for access by the MS, although it belongs to a permitted PLMN. A MS in idle mode, i.e., after it has just been switched on or after it has lost contact with the network, searches all 125 RF channels and takes readings of RXLEV on each of them. Then it tunes to the carrier with the highest RXLEV and searches for FCB in order to determine whether or not the carrier is a BCCH carrier. If it is not, then the MS tunes to the next highest carrier, and so on, until it finds a BCCH carrier, synchronizes to it and decodes the parameters BSIC, PLMN_PERMITTED and CELL_BAR_ACCESS in order to decide whether to continue the search. The MS may store the BCCH carrier frequencies used in the network accessed, in which case the search time would be reduced. Again, the process described is summarized in the flowchart of Fig. 80.10.

The adaptive power control is based on RXLEV measurements. In every SACCH multiframe the BS compares the RXLEV readings reported by the MS or obtained by the base station with a set of thresholds. The exact strategy for RF power control is determined by the network operator with the aim of providing an adequate quality of service for speech and data transmissions while keeping interferences low. Clearly, adequate quality must be achieved at the lowest possible transmitted power to keep cochannel interferences low, which implies contradictory requirements in terms of transmitted power. The criteria for reporting radio link failure are based on the measurements of RXLEV and RXQUAL performed by both the mobile and base stations, and the procedures for handling link failures result in the re-establishment or the release of the call, depending on the network operator's strategy.

The handover process involves the most complex set of procedures in the radio-link control. Handover decisions are based on results of measurements performed both by the base and mobile stations. The base station measures RXLEV, RXQUAL, DISTANCE, and also the interference level in unallocated time slots, whereas the MS measures and reports to the BS the values of RXLEV and RXQUAL for the serving

cell and RXLEV for the adjacent cells. When the MS moves away from the BS, the RXLEV and RXQUAL parameters for the serving station become lower, whereas RXLEV for one of the adjacent cells increases.

80.10 Discontinuous Transmission

Discontinuous transmission (DTX) issues are standardized in recommendation R.06.31, whereas the associated problems of voice activity detection VAD are specified by R.06.32. Assuming an average speech activity of 50% and a high number of interferers combined with frequency hopping to randomize the interference load, significant spectral efficiency gains can be achieved when deploying discontinuous transmissions due to decreasing interferences, while reducing power dissipation as well. Because of the reduction in power consumption, full DTX operation is mandatory for MSs, but in BSs, only receiver DTX functions are compulsory.

The fundamental problem in voice activity detection is how to differentiate between speech and noise, while keeping false noise triggering and speech spurt clipping as low as possible. In vehicle-mounted MSs the severity of the speech/noise recognition problem is aggravated by the excessive vehicle background noise. This problem is resolved by deploying a combination of threshold comparisons and spectral domain techniques [1,3]. Another important associated problem is the introduction of noiseless inactive segments, which is mitigated by comfort noise insertion (**CNI**) in these segments at the receiver.

80.11 Summary

Following the standardization and launch of the GSM system its salient features were summarized in this brief review. Time division multiple access (TDMA) with eight users per carrier is used at a multiuser rate of 271 kb/s, demanding a channel equalizer to combat dispersion in large cell environments. The error protected chip rate of the full-rate traffic channels is 22.8 kb/s, whereas in half-rate channels it is 11.4 kb/s. Apart from the full- and half-rate speech traffic channels, there are 5 different rate data traffic channels and 14 various control and signalling channels to support the system's operation. A moderately complex, 13 kb/s regular pulse excited speech codec with long term predictor (**LTP**) is used, combined with an embedded three-class error correction codec and multilayer interleaving to provide sensitivity-matched unequal error protection for the speech bits. An overall speech delay of 57.5 ms is maintained. Slow frequency hopping at 217 hops/s yields substantial performance gains for slowly moving pedestrians.

TABLE 80.3 Summary of GSM Features

System feature	Specification
Up-link bandwidth, MHz	890–915 = 25
Down-link bandwidth, MHz	935–960 = 25
Total GSM bandwidth, MHz	50
Carrier spacing, KHz	200
No. of RF carriers	125
Multiple access	TDMA
No. of users/carrier	8
Total No. of channels	1000
TDMA burst rate, kb/s	271
Modulation	GMSK with BT = 0.3
Bandwidth efficiency, b/s/Hz	1.35
Channel equalizer	yes
Speech coding rate, kb/s	13
FEC coded speech rate, kb/s	22.8
FEC coding	Embedded block/convolutional
Frequency hopping, hop/s	217
DTX and VAD	yes
Maximum cell radius, km	35

Constant envelope partial response GMSK with a channel spacing of 200 kHz is deployed to support 125 duplex channels in the 890–915-MHz up-link and 935–960-MHz down-link bands, respectively. At a transmission rate of 271 kb/s a spectral efficiency of 1.35-bit/s/Hz is achieved. The controlled GMSK-induced and uncontrolled channel-induced intersymbol interferences are removed by the channel equalizer. The set of standardized wideband GSM channels was introduced in order to provide bench markers for performance comparisons. Efficient power budgeting and minimum cochannel interferences are ensured by the combination of adaptive power and handover control based on weighted averaging of up to eight up-link and down-link system parameters. Discontinuous transmissions assisted by reliable spectral-domain voice activity detection and comfort-noise insertion further reduce interferences and power consumption. Because of ciphering, no unprotected information is sent via the radio link. As a result, spectrally efficient, high-quality mobile communications with a variety of services and international roaming is possible in cells of up to 35 km radius for signal-to-noise and interference ratios in excess of 10–12 dBs. The key system features are summarized in Table 80.3.

Defining Terms

A3: Authentication algorithm.
A5: Cyphering algorithm.
A8: Confidential algorithm to compute the cyphering key.
AB: Access burst.
ACCH: Associated control channel.
ADC: Administration center.
AGCH: Access grant control channel.
AUC: Authentication center.
AWGN: Additive Gaussian noise.
BCCH: Broadcast control channel.
BER: Bit error ratio.
BFI: Bad frame indicator flag.
***BN*:** Bit number.
BS: Base station.
BS-PBGT: BS powerbudget; to be evaluated for power budget motivated handovers.
BSIC: Base station identifier code.
CC: Convolutional codec.
CCCH: Common control channel.
CELL_BAR_ACCESS: Boolean flag to indicate if the MS is permitted to access the specific traffic cell.
CNC: Comfort noise computation.
CNI: Comfort noise insertion.
CNU: Comfort noise update state in the DTX handler.
DB: Dummy burst.
DL: Down link.
DSI: Digital speech interpolation to improve link efficiency.
DTX: Discontinuous transmission for power consumption and interference reduction.
EIR: Equipment identity register.
EOS: End of speech flag in the DTX handler.
FACCH: Fast associated control channel.
FCB: Frequency correction burst.
FCCH: Frequency correction channel.
FEC: Forward error correction.
FH: Frequency hopping.
***FN*:** TDMA frame number.
GMSK: Gaussian minimum shift keying.

GP: Guard space.
HGO: Handover in the VAD.
HLR: Home location register.
HO: Handover.
HOCT: Handover counter in the VAD.
HO_MARGIN: Handover margin to facilitate hysteresis.
HSN: Hopping sequence number; frequency hopping algorithm's input variable.
IMSI: International mobile subscriber identity.
ISDN: Integrated services digital network.
LAI: Location area identifier.
LAR: Logarithmic area ratio.
LTP: Long term predictor.
MA: Mobile allocation; set of legitimate RF channels, input variable in the frequency hopping algorithm.
MAI: Mobile allocation index; output variable of the FH algorithm.
MAIO: Mobile allocation index offset; initial RF channel offset, input variable of the FH algorithm.
MS: Mobile station.
MSC: Mobile switching center.
MSRN: Mobile station roaming number.
MS_TXPWR_MAX: Maximum permitted MS transmitted power on a specific traffic channel in a specific traffic cell.
MS_TXPWR_MAX(n): Maximum permitted MS transmitted power on a specific traffic channel in the *n*th adjacent traffic cell.
NB: Normal burst.
NMC: Network management center.
NUFR: Receiver noise update flag.
NUFT: Noise update flag to ask for SID frame transmission.
OMC: Operation and maintenance center.
PARCOR: Partial correlation.
PCH: Paging channel.
PCM: Pulse code modulation.
PIN: Personal identity number for MSs.
PLMN: Public land mobile network.
PLMN_PERMITTED: Boolean flag to indicate if the MS is permitted to access the specific PLMN.
PSTN: Public switched telephone network.
***QN*:** Quarter bit number.
R: Random number in the authentication process.
RA: Rural area channel impulse response.
RACH: Random access channel.
RF: Radio frequency.
RFCH: Radio frequency channel.
RFN: Reduced TDMA frame number; equivalent representation of the TDMA frame number that is used in the synchronization channel.
RNTABLE: Random number table utilized in the frequency hopping algorithm.
RPE: Regular pulse excited.
RPE-LTP: Regular pulse excited codec with long term predictor.
RS-232: Serial data transmission standard equivalent to CCITT V24. interface.
RXLEV: Received signal level; parameter used in handovers.
RXQUAL: Received signal quality; parameter used in handovers.
S: Signed response in the authentication process.
SACCH: Slow associated control channel.
SB: Synchronization burst.

SCH: Synchronization channel.
SCPC: Single channel per carrier.
SDCCH: Stand-alone dedicated control channel.
SE: Speech extrapolation.
SID: Silence identifier.
SIM: Subscriber identity module in MSs.
SPRX: Speech received flag.
SPTX: Speech transmit flag in the DTX handler.
STP: Short term predictor.
TA: Timing advance.
TB: Tailing bits.
TCH: Traffic channel.
TCH/F: Full-rate traffic channel.
TCH/F2.4: Full-rate 2.4-kb/s data traffic channel.
TCH/F4.8: Full-rate 4.8-kb/s data traffic channel.
TCH/F9.6: Full-rate 9.6-kb/s data traffic channel.
TCH/FS: Full-rate speech traffic channel.
TCH/H: Half-rate traffic channel.
TCH/H2.4: Half-rate 2.4-kb/s data traffic channel.
TCH/H4.8: Half-rate 4.8-kb/s data traffic channel.
TDMA: Time division multiple access.
TMSI: Temporary mobile subscriber identifier.
***TN*:** Time slot number.
TU: Typical urban channel impulse response.
TXFL: Transmit flag in the DTX handler.
UL: Up link.
VAD: Voice activity detection.
VE: Viterbi equalizer.
VLR: Visiting location register.

References

1. European Telecommunications Standardization Institute. Group Speciale Mobile or Global System of Mobile Communication (GSM) Recommendation, ETSI Secretariat, Sophia Antipolis Cedex, France, 1988.
2. Greenwood, D. and Hanzo, L., Characterisation of mobile radio channels, In *Mobile Radio Communications,* Steele, R. and Hanzo, L., Eds., Chap. 2, IEEE Press–John Wiley, 1999.
3. Hanzo, L. and Stefanov, J., The Pan-European digital cellular mobile radio system—known as GSM. In *Mobile Radio Communications,* Steele, R. and Hanzo, L., Eds., Chap. 8, IEEE Press–John Wiley, 1999.
4. Hanzo, L. and Steele, R., The Pan-European mobile radio system, Pts. 1 and 2, *European Trans. on Telecomm.,* 5(2), 245–276, 1994.
5. Salami, R.A., Hanzo, L., et al., Speech coding. In *Mobile Radio Communications,* Steele, R. and Hanzo, L., Eds., Chap. 3, IEEE Press–John Wiley, 1999.
6. Steele, R. and Hanzo, L., Eds., *Mobile Radio Communications,* IEEE Press–John Wiley, 1999.
7. Vary, P. and Sluyter, R.J., MATS-D speech codec: Regular-pulse excitation LPC, *Proceedings of Nordic Conference on Mobile Radio Communications,* 257–261, 1986.
8. Wong, K.H.H. and Hanzo, L., Channel coding. In *Mobile Radio Communications,* Steele, R. and Hanzo, L., Eds., Chap. 4, IEEE Press–John Wiley, 1999.

81

Speech and Channel Coding for North American TDMA Cellular Systems

Paul Mermelstein
INRS-Télécommunications
University of Québec

81.1 Introduction

The goals of this chapter are to give the reader a tutorial introduction and high-level understanding of the techniques employed for speech transmission by the IS-54 digital cellular standard. It builds on the information provided in the standards document but is not meant to be a replacement for it. Separate standards cover the control channel used for the setup of calls and their handoff to neighboring cells, as well as the encoding of data signals for transmission. For detailed implementation information the reader should consult the most recent standards document [9].

IS-54 provides for encoding bidirectional speech signals digitally and transmitting them over cellular and microcellular mobile radio systems. It retains the 30-kHz channel spacing of the earlier advanced

mobile telephone service (AMPS), which uses analog frequency modulation for speech transmission and frequency shift keying for signalling. The two directions of transmission use frequencies some 45 MHz apart in the band between 824 and 894 MHz. AMPS employs one channel per conversation in each direction, a technique known as frequency division multiple access (FDMA). IS-54 employs time division multiple access (TDMA) by allowing three, and in the future six, simultaneous transmissions to share each frequency band. Because the overall 30-kHz channelization of the allocated 25 MHz of spectrum in each direction is retained, it is also known as a FDMA-TDMA system. In contrast, the later IS-95 standard employs code division multiple access (CDMA) over bands of 1.23 MHz by combining several 30-kHz frequency channels.

Each frequency channel provides for transmission at a digital bit rate of 48.6 kb/s through use of differential quadrature-phase shift key (DQPSK) modulation at a 24.3-kBd channel rate. The channel is divided into six time slots every 40 ms. The full-rate voice coder employs every third time slot and utilizes 13 kb/s for combined speech and channel coding. The six slots provide for an eventual half-rate channel occupying one slot per 40 ms frame and utilizing only about 6.5 kb/s for each call. Thus, the simultaneous call carrying capacity with IS-54 is increased by a factor 3 (factor 6 in the future) above that of AMPS. All digital transmission is expected to result in a reduction in transmitted power. The resulting reduction in intercell interference may allow more frequent reuse of the same frequency channels than the reuse pattern of seven cells for AMPS. Additional increases in erlang capacity (the total call-carrying capacity at a given blocking rate) may be available from the increased trunking efficiency achieved by the larger number of simultaneously available channels. The first systems employing dual-mode AMPS and TDMA service were put into operation in 1993.

In 1996 the TIA introduced the IS-641 enhanced full rate codec. This codec consists of 7.4 kb/s speech coding following the algebraic code-excited linear prediction (ACELP) technique [7], and 5.6 kb/s channel coding. The 13 kb/s coded information replaces the combined 13 kb/s for speech and channel coding introduced by the IS-54 standard. The new codec provides significant enhancements in terms of speech quality and robustness to transmission errors. The quality enhancement for clear channels results from the improved modeling of the stochastic excitation by means of an algebraic **codebook** instead of the two trained VSELP codebooks. Improved robustness to transmission errors is achieved by employing predictive quantization techniques for the linear-prediction filter and gain parameters, and increasing the number of bits protected by forward error correction.

81.2 Modulation of Digital Voice and Data Signals

The modulation method used in IS-54 is $\pi/4$ shifted differentially encoded quadrature phase-shift keying (DPSK). Symbols are transmitted as changes in phase rather than their absolute values. The binary data stream is converted to two binary streams X_k and Y_k formed from the odd- and even-numbered bits, respectively. The quadrature streams I_k and Q_k are formed according to

$$\begin{aligned} I_k &= I_{k-1}\cos[\Delta\phi(X_k, Y_k)] - Q_{k-1}\sin[\Delta\phi(X_k, Y_k)] \\ Q_k &= I_{k-1}\sin[\Delta\phi(X_k, Y_k)] + Q_{k-1}\cos[\Delta\phi(X_k, Y_k)] \end{aligned}$$

where I_{k-1} and Q_{k-1} are the amplitudes at the previous pulse time. The phase change $\Delta\phi$ takes the values $\pi/4$, $3\pi/4$, $-\pi/4$, and $-3\pi/4$ for the dibit (X_k, Y_k) symbols (0,0), (0,1), (1,0), and (1,1), respectively. This results in a rotation by $\pi/4$ between the constellations for odd and even symbols. The differential encoding avoids the problem of 180° phase ambiguity that may otherwise result in estimation of the carrier phase.

The signals I_k and Q_k at the output of the differential phase encoder can take one of five values, 0, ±1, $\pm 1/\sqrt{2}$ as indicated in the constellation of Fig. 81.1. The corresponding impulses are applied to the inputs of the I and Q baseband filters, which have linear phase and square root raised cosine frequency responses. The generic modulator circuit is shown in Fig. 81.2. The rolloff factor α determines the width of the

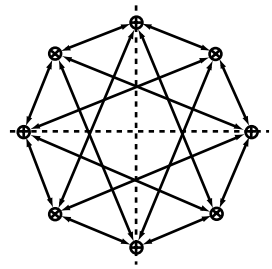


FIGURE 81.1 Constellation for $\pi/4$ shifted QPSK modulation. *Source*: TIA, 1992. Cellular System Dual-mode Mobile Station–Base Station Compatibility Standard TIA/EIA IS-54. With permission.

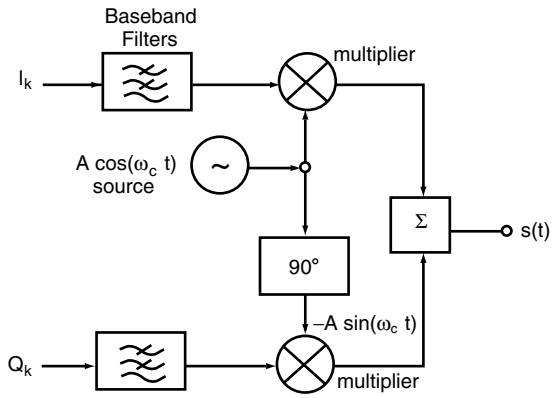


FIGURE 81.2 Generic modulation circuit for digital voice and data signals. *Source*: TIA, 1992. Cellular System Dual-mode Mobile Station–Base Station Compatibility Standard TIA/EIA IS-54.

transition band and its value is 0.35,

$$|H(f)| = \begin{cases} 1, & 0 \le f \le (1-\alpha)/2T \\ \sqrt{1/2\{1-\sin[\pi(2fT-1)/2\alpha]\}}, & (1-\alpha)/2T \le f \le (1+\alpha)/2T \\ 0, & f > (1+\alpha)/2T \end{cases}$$

81.3 Speech Coding Fundamentals

The IS-54 standard employs a vector-sum excited linear prediction (VSELP) coding technique. It represents a specific formulation of the much larger class of code-excited linear prediction (CELP) coders [2] that have proved effective in recent years for the coding of speech at moderate rates in the range 4–16 kb/s. VSELP provides reconstructed speech with a quality that is comparable to that available with frequency modulation and analog transmission over the AMPS system. The coding rate employed is 7.95 kb/s. Each of the six slots per frame carry 260 b of speech and channel coding information for a gross information rate of 13 kb/s. The 260 b correspond to 20 ms of real time speech, transmitted as a single burst.

For an excellent recent review of speech coding techniques for transmission, the reader is referred to Gersho, 1994 [3]. Most modern speech coders use a form of analysis by synthesis coding where the encoder determines the coded signal one segment at a time by feeding candidate excitation segments into a replica of a synthesis filter and selecting the segment that minimizes the distortion between the original and reproduced signals. Linear prediction coding (LPC) techniques [1] encode the speech signal by first finding an optimum linear filter to remove the short-time correlation, passing the signal through that LPC filter to obtain a residual signal, and encoding this residual using much fewer bits than would have been required to code the original signal with the same fidelity. In most cases the coding of the residual is divided into

two steps. First, the long-time correlation due to the periodic pitch excitation is removed by means of an optimum one-tap filter with adjustable gain and lag. Next, the remaining residual signal, which now closely resembles a white-noise signal, is encoded. Code-excited linear predictors use one or more codebooks from which they select replicas of the residual of the input signal by means of a closed-loop error-minimization technique. The index of the codebook entry as well as the parameters of all the filters are transmitted to allow the speech signal to be reconstructed at the receiver. Most code-excited coders use trained codebooks. Starting with a codebook containing Guassian signal segments, entries that are found to be used rarely in coding a large body of speech data are iteratively eliminated to result in a smaller codebook that is considered more effective.

The speech signal can be considered quasistationary or stationary for the duration of the speech frame, of the order of 20 ms. The parameters of the short-term filter, the LPC coefficients, are determined by analysis of the autocorrelation function of a suitably windowed segment of the input signal. To allow accurate determination of the time-varying pitch lag as well as simplify the computations, each speech frame is divided into four 5-ms subframes. Independent pitch filter computations and residual coding operations are carried out for each subframe.

The speech decoder attempts to reconstruct the speech signal from the received information as best possible. It employs a codebook identical to that of the encoder for excitation generation and, in the absence of transmission errors, would produce an exact replica of the signal that produced the minimized error at the encoder. Transmission errors do occur, however, due, to signal fading and excessive interference. Since any attempt at retransmission would incur unacceptable signal delays, sufficient error protection is provided to allow correction of most transmission errors.

81.4 Channel Coding Considerations

The sharp limitations on available bandwidth for error protection argue for careful consideration of the sensitivity of the speech coding parameters to transmission errors. Pairwise interleaving of coded blocks and convolutional coding of a subset of the parameters permit correction of a limited number of transmission errors. In addition, a cyclic redundancy check (CRC) is used to determine whether the error correction was successful. The coded information is divided into three blocks of varying sensitivity to errors. Group 1 contains the most sensitive bits, mainly the parameters of the LPC filter and frame energy, and is protected by both error detection and correction bits. Group 2 is provided with error correction only. The third group, comprising mostly the fixed codebook indices, is not protected at all.

The speech signal contains significant temporal redundancy. Thus, speech frames within which errors have been detected may be reconstructed with the aid of previously correctly received information. A bad-frame masking procedure attempts to hide the effects of short fades by extrapolating the previously received parameters. Of course, if the errors persist, the decoded signal must be muted while an attempt is made to hand off the connection to a base station to/from which the mobile may experience better reception.

81.5 VSELP Encoder

A block diagram of the VSELP speech encoder [4] is shown in Fig. 81.3. The excitation signal is generated from three components, the output of a long term or pitch filter, as well as entries from two codebooks. A weighted synthesis filter generates a synthesized approximation to the frequency-weighted input signal. The weighted mean square error between these two signals is used to drive the error minimization process. This weighted error is considered to be a better approximation to the perceptually important noise components than the unweighted mean square error. The total weighted square error is minimized by adjusting the pitch lag and the codebook indices as well as their gains. The decoder follows the encoder closely and generates the excitation signal identically to the encoder but uses an unweighted linear-prediction synthesis filter to generate the decoded signal. A spectral postfilter is added after the synthesis filter to enhance the quality of the reconstructed speech.

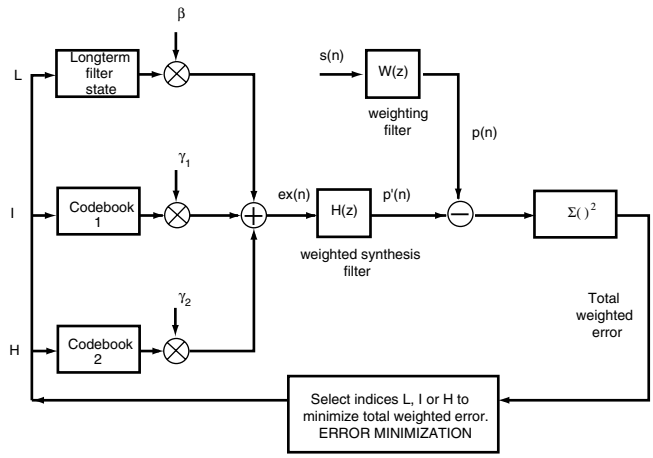


FIGURE 81.3 Block diagram of the speech encoder in VSELP. *Source*: TIA, 1992. Cellular system Dual-mode Mobile Station–Base Station Compatibility Standard TIA/EIA IS-54.

The precise data rate of the speech coder is 7950 b/s or 159 b per time slot, each corresponding to 20 ms of signal in real time. These 159 b are allocated as follows: (1) short-term filter coefficients, 38 bits; (2) frame energy, 5 bits; (3) pitch lag, 28 bits; (4) codewords, 56 bits; and (5) gain values, 32 bits.

81.6 Linear Prediction Analysis and Quantization

The purpose of the LPC analysis filter is to whiten the spectrum of the input signal so that it can be better matched by the codebook outputs. The corresponding LPC synthesis filter $A(z)$ restores the short-time speech spectrum characteristics to the output signal. The transfer function of the tenth-order synthesis filter is given by

$$A(z) = \frac{1}{1 - \sum_{i=1}^{N_p} \alpha_i z^{-i}}$$

The filter predictor parameters $\alpha_1, \ldots, \alpha_{N_p}$ are not transmitted directly. Instead, a set of **reflection coefficients** $r_1, \ldots, r_{N_p}$ are computed and quantized. The predictor parameters are determined from the reflection coefficients using a well-known backward recursion algorithm [6].

A variety of algorithms are known that determine a set of reflection coefficients from a windowed input signal. One such algorithm is the fixed point **covariance lattice**, FLAT, which builds an optimum inverse lattice stage by stage. At each stage j, the sum of the mean-squared forward and backward residuals is minimized by selection of the best reflection coefficient r_j. The analysis window used is 170 samples long, centered with respect to the middle of the fourth 5-ms subframe of the 20-ms frame. Since this centerpoint is 20 samples from the end of the frame, 65 samples from the next frame to be coded are used in computing the reflection coefficient of the current frame. This introduces a lookahead delay of 8.125 ms.

The FLAT algorithm first computes the covariance matrix of the input speech for $N_A = 170$ and $N_p = 10$,

$$\phi(i, k) = \sum_{n=N_p}^{N_A-1} s(n-i)s(n-k), \qquad 0 \le i, \quad k \le N_p$$

Define the forward residual out of stage j as $f_j(n)$ and the backward residual as $b_j(n)$. Then the autocorrelation of the initial forward residual $F_0(i, k)$ is given by $\phi(i, k)$. The autocorrelation of the initial backward residual $B_0(i, k)$ is given by $\phi(i + 1, k + 1)$ and the initial cross correlation of the two residuals is given

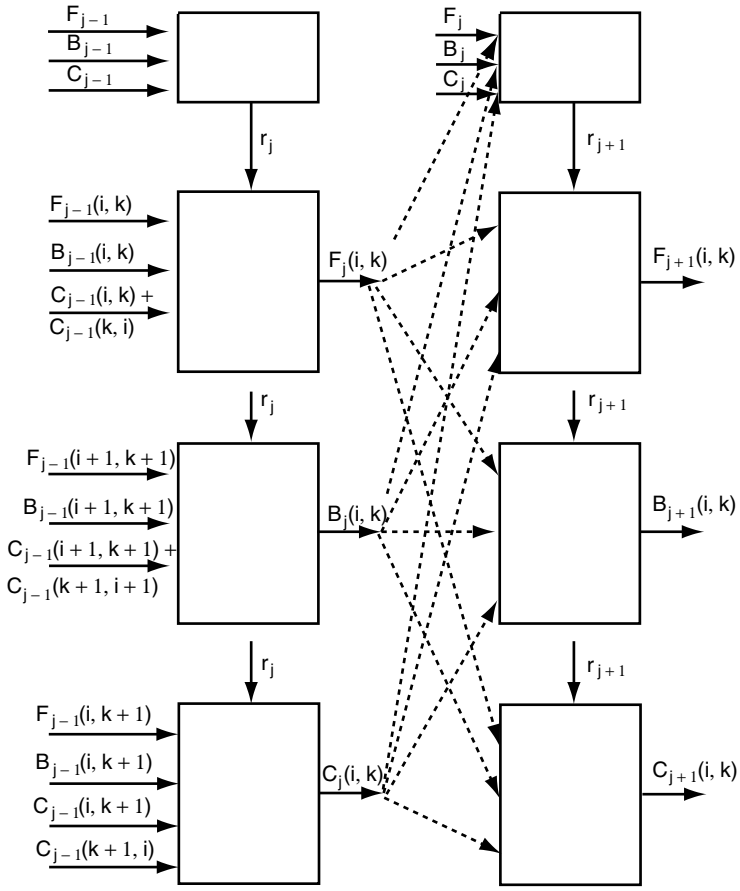


FIGURE 81.4 Block diagram for lattice covariance computations.

by $C_0(i, k) = \phi(i, k+1)$ for $0 \le i, k \le N_{p-1}$. Initially j is set to 1. The reflection coefficient at each stage is determined as the ratio of the cross correlation to the mean of the autocorrelations. A block diagram of the computations is shown in Fig. 81.4. By quantizing the reflection coefficients within the computation loops, reflection coefficients at subsequent stages are computed taking into account the quantization errors of the previous stages. Specifically,

$$\begin{aligned} C'_{j-1} &= C_{j-1}(0, 0) + C_{j-1}(N_p - j, N_p - j) \\ F'_{j-1} &= F_{j-1}(0, 0) + F_{j-1}(N_p - j, N_p - j) \\ B'_{j-1} &= B_{j-1}(0, 0) + B_{j-1}(N_p - j, N_p - j) \end{aligned}$$

and

$$r_j = \frac{-2C'_{j-1}}{F'_{j-1} + B'_{j-1}}$$

Use of two sets of correlation values separated by $N_p - j$ samples provides additional stability to the computed reflection coefficients in case the input signal changes form rapidly.

Once a quantized reflection coefficient r_j has been determined, the resulting auto- and cross correlations can be determined iteratively as

$$F_j(i, k) = F_{j-1}(i, k) + r_j[C_{j-1}(i, k) + C_{j-1}(k, i)] + r_j^2 B_{j-1}(i, k)$$
$$B_j(i, k) = B_{j-1}(i+1, k+1) + r_j[C_{j-1}(i+1, k+1) + C_{j-1}(k+1, i+1)] + r_j^2 F_{j-1}(i+1, k+1)$$

and

$$C_j(i, k) = C_{j-1}(i, k+1) + r_j[B_{j-1}(i, k+1) + F_{j-1}(i, k+1)] + r_j^2 C_{j-1}(k+1, i)$$

These computations are carried out iteratively for r_j, $j = 1,\ldots,N_p$.

81.7 Bandwidth Expansion

Poles with very narrow bandwidths may introduce undesirable distortions into the synthesized signal. Use of a binomial window with effective bandwidth of 80 Hz suffices to limit the ringing of the LPC filter and reduce the effect of the LPC filter selected for one frame on the signal reconstructed for subsequent frames. To achieve this, prior to searching for the reflection coefficients, the $\phi(i, k)$ is modified by use of a window function $w(j)$, $j = 1,\ldots,10$, as follows:

$$\phi'(i, k) = \phi(i, k)w(|i - k|)$$

81.8 Quantizing and Encoding the Reflection Coefficients

The distortion introduced into the overall spectrum by quantizing the reflection coefficients diminishes as we move to higher orders in the reflection coefficients. Accordingly, more bits are assigned to the lower order coefficients. Specifically, 6, 5, 5, 4, 4, 3, 3, 3, 3, and 2 b are assigned to $r_1,\ldots,r_{10}$, respectively. Scalar quantization of the reflection coefficients is used in IS-54 because it is particularly simple. **Vector quantization** achieves additional quantizing efficiencies at the cost of significant added complexity.

It is important to preserve the smooth time evolution of the linear prediction filter. Both the encoder and decoder linearly interpolate the coefficients α_i for the first, second, and third subframes of each frame using the coefficients determined for the previous and current frames. The fourth subframe uses the values computed for that frame.

81.9 VSELP Codebook Search

The codebook search operation selects indices for the long-term filter (pitch lag L) and the two codebooks I and H so as to minimize the total weighted error. This closed-loop search is the most computationally complex part of the encoding operation, and significant effort has been invested to minimize the complexity of these operations without degrading performance. To reduce complexity, simultaneous optimization of the codebook selections is replaced by a sequential optimization procedure, which considers the long-term filter search as the most significant and therefore executes it first. The two vector-sum codebooks are considered to contribute less and less to the minimization of the error, and their search follows in sequence. Subdivision of the total codebook into two vector sums simplifies the processing and makes the result less sensitive to errors in decoding the individual bits arising from transmission errors.

Entries from each of the two vector-sum codebooks can be expressed as the sum of basis vectors. By orthogonalizing these basis vectors to the previously selected codebook component(s), one ensures that

the newly introduced components reduce the remaining errors. The subframes over which the codebook search is carried out are 5 ms or 40 samples long. An optimal search would need exploration of a 40 dimensional space. The vector-sum approximation limits the search to 14 dimensions after the optimal pitch lag has been selected. The search is further divided into two stages of 7 dimensions each. The two codebooks are specified in terms of the fourteen, 40-dimensional basis vectors stored at the encoder and decoder. The two 7-b indices indicate the required weights on the basic vectors to arrive at the two optimum codewords.

The codebook search can be viewed as selecting the three best directions in 40-dimensional space, which when summed result in the best approximation to the weighted input signal. The gains of the three components are determined through a separate error minimization process.

81.10 Long-Term Filter Search

The long-term filter is optimized by selection of a lag value that minimizes the error between the weighted input signal $p(n)$ and the past excitation signal filtered by the current weighted synthesis filter $H(z)$. There are 127 possible coded lag values provided corresponding to lags of 20–146 samples. One value is reserved for the case when all correlations between the input and the lagged residuals are negative and use of no long term filter output would be best. To simplify the convolution operation between the impulse response of the weighted synthesis filter and the past excitation, the impulse response is truncated to 21 samples or 2.5 ms. Once the lag is determined, the untruncated impulse response is used to compute the weighted long-term lag vector.

81.11 Orthogonalization of the Codebooks

Prior to the search of the first codebook, each filtered basis vector may be made orthogonal to the long-term filter output, the zero-state response of the weighted synthesis filter $H(z)$ to the long-term prediction vector. Each orthogonalized filtered basis vector is computed by subtracting its projection onto the long-term filter output from itself.

Similarly, the basis vectors of the second codebook can be orthogonalized with respect to both the long-term filter output and the first codebook output, the zero-state response of $H(z)$ to the previously selected summation of first-codebook basis vectors. In each case the codebook excitation can be reconstituted as

$$u_{k,i}(n) = \sum_{m=1}^{M} q_{im} v_{k,m}(n)$$

where $k = 1, 2$ for the two codebooks, $i = I$ or H the 7-b code vector received, $v_{k,m}$ are the two sets of basis vectors, and $\theta_{im} = +1$ if bit m of codeword $i = 1$ and -1 if bit m of codeword $i = 0$. Orthogonalization is not required at the decoder since the gains of the codebooks outputs are determined with respect to the weighted nonorthogonalized code vectors.

81.12 Quantizing the Excitation and Signal Gains

The three codebook gain values β, γ_1, and γ_2 are transformed to three new parameters GS, $P0$, and $P1$ for quantization purposes. GS is an energy offset parameter that equalizes the input and output signal energies. It adjusts the energy of the output of the LPC synthesis filter to equal the energy computed for the same subframe at the encoder input. $P0$ is the energy contribution of the long-term prediction vector as a fraction of the total excitation energy within the subframe. Similarly, $P1$ is the energy contribution of the code vector selected from the first codebook as a fraction of the total excitation energy of the subframe.

The transformation reduces the dynamic range of the parameters to be encoded. An 8-b **vector quantizer** efficiently encodes the appropriate (GS, $P0$, $P1$) vectors by selecting the vector which minimizes the weighted error. The received and decoded values β, γ_1, and γ_2 are computed from the received (GS, $P0$, $P1$) vector and applied to reconstitute the decoded signal.

81.13 Channel Coding and Interleaving

The goals of channel coding are to reduce the impairments in the reconstructed speech due to transmission errors. The 159 b characterizing each 20-ms block of speech are divided into two classes, 77 in class 1 and 82 in class 2. Class 1 includes the bits in which errors result in a more significant impairment, whereas the speech quality is considered less sensitive to the class-2 bits. Class 1 generally includes the gain, pitch lag, and more significant reflection coefficient bits. In addition, a 7-b cyclic redundancy check is applied to the 12 most perceptually significant bits of class 1 to indicate whether the error correction was successful. Failure of the CRC check at the receiver suggests that the received information is so erroneous that it would be better to discard it than use it. The error correction coding is illustrated in Fig. 81.5.

The error correction technique used is rate 1/2 convolutional coding with a constraint length of 5 [5]. A tail of 5 b is appended to the 84 b to be convolutionally encoded to result in a 178-b output. Inclusion of the tail bits ensures independent decoding of successive time slots and no propagation of errors between slots.

Interleaving the bits to be transmitted over two time slots is introduced to diminish the effects of short deep fades and to improve the error-correction capabilities of the channel coding technique. Two speech frames, the previous and the present, are interleaved so that the bits from each speech block span two transmission time slots separated by 20 ms. The interleaving attempts to separate the convolutionally coded class-1 bits from one frame as much as possible in time by inserting noncoded class-2 bits between them.

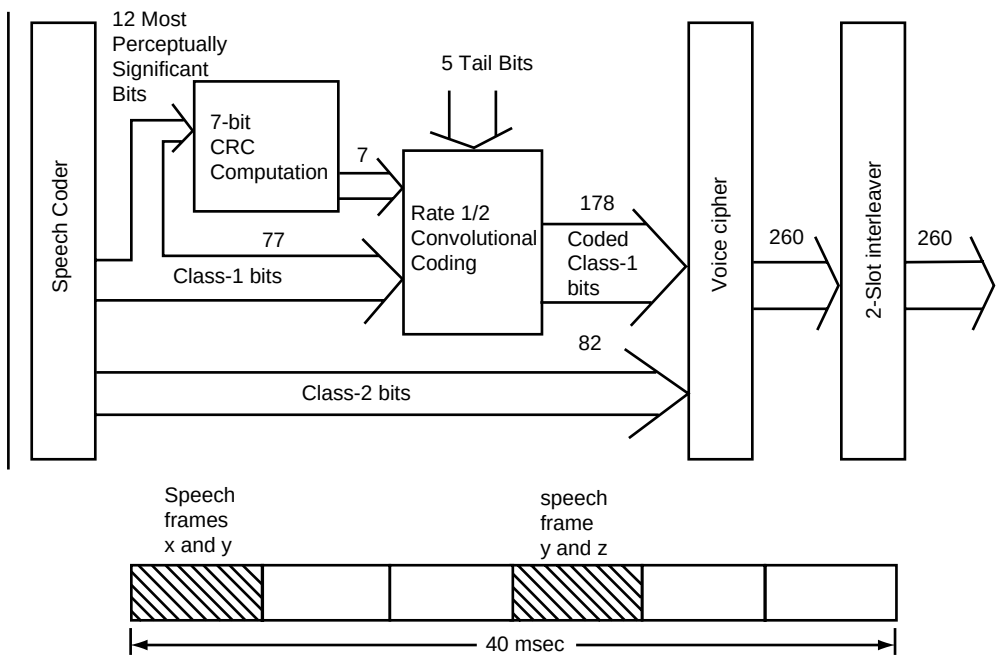


FIGURE 81.5 Error correction insertion for speech coder. *Source*: TIA, 1992. Cellular Systems Dual-mode Mobile Station–Base Station Compatibility Standards TIA/EIA IS-54. With permission.

81.14 Bad Frame Masking

A CRC failure indicates that the received data is unusable, either due to transmission errors resulting from a fade, or from pre-emption of the time slot by a control message (fast associated control channel, FACCH). To mask the effects that may result from leaving a gap in the speech signal, a masking operation based on the temporal redundancy between adjacent speech blocks has been proposed. Such masking can at best bridge over short gaps but cannot recover loss of signal of longer duration. The bad frame masking operation may follow a finite state machine where each state indicates an operation appropriate to the elapsed duration of the fade to which it corresponds. The masking operation consists of copying the previous LPC information and attenuating the gain of the signal. State 6 corresponds to error sequences exceeding 100 ms, for which the output signal is muted. The result of such a masking operation is generation of an extrapolation in the gap to the previously received signal, significantly reducing the perceptual effects of short fades. No additional delay is introduced in the reconstructed signal. At the same time, the receiver will report a high frequency of bad frames leading the system to explore handoff possibilities immediately. A quick successful handoff will result in rapid signal recovery.

81.15 ACELP Encoder

The ACELP encoder employs linear prediction analysis and quantization techniques similar to those used in VSELP and discussed in Section 81.6. The frame structure of 20 ms frames and 5 ms subframes is preserved. Linear prediction analysis is carried out for every frame. The ACELP encoder uses a long-term filter similar to the one discussed in Section 81.10 and represented as an adaptive codebook. The nonpredictable part of the LPC residual is represented in terms of ACELP codebooks, which replace the two VSELP codebooks shown in Fig. 81.3.

Instead of encoding the reflection coefficients as in VSELP, the information is transformed into line-spectral frequency pairs (LSP) [8]. The LSPs can be derived from linear prediction coefficients, a 10th order analysis generating 10 line-spectral frequencies (LSF), 5 poles, and 5 zeroes. The LSFs can be vector quantized and the LPC coefficients recalculated from the quantized LSFs. As long as the interleaved order of the poles and zeroes is preserved, quantization of the LSPs preserves the stability of the LPC synthesis filters. The LSPs of any frame can be better predicted from the values calculated and transmitted corresponding to previous frames, resulting in additional advantages. The long-term means of the LSPs are calculated for a large body of speech data and stored at both the encoder and decoder. First-order moving-average prediction is then used for the mean-removed LSPs. The time-prediction technique also permits use of predicted values for the LSPs in case uncorrectable transmissions errors are encountered, resulting in reduced speech degradation. To simplify the vector quantization operations, each LSP vector is split into 3 subvectors of dimensions 3, 3, and 4. The three subvectors are quantized with 8, 9, and 9 bits respectively, corresponding to a total bit assignment of 26 bits per frame for LPC information.

81.16 Algebraic Codebook Structure and Search

Algebraic codebooks contain relatively few pulses having nonzero values leading to rapid search of the possible innovation vectors, the vectors which together with the ACB output form the excitation of the LPC filter for the current subframe. In this implementation the 40-position innovation vector contains only four nonzero pulses and each can take on only values +1 and -1. The 40 positions are divided into four tracks and one pulse is selected from each track. The tracks are generally equally spaced but differ in their starting value, thus the first pulse can take on positions 0, 5, 10, 15, 20, 25, 30, or 35 and the second has possible positions 1, 6, 11, 16, 21, 26, 31, or 36. The first three pulse positions are coded with 3 bits and the fourth pulse position (starting positions 3 or 4) with 4 bits, resulting in a 17-bit sequence for the algebraic code of each subframe.

The algebraic codebook is searched by minimizing the mean square error between the weighted input speech and the weighted synthesized speech over the time span of each subframe. In each case the weighting

is that produced by a perceptual weighting filter that has the effect of shaping the spectrum of the synthesis error signal so that it is better masked by spectrum of the current speech signal.

81.17 Quantization of the Gains for ACELP Encoding

The adaptive codebook gain and the fixed (algebraic) codebook gains are vector quantized using a 7-bit codebook. The gain codebook search is performed by minimizing the mean-square of the weighted error between the original and the reconstructed speech, expressed as a function of the adaptive codebook gain and a fixed codebook correction factor. This correction factor represents the log energy difference between a predicted gain and an estimated gain. The predicted gain is computed using fourth-order moving-average prediction with fixed coefficients on the innovation energy of each subframe. The result is a smoothed energy profile even in the presence of modest quantization errors. As discussed above in case of the LSP quantization, the moving-average prediction serves to provide predicted values even when the current frame information is lost due to transmission errors. Degradations resulting from loss of one or two frames of information are thereby mitigated.

81.18 Channel Coding for ACELP Encoding

The channel coding and interleaving operations for ACELP speech coding are similar to those discussed in Section 81.13 for VSELP coding. The number of bits protected by both error-detection (parity) and error-correction convolutional coding is increased to 48 from 12. Rate 1/2 convolutional coding is used on the 108 more significant bits, 96 class-1 bits, 7 CRC bits and the 5 tail bits of the convolutional coder, resulting in 216 coded class-1 bits. Eight of the 216 bits are dropped by puncturing, yielding 208 coded class-1 bits which are then combined with 52 nonprotected class-2 bits. As compared to the channel coding of the VSELP encoder, the numbers of protected bits is increased and the number of unprotected bits is reduced while keeping the overall coding structure unchanged.

81.19 Conclusions

The IS-54 digital cellular standard specifies modulation and speech coding techniques for mobile cellular systems that allow the interoperation of terminals built by a variety of manufacturers and systems operated across the country by a number of different service providers. It permits speech communication with good quality in a transmission environment characterized by frequent multipath fading and significant intercell interference. Generally, the quality of the IS-54 decoded speech is better at the edges of a cell than the corresponding AMPS transmission due to the error mitigation resulting from channel coding. Near a base station or in the absence of significant fading and interference, the IS-54 speech quality is reported to be somewhat worse than AMPS due to the inherent limitations of the analysis–synthesis model in reconstructing arbitrary speech signals with limited bits. The IS-641 standard coder achieves higher speech quality, particularly at the edges of heavily occupied cells where transmission errors may be more numerous. At this time no new systems following the IS-54 standard are being introduced. Most base stations have been converted to transmit and receive on the IS-641 standard as well and use of IS-54 transmissions is dropping rapidly. At the time of its introduction in 1996 the IS-641 coder represented the state of the art in terms of toll quality speech coding near 8 kb/s, a significant improvement over the IS-54 coder introduced in 1990. These standards represent reasonable engineering compromises between high performance and complexity sufficiently low to permit single-chip implementations in mobile terminals.

Both IS-54 and IS-641 are considered second generation cellular standards. Third generation cellular systems promise higher call capacities through better exploitation of the time-varying transmission requirements of speech conversations, as well as improved modulation and coding in wider spectrum bandwidths that achieve similar bit-error ratios but reduce the required transmitted power. Until such systems are introduced, the second generation TDMA systems can be expected to provide many years of successful cellular and personal communications services.

Defining Terms

Codebook: A set of signal vectors available to both the encoder and decoder.

Covariance lattice algorithm: An algorithm for reduction of the covariance matrix of the signal consisting of several lattice stages, each stage implementing an optimal first-order filter with a single coefficient.

Reflection coefficient: A parameter of each stage of the lattice linear prediction filter that determines (1) a forward residual signal at the output of the filter-stage by subtracting from the forward residual at the input a linear function of the backward residual, also (2) a backward residual at the output of the filter stage by subtracting a linear function of the forward residual from the backward residual at the input.

Vector quantizer: A quantizer that assigns quantized vectors to a vector of parameters based on their current values by minimizing some error criterion.

References

1. Atal, B.S. and Hanauer, S.L., Speech analysis and synthesis by linear prediction of the speech wave. *J. Acoust. Soc. Am.,* 50, 637–655, 1971.
2. Atal, B.S. and Schroeder, M., Stochastic coding of speech signals at very low bit rates. *Proc. Int. Conf. Comm.,* 1610–1613, 1984.
3. Gersho, A., Advances in speech and audio compression. *Proc. IEEE,* 82, 900–918, 1994.
4. Gerson, I.A. and Jasiuk, M.A., Vector sum excited linear prediction (VSELP) speech coding at 8 kbps. *Int. Conf. Acoust. Speech and Sig. Proc.,* ICASSP90, 461–464, 1990.
5. Lin S. and Costello, D., *Error Control Coding: Fundamentals and Application,* Prentice Hall, Englewood Cliffs, NJ, 1983.
6. Makhoul, J., Linear prediction, a tutorial review. *Proc. IEEE,* 63, 561–580, 1975.
7. Salami, R., Laflamme, C., Adoul, J.P., and Massaloux, D., A toll quality 8 kb/s speech codec for the personal communication system (PCS). *IEEE Trans. Vehic. Tech.,* 43, 808–816, 1994.
8. Soong, F.K. and Juang, B.H., Line spectrum pair (LSP) and speech data compression. *Proc. ICASSP'84*, 1.10.1–1.10.4, 1984.
9. Telecommunications Industry Association, EIA/TIA Interim Standard, Cellular System Dual-mode Mobile Station–Base Station Compatibility Standard IS-54B, TIA/EIA, Washington, D.C., 1992.

Further Information

For a general treatment of speech coding for telecommunications, see N.S. Jayant and P. Noll, *Digital Coding of Waveforms*, Prentice Hall, Englewood, NJ, 1984. For a more detailed treatment of linear prediction techniques, see J. Markel and A. Gray, *Linear Prediction of Speech*, Springer–Verlag, New York, 1976.

82

The British Cordless Telephone Standard: CT-2

Lajos Hanzo
University of Southampton

82.1 History and Background

Following a decade of world-wide research and development (R&D), cordless telephones (**CT**) are now becoming widespread consumer products, and they are paving the way toward ubiquitous, low-cost personal communications networks (**PCN**) [7,8]. The two most well-known European representatives of CTs are the digital European cordless telecommunications (**DECT**) system [1,5] and the CT-2 system [2,6]. Three potential application areas have been identified, namely, domestic, business, and public access, which is also often referred to as telepoint (**TP**).

In addition to conventional voice communications, CTs have been conceived with additional data services and local area network (**LAN**) applications in mind. The fundamental difference between conventional mobile radio systems and CT systems is that CTs have been designed for small to very small cells, where typically benign low-dispersion, dominant line-of-sight (**LOS**) propagation conditions prevail. Therefore, CTs can usually dispense with channel equalizers and complex low-rate speech codecs, since the low-signal dispersion allows for the employment of higher bit rates before the effect of channel dispersion becomes a limiting factor. On the same note, the LOS propagation scenario is associated with mild fading or near-constant received signal level, and when combined with appropriate small-cell power-budget design, it ensures a high average signal-to-noise ratio (**SNR**). These prerequisites facilitate the employment of high-rate, low-complexity speech codecs, which maintain a low battery drain. Furthermore, the deployment of forward error correction codecs

can often also be avoided, which reduces both the bandwidth requirement and the power consumption of the portable station (PS).

A further difference between public land mobile radio (**PLMR**) systems [3] and CTs is that whereas the former endeavor to standardize virtually all system features, the latter seek to offer a so-called access technology, specifying the common air interface (**CAI**), access and signalling protocols, and some network architecture features, but leaving many other characteristics unspecified. By the same token, whereas PLMR systems typically have a rigid frequency allocation scheme and fixed cell structure, CTs use dynamic channel allocation (**DCA**) [4]. The DCA principle allows for a more intelligent and judicious channel assignment, where the base station (BS) and PS select an appropriate traffic channel on the basis of the prevailing traffic and channel quality conditions, thus minimizing, for example, the effect of cochannel interference or channel blocking probability.

In contrast to PLMR schemes, such as the Pan-European global system of mobile communications (GSM) system [3], CT systems typically dispense with sophisticated mobility management, which accounts for the bulk of the cost of PLMR call charges, although they may facilitate limited hand-over capabilities. Whereas in residential applications CTs are the extension of the public switched telephone network (PSTN), the concept of omitting mobility management functions, such as location update, etc., leads to telepoint CT applications where users are able to initiate but not to receive calls. This fact drastically reduces the network operating costs and, ultimately, the call charge at a concomittant reduction of the services rendered.

Having considered some of the fundamental differences between PLMR and CT systems let us now review the basic features of the CT-2 system.

82.2 The CT-2 Standard

The European CT-2 recommendation has evolved from the British standard **MPT-1375** with the aim of ensuring the compatibility of various manufacturers' systems as well as setting performance requirements, which would encourage the development of cost-efficient implementations. Further standardization objectives were to enable future evolution of the system, for example, by reserving signalling messages for future applications and to maintain a low PS complexity even at the expense of higher BS costs. The CT-2 or MPT 1375 CAI recommendation is constituted by the four following parts.

1. *Radio interface:* Standardizes the radio frequency (RF) parameters, such as legitimate channel frequencies, the modulation method, the transmitter power control, and the required receiver sensitivity as well as the carrier-to-interference ratio (CIR) and the time division duplex (**TDD**) multiple access scheme. Furthermore, the transmission burst and master/slave timing structures to be used are also laid down, along with the scrambling procedures to be applied.
2. *Signalling layers one and two:* Defines how the bandwidth is divided among signalling, traffic data, and synchronization information. The description of the first signalling layer includes the dynamic channel allocation strategy, calling channel detection, as well as link setup and establishment algorithms. The second layer is concerned with issues of various signalling message formats, as well as link establishment and re-establishment procedures.
3. *Signalling layer three:* The third signalling layer description includes a range of message sequence diagrams as regards to call setup to telepoint BSs, private BSs, as well as the call clear down procedures.
4. *Speech coding and transmission:* The last part of the standard is concerned with the algorithmic and performance features of the audio path, including frequency responses, clipping, distortion, noise, and delay characteristics.

Having briefly reviewed the structure of the CT-2 recommendations let us now turn our attention to its main constituent parts and consider specific issues of the system's operation.

82.3 The Radio Interface

Transmission Issues

In our description of the system we will adopt the terminology used in the recommendation, where the PS is called cordless portable part (**CPP**), whereas the BS is referred to as cordless fixed part (**CFP**). The channel bandwidth and the channel spacing are 100 kHz, and the allocated system bandwidth is 40 MHz, which is hosted in the range of 864.15–868.15 MHz. Accordingly, a total of 40 RF channels can be utilized by the system.

The accuracy of the radio frequency must be maintained within ±10 kHz of its nominal value for both the CFP and CPP over the entire specified supply voltage and ambient temperature range. To counteract the maximum possible frequency drift of 20 kHz, automatic frequency correction (**AFC**) may be used in both the CFP and CPP receivers. The AFC may be allowed to control the transmission frequency of only the CPP, however, in order to prevent the misalignment of both transmission frequencies.

Binary frequency shift keying (FSK) is proposed, and the signal must be shaped by an approximately Gaussian filter in order to maintain the lowest possible frequency occupancy. The resulting scheme is referred to as Gaussian frequency shift keying (**GFSK**), which is closely related to Gaussian minimum shift keying (GMSK) [7] used in the DECT [1] and GSM [3] systems.

Suffice to say that in M-arry FSK modems the carrier's frequency is modulated in accordance with the information to be transmitted, where the modulated signal is given by

$$S_i(t) = \sqrt{\frac{2E}{T}}\cos[\omega_i t + \Phi] \quad i = 1,\ldots,M$$

and E represents the bit energy, T the signalling interval length, ω_i has M discrete values, whereas the phase Φ is constant.

Multiple Access and Burst Structure

The so-called TDD multiple access scheme is used, which is demonstrated in Fig. 82.1. The simple principle is to use the same radio frequency for both uplink and downlink transmissions between the CPP and the CFP, respectively, but with a certain staggering in time. This figure reveals further details of the burst structure, indicating that 66 or 68 b per TDD frame are transmitted in both directions.

There is a 3.5- or 5.5-b duration guard period (**GP**) between the uplink and downlink transmissions, and half of the time the CPP (the other half of the time the CFP) is transmitting with the other part listening, accordingly. Although the guard period wastes some channel capacity, it allows a finite time for both the CPP and CFP for switching from transmission to reception and vice versa. The burst structure of Fig. 82.1 is used during normal operation across an established link for the transmission of adaptive differential pulse code modulated (ADPCM) speech at 32 kb/s according to the CCITT G721 standard in

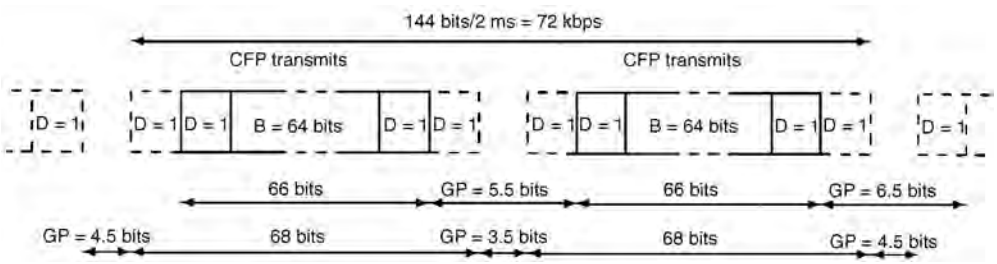


FIGURE 82.1 M1 burst and TDD frame structure.

a so-called B channel or bearer channel. The D channel, or signalling channel, is used for the transmission of link control signals. This specific burst structure is referred to as a multiplex one (**M1**) frame.

Since the speech signal is encoded according to the CCITT G721 recommendation at 32 kb/s the TDD bit rate must be in excess of 64 kb/s in order to be able to provide the idle guard space of 3.5- or 5.5-b interval duration plus some signalling capacity. This is how channel capacity is sacrificed to provide the GP. Therefore, the transmission bit rate is stipulated to be 72 kb/s and the transmission burst length is 2 ms, during which 144-b intervals can be accommodated. As it was demonstrated in Fig. 82.1, 66 or 68 b are transmitted in both the uplink and downlink burst, and taking into account the guard spaces, the total transmission frame is constituted by $(2 \cdot 68) + 3.5 + 4.5 = 144$ b or equivalently, by $(2 \cdot 66) + 5.5 + 4.5 = 144$ b. The 66-b transmission format is compulsory, whereas the 68-b format is optional. In the 66-b burst there is one D bit dedicated to signalling at both ends of the burst, whereas in the 68-b burst the two additional bits are also assigned to signalling. Accordingly, the signalling rate becomes 2 b/2 ms or 4 b/2 ms, corresponding to 1 kb/s or 2 kb/s signalling rates.

Power Ramping, Guard Period, and Propagation Delay

As mentioned before and suggested in Fig. 82.1, there is a 3.5- or 5.5-b interval duration GP between transmitted and received bursts. Since the signalling rate is 72 kb/s, the bit interval becomes about $1/(72\text{ kb/s}) \approx 13.9\ \mu$s and, hence, the GP duration is about 49 μs or 76 μs. This GP serves a number of purposes. Primarily, the GP allows the transmitter to ramp up and ramp down the transmitted signal level smoothly over a finite time interval at the beginning and end of the transmitted burst. This is necessary, because if the transmitted signal is toggled instantaneously, that is equivalent to multiplying the transmitted signal by a rectangular time-domain window function, which corresponds in the frequency domain to convolving the transmitted spectrum with a sinc function. This convolution would result in spectral side-lobes over a very wide frequency range, which would interfere with adjacent channels. Furthermore, due to the introduction of the guard period, both the CFP and CPP can tolerate a limited propagation delay, but the entire transmitted burst must arrive within the receivers' window, otherwise the last transmitted bits cannot be decoded.

Power Control

In order to minimize the battery drain and the cochannel interference load imposed upon cochannel users, the CT-2 system provides a power control option. The CPPs must be able to transmit at two different power levels, namely, either between 1 and 10 mW or at a level between 12 and 20 dB lower. The mechanism for invoking the lower CPP transmission level is based on the received signal level at the CFP. If the CFP detects a received signal strength more than 90 dB relative to 1 μV/m, it may instruct the CPP to drop its transmitted level by the specified 12–20 dB. Since the 90-dB gain factor corresponds to about a ratio of 31,623, this received signal strength would be equivalent for a 10-cm antenna length to an antenna output voltage of about 3.16 mV. A further beneficial ramification of using power control is that by powering down CPPs that are in the vicinity of a telepoint-type multiple-transceiver CFP, the CFP's receiver will not be so prone to being desensitised by the high-powered close-in CPPs, which would severely degrade the reception quality of more distant CPPs.

82.4 Burst Formats

As already mentioned in the previous section on the radio interface, there are three different subchannels assisting the operation of the CT-2 system, namely, the *voice/data channel* or *B channel*, the *signalling channel* or *D channel*, and the *burst synchronization channel* or **SYN** *channel*. According to the momentary system requirements, a variable fraction of the overall channel capacity or, equivalently, a variable fraction of the bandwidth can be allocated to any of these channels. Each different channel capacity or bandwidth allocation mode is associated with a different burst structure and accordingly bears a different name. The corresponding burst structures are termed as multiplex one (M1), multiplex two (**M2**), and multiplex three (**M3**), of which multiplex one used during the normal operation of established links has already

TABLE 82.1 CT-2 Synchronization Patterns

	MSB (sent last)					LSB (sent first)
CHMF	1011	1110	0100	1110	0101	0000
CHMP	0100	0001	1011	0001	1010	1111
SYNCF	1110	1011	0001	1011	0000	0101
SYNCP	0001	0100	1110	0100	1111	1010

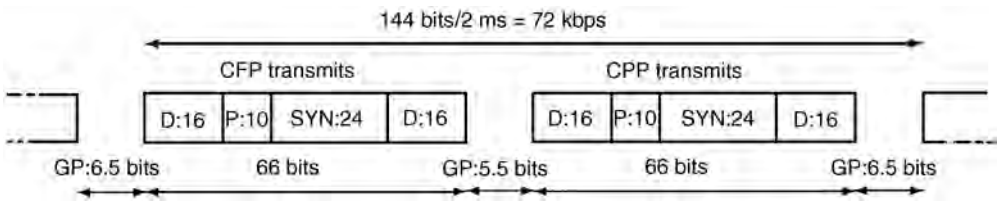


FIGURE 82.2 CT-2 multiplex two burst structure.

been described in the previous section. Multiplex two and three will be extensively used during link setup and establishment in subsequent sections, as further details of the system's operation are unravelled.

Signalling layer one (L1) defines the burst formats multiplex one–three just mentioned, outlines the calling channel detection procedures, as well as link setup and establishment techniques. *Layer two* (**L2**) deals with issues of acknowledged and unacknowledged information transfer over the radio link, error detection and correction by retransmission, correct ordering of messages, and link maintenance aspects.

The burst structure multiplex two is shown in Fig. 82.2. It is constituted by two 16-b D-channel segments at both sides of the 10-b *preamble (P)* and the 24-b frame synchronization pattern (SYN), and its signalling capacity is 32 b/2 ms = 16 kb/s. Note that the M2 burst does not carry any B-channel information, it is dedicated to synchronization purposes. The 32-b D-channel message is split in two 16-b segments in order to prevent that any 24-b fraction of the 32-b word emulates the 24-b SYN segment, which would result in synchronization misalignment.

Since the CFP plays the role of the master in a telepoint scenario communicating with many CPPs, all of the CPP's actions must be synchronized to those of the CFP. Therefore, if the CPP attempts to initiate a call, the CFP will reinitiate it using the M2 burst, while imposing its own timing structure. The 10-b preamble consists of an alternate zero/one sequence and assists in the operation of the clock recovery circuitry, which has to be able to recover the clock frequency before the arrival of the SYN sequence, in order to be able to detect it. The SYN sequence is a unique word determined by computer search, which has a sharp autocorrelation peak, and its function is discussed later. The way the M2 and M3 burst formats are used for signalling purposes will be made explicit in our further discussions when considering the link setup procedures.

The specific SYN sequences used by the CFP and the CPP are shown in Table 82.1 along with the so-called *channel marker* (**CHM**) sequences used for synchronization purposes by the M3 burst format. Their differences will be made explicit during our further discourse. Observe from the table that the sequences used by the CFP and CPP, namely, SYNF, **CHMF** and SYNP, **CHMP**, respectively, are each other's bit-wise inverses. This was introduced in order to prevent CPPs and CFPs from calling each other directly. The CHM sequences are used, for instance, in residential applications, where the CFP can issue an M2 burst containing a 24-b CHMF sequence and a so-called poll message mapped on to the D-channel bits in order to wake up the specific CPP called. When the called CPP responds, the CFP changes the CHMF to SYNF in order to prevent waking up further CPPs unnecessarily.

Since the CT-2 system does not entail mobility functions, such as registration of visiting CPPs in other than their own home cells, in telepoint applications all calls must be initiated by the CPPs. Hence, in this scenario when the CPP attempts to set up a link, it uses the so-called multiplex three

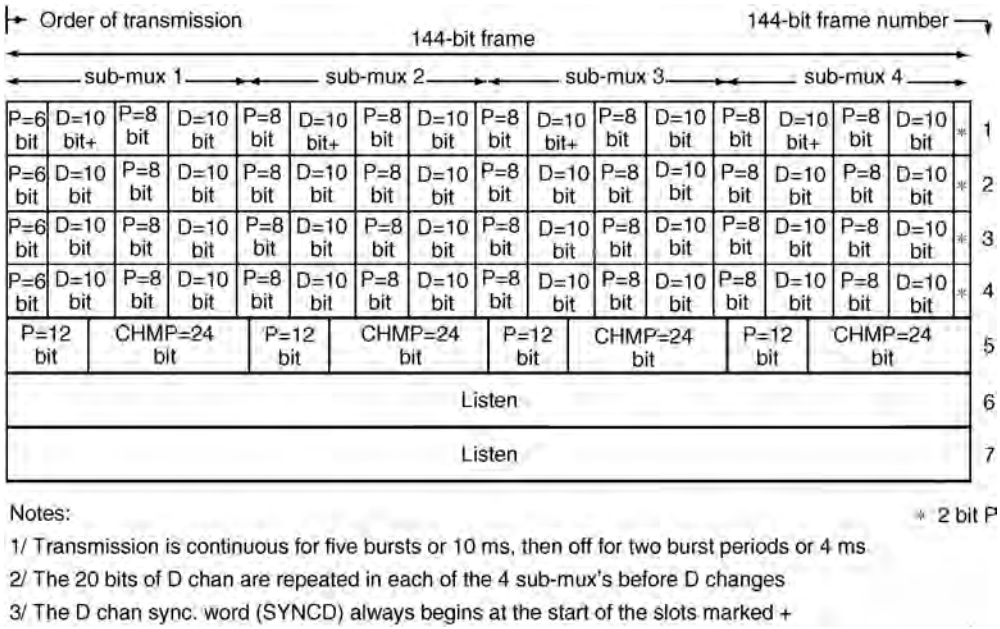


FIGURE 82.3 CT-2 multiplex three burst structure.

burst format displayed in Fig. 82.3. The design of the M3 burst reflects that the CPP initiating the call is oblivious of the timing structure of the potentially suitable target CFP, which can detect access attempts only during its receive window, but not while the CFP is transmitting. Therefore, the M3 format is rather complex at first sight, but it is well structured, as we will show in our further discussions. Observe in the figure that in the M3 format there are five consecutive 2-ms long 144-b transmitted bursts, followed by two idle frames, during which the CPP listens in order to determine whether its 24-b CHMP sequence has been detected and acknowledged by the CFP. This process can be followed by consulting Fig. 82.6, which will be described in depth after considering the detailed construction of the M3 burst.

The first four of the five 2-ms bursts are identical D-channel bursts, whereas the fifth one serves as a synchronization message and has a different construction. Observe, furthermore, that both the first four 144-b bursts as well as the fifth one contain four so-called submultiplex segments, each of which hosts a total of (6 + 10 + 8 + 10 + 2) = 36 b. In the first four 144-b bursts there are (6 + 8 + 2) = 16 one/zero clock-synchronizing P bits and (10 + 10) = 20 D bits or signalling bits. Since the D-channel message is constituted by two 10-b half-messages, the first half of the D-message is marked by the + sign in Fig. 82.3. As mentioned in the context of M2, the D-channel bits are split in two halves and interspersed with the preamble segments in order to ensure that these bits do not emulate valid CHM sequences. Without splitting the D bits this could happen upon concatenating the one/zero P bits with the D bits, since the tail of the SYNF and SYNP sequences is also a one/zero segment. In the fifth 144-b M3 burst, each of the four submultiplex segments is constituted by 12 preamble bits and 24 CPP channel marker (CHMP) bits.

The four-fold submultiplex M3 structure ensures that irrespective of how the CFP's receive window is aligned with the CPP's transmission window, the CFP will be able to capture one of the four submultiplex segments of the fifth M3 burst, establish clock synchronization during the preamble, and lock on to the CHMP sequence. Once the CFP has successfully locked on to one of the CHMP words, the corresponding D-channel messages comprising the CPP identifier can be decoded. If the CPP identifier has been recognized, the CFP can attempt to reinitialize the link using its own master synchronization.

82.5 Signalling Layer Two (L2)

General Message Format

The signalling L2 is responsible for acknowledged and un-acknowledged information transfer over the air interface, error detection and correction by retransmission, as well as for the correct ordering of messages in the acknowledged mode. Its further functions are the link end-point identification and link maintenance for both CPP and CFP, as well as the definition of the L2 and **L3** interface.

Compliance with the L2 specifications will ensure the adequate transport of messages between the terminals of an established link. The L2 recommendations, however, do not define the meaning of messages, this is specified by L3 messages, albeit some of the messages are undefined in order to accommodate future system improvements.

The L3 messages are broken down to a number of standard packets, each constituted by one or more codewords (CW), as shown in Fig. 82.4. The codewords have a standard length of eight octets, and each packet contains up to six codewords. The first codeword in a packet is the so-called address codeword (ACW) and the subsequent ones, if present, are data codewords (**DCW**). The first octet of the ACW of each packet contains a variety of parameters, of which the binary flag L3_END is indicated in Fig. 82.4, and it is set to zero in the last packet. If the L3 message transmitted is mapped onto more than one packet, the packets must be numbered up to *N*. The address codeword is always preceded by a 16-b D-channel frame synchronization word **SYNCD**. Furthermore, each eight-octet CW is protected by a 16-b parity-check word occupying its last two octets. The binary Bose–Chaudhuri–Hocquenghem BCH (63,48) code is used to encode the first six octets or 48 b by adding 15 parity b to yield 63 b. Then bit 7 of octet 8 is inverted and bit 8 of octet 8 added such that the 64-b codeword has an even parity. If there are no D-channel packets to send, a 3-octet idle message IDLE_D constituted by zero/one reversals is transmitted. The 8-octet format of the ACWs and DCWs is made explicit in Fig. 82.5, where the two parity check octets occupy octets 7 and 8. The first octet hosts a number of control bits. Specifically, bit 1 is set to logical one for an ACW and to zero for a DCW, whereas bit 2 represents the so-called format type (**FT**) bit. $FT = 1$ indicates that variable length packet format is used for the transfer of L3 messages, whereas $FT = 0$ implies that a fixed length link setup is used for link end point addressing end service requests. *FT* is only relevant to ACWs, and in DCWs it has to be set to one.

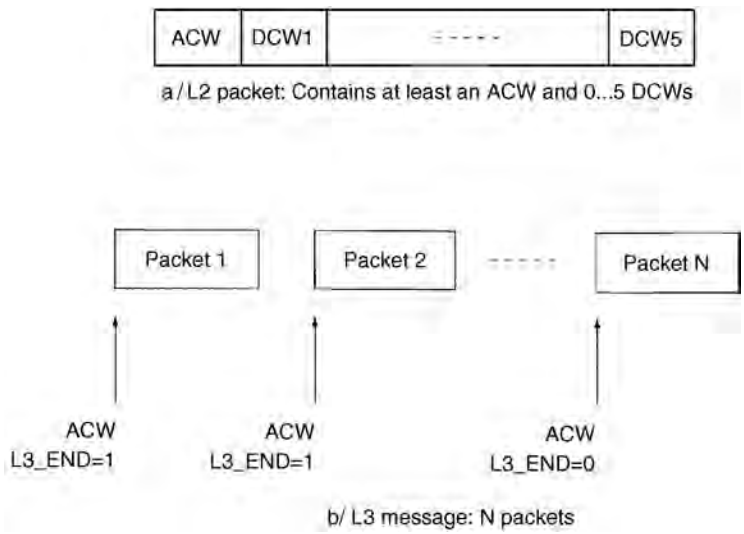


FIGURE 82.4 General L2 and L3 message format.

TABLE 82.2 Encoding of Link Status Messages

LS1	LS0	Message
0	0	Link_request
0	1	Link_grant
1	0	ID_OK
1	1	ID_lost

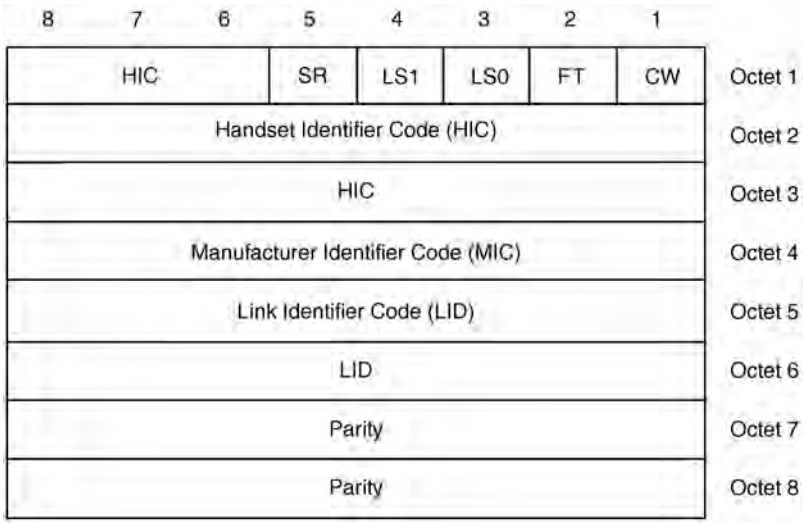


FIGURE 82.5 Fixed format packets mapped on M1, M2, and M3 during link initialization and on M1 and M2 during handshake.

Fixed Format Packet

As an example, let us focus our attention on the fixed format scenario associated with $FT = 0$. The corresponding codeword format defined for use in M1, M2, and M3 for link initiation and in M1 and M2 for handshaking is displayed in Fig. 82.5. Bits 1 and 2 have already been discussed, whereas the 2-bit link status (**LS**) field is used during call setup and handshaking. The encoding of the four possible LS messages is given in Table 82.2. The aim of these LS messages will become more explicit during our further discussions with reference to Fig. 82.6 and Fig. 82.7. Specifically, link_request is transmitted from the CPP to the CFP either in an M3 burst as the first packet during CPP-initiated call setup and link re-establishment, or returned as a poll response in an M2 burst from the CPP to the CFP, when the CPP is responding to a call. Link_grant is sent by the CFP in response to a link_request originating from the CPP. In octets 5 and 6 it hosts the so-called link identification (**LID**) code, which is used by the CPP, for example, to address a specific CFP or a requested service. The LID is also used to maintain link reference during handshake exchanges and link re-establishment. The two remaining link status handshake messages, namely, ID_OK and ID_lost, are used to report to the far end whether a positive confirmation of adequate link quality has been received within the required time-out period. These issues will be revisited during our further elaborations. Returning to Fig. 82.5, we note that the fixed packet format ($FT = 0$) also contains a 19-b handset identification code (**HIC**) and an 8-b manufacturer identification code (**MIC**). The concatenated HIC and MIC fields jointly from the unique 27-b portable identity code (**PIC**), serving as a link end-point identifier. Lastly, we have to note that bit 5 of octet 1 represents the signalling rate (**SR**) request/response bit, which is used by the calling party to specify the choice of the 66- or 68-b M1 format. Specifically, $SR = 1$ represents the four bit/burst M1 signalling format. The first 6 octets are then protected by the parity check information contained in octets 7 and 8.

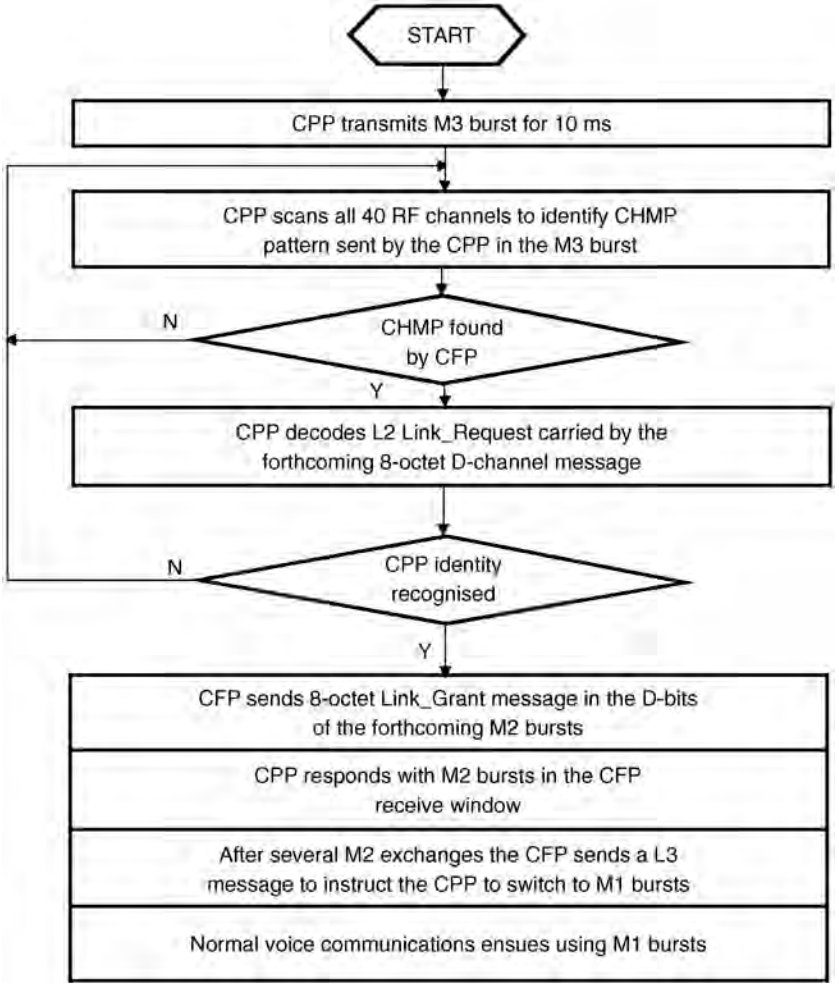


FIGURE 82.6 Flowchart of the CT-2 link initialization by the CPP.

82.6 CPP-Initiated Link Setup Procedures

Calls can be initiated at both the CPP and CFP, and the call initiation and detection procedures invoked depend on which party initiated the call. Let us first consider calling channel detection at the CFP, which ensues as follows. Under the instruction of the CFP control scheme, the RF synthesizer tunes to a legitimate RF channel and after a certain settling time commences reception. Upon receiving the M3 bursts from the CPP, the automatic gain control (AGC) circuitry adjusts its gain factor, and during the 12-b preamble in the fifth M3 burst, bit synchronization is established. This specific 144-b M3 burst, is transmitted every 14 ms, corresponding to every seventh 144-b burst. Now the CFP is ready to bit-synchronously correlate the received sequences with its locally stored CHMP word in order to identify any CHMP word arriving from the CPP. If no valid CHMP word is detected, the CFP may retune itself to the next legitimate RF channel, etc.

As mentioned, the call identification and link initialization process is shown in the flowchart of Fig. 82.6. If a valid 24-b CHMP word is identified, D-channel frame synchronization can take place using the 16-b SYNCD sequence and the next 8-octet L2 D-channel message delivering the link_request handshake portrayed earlier in Fig. 82.5 and Table 82.2 is decoded by the CFP. The required 16 + 64 = 80 D bits are accommodated in this scenario by the $4 \cdot 20 = 80$ D bits of the next four 144-b bursts of

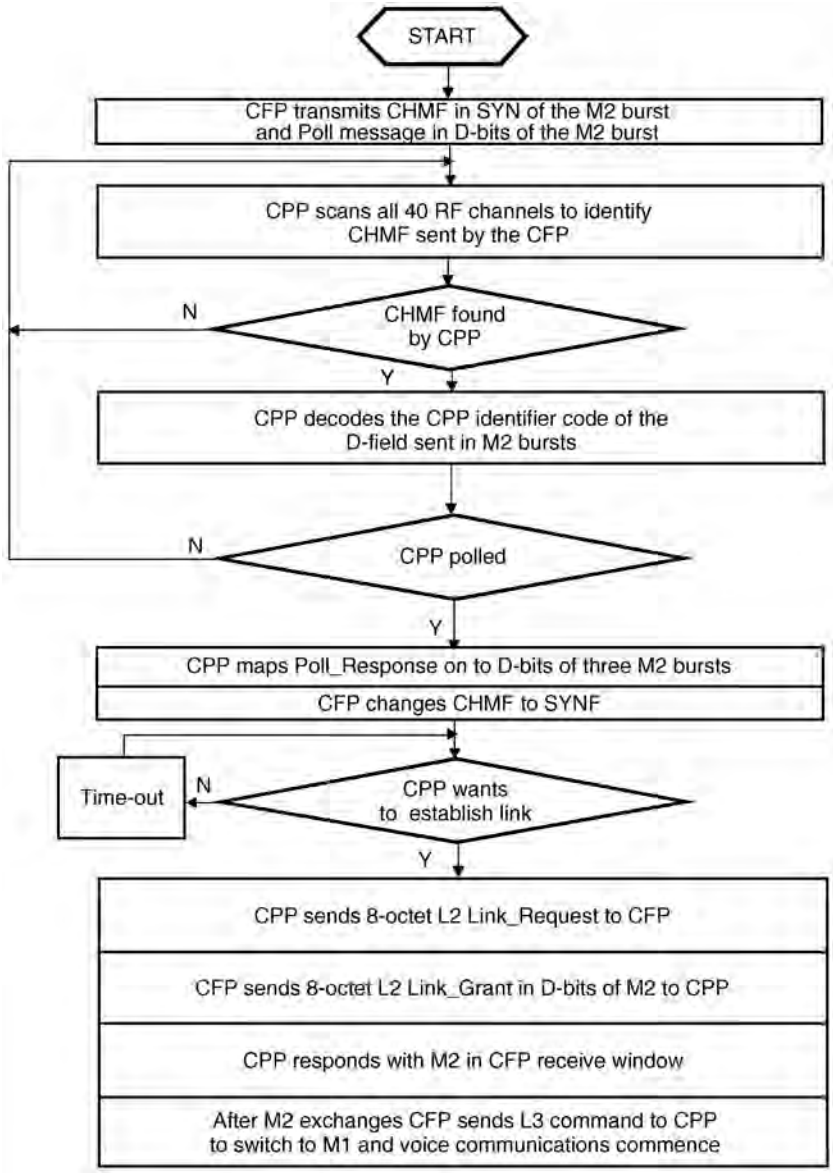


FIGURE 82.7 Flowchart of the CT-2 link initialization by the CFP.

the M3 structure, where the 20 D bits of the four submultiplex segments are transmitted four times within the same burst before the D message changes. If the decoded LID code of Fig. 82.5 is recognized by the CFP, the link may be reinitialized based on the master's timing information using the M2 burst associated with SYNF and containing the link_grant message addressed to the specific CPP identified by its PID.

Otherwise the CFP returns to its scanning mode and attempts to detect the next CHMP message. The reception of the CFP's 24-b SYNF segment embedded in the M2 message shown previously in Fig. 82.2 allows the CPP to identify the position of the CFP's transmit and receive windows and, hence, the CPP now can respond with another M2 burst within the receive window of the CFP. Following a number of M2 message exchanges, the CFP then sends a L3 message to instruct the CPP to switch to M1 bursts, which marks the commencement of normal voice communications and the end of the link setup session.

82.7 CFP-Initiated Link Setup Procedures

Similar procedures are followed when the CPP is being polled. The CFP transmits the 24-b CHMF words hosted by the 24-b SYN segment of the M2 burst shown in Fig. 82.2 in order to indicate that one or more CPPs are being paged. This process is displayed in the flowchart of Fig. 82.7, as well as in the timing diagram displayed in Fig. 82.8. The M2 D-channel messages convey the identifiers of the polled CPPs.

The CPPs keep scanning all 40 legitimate RF channels in order to pinpoint any 24-b CHMF words. Explicitly, the CPP control scheme notifies the RF synthesizer to retune to the next legitimate RF channel if no CHMF words have been found on the current one. The synthesizer needs a finite time to settle on the new center frequency and then starts receiving again. Observe in Fig. 82.8 that at this stage only the CFP is transmitting the M2 bursts; hence, the uplink-half of the 2-ms TDD frame is unused.

Since the M2 burst commences with the D-channel bits arriving from the CFP, the CPP receiver's AGC will have to settle during this 16-b interval, which corresponds to about $16 \cdot 1/[72\ \text{kb/s}] \approx 0.22$ ms. Upon the arrival of the 10 alternating one–zero preamble bits, bit synchronization is established. Now the CPP is ready to detect the CHMF word using a simple correlator circuitry, which establishes the appropriate frame synchronization. If, however, no CHMF word is detected within the receive window, the synthesizer will be retuned to the next RF channel, and the same procedure is repeated, until a CHMF word is detected.

When a CHMF word is correctly decoded by the CPP, the CPP is now capable of frame and bit synchronously decoding the D-channel bits. Upon decoding the D-channel message of the M2 burst, the CPP identifier (**ID**) constituted by the LID and PID segments of Fig. 82.5 is detected and compared to the CPP's own ID in order to decide as to whether the call is for this specific CPP. If so, the CPP ID is reflected back to the CFP along with a SYNP word, which is included in the SYN segment of an uplink M2 burst. This channel scanning and retuning process continues until a legitimate incoming call is detected or the CPP intends to initiate a call.

More precisely, if the specific CPP in question is polled and its own ID is recognized, the CPP sends its poll_response message in three consecutive M2 bursts, since the capacity of a single M2 burst is 32 D bits only, while the handshake messages of Fig. 82.5 and Table 82.2 require 8 octets preceded by a 16-b SYNCD segment. If by this time all paged CPPs have responded, the CFP changes the CHMF word to a SYNF word, in order to prevent activating dormant CPPs who are not being paged. If any of the paged CPPs intends to set up the link, then it will change its poll_response to a L2 link_request message, in response to which the CFP will issue an M2 link_grant message, as seen in Fig. 82.7, and from now on the procedure is identical to that of the CPP-initiated link setup portrayed in Fig. 82.6.

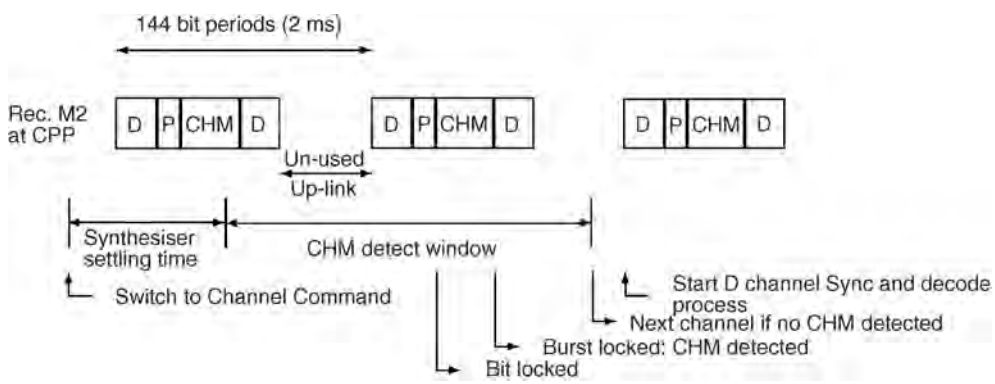


FIGURE 82.8 CT-2 call detection by the CPP.

82.8 Handshaking

Having established the link, voice communications is maintained using M1 bursts, and the link quality is monitored by sending handshaking (**HS**) signalling messages using the D-channel bits. The required frequency of the handshaking messages must be between once every 400 ms and 1000 ms. The CT-2 codewords ID_OK, ID_lost, link_request and link_grant of Table 82.2 all represent valid handshakes. When using M1 bursts, however, the transmission of these 8-octet messages using the 2- or 4-b/2 ms D-channel segment must be spread over 16 or 32 M1 bursts, corresponding to 32 or 64 ms.

Let us now focus our attention on the *handshake protocol* shown in Fig. 82.9. Suppose that the CPP's handshake interval of Thtx_p = 0.4 s since the start of the last transmitted handshake has expired, and hence the CPP prepares to send a handshake message HS_p. If the CPP has received a valid HS_f message from the CFP within the last Thrx_p = 1s, the CPP sends an HS_p = ID_OK message to the CFP, otherwise an ID_Lost HS_p. Furthermore, if the valid handshake was HS_f = ID_OK, the CPP will reset its HS_f lost timer Thlost_p to 10 s. The CFP will maintain a 1-s timer referred to as Thrx_f, which is reset to its initial value upon the reception of a valid HS_p from the CPP.

The CFP's actions also follow the structure of Fig. 82.9 upon simply interchanging CPP with CFP and the descriptor _p with _f. If the Thrx_f = 1 s timer expires without the reception of a valid HS_p from

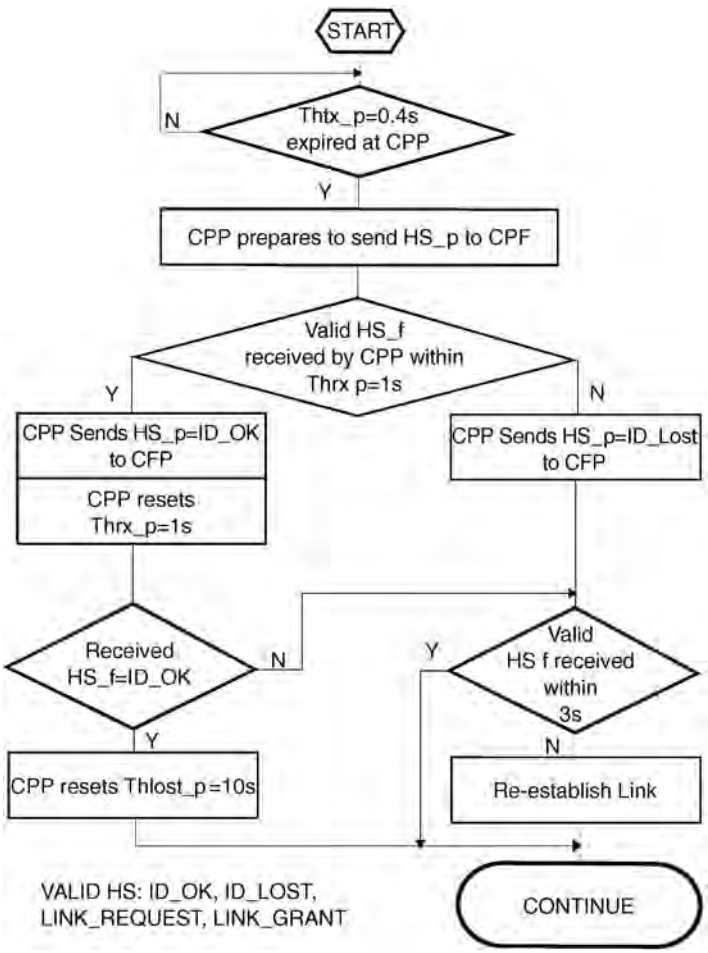


FIGURE 82.9 CT-2 handshake algorithms.

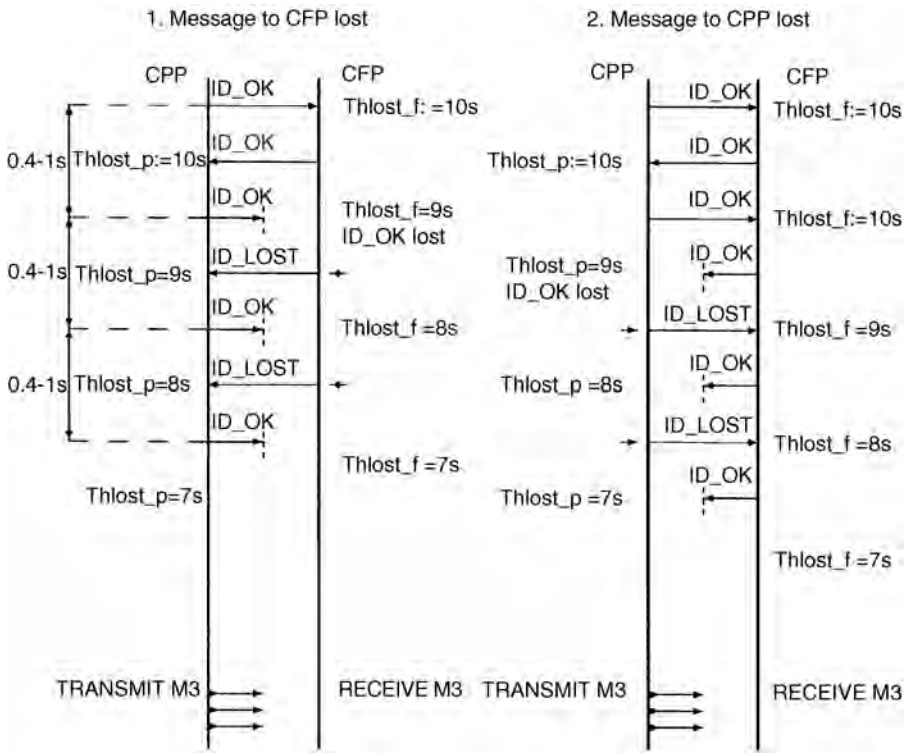


FIGURE 82.10 Handshake loss scenarios.

the CPP, then the CFP will send its ID_Lost HS_f message to the CPP instead of the ID_OK message and will not reset the Thlost_f = 10 s timer. If, however, the CFP happens to detect a valid HS_p, which can be any of the ID_OK, ID_Lost, link_request and link_grant messages of Table 82.2, arriving from the CPP, the CFP will reset its Thrx_f = 1 s timer and resumes transmitting the ID_OK HS_f message instead of the ID_Lost. Should any of the HS messages go astray for more than 3 s, the CPP or the CFP may try and re-establish the link on the current or another RF channel. Again, although any of the ID_OK, ID_Lost, link_request and link_grant represent valid handshakes, only the reception of the ID_OK HS message is allowed to reset the Thlost = 10 s timer at both the CPP and CFP. If this timer expires, the link will be relinquished and the call dropped.

The handshake mechanism is further augmented by referring to Fig. 82.10, where two different scenarios are examplified, portraying the situation when the HS message sent by the CPP to the CFP is lost or, conversely, that transmitted by the CFP is corrupted.

Considering the first scenario, during error-free communications the CPP sends HS_p = ID_OK, and upon receiving it the CFP resets its Thlost_f timer to 10 s. In due course it sends an HS_f = ID_OK acknowledgement, which also arrives free from errors. The CPP resets the Thlost_f timer to 10 s and, after the elapse of the 0.4–1 s handshake interval, issues an HS_p = ID_OK message, which does not reach the CFP. Hence, the Thlost_f timer is now reduced to 9 s and an HS_f = ID_Lost message is sent to the CPP. Upon reception of this, the CPP now cannot reset its Thlost_p timer to 10 s but can respond with an HS_p = ID_OK message, which again goes astray, forcing the CFP to further reduce its Thlost_f timer to 8 s. The CFP issues the valid handshake HS_f = ID_Lost, which arrives at the CPP, where the lack of HS_f = ID_OK reduces Thlost_p to 8 s. Now the corruption of the issued HS_p = ID_OK reduces Thlost_f to 7 s, in which event the link may be reinitialized using the M3 burst. The portrayed second example of Fig. 82.10 can be easily followed in case of the scenario when the HS_f message is corrupted.

82.9 Main Features of the CT-2 System

In our previous discourse we have given an insight in the algorithmic procedures of the CT-2 MPT 1375 recommendation. We have briefly highlighted the four-part structure of the standard dealing with the radio interface, signalling layers 1 and 2, signalling layer 3, and the speech coding issues, respectively. There are forty 100-kHz wide RF channels in the band 864.15–868.15 MHz, and the 72 kb/s bit stream modulates a Gaussian filtered FSK modem. The multiple access technique is TDD, transmitting 2-ms duration, 144-b M1 bursts during normal voice communications, which deliver the 32-kb/s ADPCM-coded speech signal. During link establishment the M2 and M3 bursts are used, which were also portrayed in this treatise, along with a range of handshaking messages and scenarios.

Defining Terms

AFC: Automatic frequency correction.
CAI: Common air interface.
CFP: Cordless fixed part.
CHM: Channel marker sequence.
CHMF: CFP channel marker.
CHMP: CPP channel marker.
CPP: Cordless portable part.
CT: Cordless telephone.
DCA: Dynamic channel allocation.
DCW: Data code word.
DECT: Digital European cordless telecommunications system.
FT: Frame format type bit.
GFSK: Gaussian frequency shift keying.
GP: Guard period.
HIC: Handset identification code.
HS: Handshaking.
ID: Identifier.
L2: Signalling layer 2.
L3: Signalling layer 3.
LAN: Local area network.
LID: Link identification.
LOS: Line of sight.
LS: Link status.
M1: Multiplex one burst format.
M2: Multiplex two burst format.
M3: Multiplex three burst format.
MIC: Manufacturer identification code.
MPT-1375: British CT2 standard.
PCN: Personal communications network .
PIC: Portable identification code.
PLMR: Public land mobile radio.
SNR: Signal-to-noise ratio.
SR: Signalling rate bit.
SYN: Synchronization sequence.
SYNCD: 16-b D-channel frame synchronization word.
TDD: Time division duplex multiple access scheme.
TP: Telepoint.

References

1. Asghar, S., Digital European cordless telephone (DECT), In *The Mobile Communications Handbook,* Chap. 30, CRC Press, Boca Raton, FL, 1995.
2. Gardiner, J.G., Second generation cordless (CT-2) telephony in the UK: telepoint services and the common air-interface, *Elec. & Comm. Eng. J.,* 71–78, Apr. 1990.
3. Hanzo, L., The Pan-European mobile radio system, In *The Mobile Communications Handbook,* Chap. 25, CRC Press, Boca Raton, FL, 1995.
4. Jabbari, B., Dynamic channel assignment, In *The Mobile Communications Handbook,* Chap. 21, CRC Press, Boca Raton, FL, 1995.
5. Ochsner, H., The digital European cordless telecommunications specification DECT. In *Cordless Telecommunication in Europe,* Tuttlebee, W.H.M., Ed., 273–285, Springer-Verlag, Berlin, 1990.
6. Steedman, R.A.J., The Common Air Interface MPT 1375. In *Cordless Telecommunication in Europe,* Tuttlebee, W.H.W. Ed., 261–272, Springer-Verlag, Berlin, 1990.
7. Steele, R. and Hanzo, L., Eds., *Mobile Radio Communications,* IEEE Press and John Wiley, 1999.
8. Tuttlebee, W.H.W., Ed., *Cordless Telecommunication in Europe,* Springer-Verlag, Berlin, 1990.

83

Half-Rate Standards

Wai-Yip Chan
Illinois Institute of Technology

Ira Gerson
Auvo Technologies, Inc.

Toshio Miki
NTT Mobile Communications Network, Inc.

83.1 Introduction

A half-rate speech coding standard specifies a procedure for digital transmission of speech signals in a digital cellular radio system. The speech processing functions that are specified by a half-rate standard are depicted in Fig. 83.1. An input speech signal is processed by a *speech encoder* to generate a digital representation at a *net bit rate* of R_s bits per second. The encoded bit stream representing the input speech signal is processed by a *channel encoder* to generate another bit stream at a *gross bit rate* of R_c bits per second, where $R_c > R_s$. The channel encoded bit stream is organized into data frames, and each frame is transmitted as payload data by a radio-link access controller and modulator. The net bit rate R_s counts the number of bits used to describe the speech signal, and the difference between the gross and net bit rates (R_c - R_s) counts the number of error protection bits needed by the *channel decoder* to correct and detect transmission errors. The output of the channel decoder is given to the *speech decoder* to generate a *quantized* version of the speech encoder's input signal. In current digital cellular radio systems that use time-division multiple access (TDMA), a voice connection is allocated a fixed transmission rate (i.e., R_c is a constant). The operations performed by the speech and channel encoders and decoders and their input and output data formats are governed by the half-rate standards.

Globally, three major TDMA cellular radio systems have been developed and deployed. The initial digital speech services offered by these cellular systems were governed by *full-rate standards*. Because of the rapid growth in demand for cellular services, the available transmission capacity in some areas is frequently saturated, eroding customer satisfaction. By providing essentially the same voice quality but at half the gross bit rates of the full-rate standards, half-rate standards can readily double the number of callers that can be serviced by the cellular systems. The gross bit rates of the full-rate and half-rate standards for the European Groupe Speciale Mobile (GSM), Japanese Personal Digital Cellular[1] (PDC), and North American cellular (IS-54) systems are listed in Table 83.1. The three systems were developed

[1]Personal Digital Cellular was formerly Japanese Digital Cellular (JDC).

TABLE 83.1 Gross Bit Rates Used for Digital Speech Transmission in Three TDMA Cellular Radio Systems

Standard Organization and Digital Cellular System	Gross Bit Rate, b/s	
	Full Rate	Half Rate
European Telecommunications Standards Institute (ETSI), GSM	22,800	11,400
Research & Development Center for Radio Systems (RCR), PDC	11,200	5,600
Telecommunication Industries Association (TIA), IS-54	13,000	6,500

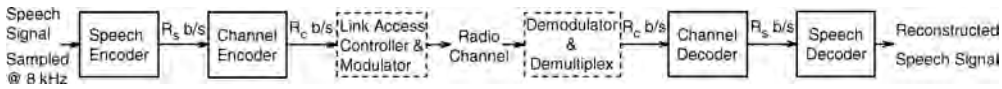


FIGURE 83.1 Digital speech transmission for digital cellular radio. Boxes with solid outlines represent processing modules that are specified by the half-rate standards.

and deployed under different time tables. Their disparate full- and half-bit rates partly reflect this difference. At the time of writing (January, 1995), the European and the Japanese systems have each selected an algorithm for their respective half-rate **codec**. Standardization of the North American half-rate codec has not reached a conclusion as none of the candidate algorithms has fully satisfied the standard's requirements. Thus, we focus here on the Japanese and European half-rate standards and will only touch upon the requirements of the North American standard.

83.2 Speech Coding for Cellular Mobile Radio Communications

Unlike the relatively benign transmission media commonly used in the public-switched telephone network (PSTN) for analog and digital transmission of speech signals, mobile radio channels are impaired by various forms of fading and interference effects. Whereas proper engineering of the radio link elements (modulation, power control, diversity, equalization, frequency allocation, etc.) ameliorates fading effects, burst and isolated bit errors still occur frequently. The net effect is such that speech communication may be required to be operational even for bit-error rates greater than 1%. In order to furnish reliable voice communication, typically half of the transmitted payload bits are devoted to error correction and detection.

It is common for low-bit-rate speech codecs to process samples of the input speech signal one frame at a time, e.g., 160 samples processed once every 20 ms. Thus, a certain amount of time is required to gather a block of speech samples, encode them, perform channel encoding, transport the encoded data over the radio channel, and perform channel decoding and speech synthesis. These processing steps of the speech codec add to the overall end-to-end transmission delay. Long transmission delay hampers conversational interaction. Moreover, if the cellular system is interconnected with the PSTN and a four-wire to two-wire (analog) circuit conversion is performed in the network, feedbacks called *echoes* may be generated across the conversion circuit. The echoes can be heard by the originating talker as a delayed and distorted version of his/her speech and can be quite annoying. The annoyance level increases with the transmission delay and may necessitate (at additional costs) the deployment of **echo cancellers**.

A consequence of user mobility is that the level and other characteristics of the acoustic background noise can be highly variable. Though acoustic noise can be minimized through suitable acoustic transduction design and the use of adaptive filtering/cancellation techniques [9,13,15], the speech encoding algorithm still needs to be robust against background noise of various levels and kinds (e.g., babble, music, noise bursts, and colored noise).

Processing complexity directly impacts the viability of achieving a circuit realization that is compact and has low-power consumption, two key enabling factors of equipment portability for the end user.

Factors that tend to result in low complexity are fixed-point instead of floating-point computation, lack of complicated arithmetic operations (division, square roots, transcendental functions), regular algorithm structure, small data memory, and small program memory. Since, in general, better speech quality can be achieved with increasing speech and channel coding delay and complexity, the digital cellular mobile-radio environment imposes conflicting and challenging requirements on the speech codec.

83.3 Codec Selection and Performance Requirements

The half-rate speech coding standards are drawn up through competitive testing and selection. From a set of candidate codec algorithms submitted by contending organizations, the one algorithm that meets basic selection criteria and offers the best performance is selected to form the standard. The codec performance measures and codec testing and selection procedures are set out in a test plan under the auspices of the organization (Table 83.1) responsible for the standardization process (see, e.g., [16]). Major codec characteristics evaluated are speech quality, delay, and complexity. The full-rate codec is also evaluated as a *reference codec*, and its evaluation scores form part of the selection criteria for the codec candidates.

The speech quality of each candidate codec is evaluated through listening tests. To conduct the tests, each candidate codec is required to process speech signals and/or encoded bit streams that have been preprocessed to simulate a range of operating conditions: variations in speaker voice and level, acoustic background noise type and level, channel error rate, and stages of **tandem coding**. During the tests, subjects listen to processed speech signals and judge their quality levels or annoyance levels on a five-point opinion scale. The opinion scores collected from the tests are suitably averaged over all trials and subjects for each test condition (see [11], for mean opinion score (MOS) and degradation mean opinion score). The categorical opinion scales of the subjects are also calibrated using *modulated noise reference units* (*MNRUs*) [3]. Modulated noise better resembles the distortions created by speech codecs than noise that is uncorrelated with the speech signal. Modulated noise is generated by multiplying the speech signal with a noise signal. The resultant modulated noise is scaled to a desired power level and then added to the uncoded (clean) speech signal. The ratio between the power level of the speech signal and that of the modulated noise is expressed in decibels and given the notation *dBQ*. Under each test condition, subjects are presented with speech signals processed by the codecs as well as speech signals corrupted by modulated noise. Through presenting a range of modulated-noise levels, the subjects' opinions are calibrated on the dBQ scale. Thereafter, the mean opinion scores obtained for the codecs can also be expressed on that scale.

For each codec candidate, a profile of scores is compiled, consisting of speech quality scores, delay measurements, and complexity estimates. Each candidate's score profile is compared with that of the reference codec, ensuring that basic requirements are satisfied [see, e.g., [12]). An overall figure of merit for each candidate is also computed from the profile. The candidates, if any, that meet the basic requirements then compete on the basis of maximizing the figure of merit.

Basic performance requirements for each of the three half-rate standards are summarized in Table 83.2. In terms of speech quality, the GSM and PDC half-rate codecs are permitted to underperform their respective full-rate codecs by no more than 1 dBQ averaging over all test conditions and no more than

TABLE 83.2 Basic Performance Requirements for the Three Half-Rate Standards

	Basic Performance Requirements		
Digital Cellular Systems	Min. Speech Quality, dBQ Rel. to Full Rate	Max. Delay, ms	Max. Complexity Rel. to Full Rate
Japanese (PDC)	-1 average, -3 maximum	94.8	3×
European (GSM)	-1 average, -3 maximum	90	4×
North American (IS-54)	Statistically equivalent	100	4×

3 dBQ within each test condition. More stringently, the North American half-rate codec is required to furnish a speech-quality profile that is statistically equivalent to that of the North American full-rate codec as determined by a specific statistical procedure for multiple comparisons [16]. Since various requirements on the half-rate standards are set relative to their full-rate counterparts, an indication of the *relative* speech quality between the three half-rate standards can be deduced the test results of De Martino [2] comparing the three full-rate codecs. The maximum delays in Table 83.2 apply to the total of the delays through the speech and channel encoders and decoders (Fig. 83.1). Codec complexity is computed using a formula that counts the computational operations and memory usage of the codec algorithm. The complexity of the half-rate codecs is limited to 3 or 4 times that of their full-rate counterparts.

83.4 Speech Coding Techniques in the Half-Rate Standards

Existing half-rate and full-rate standard coders can be characterized as *linear-prediction based analysis-by-synthesis* (LPAS) speech coders [4]. LPAS coding entails using a time-varying all-pole filter in the decoder to synthesize the quantized speech signal. A short segment of the signal is synthesized by driving the filter with an *excitation* signal that is either *quasiperiodic* (for *voiced* speech) or *random* (for *unvoiced* speech). In either case, the excitation signal has a *spectral envelope* that is relatively flat. The synthesis filter serves to shape the spectrum of the excitation input so that the spectral envelope of the synthesized output resembles the filter's magnitude frequency response. The magnitude response often has prominent peaks; they render the *formants* that give a speech signal its phonetic character. The synthesis filter has to be adapted to the current frame of input speech signal. This is accomplished with the encoder performing a linear prediction (LP) analysis of the frame: the inverse of the all-pole synthesis filter is applied as an LP *error filter* to the frame, and the values of the filter parameters are computed to minimize the energy of the filter's output error signal. The resultant filter parameters are quantized and conveyed to the decoder for it to update the synthesis filter.

Having executed an LP analysis and quantized the synthesis filter parameters, the LPAS encoder performs analysis-by-synthesis (ABS) on the input signal to find a suitable excitation signal. An ABS encoder maintains a *copy* of the decoder. The encoder examines the possible outputs that can be produced by the decoder copy in order to determine how best to instruct (using transmitted information) the actual decoder so that it would output (synthesize) a good approximation of the input speech signal. The decoder copy tracks the state of the actual decoder, since the latter evolves (under ideal channel conditions) according to information received from the encoder. The details of the ABS procedure vary with the particular excitation model employed in a specific coding scheme. One of the earliest seminal LPAS schemes is *code excited linear prediction* (*CELP*) [4]. In CELP, the excitation signal is obtained from a **codebook** of *code vectors*, each of which is a candidate for the excitation signal. The encoder searches the codebook to find the one code vector that would result in a best match between the resultant synthesis output signal and the encoder's input speech signal. The matching is considered best when the energy of the difference between the two signals being matched is minimized. A *perceptual weighting filter* is usually applied to the difference signal (prior to energy integration) to make the minimization more relevant to human perception of speech fidelity. Regions in the frequency spectrum where human listeners are more sensitive to distortions are given relatively stronger weighting by the filter and vice versa. For instance, the concentration of spectral energy around the formant frequencies gives rise to stronger *masking* of coder noise (i.e., rendering the noise less audible) and, therefore, weaker weighting can be applied to the formant frequency regions. For masking to be effective, the weighting filter has to be adapted to the time-varying speech spectrum. Adaptation is achieved usually by basing the weighting filter parameters on the synthesis filter parameters.

The CELP framework has evolved to form the basis of a great variety of speech coding algorithms, including all existing full- and half-rate standard algorithms for digital cellular systems. We outline next the basic CELP encoder-processing steps, in a form suited to our subsequent detailed descriptions of the PDC and GSM half-rate coders. These steps have accounted for various computational efficiency

considerations and may, therefore, deviate from a conceptual functional description of the encoder constituents.

1. LP analysis on the current frame of input speech to determine the coefficients of the all-pole synthesis filter;
2. quantization of the LP filter parameters;
3. determination of the open-loop **pitch period** or lag;
4. adapting the perceptual weighting filter to the current LP information (and also pitch information when appropriate) and applying the adapted filter to the input speech signal;
5. formation of a filter cascade (which we shall refer to as *perceptually weighted synthesis filter*) consisting of the LP synthesis filter, as specified by the quantized parameters in step 2, followed by the perceptual weighting filter;
6. subtraction of the *zero-input response* of the perceptually weighted synthesis filter (the filter's decaying response due to past input) from the perceptually weighted input speech signal obtained in step 4;
7. an *adaptive codebook* is searched to find the most suitable periodic excitation, i.e., when the perceptually weighted synthesis filter is driven by the best code vector from the adaptive codebook, the output of the filter cascade should best match the difference signal obtained in step 6;
8. one or more nonadaptive excitation codebooks are searched to find the most suitable random excitation vectors that, when added to the best periodic excitation as determined in step 7 and with the resultant sum signal driving the filter cascade, would result in an output signal best matching the difference signal obtained in step 6.

Steps 1–6 are executed once per frame. Steps 7 and 8 are executed once for each of the *subframes* that together constitute a frame. Step 7 may be skipped depending on the pitch information from step 3, or if step 7 were always executed, a *nonperiodic excitation* decision would be one of the possible outcomes of the search process in step 7. Integral to steps 7 and 8 is the determination of gain (scaling) parameters for the excitation vectors. For each frame of input speech, the filter and excitation and gain parameters determined as outlined are conveyed as encoded bits to the speech decoder.

In a properly designed system, the data conveyed by the channel decoder to the speech decoder should be free of errors most of the time, and the speech signal synthesized by the speech decoder would be identical to that as determined in the speech encoder's ABS operation. It is common to enhance the quality of the synthesized speech by using an adaptive *postfilter* to attenuate coder noise in the perceptually sensitive regions of the spectrum. The postfilter of the decoder and the perceptual weighting filter of the encoder may seem to be functionally identical. The weighting filter, however, influences the selection of the best excitation among available choices, whereas the postfilter actually shapes the spectrum of the synthesized signal. Since postfiltering introduces its own distortion, its advantage may be diminished if tandem coding occurs along the end-to-end communication path. Nevertheless, proper design can ensure that the net effect of postfiltering is a reduction in the amount of audible codec noise [1]. Excepting postfiltering, all other speech synthesis operations of an LPAS decoder are (effectively) duplicated in the encoder (though the converse is not true). Using this fact, we shall illustrate each coder in the sequel by exhibiting only a block diagram of its encoder or decoder but not both.

83.5 Channel Coding Techniques in the Half-Rate Standards

Crucial to the maintenance of quality speech communication is the ability to transport coded speech data across the radio channel with minimal errors. Low-bit-rate LPAS coders are particularly sensitive to channel errors; errors in the bits representing the LP parameters in one frame, for instance, could result in the synthesis of nonsensical sounds for longer than a frame duration. The error rate of a digital cellular radio channel with no channel coding can be catastrophically high for LPAS coders. The amount of tolerable transmission delay is limited by the requirement of interactive communication and, consequently, *forward error control* is used to remedy transmission errors. "Forward" means that channel errors are remedied in

the receiver, with no additional information from the transmitter and, hence, no additional transmission delay. To enable the channel decoder to correct channel errors, the channel encoder conveys more bits than the amount generated by the speech encoder. The additional bits are for error *protection*, as errors may or may not occur in any particular transmission epoch. The ratio of the number of encoder input (information) bits to the number of encoder output (code) bits is called the (channel) *coding rate*. This is a number no more than one and generally decreases as the error protection power increases. Though a lower channel coding rate gives more error protection, fewer bits will be available for speech coding. When the channel is in good condition and, hence, less error protection is needed, the received speech quality could be better if bits devoted to channel coding were used for speech coding. On the other hand, if a high channel coding rate were used, there would be uncorrected errors under poor channel conditions and speech quality would suffer. Thus, when nonadaptive forward error protection is used over channels with nonstationary statistics, there is an inevitable tradeoff between quality degradation due to uncorrected errors and that due to expending bits on error protection (instead of on speech encoding).

Both the GSM and PDC half-rate coders use *convolutional coding* [14] for error correction. Convolutional codes are sliding or sequential codes. The encoder of a rate m/n, $m < n$ convolutional code can be realized using m shift registers. For every m information bits input to the encoder (one bit to each of the m shift registers), n code bits are output to the channel. Each code bit is computed as a modulo-2 sum of a subset of the bits in the shift registers. Error protection overhead can be reduced by exploiting the unequal sensitivity of speech quality to errors in different positions of the encoded bit stream. A family of *rate-compatible punctured convolutional codes* (RCPCCs) [10] is a collection of related convolutional codes; all of the codes in the collection except the one with the lowest rate are derived by *puncturing* (dropping) code bits from the convolutional code with the lowest rate. With an RCPCC, the channel coding rate can be varied on the fly (i.e., variable-rate coding) while a sequence of information bits is being encoded through the shift registers, thereby imparting on different segments in the sequence different degrees of error protection.

For decoding a convolutional coded bit stream, the *Viterbi algorithm* [14] is a computationally efficient procedure. Given the output of the demodulator, the algorithm determines the most likely sequence of data bits sent by the channel encoder. To fully utilize the error correction power of the convolutional code, the amplitude of the demodulated *channel symbol* can be quantized to more bits than the minimum number required, i.e., for subsequent *soft decision decoding*. The minimum number of bits is given by the number of channel-coded bits mapped by the modulator onto each channel symbol; decoding based on the minimum-rate bit stream is called *hard decision* decoding. Although soft decoding gives better error protection, decoding complexity is also increased.

Whereas convolutional codes are most effective against randomly scattered bit errors, errors on cellular radio channels often occur in bursts of bits. These bursts can be broken up if the bits put into the channel are rearranged after demodulation. Thus, in *block interleaving*, encoded bits are read into a matrix by row and then read out of the matrix by column (or vice versa) and then passed on to the modulator; the reverse operation is performed by a *deinterleaver* following demodulation. Interleaving increases the transmission delay to the extent that enough bits need to be collected in order to fill up the matrix.

Owing to the severe nature of the cellular radio channel and limited available transmission capacity, uncorrected errors often remain in the decoded data. A common countermeasure is to append an error detection code to the speech data stream prior to channel coding. When residual channel errors are detected, the speech decoder can take various remedial measures to minimize the negative impact on speech quality. Common measures are repetition of speech parameters from the most recent good frames and gradual muting of the possibly corrupted synthesized speech.

The PDC and GSM half-rate standard algorithms together embody some of the latest advances in speech coding techniques, including: *multimodal coding* where the coder configuration and bit allocation change with the type of speech input; *vector quantization* (VQ) [5] of the LP filter parameters; higher precision and improved coding efficiency for pitch-periodicg excitation; and postfiltering with improved tandeming performance. We next explore the more distinctive features of the PDC and GSM speech coders.

TABLE 83.3 Bit Allocations for the PSI-CELP Half-Rate PDC Speech Coder

Parameter	Bits	Error Protected Bits
LP synthesis filter	31	15
Frame energy	7	7
Periodic excitation	8 × 4	8 × 4
Stochastic excitation	10 × 4	0
Gain	7 × 4	3 × 4
Total	138	66

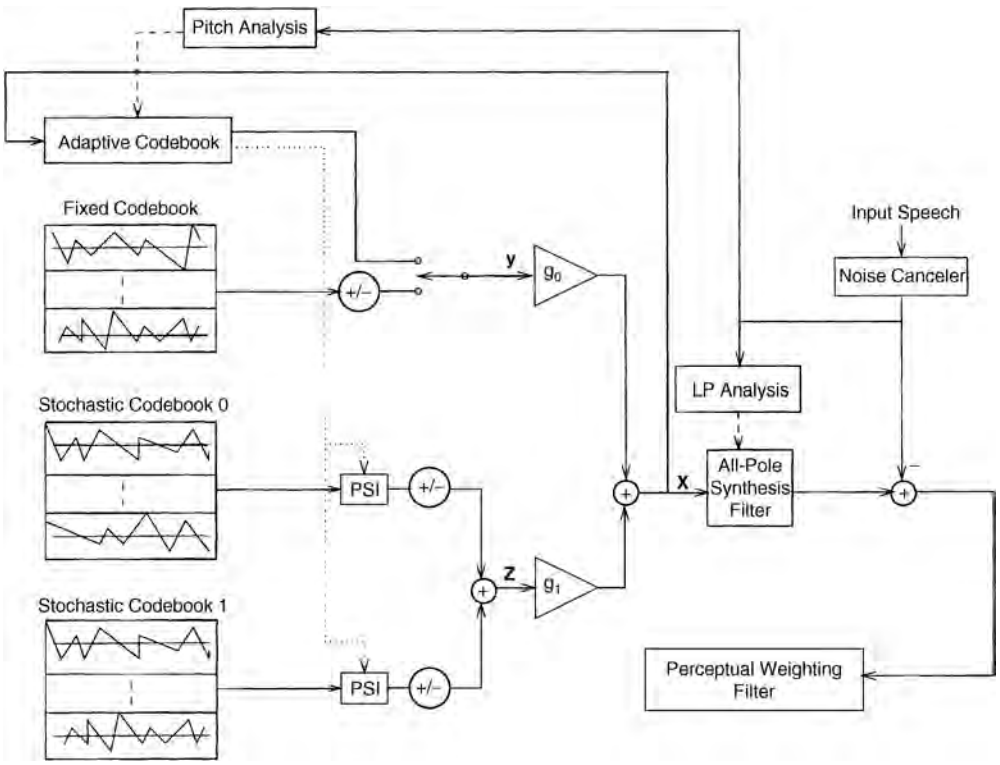


FIGURE 83.2 Basic structure of the PSI-CELP encoder.

83.6 The Japanese Half-Rate Standard

An algorithm was selected for the Japanese half-rate standard in April 1993, following the evaluation of 12 submissions in a first round, and four final candidates in a second round [12]. The selected algorithm, called pitch synchronous innovation CELP[2] (PSI-CELP), met all of the basic selection criteria and scored the highest among all candidates evaluated. A block diagram of the PSI-CELP encoder is shown in Fig. 83.2 and bit allocations are summarized in Table 83.3. The complexity of the coder is estimated to be approximately 2.4 times that of the PDC full-rate coder. The frame size of the coder is 40 ms, and its subframe size is 10 ms. These sizes are longer than those used in most existing CELP-type standard coders. However, LP analysis is performed twice per frame in the PSI-CELP coder.

[2]There were two candidate algorithms named PSI-CELP in the PDC half-rate competition. The algorithm described here was contributed by NTT Mobile Communications Network, Inc. (NTT DoCoMo).

A distinctive feature of the PSI-CELP coder is the use of an adaptive noise canceller [13,15] to suppress noise in the input signal prior to coding. The input signal is classified into various modes, depending on the presence or absence of background noise and speech and their relative power levels. The current active mode determines whether *Kalman filtering* [9] is applied to the input signal and whether the parameters of the Kalman filter are adapted. Kalman filtering is applied when a significant amount of background noise is present or when both background noise and speech are strongly present. The filter parameters are adapted to the statistics of the speech and noise signals in accordance with whether they are both present or only noise is present.

The LP filter parameters in the PSI-CELP coder are encoded using VQ. A tenth-order LP analysis is performed every 20 ms. The resultant filter parameters are converted to 10 *line spectral frequencies* (LSFs).[3] The LSF parameters have a naturally increasing order, and together are treated as the ordered components of a vector. Since the speech spectral envelope tends to evolve slowly with time, there is intervector dependency between adjacent LSF vectors that can be exploited. Thus, the two LSF vectors for each 40-ms frame are paired together and jointly encoded. Each LSF vector in the pair is split into three subvectors. The pair of subvectors that cover the same vector component indexes are combined into one composite vector and vector quantized. Altogether, 31 b are used to encode a pair of LSF vectors. This three-way *split VQ*[4] scheme embodies a compromise between the prohibitively high complexity of using a large vector dimension and the performance gain from exploiting intra- and intervector dependency.

The PSI-CELP encoder uses a perceptual weighting filter consisting of a cascade of two filter sections. The sections exploit the pitch-harmonic structure and the LP spectral-envelope structure of the speech signal, respectively. The pitch-harmonic section has four parameters, a pitch lag and three coefficients, whose values are determined from an analysis of the periodic structure of the input speech signal. Pitch-harmonic weighting reduces the amount of noise in between the pitch harmonics by aggregating coder noise to be closer to the harmonic frequencies of the speech signal. In high-pitched voice, the harmonics are spaced relatively farther apart, and pitch-harmonic weighting becomes correspondingly more important.

The excitation vector $\boldsymbol{x}$ (Fig. 83.2) is updated once every subframe interval (10 ms) and is constructed as a *linear combination* of two vectors

$$\boldsymbol{x} = g_0\boldsymbol{y} + g_1\boldsymbol{z} \tag{83.1}$$

where g_0 and g_1 are scalar gains, $\boldsymbol{y}$ is labeled as the *periodic* component of the excitation and $\boldsymbol{z}$ as the *stochastic* or *random* component. When the input speech is voiced, the ABS operation would find a value for $\boldsymbol{y}$ from the *adaptive codebook* (Fig. 83.2). The codebook is constructed out of past samples of the excitation signal $\boldsymbol{x}$; hence, there is a feedback path into the adaptive codebook in Fig. 83.2. Each code vector in the adaptive codebook corresponds to one of the 192 possible pitch lag L values available for encoding; the code vector is populated with samples of $\boldsymbol{x}$ beginning with the Lth sample backward in time. L is not restricted to be an integer, i.e., *fractional pitch period* is permitted. Successive values of L are more closely spaced for smaller values of L; short, medium, and long lags are quantized to one-quarter, one-half, and one sampling-period resolution, respectively. As a result, the *relative* quantization error in the encoded pitch frequency (which is the reciprocal of the encoded pitch lag) remains roughly constant with increasing pitch frequency. When the input speech is unvoiced, $\boldsymbol{y}$ would be obtained from the fixed codebook (Fig. 83.2). To find the best value for $\boldsymbol{y}$, the encoder searches through the aggregate of 256 code vectors from both the adaptive and fixed codebooks. The code vector that results in a synthesis output most resembling the input speech is selected. The best code vector thus chosen also implicitly determines the voicing condition (voiced/unvoiced) and the pitch lag value L^* most appropriate to the current subframe of input speech. These parameters are said to be determined in a *closed-loop* search.

The stochastic excitation $\boldsymbol{z}$ is formed as a sum of two code vectors, each selected from a *conjugate codebook* (Fig. 83.2) [13]. Using a pair of conjugate codebooks each of size 16 code vectors (4 b) has been

[3]Also known as line spectrum pairs (LSPs).

[4]Matrix quantization is another possible description.

found to improve robustness against channel errors, in comparison with using one single codebook of size 256 code vectors (8 b). The synthesis output due to $\boldsymbol{z}$ can be decomposed into a sum of two orthogonal components, one of which points in the same direction as the synthesis output due to the periodic excitation $\boldsymbol{y}$ and the other component points in a direction orthogonal to the synthesis output due to $\boldsymbol{y}$. The latter synthesis output component of $\boldsymbol{z}$ is kept, whereas the former component is discarded. Such decomposition enables the two gain factors g_0 and g_1 to be separately quantized. For voiced speech, the conjugate code vectors are preprocessed to produce a set of *pitch synchronous innovation* (PSI) vectors. The first L^* samples of each code vector are treated as a fundamental period of samples. The fundamental period is replicated until there are enough samples to populate a subframe. If L^* is not an integer, interpolated samples of the code vectors are used (upsampled versions of the code vectors can be precomputed). PSI has been found to reinforce the periodicity and substantially improve the quality of synthesized voiced speech.

The postfilter in the PSI-CELP decoder has three sections, for enhancing the formants, the pitch harmonics, and the high frequencies of the synthesized speech, respectively. Pitch-harmonic enhancement is applied only when the adaptive codebook has been used. Formant enhancement makes use of the decoded LP synthesis filter parameters, whereas a refined pitch analysis is performed on the synthesized speech to obtain the values for the parameters of the pitch-harmonic section of the postfilter. A first-order high-pass filter section compensates for the low-pass spectral tilt [1] of the formant enhancement section.

Of the 138 speech data bits generated by the speech encoder every 40-ms frame, 66 b (Table 83.3) receive error protection and the remaining 72 speech data bits of the frame are not error protected. An error detection code of 9 *cyclic redundancy check* (CRC) bits is appended to the 66 b and then submitted to a rate 1/2, punctured convolutional encoder to generate a sequence of 152 channel coded bits. Of the unprotected 72 b, the 40 b that index the excitation codebooks (Table 83.3) are remapped or *pseudo-Gray coded* [17] so as to equalize their channel error sensitivity. As a result, a bit error occurring in an index word is likely to cause about the same amount of degradation regardless of the bit error position in the index word. For each speech frame, the channel encoder emits 224 b of payload data. The payload data from two adjacent frames are interleaved before transmission over the radio link.

Uncorrected errors in the most critical 66 b are detected with high probability as a CRC error. A finite state machine keeps track of the recent history of CRC errors. When a sequence of CRC errors is encountered, the power level of the synthesized speech is progressively suppressed, so that muting is reached after four consecutive CRC errors. Conversely, following the cessation of a sequence of CRC errors, the power level of the synthesized speech is ramped up gradually.

83.7 The European GSM Half-Rate Standard

A *vector sum excited linear prediction* (VSELP) coder, contributed by Motorola, Inc., was selected in January 1994 by the main GSM technical committee as a basis for the GSM half-rate standard. The standard was finally approved in January 1995. VSELP is a generic name for a family of algorithms from Motorola; the North American full-rate and the Japanese full-rate standards are also based on VSELP. All VSELP coders make use of the basic idea of representing the excitation signal by a linear combination of *basis vectors* [6]. This representation renders the excitation codebook search procedure very computationally efficient. A block diagram of the GSM half-rate decoder is depicted in Fig. 83.3 and bit allocations are tabulated in Table 83.4. The coder's frame size is 20 ms, and each frame comprises four subframes of 5 ms each. The coder has been optimized for execution on a processor with 16-b word length and 32-b accumulator. The GSM standard is a *bit exact* specification: in addition to specifying the codec's processing steps, the numerical formats and precisions of the codec's variables are also specified.

The synthesis filter coefficients in GSM VSELP are encoded using the *fixed point lattice technique* (FLAT)[8] and vector quantization. FLAT is based on the *lattice filter* representation of the linear prediction error filter. The tenth-order lattice filter has 10 stages, with the ith stage, $i \in \{1, \ldots, 10\}$, containing a *reflection coefficient parameter* r_i. The lattice filter has an *order-recursion* property such that the

TABLE 83.4 Bit Allocations for the VSELP Half-Rate GSM Coder

Parameter	Bits/Subframe	Bits/Frame
LP synthesis filter		28
Soft interpolation		1
Frame energy		5
Mode selection		2
Mode 0		
Excitation code I	7	28
Excitation code H	7	28
Gain code G_s, P_0	5	20
Mode 1, 2, and 3		
Pitch lag L (first subframe)		8
Difference lag (subframes 2, 3, 4)	4	12
Excitation code J	9	36
Gain code G_s, P_0	5	20
Total		112

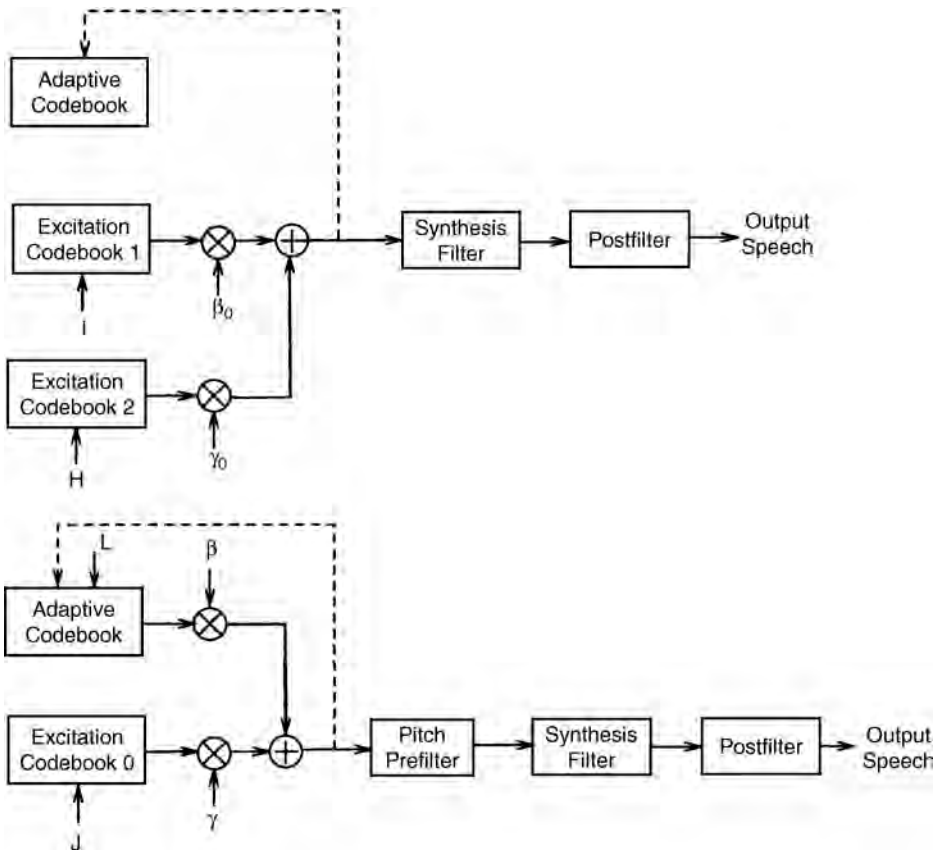


FIGURE 83.3 Basic structure of the GSM VSELP decoder. Top is for mode 0 and bottom is for modes 1, 2, and 3.

best prediction error filters of all orders less than ten are all embedded in the best tenth-order lattice filter. This means that once the values of the lower order reflection coefficients have been optimized, they do not have to be reoptimized when a higher order predictor is desired; in other words, the coefficients can be optimized sequentially from low to high orders. On the other hand, if the lower order coefficients were suboptimal (as in the case when the coefficients are quantized), the higher order coefficients could

still be selected to minimize the prediction *residual* (or error) energy at the output of the higher order stages; in effect, the higher order stages can compensate for the suboptimality of lower order stages.

In the GSM VSELP coder, the ten reflection coefficients $\{r_1, \ldots, r_{10}\}$ that have to be encoded for each frame are grouped into three coefficient vectors $\boldsymbol{v}_1 = [r_1 r_2 r_3]$, $\boldsymbol{v}_2 = [r_4 r_5 r_6]$, $\boldsymbol{v}_3 = [r_7 r_8 r_9 r_{10}]$. The vectors are quantized sequentially, from $\boldsymbol{v}_1$ to $\boldsymbol{v}_3$, using a b_i-bit VQ codebook C_i for $\boldsymbol{v}_i$, where b_i, $i = 1, 2, 3$ are 11, 9, and 8 b, respectively. The vector $\boldsymbol{v}_i$ is quantized to minimize the prediction error at the energy output of the jth stage of the lattice filter where r_j is the highest order coefficient in the vector $\boldsymbol{v}_i$. The computational complexity associated with quantizing $\boldsymbol{v}_i$ is reduced by searching only a small subset of the code vectors in C_i. The subset is determined by first searching a *prequantizer* codebook of size c_i bits, where c_i, $i = 1, \ldots, 3$ are 6, 5, and 4 b, respectively. Each code vector in the prequantizer codebook is associated with $2^{b_i - c_i}$ code vectors in the target codebook. The subset is obtained by pooling together all of the code vectors in C_i that are associated with the top few best matching prequantizer code vectors. In this way, a factor of reduction in computational complexity of nearly $2^{b_i - c_i}$ is obtained for the quantization of $\boldsymbol{v}_i$.

The half-rate GSM coder changes its configuration of excitation generation (Fig. 83.3) in accordance with a *voicing mode* [7]. For each frame, the coder selects one of four possible voicing modes depending on the values of the *open-loop* pitch-prediction gains computed for the frame and its four subframes. Open loop refers to determining the pitch lag and the pitch-predictor coefficient(s) via a direct analysis of the input speech signal or, in the case of the half-rate GSM coder, the perceptually weighted (LP-weighting only) input signal. Open-loop analysis can be regarded as the opposite of closed-loop analysis, which in our context is synonymous with ABS. When the pitch-prediction gain for the frame is weak, the input speech signal is deemed to be unvoiced and mode 0 is used. In this mode, two 7-b *trained* codebooks (excitation codebooks 1 and 2 in Fig. 83.3) are used, and the excitation signal for each subframe is formed as a linear combination of two code vectors, one from each of the codebooks. A trained codebook is one designed by applying the coder to a representative set of speech signals while optimizing the codebook to suit the set. Mode 1, 2, or 3 is chosen depending on the strength of the pitch-prediction gains for the frame and its subframes. In these modes, the excitation signal is formed as a linear combination of a code vector from an 8-b adaptive codebook and a code vector from a 9-b trained codebook (Fig. 83.3). The code vectors that are summed together to form the excitation signal for a subframe are each scaled by a gain factor (β and γ in Fig. 83.3). Each mode uses a gain VQ codebook specific to that mode.

As depicted in Fig. 83.3, the decoder contains an adaptive pitch prefilter for the voiced modes and an adaptive postfilter for all modes. The filters enhance the perceptual quality of the decoded speech and are not present in the encoder. It is more conventional to locate the pitch prefilter as a section of the postfilter; the distinctive placement of the pitch prefilter in VSELP was chosen to reduce artifacts caused by the time-varying nature of the filter. In mode 0, the encoder uses an LP spectral weighting filter in its ABS search of the two excitation codebooks. In the other modes, the encoder uses a pitch-harmonic weighting filter in cascade with an LP spectral weighting filter for searching excitation codebook 0, whereas only LP spectral weighting is used for searching the adaptive codebook. The pitch-harmonic weighting filter has two parameters, a pitch lag and a coefficient, whose values are determined in the aforementioned open-loop pitch analysis.

A code vector in the 8-b adaptive codebook has a dimension of 40 (the duration of a subframe) and is populated with past samples of the excitation signal beginning with the Lth sample back from the present time. L can take on one of 256 different integer and fractional values. The best adaptive code vector for each subframe can be selected via a complete ABS; the required exhaustive search of the adaptive codebook is, however, computationally expensive. To reduce computation, the GSM VSELP coder makes use of the aforementioned open-loop pitch analysis to produce a list of *candidate lag values*. The open-loop pitch-prediction gains are ranked in decreasing order, and only the lags corresponding to top-ranked gains are kept as candidates. The final decisions for the four L values of the four subframes in a frame are made jointly. By assuming that the four L values can not vary over the entire range of all possible 256 values in the short duration of a frame, the L of the first subframe is coded using 8 b, and the L of each of the other three subframes is coded *differentially* using 4 b. The 4 b represent 16 possible values of deviation relative to the lag of the previous subframe. The four lags in a frame trace out a *trajectory* where the change

from one time point to the next is restricted; consequently, only 20 b are needed instead of 32 b for encoding the four lags. Candidate trajectories are constructed by linking top ranked lags that are commensurate with differential encoding. The best trajectory among the candidates is then selected via ABS.

The trained excitation codebooks of VSELP have a special vector sum structure that facilitates fast searching [6]. Each of the 2^b code vectors in a b-bit trained codebook is formed as a linear combination of b *basis vectors*. Each of the b scalar weights in the linear combination is restricted to have a binary value of either 1 or -1. The 2^b code vectors in the codebook are obtained by taking all 2^b possible combinations of values of the weights. A substantial storage saving is incurred by storing only b basis vectors instead of 2^b code vectors. Computational saving is another advantage of the vector-sum structure. Since filtering is a linear operation, the synthesis output due to each code vector is a linear combination of the synthesis outputs due to the individual basis vectors, where the same weight values are used in the output linear combination as in forming the code vector. A vector sum codebook can be searched by first performing synthesis filtering on its b basis vectors. If, for the present subframe, another trained codebook (mode 0) or an adaptive codebook (mode 1, 2, 3) had been searched, the filtered basis vectors are further orthogonalized with respect to the signal synthesized from that codebook, i.e., each filtered basis vector is replaced by its own component that is orthogonal to the synthesized signal. Further complexity reduction is obtained by examining the code vectors in a sequence such that two successive code vectors differ in only one of the b scalar weight values; that is, the entire set of 2^b code vectors is searched in a *Gray coded* sequence. With successive code vectors differing in only one term in the linear combination, it is only necessary in the codebook search computation to progressively track the difference [6].

The total energy of a speech frame is encoded with 5 b (Table 83.4). The two gain factors (β and γ in Fig. 83.3) for each subframe are computed after the excitation codebooks have been searched and are then transformed to parameters G_s and P_0 to be vector quantized. Each mode has its own 5-b gain VQ codebook. G_s represents the energy of the subframe relative to the total frame energy, and P_0 represents the fraction of the subframe energy due to the first excitation source (excitation codebook 1 in mode 0, or the adaptive codebook in the other modes).

An *interpolation bit* (Table 83.4) transmitted for each frame specifies to the decoder whether the LP synthesis filter parameters for each subframe should be obtained from interpolating between the decoded filter parameters for the current and the previous frames. The encoder determines the value of this bit according to whether interpolation or no interpolation results in a lower prediction residual energy for the frame. The postfilter in the decoder operates in concordance with the actual LP parameters used for synthesis.

The speech encoder generates 112 b of encoded data (Table 83.4) for every 20-ms frame of the speech signal. These bits are processed by the channel encoder to improve, after channel decoding at the receiver, the uncoded bit-error rate and the detectability of uncorrected errors. Error detection coding in the form of 3 CRC bits is applied to the most critical 22 data bits. The combined 25 b plus an additional 73 speech data bits and 6 *tail bits* are input to an RCPCC encoder (the tail bits serve to bring the channel encoder and decoder to a fixed terminal state at the end of the payload data stream). The 3 CRC bits are encoded at rate 1/3 and the other 101 b are encoded at rate 1/2, generating a total of 211 channel coded bits. These are finally combined with the remaining 17 (uncoded) speech data bits to form a total of 228 b for the payload data of a speech frame. The payload data from two speech frames are interleaved for transmission over four timeslots of the GSM TDMA channel.

With the Viterbi algorithm, the channel decoder performs soft decision decoding on the demodulated and deinterleaved channel data. Uncorrected channel errors may still be present in the decoded speech data after Viterbi decoding. Thus, the channel decoder classifies each frame into three integrity categories: bad, unreliable, and reliable, in order to assist the speech decoder in undertaking error concealment measures. A frame is considered bad if the CRC check fails or if the received channel data is close to more than one candidate sequence. The latter evaluation is based on applying an adaptive threshold to the metric values produced by the Viterbi algorithm over the course of decoding the most critical 22 speech data bits and their 3 CRC bits. Frames that are not bad may be classified as unreliable, depending on the metric values produced by the Viterbi algorithm and on channel reliability information supplied by the demodulator.

Depending on the recent history of decoded data integrity, the speech decoder can take various error concealment measures. The onset of bad frames is concealed by repetition of parameters from previous reliable frames, whereas the persistence of bad frames results in power attenuation and ultimately muting of the synthesized speech. Unreliable frames are decoded with normality constraints applied to the energy of the synthesized speech.

83.8 Conclusions

The half-rate standards employ some of the latest techniques in speech and channel coding to meet the challenges posed by the severe transmission environment of digital cellular radio systems. By halving the bit rate, the voice transmission capacity of existing full-rate digital cellular systems can be doubled. Although advances are still being made that can address the needs of quarter-rate speech transmission, much effort is currently devoted to enhancing the speech quality and robustness of full-rate (GSM and IS-54) systems, aiming to be closer to *toll quality*. On the other hand, the imminent introduction of competing wireless systems that use different modulation schemes [e.g., coded division multiple access (CDMA)] and/or different radio frequencies [e.g., personal communications systems (PCS)] is poised to alleviate congestion in high-user-density areas.

Defining Terms

Codebook: An ordered collection of all possible values that can be assigned to a scalar or vector variable. Each unique scalar or vector value in a codebook is called a *codeword*, or *code vector* where appropriate.

Codec: A contraction of (*en*)*coder–decoder*, used synonymously with the word *coder*. The encoder and decoder are often designed and deployed as a pair. A half-rate standard codec performs speech as well as channel coding.

Echo canceller: A signal processing device that, given the source signal causing the echo signal, generates an estimate of the echo signal and subtracts the estimate from the signal being interfered with by the echo signal. The device is usually based on a discrete-time adaptive filter.

Pitch period: The fundamental period of a voiced speech waveform that can be regarded as periodic over a short-time interval (quasiperiodic). The reciprocal of pitch period is *pitch frequency* or simply, *pitch*.

Tandem coding: Having more than one encoder–decoder pair in an end-to-end transmission path. In cellular radio communications, having a radio link at each end of the communication path could subject the speech signal to two passes of speech encoding–decoding. In general, repeated encoding and decoding increases the distortion.

Acknowledgment

The authors would like to thank Erdal Paksoy and Mark A. Jasiuk for their valuable comments.

References

1. Chen, J.-H. and Gersho, A., Adaptive postfiltering for quality enhancement of coded speech. *IEEE Trans. Speech & Audio Proc.,* 3(1), 59–71, 1995.
2. De Martino, E., Speech quality evaluation of the European, North-American and Japanese speech codec standards for digital cellular systems. In *Speech and Audio Coding for Wireless and Network Applications,* Atal, B.S., Cuperman, V., and Gersho, A., Eds., 55–58, Kluwer Academic Publishers, Norwell, MA, 1993.
3. Dimolitsas, S., Corcoran, F.L., and Baraniecki, M.R., Transmission quality of North American cellular, personal communications, and public switched telephone networks. *IEEE Trans. Veh. Tech.,* 43(2), 245–251, 1994.
4. Gersho, A., Advances in speech and audio compression. *Proc. IEEE,* 82(6), 900–918, 1994.

5. Gersho, A. and Gray, R.M., *Vector Quantization and Signal Compression,* Kluwer Academic Publishers, Norwell, MA, 1991.
6. Gerson, I.A. and Jasiuk, M.A., Vector sum excited linear prediction (VSELP) speech coding at 8 kbps. In *Proceedings, IEEE Intl. Conf. Acoustics, Speech, & Sig. Proc.,* 461–464, April, 1990.
7. Gerson, I.A. and Jasiuk, M.A., Techniques for improving the performance of CELP-type speech coders. *IEEE J. Sel. Areas Comm.,* 10(5), 858–865, 1992.
8. Gerson, I.A., Jasiuk, M.A., Nowack, J.M., Winter, E.H., and Müller, J.-M., Speech and channel coding for the half-rate GSM channel. In *Proceedings, ITG-Report 130 on Source and Channel Coding,* 225–232. Munich, Germany, Oct., 1994.
9. Gibson, J.D., Koo, B., and Gray, S.D., Filtering of colored noise for speech enhancement and coding. *IEEE Trans. Sig. Proc.,* 39(8), 1732–1742, 1991.
10. Hagenauer, J., Rate-compatible punctured convolutional codes (RCPC codes) and their applications. *IEEE Trans. Comm.,* 36(4), 389–400, 1988.
11. Jayant, N.S. and Noll, P., *Digital Coding of Waveforms,* Prentice-Hall, Englewood Cliffs, NJ, 1984.
12. Masui, F. and Oguchi, M., Activity of the half rate speech codec algorithm selection for the personal digital cellular system. *Tech. Rept. of IEICE,* RCS93-77(11), 55–62 (in Japanese), 1993.
13. Ohya, T., Suda, H., and Miki, T., 5.6 kbits/s PSI-CELP of the half-rate PDC speech coding standard. In *Proceedings, IEEE Veh. Tech. Conf.,* 1680–1684, June, 1994.
14. Proakis, J.G., *Digital Communications,* 3rd ed., McGraw-Hill, New York, 1995.
15. Suda, H., Ikeda, K., and Ikedo, J., Error protection and speech enhancement schemes of PSI-CELP, *NTT R&D.* (Special issue on PSI-CELP speech coding system for mobile communications), 43(4), 373–380, (in Japanese), 1994.
16. Telecommunication Industries Association (TIA). Half-rate speech codec test plan V6.0. TR45.3.5/93.05.19.01, 1993.
17. Zeger, K. and Gersho, A., Pseudo-Gray coding. *IEEE Trans. Comm.,* 38(12), 2147–2158, 1990.

Further Information

Additional technical information on speech coding can be found in the books, periodicals, and conference proceedings that appear in the list of references. Other relevant publications not represented in the list are *Speech Communication,* Elsevier Science Publishers; *Advances in Speech Coding,* B. S. Atal, V. Cuperman, and A. Gersho, Eds., Kluwer Academic Publishers; and *Proceedings of the IEEE Workshop on Speech Coding.*

84

Wireless Video Communications

Madhukar Budagavi
Texas Instruments

Raj Talluri
Texas Instruments

84.1 Introduction

Recent advances in technology have resulted in a rapid growth in mobile communications. With this explosive growth, the need for reliable transmission of mixed media information—audio, video, text, graphics, and speech data—over wireless links is becoming an increasingly important application requirement. The bandwidth requirements of raw video data are very high (a 176 × 144 pixels, color video sequence requires over 8 Mb/s). Since the amount of bandwidth available on current wireless channels is limited, the video data has to be compressed before it can be transmitted on the wireless channel. The techniques used for video compression typically utilize predictive coding schemes to remove redundancy in the video signal. They also employ variable length coding schemes, such as Huffman codes, to achieve further compression.

The wireless channel is a noisy fading channel characterized by long bursts of errors [8]. When compressed video data is transmitted over wireless channels, the effect of channel errors on the video can be severe. The variable length coding schemes make the compressed bitstream sensitive to channel errors. As a result, the video decoder that is decoding the corrupted video bitstream can easily lose synchronization with the encoder. Predictive coding techniques, such as **block motion compensation**, which are used in current video compression standards, make the matter worse by quickly propagating the effects of channel errors across the video sequence and rapidly degrading the video quality. This may render the video sequence totally unusable.

Error control coding [5], in the form of **Forward Error Correction (FEC)** and/or **Automatic Repeat reQuest (ARQ)**, is usually employed on wireless channels to improve the channel conditions.

TABLE 84.1 List of Relevant Standards

ISO/IEC 14496-2 (MPEG-4)	Information Technology—Coding of Audio-Visual Objects: Visual
H.263 (Version 1 and Version 2)	Video coding for low bitrate communication
H.261	Video codec for audiovisual services at p X 64 kbit/s
H.223	Multiplexing protocol for low bitrate multimedia communication
H.324	Terminal for low bitrate multimedia communication
H.245	Control protocol for multimedia communication
G.723.1	Dual rate speech coder for multimedia communication transmitting at 5.3 and 6.3 kbit/s

FEC techniques prove to be quite effective against random bit errors, but their performance is usually not adequate against longer duration burst errors. FEC techniques also come with an increased overhead in terms of the overall bitstream size; hence, some of the coding efficiency gains achieved by video compression are lost. ARQ techniques typically increase the delay and, therefore, might not be suitable for real-time videoconferencing. Thus, in practical video communication schemes, error control coding is typically used only to provide a certain level of error protection to the compressed video bitstream, and it becomes necessary for the video coder to accept some level of errors in the video bitstream. Error-resilience tools are introduced in the video codec to handle these residual errors that remain after error correction.

The emphasis in this chapter is on discussing relevant international standards that are making wireless video communications possible. We will concentrate on both the error control and source coding aspects of the problem. In the next section, we give an overview of a wireless video communication system that is a part of a complete wireless multimedia communication system. The International Telecommunication Union—Telecommunications Standardization Sector (ITU-T) H.223 [1] standard that describes a method of providing error protection to the video data before it is transmitted is also described. It should be noted that the main function of H.223 is to multiplex/demultiplex the audio, video, text, graphics, etc., which are typically communicated together in a videoconferencing application—error protection of the transmitted data becomes a requirement to support this functionality on error-prone channels. In Section 84.3, an overview of error-resilient video coding is given. The specific tools adopted into the International Standards Organization (ISO)/International Electrotechnical Commission (IEC) Motion Picture Experts Group (MPEG) v.4 (i.e., MPEG-4) [7] and the ITU-T H.263 [3] video coding standards to improve the error robustness of the video coder are described in Sections 84.4 and 84.5, respectively.

Table 84.1 provides a listing of some of the standards that are described or referred to in this chapter.

84.2 Wireless Video Communications

Figure 84.1 shows the basic block diagram of a wireless video communication system [10]. Input video is compressed by the video encoder to generate a compressed bitstream. The transport coder converts the compressed video bitstream into data units suitable for transmission over wireless channels. Typical operations carried out in the transport coder include channel coding, framing of data, modulation, and control operations required for accessing the wireless channel. At the receiver side, the inverse operations are performed to reconstruct the video signal for display.

In practice, the video communication system is part of a complete multimedia communication system and needs to interact with other system components to achieve the desired functionality. Hence, it becomes necessary to understand the other components of a multimedia communication system in order to design a good video communication system. Figure 84.2 shows the block diagram of a wireless multimedia terminal based on the ITU-T H.324 set of standards [4]. We use the H.324 standard as an example because it is the first videoconferencing standard for which mobile extensions were added to facilitate use on wireless channels. The system components of a multimedia terminal can be grouped into three processing blocks: (1) audio, video, and data (the word *data* is used here to mean still images/slides, shared files, documents, etc.), (2) control, and (3) multiplex-demultiplex blocks.

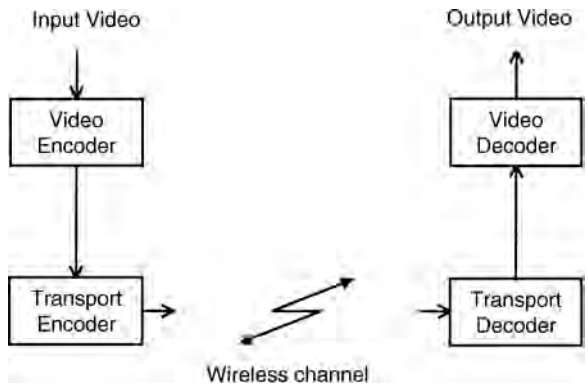


FIGURE 84.1 A wireless video communication system.

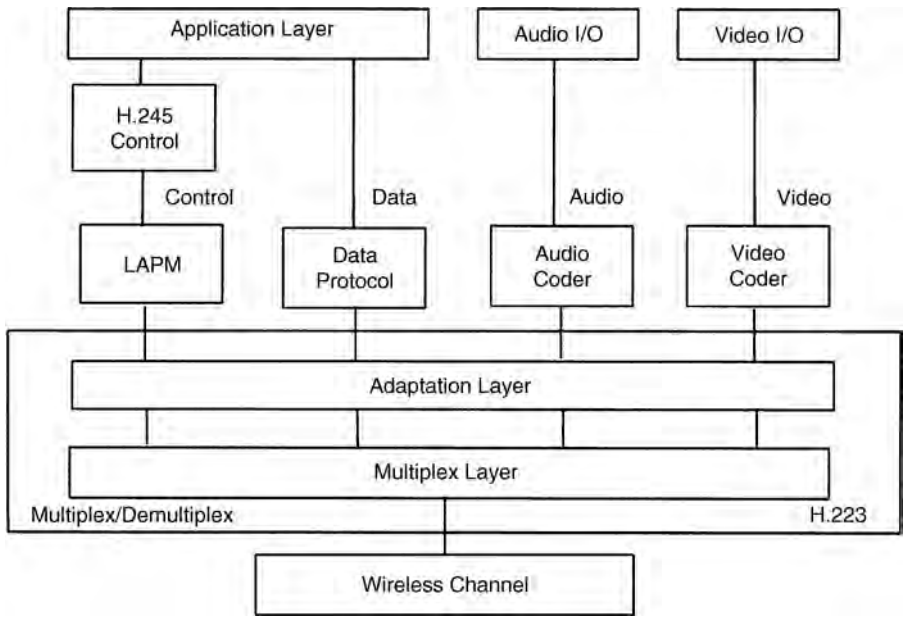


FIGURE 84.2 Configuration of a wireless multimedia terminal.

1. Audio, video, and data processing blocks—These blocks basically produce/consume the multimedia information that is communicated. The aggregate bitrate generated by these blocks is restricted due to limitations of the wireless channel and, therefore, the total rate allowed has to be judiciously allocated among these blocks. Typically, the video blocks use up the highest percentage of the aggregate rate, followed by audio and then data. H.324 specifies the use of H.261/H.263 for video coding and G.723.1 for audio coding.
2. Control block—This block has a wide variety of responsibilities all aimed at setting up and maintaining a multimedia call. The control block facilitates the set-up of compression methods and preferred bitrates for audio, video, and data to be used in the multimedia call. It is also responsible for end-to-network signalling for accessing the network and end-to-end signalling for reliable operation of the multimedia call. H.245 is the control protocol in the H.324 suite of standards that specifies the control messages to achieve the above functionality.

3. Multiplex-Demultiplex (MUX) block—This block multiplexes the resulting audio, video, data, and control signals into a single stream before transmission on the network. Similarly, the received bitstream is demultiplexed to obtain the audio, video, data, and control signals, which are then passed to their respective processing blocks. The MUX block accesses the network through a suitable network interface. The H.223 standard is the multiplexing scheme used in H.324.

Proper functioning of the MUX is crucial to the operation of the video communication system, as all the multimedia data/signals flow through it. On wireless channels, transmission errors can lead to a breakdown of the MUX resulting in, for example, nonvideo data being channeled to the video decoder or corrupted video data being passed on to the video decoder. Three annexes were specifically added to H.223 to enable its operation in error-prone environments. Below, we give a more detailed overview of H.223 and point out the levels of error protection provided by H.223 and its three annexes. It should also be noted that MPEG-4 does not specify a lower-level MUX like H.223, and thus H.223 can also be used to transmit MPEG-4 video data.

Recommendation H.223

Video, audio, data, and control information is transmitted in H.324 on distinct logical channels. H.223 determines the way in which the logical channels are mixed into a single bitstream before transmission over the physical channel (e.g., the wireless channel). The H.223 multiplex consists of two layers—the multiplex layer and the adaptation layer, as shown in Fig. 84.2. The multiplex layer is responsible for multiplexing the various logical channels. It transmits the multiplexed stream in the form of packets. The adaptation layer adapts the information stream provided by the applications above it to the multiplex layer below it by adding, where appropriate, additional octets for the purposes of error control and sequence numbering. The type of error control used depends on the type of information (audio/video/data/control) being conveyed in the stream. The adaptation layer provides error control support in the form of both FEC and ARQ.

H.223 was initially targeted for use on the benign general switched telephone network (GSTN). Later on, to enable its use on wireless channels, three annexes (referred to as Levels 1–3, respectively), were defined to provide improved levels of error protection. The initial specification of H.223 is referred to as Level 0. Together, Levels 0–3 provide for a trade-off of error robustness against the overhead required, with Level 0 being the least robust and using the least amount of overhead and Level 3 being the most robust and also using the most amount of overhead.

1. H.223 Level 0—Default mode. In this mode the transmitted packet sizes are of variable length and are delimited by an 8-bit HDLC (High-level Data Link Control) flag (**01111110**). Each packet consists of a 1-octet header followed by the payload, which consists of a variable number of information octets. The header octet includes a Multiplex Code (MC) which specifies, by indexing to a multiplex table, the logical channels to which each octet in the information field belongs. To prevent emulation of the HDLC flag in the payload, bitstuffing is adopted.
2. H.223 Level 1 (Annex A)—Communication over low error-prone channels. The use of bitstuffing leads to poor performance in the presence of errors; therefore in Level 1, bitstuffing is not performed. The other improvement incorporated in Level 1 is the use of a longer 16-bit pseudo-noise synchronization flag to allow for more reliable detection of packet boundaries. The input bitstream is correlated with the synchronization flag and the output of the correlator is compared with a correlation threshold. Whenever the correlator output is equal to or greater than the threshold, a flag is detected. Since bitstuffing is not performed, it is possible to have this flag emulated in the payload. However, the probability of such an emulation is low and is outweighed by the improvement gained by not using bitstuffing over error-prone channels.
3. H.223 Level 2 (Annex B)—Communication over moderately error-prone channels. When compared to the Level 1 operation, Level 2 increases the protection on the packet header. A Multiplex Payload Length (MPL) field, which gives the length of the payload in bytes, is introduced into the header to provide additional redundancy for detecting the length of the video packet. A (24,12,8)

extended Golay code is used to protect the MC and the MPL fields. Use of error protection in the header enables robust delineation of packet boundaries. Note that the payload data is not protected in Level 2.

4. H.223 Level 3 (Annex C)—Communication over highly error-prone channels. Level 3 goes one step above Level 2 and provides for protection of the payload data. Rate Compatible Punctured Convolutional (RCPC) codes, various CRC polynomials, and ARQ techniques are used for protection of the payload data. Level 3 allows for the payload error protection overhead to vary depending on the channel conditions. RCPC codes are used for achieving this adaptive level of error protection because RCPC codes use the same channel decoder architecture for all the allowed levels of error protection, thereby reducing the complexity of the MUX.

84.3 Error Resilient Video Coding

Even after error control and correction, some amount of residual errors still exist in the compressed bitstream fed to the video decoder in the receiver. Therefore, the video decoder should be robust to these errors and should provide acceptable video quality even in the presence of some residual errors. In this section, we first describe a standard video coder configuration that is the basis of many international standards and also highlight the potential problems that are encountered when compressed video from these systems is transmitted over wireless channels. We then give an overview of the strategies that can be adopted to overcome these problems. Most of these strategies are incorporated in the MPEG-4 video coding standard and the H.263 (Version 2) video coding standard [3]. The original H.263 standard [2] which was standardized in 1996 for use in H.324 terminals connected to GSTN is referred to as Version 1. Version 2 of the H.263 standard provides additional improvements and functionalities (which include error-resilience tools) over the Version 1 standard. We will use H.263 to refer to both Version 1 and Version 2 standards and a distinction will be made only when required.

A Standard Video Coder

Redundancy exists in video signals in both spatial and temporal dimensions. Video coding techniques exploit this redundancy to achieve compression. A plethora of video compression techniques have been proposed in the literature, but a hybrid coding technique consisting of block motion compensation (BMC) and discrete cosine transforms (DCT) has been found to be very effective in practice. In fact, most of the current video coding standards such as H.263 and MPEG-4, which provide state-of-the-art compression performance, are all based on this hybrid coding technique. In this hybrid BMC/DCT coding technique, BMC is used to exploit temporal redundancy and the DCT is used to reduce spatial redundancy.

Figure 84.3 illustrates a standard hybrid BMC/DCT video coder configuration. Pictures are coded in either of two modes—interframe (INTER) or intraframe (INTRA) mode. In intraframe coding, the video image is encoded without any relation to the previous image, whereas in interframe coding, the current image is predicted from the previous image using BMC, and the difference between the current image and the predicted image, called the residual image, is encoded. The basic unit of data which is operated on is called a macroblock (MB) and is the data (both **luminance and chrominance** components) corresponding to a block of 16×16 pixels. The input image is split into disjoint macroblocks and the processing is done on a macroblock basis. Motion information, in the form of **motion vectors**, is calculated for each macroblock. The motion compensated prediction residual error is then obtained by subtracting each pixel in the macroblock with its motion shifted counterpart in the previous frame. Depending on the mode of coding used for the macroblock, either the image macroblock or the corresponding residual image macroblock is split into blocks of size 8×8 and an 8×8 DCT is applied to each of these 8×8 blocks. The resulting DCT coefficients are then quantized. Depending on the quantization step-size, this will result in a significant number of zero-valued coefficients. To efficiently encode the DCT coefficients that remain nonzero after quantization, the DCT coefficients are zig-zag scanned, and run-length encoded and the run-lengths are variable length encoded before transmission. Since a significant amount of

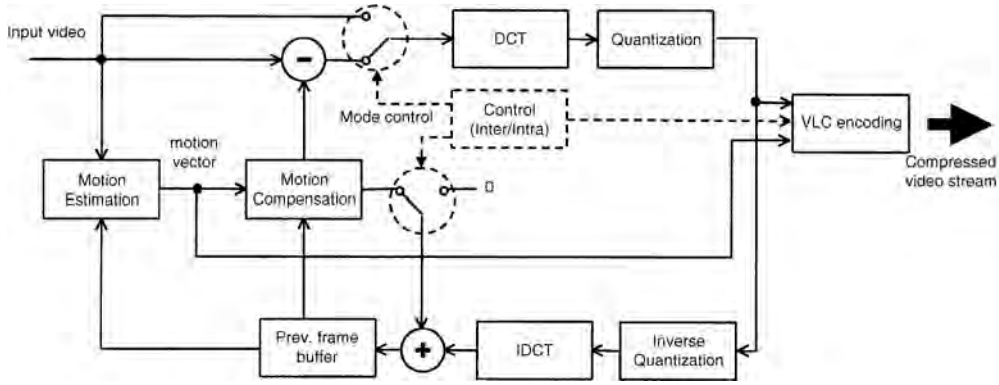


FIGURE 84.3 A standard video coder.

correlation exists between the neighboring macroblocks' motion vectors, the motion vectors are themselves predicted from already transmitted motion vectors and the motion vector prediction error is encoded. The motion vector prediction error and the mode information are also variable length coded before transmission to achieve efficient compression.

The decoder uses a reverse process to reconstruct the macroblock at the receiver. The variable length codewords present in the received video bitstream are decoded first. For INTER macroblocks, the pixel values of the prediction error are reconstructed by inverse quantization and inverse DCT and are then added to the motion compensated pixels from the previous frame to reconstruct the transmitted macroblock. For INTRA macroblocks, inverse quantization and inverse DCT directly result in the transmitted macroblock. All macroblocks of a given picture are decoded to reconstruct the whole picture.

Error Resilient Video Decoding

The use of predictive coding and variable length coding (VLC), though very effective from a compression point of view, makes the video decoding process susceptible to transmission errors. In VLC, the boundary between codewords is implicit. The compressed bitstream has to be read until a full codeword is encountered; the codeword is then decoded to obtain the information encoded in the codeword. When there are transmission errors, the implicit nature of the boundary between codewords typically leads to an incorrect number of bits being used in VLC decoding and, thus, subsequently results in a loss of synchronization with the encoder. In addition, the use of predictive coding leads to the propagation of these transmission errors to neighboring spatial blocks and to subsequently decoded frames, which leads to a rapid degradation in the reconstructed video quality.

To minimize the disastrous impact that transmission errors can have on the video decoding process, the following stages are incorporated in the video decoder to make it more robust:

- Error detection and localization
- Resynchronization
- Data recovery
- Error concealment

Figure 84.4 shows an error resilient video decoder configuration. The first step involved in robust video coding is the detection of errors in the bitstream. The presence of errors in the bitstream can be signaled by the FEC used in the multiplex layer. The video coder can also detect errors whenever illegal VLC codewords are encountered in the bitstream or when the decoding of VLC codewords leads to an illegal value of the decoded information (e.g., occurrence of more than 64 DCT coefficients for an 8×8 DCT block). Accurate detection of errors in the bitstream is a very important step, since most of the other error resilience techniques can only be invoked if an error is detected.

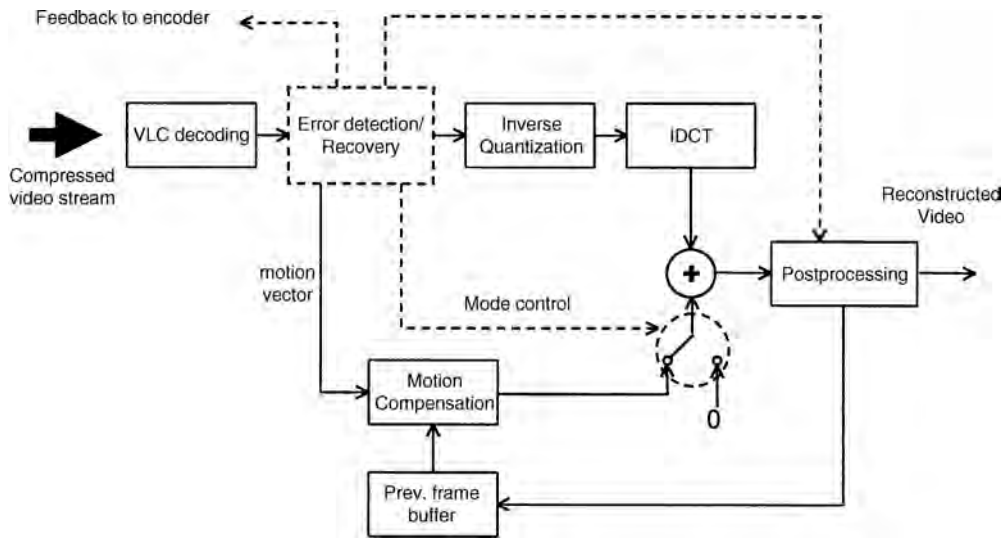


FIGURE 84.4 Error resilient video decoder.

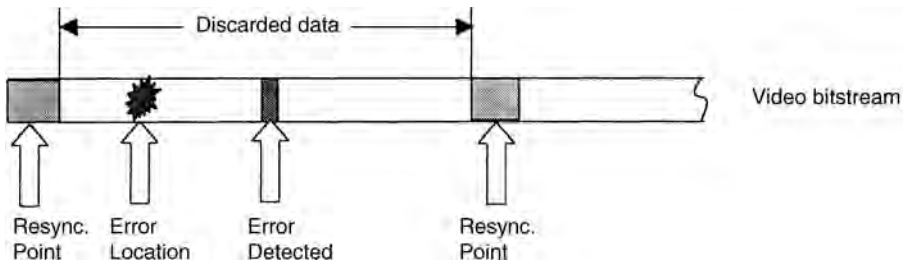


FIGURE 84.5 At the decoder, it is usually not possible to detect the error at the actual error occurrence location; hence, all the data between the two resynchronization points may need to be discarded.

Due to the use of VLC, the location in the bitstream where the decoder detects an error is not the same location where the error has actually occurred but some undetermined distance away from it. This is shown in Fig. 84.5. Once an error is detected, it also implies that the decoder is not in synchronization with the encoder. Resynchronization schemes are then employed for the decoder to fall back into lock step with the encoder. While constructing the bitstream, the encoder inserts unique resynchronization words into the bitstream at approximately equally spaced intervals. These resynchronization words are chosen such that they are unique from the valid video bitstream. That is, no valid combination of the video algorithm's VLC tables can produce these words. The decoder, upon detection of an error, seeks forward in the bitstream looking for this known resynchronization word. Once this word is found, the decoder then falls back in synchronization with the encoder. At this point, the decoder has detected an error, regained synchronization with the encoder, and isolated the error to be between the two resynchronization points. Since the decoder can only isolate the error to be somewhere between the resynchronization points but not pinpoint its exact location, all of the data that corresponds to the macroblocks between these two resynchronization points needs to be discarded. Otherwise, the effects of displaying an image reconstructed from erroneous data can cause highly annoying visual artifacts.

Some data recovery techniques, such as "reversible decoding," enable the decoder to salvage some of the data between the two resynchronization points. These techniques advocate the use of a special kind of VLC table at the encoder in coding the DCTs and motion vector information. These special VLCs have the property that they can be decoded both in the forward and reverse directions. By comparing the

forward and reverse decoded data, the exact location of the error in the bit stream can be localized more precisely and some of the data between the two resynchronization points can be salvaged. The use of these reversible VLCs (RVLCs) is part of the MPEG-4 standard and will be described in greater detail in the following sections.

After data recovery, the impact of the data that is deemed to be in error needs to be minimized. This is the error concealment stage. One simple error concealment strategy is to simply replace the luminance and chrominance components of the erroneous macroblocks with the luminance and chrominance of the corresponding macroblocks in the previous frame of the video sequence. While this technique works fairly well and is simple to implement, more complex techniques use some type of estimation strategies to exploit the local correlation that exists within a frame of video data to come up with a better estimate of the missing or erroneous data. These error concealment strategies are essentially postprocessing algorithms and are not mandated by the video coding standards. Different implementations of the wireless video systems utilize different kinds of error concealment strategies based on the available computational power and the quality of the channel.

If there is support for a decoder feedback path to the encoder as shown in Fig. 84.3, this path can be used to signal detected errors. The feedback information from the decoder can be used to retransmit data or to influence future encoder action so as to stop the propagation of detected errors in the decoder. Note that for the feedback to take place, the network must support a back channel.

Classification of Error-Resilience Techniques

In general, techniques to improve the robustness of the video coder can be classified into three categories based on whether the encoder or the decoder plays a primary part in improving the error robustness [10]. *Forward error resilience* techniques refer to those techniques where the encoder plays the primary part in improving the error robustness, typically by introducing redundancy in the transmitted information. In *postprocessing* techniques, the decoder plays the primary part and does concealment of errors by estimation and interpolation (e.g., spatial-temporal filtering) using information it has already received. In *interactive error* resilience techniques, the decoder and the encoder interact to improve the error resilience of the video coder. Techniques that use decoder feedback come under this category.

84.4 MPEG-4 Error Resilience Tools

MPEG-4 is an ISO/IEC standard being developed by the Motion Pictures Expert Group. Initially MPEG was aimed primarily at low-bit-rate communications; however, its scope was later expanded to be much more of a multimedia coding standard [7]. The MPEG-4 video coding standard is the first video coding standard to address the problem of efficient representation of visual objects of arbitrary shape. MPEG-4 was also designed to provide "universal accessibility," i.e., the ability to access audio-visual information over a wide range of storage and transmission media. In particular, because of the proliferation of wireless communications, this implied development of specific tools to enable error-resilient transmission of compressed data over noisy communication channels.

A number of tools have been incorporated into the MPEG-4 video coder to make it more error resilient. All these tools are basically forward error resilience tools. We describe below each of these tools and its advantages.

Resynchronization

As mentioned earlier, a video decoder that is decoding a corrupted bitstream may lose synchronization with the encoder (i.e., it is unable to identify the precise location in the image where the current data belongs). If remedial measures are not taken, the quality of the decoded video rapidly degrades and becomes unusable. One approach is for the encoder to introduce resynchronization markers in the bitstream at various locations. When the decoder detects an error, it can then look for this resynchronization marker and regain synchronization.

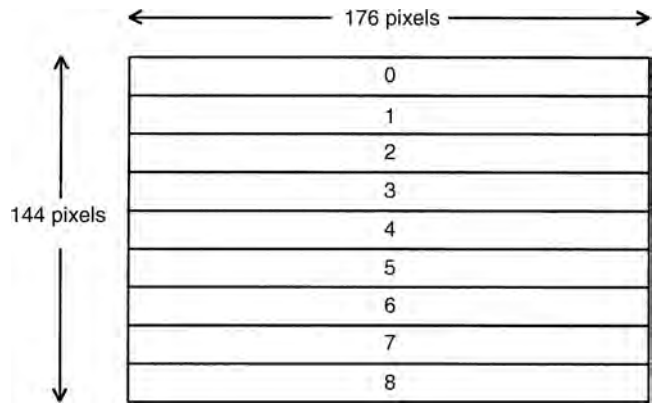


FIGURE 84.6 H.263 GOB numbering for a QCIF image.

Previous video coding standards such as H.261 and H.263 (Version 1) logically partition each of the images to be encoded into rows of macroblocks called Group Of Blocks (GOBs). These GOBs correspond to a horizontal row of macroblocks for **QCIF** images. Figure 84.6 shows the GOB numbering scheme for H.263 (Version 1) for QCIF resolution. For error resilience purposes, H.263 (Version 1) provides the encoder an option of inserting resynchronization markers at the beginning of each of the GOBs. Hence, for QCIF images these resynchronization markers are allowed to occur only at the left edge of the images. The smallest region that the error can be isolated to and concealed in this case is thus one row of macroblocks.

In contrast, the MPEG-4 encoder is not restricted to inserting the resynchronization markers only at the beginning of each row of macroblocks. The encoder has the option of dividing the image into video packets. Each video packet is made up of an integer number of consecutive macroblocks in raster scan order. These macroblocks can span several rows of macroblocks in the image and can even include partial rows of macroblocks. One suggested mode of operation for the MPEG-4 encoder is for it to insert a resynchronization marker periodically at approximately every K bits. Note that resynchronization markers can only be placed at a macroblock boundary and, hence, the video packet length cannot be constrained to be exactly equal to K bits. When there is a significant activity in one part of the image, the macroblocks corresponding to these areas generate more bits than other parts of the image. If the MPEG-4 encoder inserts the resynchronization markers at uniformly spaced bit intervals, the macroblock interval between the resynchronization markers is a lot closer in the high activity areas and a lot farther apart in the low activity areas. Thus, in the presence of a short burst of errors, the decoder can quickly localize the error to within a few macroblocks in the important high activity areas of the image and preserve the image quality in these important areas. In the case of H.263 (Version 1), where the resynchronization markers are restricted to be at the beginning of the GOBs, it is only possible for the decoder to isolate the errors to a fixed GOB independent of the image content. Hence, effective coverage of the resynchronization marker is reduced when compared to the MPEG-4 scheme. The recommended spacing of the resynchronization markers in MPEG-4 is based on the bitrates. For 24 Kb/s, it is recommended to insert them at intervals of 480 bits and for bitrates between 25 Kb/s to 48 Kb/s, it is recommended to place them at every 736 bits. Figures 84.7(a) and (b) illustrate the placement of resynchronization markers for H.263 (Version 1) and MPEG-4.

Note that in addition to inserting the resynchronization markers at the beginning of each video packet, the encoder also needs to remove all data dependencies that exist between the data belonging to two different video packets within the same image. This is required so that even if one of the video packets in the current image is corrupted due to errors, the other packets can be decoded and utilized by the decoder. In order to remove these data dependencies, the encoder inserts two additional fields in addition to the resynchronization marker at the beginning of each video packet, as shown in Fig. 84.8. These are (1) the absolute macroblock number of the first macroblock in the video packet, *Mb. No.*,

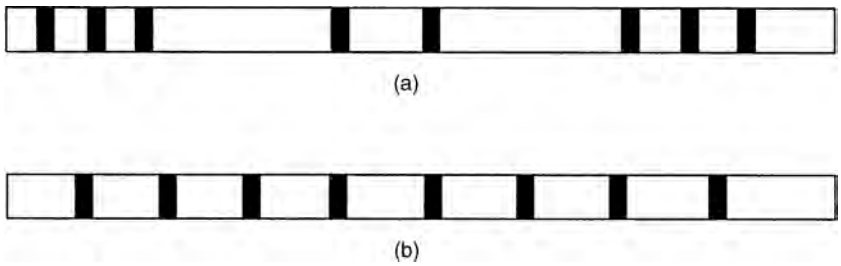


FIGURE 84.7 Position of resynchronization markers in the bitstream for (a) H.263 (Version 1) encoder with GOB headers and for (b) an MPEG-4 encoder with video packets.

FIGURE 84.8 An MPEG-4 video packet.

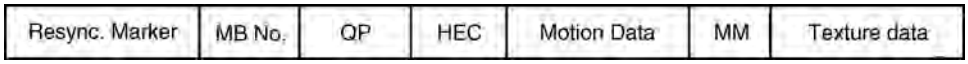


FIGURE 84.9 A data partitioned MPEG-4 video packet.

(which indicates the spatial location of the macroblock in the current image), and (2) the quantization parameter, *QP*, which denotes the initial quantization parameter used to quantize the DCT coefficients in the video packet. The encoder also modifies the predictive encoding method used for coding the motion vectors such that there are no predictions across the video packet boundaries. Also shown in Fig. 84.8 is a third field, labeled *HEC*. Its use is discussed in a later section.

Data Partitioning

Data partitioning in MPEG-4 provides enhanced error localization and error concealment capabilities. The data partitioning mode partitions the data within a video packet into a motion part and a texture part (DCT coefficients) separated by a unique Motion Marker (MM), as shown in Fig. 84.9. All the syntactic elements of the video packet that have motion-related information are placed in the motion partition and all the remaining syntactic elements that relate to the DCT data are placed in the texture partition. If the texture information is lost, data partitioning enables the salvation of motion information, which can then be used to conceal the errors in a more effective manner.

The motion marker is computed from the motion VLC tables using a search program such that it is Hamming distance 1 from any possible valid combination of the motion VLC tables [9]. The motion marker is uniquely decodable from the motion VLC tables, and it indicates to the decoder the end of the motion information and the beginning of the DCT information. The number of macroblocks in the video packet is implicitly known after encountering the motion marker. Note that the motion marker is only computed once based on the VLC tables and is fixed in the standard. Based on the VLC tables in MEPG-4, the motion marker is a 17-bit word whose value is **1 1111 0000 0000 0001**.

Reversible Variable Length Codes (RVLCs)

As was shown in Fig. 84.5, if the decoder detects an error during the decoding of VLC codewords, it loses synchronization and hence typically has to discard all the data up to the next resynchronization point. RVLCs are designed such that they can be instantaneously decoded both in the forward and the backward direction. When the decoder detects an error while decoding the bitstream in the forward

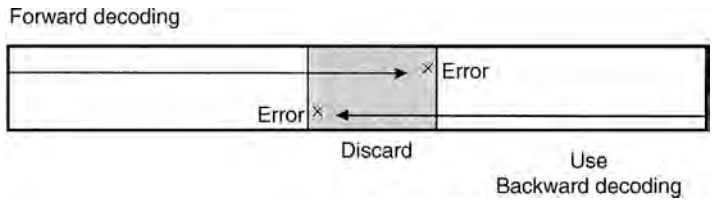


FIGURE 84.10 Use of reversible variable length codes.

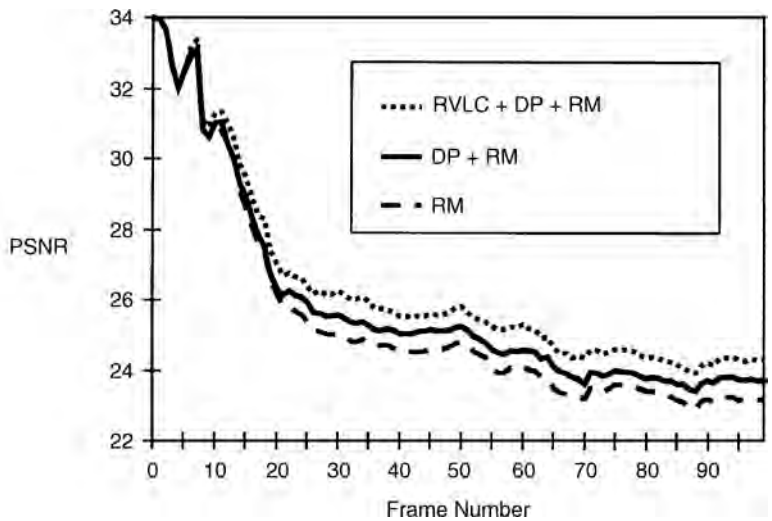


FIGURE 84.11 Performance comparison of resynchronization, data partitioning, and RVLC over a bursty channel simulated by a 2-state Gilbert model. Burst durations are 1 ms long and the probability of occurrence of a burst is 10^{-2}. Legend: RM—resynchronization marker; DP—data partitioning; RVLC—reversible variable length codes.

direction, it jumps to the next resynchronization marker and decodes the bitstream in the backward direction until it encounters an error. Based on the two error locations, the decoder can recover some of the data that would have otherwise been discarded. This is shown in Fig. 84.10, which shows only the texture part of the video packet—only data in the shaded area is discarded. Note that if RVLCs were not used, all the data in the texture part of the video packet would have to be discarded. RVLCs thus enable the decoder to better isolate the error location in the bitstream.

Figure 84.11 shows the comparison of performance of resynchronization, data partitioning, and RVLC techniques for 24 Kb/s QCIF video data. The experiments involved transmission of three video sequences, each of duration 10 s, over a bursty channel simulated by a 2-state Gilbert model [6]. The burst duration on the channel is 1 ms and the burst occurrence probability is 10^{-2}. Figure 84.11, which plots the average peak signal-to-noise ratios of the received video frames, shows that data partitioning and RVLC provide improved performance when compared to using only resynchronization markers.

Header Extension Code (HEC)

Some of the most important information that the decoder needs in order to decode the video bitstream is in the video frame header data. This data includes information about the spatial dimensions of the video data, the time stamps associated with the decoding and the presentation of this video data, and the type of the current frame (INTER/INTRA). If some of this information is corrupted due to channel errors, the decoder has no other recourse but to discard all the information belonging to the current

video frame. In order to reduce the sensitivity of this data, a technique called Header Extension Code (HEC) was introduced into the MPEG-4 standard. In each video packet, a 1-bit field called HEC is present. The location of HEC in the video packet is shown in Fig. 84.8. For each video packet, when HEC is set, the important header information that describes the video frame is repeated in the bits following the HEC. This information can be used to verify and correct the header information of the video frame. The use of HEC significantly reduces the number of discarded video frames and helps achieve a higher overall decoded video quality.

Adaptive Intra Refresh (AIR)

Whenever an INTRA macroblock is received, it basically stops the temporal propagation of errors at its corresponding spatial location. The procedure of forcefully encoding some macroblocks in a frame in INTRA mode to flush out possible errors is called INTRA refreshing. INTRA refresh is very effective in stopping the propagation of errors, but it comes at the cost of a large overhead. Coding a macroblock in INTRA mode typically requires many more bits when compared to coding the macroblock in INTER mode. Hence, the INTRA refresh technique has to be used judiciously.

For areas with low motion, simple error concealment by just copying the previous frame's macroblocks works quite effectively. For macroblocks with high motion, error concealment becomes very difficult. Since the high motion areas are perceptually the most significant, any persistent error in the high motion area becomes very noticeable. The AIR technique of MPEG-4 makes use of the above facts and INTRA refreshes the motion areas more frequently, thereby allowing the corrupted high motion areas to recover quickly from errors.

Depending on the bitrate, the AIR approach only encodes a fixed and predetermined number of macroblocks in a frame in INTRA mode (the exact number is not standardized by MPEG-4). This fixed number might not be enough to cover all the macroblocks in the motion area; hence, the AIR technique keeps track of the macroblocks that have been refreshed (using a "refresh map") and in subsequent frames refreshes any macroblocks in the motion areas that might have been left out.

84.5 H.263 Error Resilience Tools

In this section, we discuss four error resilience techniques which are part of the H.263 standard—*slice structure mode* and *independent segment decoding,* which are forward error resilience features, and *error tracking* and *reference picture selection*, which are interactive error resilience techniques. Error tracking was introduced in H.263 (Version 1) as an appendix, whereas the remaining three techniques were introduced in H.263 (Version 2) as annexes.

Slice Structure Mode (Annex K)

The slice structured mode of H.263 is similar to the video packet approach of MPEG-4 with a slice denoting a video packet. The basic functionality of a slice is the same as that of a video packet—providing periodic resynchronization points throughout the bistream. The structure of a slice is shown in Fig. 84.12. Like an MPEG-4 video packet, the slice consists of a header followed by the macroblock data. The SSC is the slice start code and is identical to the resynchronization marker of MPEG-4. The MBA field, which

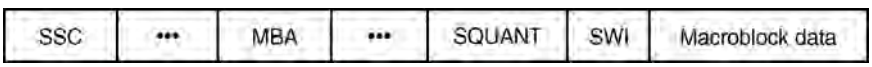


FIGURE 84.12 Structure of a slice in H.263/Annex K.

denotes the starting macroblock number in the slice, and the SQUANT field, which is the quantizer scale coded nonpredictively, allow for the slice to be coded independently.

The slice structured mode also contains two submodes which can be used to provide additional functionality. The submodes are

- Rectangular slice submode (RSS)—This allows for rectangular shaped slices. The rectangular region contained in the slice is specified by SWI+1 (See Fig. 84.12 for the location of the SWI field in the slice header), which gives the width of the rectangular region, and MBA, which specifies the upper left macroblock of the slice. Note that the height of the rectangular region gets specified by the number of macroblocks contained in the slice. This mode can be used, for example, to subdivide images into rectangular regions of interest for region-based coding.
- Arbitrary slice submode (ASO)—The default order of transmission of slices is such that the MBA field is strictly increasing from one slice to the next transmitted slice. When ASO is used, the slices may appear in any order within the bitstream. This mode is useful when the wireless network supports prioritization of slices which might result in out-of-order arrival of video slices at the decoder.

Independent Segment Decoding (ISD) (Annex R)

Even though the slice structured mode eliminates decoding dependency between neighboring slices, errors in slices can spatially propagate to neighboring slices in subsequent frames due to motion compensation. This happens because motion vectors in a slice can point to macroblocks of neighboring slices in the reference picture. Independent segment decoding eliminates this from happening by restricting the motion vectors within a predefined segment of the picture from pointing to other segments in the picture, thereby helping to contain the error to be within the erroneous segment. This improvement in the localization of errors, however, comes at a cost of a loss of coding efficiency. Because of this restriction on the motion vectors, the motion compensation is not as effective, and the residual error images use more bits.

For ease of implementation, the ISD mode puts restrictions on segment shapes and on the changes of segment shapes from picture to picture. The ISD mode cannot be used with the slice structured mode (Annex K) unless the rectangular slice submode of Annex K is active. This prevents the need for treating awkward shapes of slices that can otherwise arise when Annex K is not used with rectangular slice submode. The segment shapes are not allowed to change from picture to picture unless an INTRA frame is being coded.

Error Tracking (Appendix I)

The error tracking approach is an INTRA refresh technique but uses decoder feedback of errors to decide which macroblocks in the current image to code in INTRA mode to prevent the propagation of these errors. When there are no errors on the channel, normal coding (which usually results in the bit-efficient INTER mode being selected most of the time) is used. The use of decoder feedback allows the system to adapt to varying channel conditions and minimizes the use of forced INTRA updates to situations when there are channel errors.

Because of the time delay involved in the decoder feedback, the encoder has to track the propagation of an error from its original occurrence to the current frame to decide which macroblocks should be INTRA coded in the current frame. A low complexity algorithm was proposed in Appendix I of H.263 to track the propagation of errors. However, it should be noted that the use of this technique is not mandated by H.263. Also, H.263 itself does not standardize the mechanism by which the decoder feedback of error can be sent. Typically, H.245 control messages are used to signal the decoder feedback for error tracking purposes.

Reference Picture Selection (Annex N)

The Reference Picture Selection (RPS) mode of H.263 also relies on decoder feedback to efficiently stop the propagation of errors. The back channel used in RPS mode can be a separate logical channel (e.g., by using H.245), or if two-way communication is taking place, the back channel messages can be sent multiplexed with the encoded video data. In the presence of errors, the RPS mode allows the encoder to be instructed to select one of the several previously correctly received and decoded frames as the reference picture for motion compensation of the current frame being encoded. This effectively stops the propagation of error. Note that the use of RPS requires the use of multiple frame buffers at both the encoder and the decoder to store previously decoded frames. Hence, the improvement in performance in the RPS mode has come at the cost of increased memory requirements.

84.6 Discussion

In this chapter we presented a broad overview of the various techniques that enable wireless video transmission. Due to the enormous amount of bandwidth required, video data is typically compressed before being transmitted, but the errors introduced by the wireless channels have a severe impact on the compressed video information. Hence, special techniques need to be employed to enable robust video transmission. International standards play a very important role in communications applications. The two current standards that are most relevant to video applications are ISO MPEG-4 and ITU H.263. In this chapter, we detailed these two standards and explained the error resilient tools that are part of these standards to enable robust video communication over wireless channels. A tutorial overview of these tools has been presented and the performance of these tools has been described.

There are, however, a number of other methods that further improve the performance of a wireless video codec that the standards do not specify. If the encoder and decoder are aware of the limitations imposed by the communication channel, they can further improve the video quality by using these methods. These methods include encoding techniques such as rate control to optimize the allocation of the effective channel bit rate between various parts of video to be transmitted and intelligent decisions on when and where to place INTRA refresh macroblocks to limit the error propagation. Decoding methods such as superior error concealment strategies that further conceal the effects of erroneous macroblocks by estimating them from correctly decoded macroblocks in the spatiotemporal neighborhood can also significantly improve the effective video quality.

This chapter has mainly focused on the error resilience aspects of the video layer. There are a number of error detection and correction strategies, such as Forward Error Correction (FEC), that can further improve the reliability of the transmitted video data. These FEC codes are typically provided in the systems layer and the underlying network layer. If the video transmission system has the ability to monitor the dynamic error characteristics of the communication channel, joint source-channel coding techniques can also be effectively employed. These techniques enable the wireless communication system to perform optimal trade-offs in allocating the available bits between the source coder (video) and the channel coder (FEC) to achieve superior performance.

Current video compression standards also support *layered* coding methods. In this approach, the compressed video information can be separated into multiple layers. The *base* layer, when decoded, provides a certain degree of video quality and the *enhancement* layer, when received and decoded, then adds to the base layer to further improve the video quality. In wireless channels, these base and enhancement layers give a natural method of partitioning the video data into more important and less important layers. The base layer can be protected by a stronger level of error protection (higher overhead channel coder) and the enhancement layer by a lesser strength coder. Using this Unequal Error Protection (UEP) scheme, the communication system is assured of a certain degree of performance most of the time through the base layer, and when the channel is not as error prone and the decoder receives the enhancement layer, this scheme provides improved quality.

Given all these advances in video coding technology, coupled with the technological advances in processor technology, memory devices, and communication systems, wireless video communications is fast

becoming a very compelling application. With the advent of higher bandwidth third generation wireless communication systems, it will be possible to transmit compressed video in many wireless applications, including mobile videophones, videoconferencing systems, PDAs, security and surveillance applications, mobile Internet terminals, and other multimedia devices.

Defining Terms

Automatic Repeat reQuest (ARQ): An error control system in which notification of erroneously received messages is sent to the transmitter which then simply retransmits the message. The use of ARQ requires a feedback channel and the receiver must perform error detection on received messages. Redundancy is added to the message before transmission to enable error detection at the receiver.

Block motion compensation (BMC): Motion compensated prediction that is done on a block basis; that is, blocks of pixels are assumed to be displaced spatially in a uniform manner from one frame to another.

Forward Error Correction: Introduction of redundancy in data to allow for correction of errors without retransmission.

Luminance and chrominance: Luminance is the brightness information in a video image, whereas chrominance is the corresponding color information.

Motion vectors: Specifies the spatial displacement of a block of pixels from one frame to another.

QCIF: Quarter Common Intermediate Format (QCIF) is a standard picture format that defines the image dimensions to be 176×144 (pixels per line $\times$ lines per picture) for luminance and 88×72 for chrominance.

References

1. International Telecommunications Union—Telecommunications Standardization Sector, Recommendation H.223: Multiplexing protocol for low bitrate multimedia communications, Geneva, 1996.
2. International Telecommunications Union—Telecommunications Standardization Sector, Recommendation H.263: Video coding for low bitrate communication, Geneva, 1996.
3. International Telecommunications Union—Telecommunications Standardization Sector, Draft Recommendation H.263 (Version 2): video coding for low bitrate communication, Geneva, 1998.
4. International Telecommunications Union—Telecommunications Standardization Sector, Recommendation H.324: terminal for low bit rate multimedia communications, Geneva, 1996.
5. Lin, S. and Costello, D.J., Jr., *Error Control Coding: Fundamentals and Applications,* Prentice-Hall, Englewood Cliffs, NJ, 1983.
6. Miki, T., et al., Revised error pattern generation programs for core experiments on error resilience, *ISO/IEC JTC1/SC29/WG11 MPEG96/1492,* Maceio, Brazil, Nov. 1996.
7. International Organization for Standardization, Committee draft of Tokyo (N2202): information technology—coding of audio-visual objects: visual, *ISO/IEC 14496-2,* Mar. 1998.
8. Sklar, B., Rayleigh fading channels in mobile digital communication systems, Pt. I: Characterization, *IEEE Commun. Mag.,* 35, 90–100, 1997.
9. Talluri, R., et al., Error concealment by data partitioning, to appear in *Signal Processing: Image Commun.,* 1998.
10. Wang, Y. and Zhu, Q., Error control and concealment for video communication: a review, *IEEE Trans. Circuits Syst. Video Technol.,* 86(5), 974–997, May 1998.

Further Information

A broader overview of wireless video can be found in the special issue of *IEEE Communications Magazine,* June 1998. Wang and Zhu [10] provide an exhaustive review of error concealment techniques for video communications. More details on MPEG-4 and ongoing Version 2 activities in MPEG-4 can be found

on the web page http://drogo.cselt.it/mpeg/standards/mpeg-4/mpeg-4.htm. H.263 (Version 2) activities are tracked on the web page http://www.ece.ubc.ca/spmg/research/motion/h263plus/. Most of the ITU-T recommendations can be obtained from the web site http://www.itu.org. The special issue of *IEEE Communications Magazine,* December 1996, includes articles on H.324 and H.263.

Current research relevant to wireless video communications is reported in a number of journals including *IEEE Transactions on Circuits and Systems for Video Technology, IEEE Transactions on Image Processing, IEEE Transactions on Vehicular Technology, Signal Processing: Image Communication*. The *IEEE Communications Magazine* regularly reports review articles relevant to wireless video communications. Conferences of interest include the IEEE International Conference on Image Processing (ICIP), IEEE Vehicular Technology Conference (VTC), and IEEE International Conference on Communications (ICC).

85

Wireless LANs

Suresh Singh
Portland State University

85.1 Introduction

A proliferation of high-performance portable computers combined with end-user need for communication is fueling a dramatic growth in wireless **local area network** (LAN) technology. Users expect to have the ability to operate their portable computer globally while remaining connected to communications networks and service providers. Wireless LANs and cellular networks, connected to high-speed networks, are being developed to provide this functionality.

Before delving deeper into issues relating to the design of wireless LANs, it is instructive to consider some scenarios of user mobility.

1. A simple model of user mobility is one where a computer is physically moved while retaining network connectivity at either end. For example, a move from one room to another as in a hospital where the computer is a hand-held device displaying patient charts and the nurse using the computer moves between wards or floors while accessing patient information.
2. Another model situation is where a group of people (at a conference, for instance) set up an ad-hoc LAN to share information as in Fig. 85.1.
3. A more complex model is one where several computers in constant communication are in motion and continue to be networked. For example, consider the problem of having robots in space collaborating to retrieve a satellite.

A great deal of research has focused on dealing with physical layer and **medium access control (MAC)** layer protocols. In this chapter we first summarize standardization efforts in these areas. The remainder of the chapter is then devoted to a discussion of networking issues involved in wireless LAN design. Some of the issues discussed include routing in wireless LANs (i.e., how does data find its destination when the destination is mobile?) and the problem of providing service guarantees to end users (e.g., error-free data transmission or bounded delay and bounded bandwidth service, etc.).

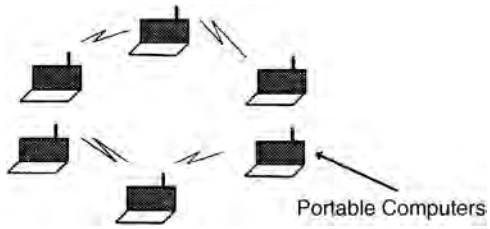


FIGURE 85.1 Ad-hoc wireless LAN.

85.2 Physical Layer Design

Two media are used for transmission over wireless LANs, infrared and radio frequency. RF LANs are typically implemented in the industrial, scientific, and medical (ISM) frequency bands 902–928 MHz, 2400–2483.5 MHz, and 5725–5850 MHz. These frequencies do not require a license allowing the LAN product to be portable, i.e., a LAN can be moved without having to worry about licensing.

IR and RF technologies have different design constraints. IR receiver design is simple (and thus inexpensive) in comparison to RF receiver design because IR receivers only detect the amplitude of the signal not the frequency or phase. Thus, a minimal of filtering is required to reject interference. Unfortunately, however, IR shares the electromagnetic spectrum with the sun and incandescent or fluorescent light. These sources of modulated infrared energy reduce the signal-to-noise ratio of IR signals and, if present in extreme intensity, can make the IR LANs inoperable. There are two approaches to building IR LANs.

1. The transmitted signal can be focused and aimed. In this case the IR system can be used outdoors and has an area of coverage of a few kilometers.
2. The transmitted signal can be bounced off the ceiling or radiated omnidirectionally. In either case, the range of the IR source is 10–20 m (i.e., the size of one medium-sized room).

RF systems face harsher design constraints in comparison to IR systems for several reasons. The increased demand for RF products has resulted in tight regulatory constraints on the allocation and use of allocated bands. In the U.S., for example, it is necessary to implement spectrum spreading for operation in the ISM bands. Another design constraint is the requirement to confine the emitted spectrum to a band, necessitating amplification at higher carrier frequencies, frequency conversion using precision local oscillators, and selective components. RF systems must also cope with environmental noise that is either naturally occurring, for example, atmospheric noise or man-made, for example, microwave ovens, copiers, laser printers, or other heavy electrical machinery. RF LANs operating in the ISM frequency ranges also suffer interference from amateur radio operators.

Operating LANs indoors introduces additional problems caused by multipath propagation, Rayleigh fading, and absorption. Many materials used in building construction are opaque to IR radiation resulting in incomplete coverage within rooms (the coverage depends on obstacles within the room that block IR) and almost no coverage outside closed rooms. Some materials, such as white plasterboard, can also cause reflection of IR signals. RF is relatively immune to absorption and reflection problems. Multipath propagation affects both IR and RF signals. The technique to alleviate the effects of multipath propagation in both types of systems is the same use of aimed (directional) systems for transmission enabling the receiver to reject signals based on their angle of incidence. Another technique that may be used in RF systems is to use multiple antennas. The phase difference between different paths can be used to discriminate between them.

Rayleigh fading is a problem in RF systems. Recall that Rayleigh fading occurs when the difference in path length of the same signal arriving along different paths is a multiple of half a wavelength. This causes the signal to be almost completely canceled out at the receiver. Because the wavelengths used in IR are so small, the effect of Rayleigh fading is not noticeable in those systems. RF systems, on the other hand,

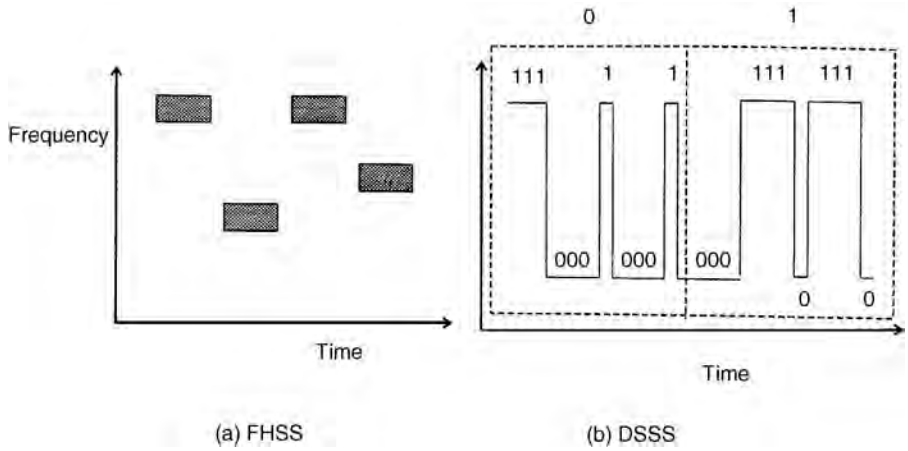


FIGURE 85.2 Spread spectrum.

use wavelengths of the order of the dimension of a laptop. Thus, moving the computer a small distance could increase/decrease the fade significantly.

Spread spectrum transmission technology is used for RF-based LANs and it comes in two varieties: direct-sequence spread spectrum (DSSS) and frequency-hopping spread spectrum (FHSS). In a FHSS system, the available band is split into several channels. The transmitter transmits on one channel for a fixed time and then hops to another channel. The receiver is synchronized with the transmitter and hops in the same sequence; see Fig. 85.2(a). In DSSS systems, a random binary string is used to modulate the transmitted signal. The relative rate between this sequence and user data is typically between 10 and 100; see Fig. 85.2(b).

The key requirements of any transmission technology is its robustness to noise. In this respect DSSS and FHSS show some differences. There are two possible sources of interference for wireless LANs: the presence of other wireless LANs in the same geographical area (i.e., in the same building, etc.) and interference due to other users of the ISM frequencies. In the latter case, FHSS systems have a greater ability to avoid interference because the hopping sequence could be designed to prevent potential interference. DSSS systems, on the other hand, do exhibit an ability to recover from interference because of the use of the spreading factor [Fig. 85.2(b)].

It is likely that in many situations several wireless LANs may be collocated. Since all wireless LANs use the same ISM frequencies, there is a potential for a great deal of interference. To avoid interference in FHSS systems, it is necessary to ensure that the hopping sequences are orthogonal. To avoid interference in DSSS systems, on the other hand, it is necessary to allocate different channels to each wireless LAN. The ability to avoid interference in DSSS systems is, thus, more limited in comparison to FHSS systems because FHSS systems use very narrow subchannels (1 MHz) in comparison to DSSS systems that use wider subchannels (for example, 25 MHz), thus, limiting the number of wireless LANs that can be collocated. A summary of design issues can be found in [1].

85.3 MAC Layer Protocols

MAC protocol design for wireless LANs poses new challenges because of the in-building operating environment for these systems. Unlike wired LANs (such as the ethernet or token ring), wireless LANs operate in strong multipath fading channels where channel characteristics can change in very short distances resulting in unreliable communication and unfair channel access due to capture. Another feature of the wireless LAN environment is that carrier sensing takes a long time in comparison to wired LANs; it typically takes between 30 and 50 μs (see [4]), which is a significant portion of the packet transmission time. This results in inefficiencies if the CSMA family of protocols is used without any modifications.

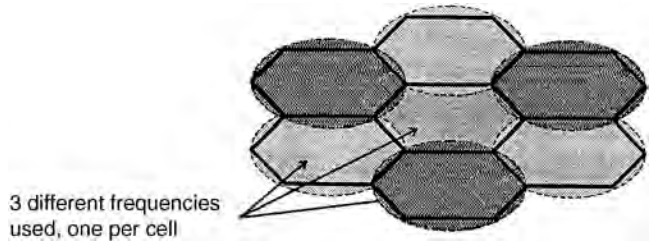


FIGURE 85.3 Cellular structure for wireless LANs (note frequency reuse).

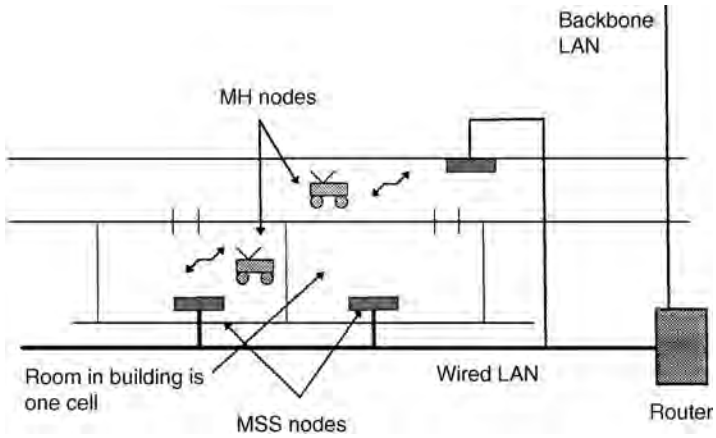


FIGURE 85.4 In-building LAN (made up of several wireless LANs).

Other differences arise because of the mobility of users in wireless LAN environments. To provide a building (or any other region) with wireless LAN coverage, the region to be covered is divided into cells as shown in Fig. 85.3. Each cell is one wireless LAN, and adjacent cells use different frequencies to minimize interference. Within each cell there is an access point called a **mobile support station** (MSS) or base station that is connected to some wired network. The mobile users are called **mobile hosts (MH)**. The MSS performs the functions of channel allocation and providing connectivity to existing wired networks; see Fig. 85.4. Two problems arise in this type of an architecture that are not present in wired LANs.

1. The number of nodes within a cell changes dynamically as users move between cells. How can the channel access protocol dynamically adapt to such changes efficiently?
2. When a user moves between cells, the user has to make its presence known to the other nodes in the cell. How can this be done without using up too much bandwidth? The protocol used to solve this problem is called a handoff protocol and works along the following lines: A switching station (or the MSS nodes working together, in concert) collects signal strength information for each mobile host within each cell. Note that if a mobile host is near a cell boundary, the MSS node in its current cell as well as in the neighboring cell can hear its transmissions and determine signal strengths. If the mobile host is currently under the coverage of MSS M1 but its signal strength at MSS M2 becomes larger, the switching station initiates a handoff whereby the MH is considered as part of M2's cell (or network).

The mode of communication in wireless LANs can be broken in two: communication from the mobile to the MSS (called *uplink* communication) and communication in the reverse direction (called *downlink* communication). It is estimated that downlink communication accounts for about 70–80% of the total consumed bandwidth. This is easy to see because most of the time users request files or data in other forms (image data, etc.) that consume much more transmission bandwidth than the

requests themselves. In order to make efficient use of bandwidth (and, in addition, guarantee service requirements for real-time data), most researchers have proposed that the downlink channel be controlled entirely by the MSS nodes. These nodes allocate the channel to different mobile users based on their current requirements using a protocol such as **time division multiple access** (TDMA). What about uplink traffic? This is a more complicated problem because the set of users within a cell is dynamic, thus making it infeasible to have a static channel allocation for the uplink. This problem is the main focus of MAC protocol design.

What are some of the design requirements of an appropriate MAC protocol? The IEEE 802.11 recommended standard for wireless LANs has identified almost 20 such requirements, some of which are discussed here (the reader is referred to [3], for further details). Clearly any protocol must maximize throughput while minimizing delays and providing fair access to all users. In addition to these requirements, however, mobility introduces several new requirements.

1. The MAC protocol must be independent of the underlying physical layer transmission technology adopted (be it DSSS, FHSS or IR).
2. The maximum number of users can be as high as a few hundred in a wireless LAN. The MAC protocol must be able to handle many users without exhibiting catastrophic degradation of service.
3. The MAC protocols must provide secure transmissions because the wireless medium is easy to tap.
4. The MAC protocol needs to work correctly in the presence of collocated networks.
5. It must have the ability to support ad-hoc networking (as in Fig. 85.1).
6. Other requirements include the need to support priority traffic, preservation of packet order, and an ability to support multicast.

Several contention-based protocols currently exist that could be adapted for use in wireless LANs. The protocols currently being looked at by IEEE 802.11 include protocols based on **carrier sense multiple access (CSMA)**, polling, and TDMA. Protocols based on **code division multiple access** (CDMA) and **frequency division multiple access** (FDMA) are not considered because the processing gains obtained using these protocols are minimal while, simultaneously, resulting in a loss of flexibility for wireless LANs.

It is important to highlight an important difference between networking requirements of ad-hoc networks (as in Fig. 85.1) and networks based on cellular structure. In cellular networks, all communication occurs between the mobile hosts and the MSS (or base station) within that cell. Thus, the MSS can allocate channel bandwidth according to requirements of different nodes, i.e., we can use centralized channel scheduling for efficient use of bandwidth. In ad-hoc networks there is no such central scheduler available. Thus, any multiaccess protocol will be contention based with little explicit scheduling. In the remainder of this section we focus on protocols for cell-based wireless LANs only.

All multiaccess protocols for cell-based wireless LANs have a similar structure; see [3].

1. The MSS announces (explicitly or implicitly) that nodes with data to send may contend for the channel.
2. Nodes interested in sending data contend for the channel using protocols such as CSMA.
3. The MSS allocates the channel to successful nodes.
4. Nodes transmit packets (contention-free transmission).
5. MSS sends an explicit acknowledgment (ACK) for packets received.

Based on this model we present three MAC protocols.

Reservation-TDMA (R-TDMA)

This approach is a combination of TDMA and some contention protocol (see PRMA in [7]). The MSS divides the channel into slots (as in TDMA), which are grouped into frames. When a node wants to transmit it needs to reserve a slot that it can use in every consecutive frame as long as it has data to transmit. When it has completed transmission, other nodes with data to transmit may contend for that

free slot. There are four steps to the functioning of this protocol.

a. At the end of each frame the MSS transmits a feedback packet that informs nodes of the current reservation of slots (and also which slots are free). This corresponds to steps 1 and 3 from the preceding list.
b. During a frame, all nodes wishing to acquire a slot transmit with a probability ρ during a free slot. If a node is successful it is so informed by the next feedback packet. If more than one node transmits during a free slot, there is a collision and the nodes try again during the next frame. This corresponds to step 2.
c. A node with a reserved slot transmits data during its slot. This is the contention-free transmission (step 4).
d. The MSS sends ACKs for all data packets received correctly. This is step 5.

The R-TDMA protocol exhibits several nice properties. First and foremost, it makes very efficient use of the bandwidth, and average latency is half the frame size. Another big benefit is the ability to implement power conserving measures in the portable computer. Since each node knows when to transmit (nodes transmit during their reserved slot only) it can move into a power-saving mode for a fixed amount of time, thus increasing battery life. This feature is generally not available in CSMA-based protocols. Furthermore, it is easy to implement priorities because of the centralized control of scheduling. One significant drawback of this protocol is that it is expensive to implement (see [2]).

Distributed Foundation Wireless MAC (DFWMAC)

The CSMA/CD protocol has been used with great success in the ethernet. Unfortunately, the same protocol is not very efficient in a wireless domain because of the problems associated with cell interference (i.e., interference from neighboring cells), the relatively large amount of time taken to sense the channel (see [6]) and the hidden terminal problem (see [12,13]). The current proposal is based on a CSMA/collision avoidance (CA) protocol with a four-way handshake; see Fig. 85.5.

The basic operation of the protocol is simple. All MH nodes that have packets to transmit compete for the channel by sending ready to transmit (RTS) messages using nonpersistent CSMA. After a station succeeds in transmitting a RTS, the MSS sends a clear to transmit (CTS) to the MH. The MH transmits its data and then receives an ACK. The only possibility of collision that exists is in the RTS phase of the protocol and inefficiencies occur in the protocol, because of the RTS and CTS stages. Note that unlike R-TDMA it is harder to implement power saving functions. Furthermore, latency is dependent on system load making it harder to implement real-time guarantees. Priorities are also not implemented. On the positive side, the hardware for this protocol is very inexpensive.

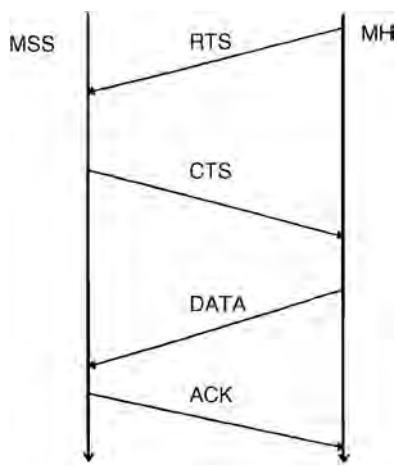


FIGURE 85.5 CSMA/CA and four-way handshaking protocol.

Randomly Addressed Polling (RAP)

In this scheme, when a MSS is ready to collect uplink packets it transmits a READY message. At this point all nodes with packets to send attempt to grab the channel as follows.

a. Each MH with a packet to transmit generates a random number between 0 and P.
b. All active MH nodes simultaneously and orthogonally transmit their random numbers (using CDMA or FDMA). We assume that all of these numbers are received correctly by the MSS. Remember that more than one MH node may have selected the same random number.
c. Steps a and b are repeated L times.
d. At the end of L stages, the MSS determines a stage (say, k) where the total number of distinct random numbers was the largest. The MSS polls each distinct each random number in this stage in increasing order. All nodes that had generated the polled random number transmit packets to the MSS.
e. Since more than one node may have generated the same random number, collisions are possible. The MSS sends a ACK or NACK after each such transmission. Unsuccessful nodes try again during the next iteration of the protocol.

The protocol is discussed in detail in [4] and a modified protocol called GRAP (for group RAP) is discussed in [3]. The authors propose that GRAP can also be used in the contention stage (step 2) for TDMA- and CSMA-based protocols.

85.4 Network Layer Issues

An important goal of wireless LANs is to allow users to move about freely while still maintaining all of their connections (network resources permitting). This means that the network must route all packets destined for the mobile user to the MSS of its current cell in a transparent manner. Two issues need to be addressed in this context.

- How can users be addressed?
- How can active connections for these mobile users be maintained?

Ioanidis, Duchamp, and Maguire [8] propose a solution called the IPIP (IP-within-IP) protocol. Here each MH has a unique **internet protocol** (IP) address called its home address. To deliver a packet to a remote MH, the source MSS first broadcasts an address resolution protocol (ARP) request to all other MSS nodes to locate the MH. Eventually some MSS responds. The source MSS then encapsulates each packet from the source MH within another packet containing the IP address of the MSS in whose cell the MH is located. The destination MSS extracts the packet and delivers it to the MH. If the MH has moved away in the interim, the new MSS locates the new location of the MH and performs the same operation. This approach suffers from several problems as discussed in [11]. Specifically, the method is not scaleable to a network spanning areas larger than a campus for the following reasons.

1. IP addresses have a prefix identifying the campus subnetwork where the node lives; when the MH moves out of the campus, its IP address no longer represents this information.
2. The MSS nodes serve the function of routers in the mobile network and, therefore, have the responsibility of tracking all of the MH nodes globally causing a lot of overhead in terms of message passing and packet forwarding; see [5].

Teraoka and Tokoro [11], have proposed a much more flexible solution to the problem called virtual IP (VIP). Here every mobile host has a virtual IP address that is unchanging regardless of the location of the MH. In addition, hosts have physical network addresses (traditional IP addresses) that may change as the host moves about. At the transport layer, the target node is always specified by its VIP address only. The address resolution from the VIP address to the current IP address takes place either at the network layer of the same machine or at a gateway. Both the host machines and the gateways maintain

a cache of VIP to IP mappings with associated timestamps. This information is in the form of a table called *address mapping table* (AMT). Every MH has an associated *home gateway*. When a MH moves into a new subnetwork, it is assigned a new IP address. It sends this new IP address and its VIP address to its home gateway via a *VipConn* control message. All intermediate gateways that relay this message update their AMT tables as well. During this process of updating the AMT tables, all packets destined to the MH continue to be sent to the old location. These packets are returned to the sender, who then sends them to the home gateway of the MH. It is easy to see that this approach is easily scaleable to large networks, unlike the IPIP approach.

Alternative View of Mobile Networks

The approaches just described are based on the belief that mobile networks are merely an extension of wired networks. Other authors [10] disagree with this assumption because there are fundamental differences between the mobile domain and the fixed wired network domain. Two examples follow.

1. The available bandwidth at the wireless link is small; thus, end-to-end packet retransmission for transmission control protocol (TCP)–like protocols (implemented over datagram networks) is a bad idea. This leads to the conclusion that transmission within the mobile network must be connection oriented. Such a solution, using virtual circuits (VC), is proposed in [5].
2. The bandwidth available for a MH with open connections changes dynamically since the number of other users present in each cell varies randomly. This is a feature not present in fixed high-speed networks where, once a connection is set up, its bandwidth does not vary much. Since bandwidth changes are an artifact of mobility and are dynamic, it is necessary to deal with the consequences (e.g., buffer overflow, large delays, etc.) locally to both, i.e., shield fixed network hosts from the idiosyncrasies of mobility as well as to respond to changing bandwidth quickly (without having to rely on end-to-end control). Some other differences are discussed in [10].

A Proposed Architecture

Keeping these issues in mind, a more appropriate architecture has been proposed in Ghai and Singh [5], and Singh [10]. Mobile networks are considered to be different and separate from wired networks. Within a mobile network is a three-layer hierarchy; see Fig. 85.6. At the bottom layer are the MHs. At the next level are the MSS nodes (one per cell). Finally, several MSS nodes are controlled by a **supervisor host (SH)** node (there may be one SH node per small building). The SH nodes are responsible for flow control for all MH connections within their domain; they are also responsible for tracking MH nodes and forwarding packets as MH nodes roam. In addition, the SH nodes serve as a *gateway* to the wired networks. Thus, any connection setup from a MH to a fixed host is broken in two, one from the MH

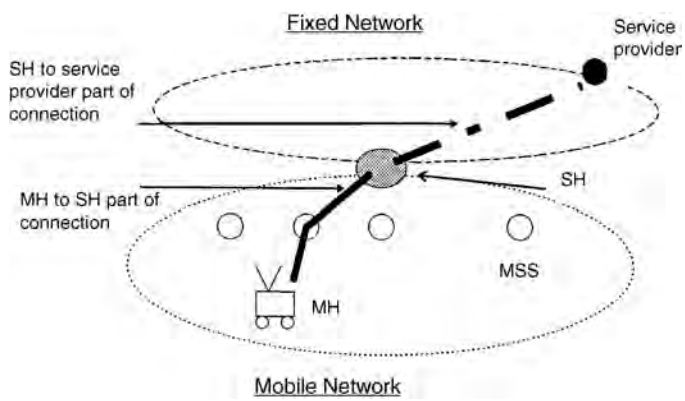


FIGURE 85.6 Proposed architecture for wireless networks.

to the SH and another from the SH to the fixed host. The MSS nodes in this design are simply connection endpoints for MH nodes. Thus, they are simple devices that implement the MAC protocols and little else. Some of the benefits of this design are as follows.

1. Because of the large coverage of the SH (i.e., a SH controls many cells) the MH remains in the domain of one SH much longer. This makes it easy to handle the consequences of dynamic bandwidth changes locally. For instance, when a MH moves into a crowded cell, the bandwidth available to it is reduced. If it had an open ftp connection, the SH simply buffers undelivered packets until they can be delivered. There is no need to inform the other endpoint of this connection of the reduced bandwidth.
2. When a MH node sets up a connection with a service provider in the fixed network, it negotiates some quality of service (QOS) parameters such as bandwidth, delay bounds, etc. When the MH roams into a crowded cell, these QOS parameters can no longer be met because the available bandwidth is smaller. If the traditional view is adopted (i.e., the mobile networks are extensions of fixed networks) then these QOS parameters will have to be renegotiated each time the bandwidth changes (due to roaming). This is a very expensive proposition because of the large number of control messages that will have to be exchanged. In the approach of Singh [10], the service provider will never know about the bandwidth changes since it deals only with the SH that is accessed via the wired network. The SH bears the responsibility of handling bandwidth changes by either buffering packets until the bandwidth available to the MH increases (as in the case of the ftp example) or it could discard a fraction of real-time packets (e.g., a voice connection) to ensure delivery of most of the packets within their deadlines. The SH could also instruct the MSS to allocate a larger amount of bandwidth to the MH when the number of buffered packets becomes large. Thus, the service provider in the fixed network is shielded from the mobility of the user.

Networking Issues

It is important for the network to provide connection-oriented service in the mobile environment (as opposed to connectionless service as in the internet) because bandwidth is at a premium in wireless networks, and it is, therefore, inadvisable to have end-to-end retransmission of packets (as in TCP). The proposed architecture is well suited to providing connection-oriented service by using VCs.

In the remainder of this section we look at how virtual circuits are used within the mobile network and how routing is performed for connections to mobile hosts. Every connection set up with one or more MH nodes as a connection endpoint is routed through the SH nodes and each connection is given a unique VC number. The SH node keeps track of all MH nodes that lie within its domain. When a packet needs to be delivered to a MH node, the SH first buffers the packet and then sends it to the MSS at the current location of the MH or to the predicted location if the MH is currently between cells. The MSS buffers all of these packets for the MH and transmits them to the MH if it is in its cell. The MSS discards packets after transmission or if the SH asks it to discard the packets. Packets are delivered in the correct order to the MH (without duplicates) by having the MH transmit the expected sequence number (for each VC) during the initial handshake (i.e., when the MH first enters the cell). The MH sends ACKs to the SH for packets received. The SH discards all packets that have been acknowledged. When a MH moves from the domain of SH1 into the domain of SH2 while having open connections, SH1 continues to forward packets to SH2 until either the connections are closed or until SH2 sets up its own connections with the other endpoints for each of MH's open connections (it also gives new identifiers to all these open connections). The detailed protocol is presented in [5].

The SH nodes are all connected over the fixed (wired) network. Therefore, it is necessary to route packets between SH nodes using the protocol provided over the fixed networks. The VIP protocol appears to be best suited to this purpose. Let us assume that every MH has a globally unique VIP address. The SHs have both a VIP as well as a fixed IP address. When a MH moves into the domain of a SH, the IP address affixed to this MH is the IP address of the SH. This ensures that all packets sent to the MH are routed through the correct SH node. The SH keeps a list of all VIP addresses of MH nodes within its

domain and a list of open VCs for each MH. It uses this information to route the arriving packets along the appropriate VC to the MH.

85.5 Transport Layer Design

The transport layer provides services to higher layers (including the application layer), which include connectionless services like UDP or connection-oriented services like TCP. A wide variety of new services will be made available in the high-speed networks, such as continuous media service for real-time data applications such as voice and video. These services will provide bounds on delay and loss while guaranteeing some minimum bandwidth.

Recently variations of the TCP protocol have been proposed that work well in the wireless domain. These proposals are based on the traditional view that wireless networks are merely extensions of fixed networks. One such proposal is called I-TCP [2] for indirect TCP. The motivation behind this work stems from the following observation. In TCP the sender times out and begins retransmission after a timeout period of several hundred milliseconds. If the other endpoint of the connection is a mobile host, it is possible that the MH is disconnected for a period of several seconds (while it moves between cells and performs the initial greeting). This results in the TCP sender timing out and transmitting the same data several times over, causing the effective throughput of the connection to degrade rapidly. To alleviate this problem, the implementation of I-TCP separates a TCP connection into two pieces—one from the fixed host to another fixed host that is near the MH and another from this host to the MH (note the similarity of this approach with the approach in Fig. 85.6). The host closer to the MH is aware of mobility and has a larger timeout period. It serves as a type of gateway for the TCP connection because it sends ACKs back to the sender before receiving ACKs from the MH. The performance of I-TCP is far superior to traditional TCP for the mobile networks studied.

In the architecture proposed in Fig. 85.6, a TCP connection from a fixed host to a mobile host would terminate at the SH. The SH would set up another connection to the MH and would have the responsibility of transmitting all packets correctly. In a sense this is a similar idea to I-TCP except that in the wireless network VCs are used rather than datagrams. Therefore, the implementation of TCP service is made much easier.

A problem that is unique to the mobile domain occurs because of the unpredictable movement of MH nodes (i.e., a MH may roam between cells resulting in a large variation of available bandwidth in each cell). Consider the following example. Say nine MH nodes have opened 11-kb/s connections in a cell where the available bandwidth is 100 kb/s. Let us say that a tenth mobile host M10, also with an open 11-kb/s connection, wanders in. The total requested bandwidth is now 110 kb/s while the available bandwidth is only 100 kb/s. What is to be done? One approach would be to deny service to M10. However, this seems an unfair policy. A different approach is to penalize all connections equally so that each connection has 10-kb/s bandwidth allocated.

To reduce the bandwidth for each connection from 11 kb/s to 10 kb/s, two approaches may be adopted:

1. Throttle back the sender for each connection by sending control messages.
2. Discard 1-kb/s data for each connection at the SH. This approach is only feasible for applications that are tolerant of data loss (e.g., real-time video or audio).

The first approach encounters a high overhead in terms of control messages and requires the sender to be capable of changing the data rate dynamically. This may not always be possible; for instance, consider a teleconference consisting of several participants where each mobile participant is subject to dynamically changing bandwidth. In order to implement this approach, the data (video or audio or both) will have to be compressed at different ratios for each participant, and this compression ratio may have to be changed dynamically as each participant roams. This is clearly an unreasonable solution to the problem. The second approach requires the SH to discard 1-kb/s of data for each connection. The question is, how should this data be discarded? That is, should the 1 kb of discarded data be consecutive (or clustered) or uniformly spread out over the data stream every 1 s? The way in which the data is discarded has

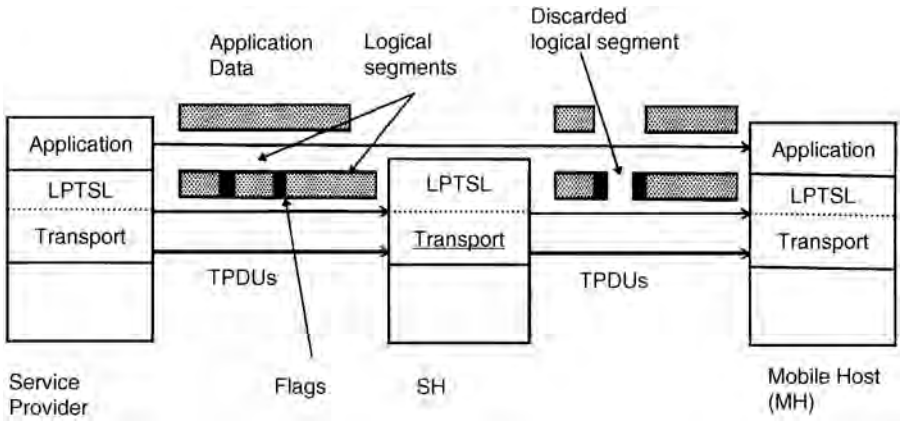


FIGURE 85.7 LPTSL, an approach to handle dynamic bandwidth variations.

an effect on the final perception of the service by the mobile user. If the service is audio, for example, a random uniform loss is preferred to a clustered loss (where several consecutive words are lost). If the data is compressed video, the problem is even more serious because most random losses will cause the encoded stream to become unreadable resulting in almost a 100% loss of video at the user.

A solution to this problem is proposed in Seal and Singh [9], where a new sublayer is added to the transport layer called the *loss profile transport sublayer (LPTSL)*. This layer determines how data is to be discarded based on special transport layer markers put by application calls at the sender and based on negotiated loss functions that are part of the QOS negotiations between the SH and service provider. Figure 85.7 illustrates the functioning of this layer at the service provider, the SH, and the MH. The original data stream is broken into *logical segments* that are separated by markers (or flags). When this stream arrives at the SH, the SH discards entire logical segments (in the case of compressed video, one logical segment may represent one frame) depending on the bandwidth available to the MH. The purpose of discarding entire logical segments is that discarding a part of such a segment of data makes the rest of the data within that segment useless—so we might as well discard the entire segment. Observe also that the flags (to identify logical segments) are inserted by the LPTSL via calls made by the application layer. Thus, the transport layer or the LPTSL does not need to know encoding details of the data stream. This scheme is currently being implemented at the University of South Carolina by the author and his research group.

85.6 Conclusions

The need for wireless LANs is driving rapid development in this area. The IEEE has proposed standards (802.11) for the physical layer and MAC layer protocols. A great deal of work, however, remains to be done at the network and transport layers. There does not appear to be a consensus regarding subnet design for wireless LANs. Our work has indicated a need for treating wireless LAN subnetworks as being fundamentally different from fixed networks, thus resulting in a different subnetwork and transport layer designs. Current efforts are under way to validate these claims.

Defining Terms

Carrier sense multiple access (CSMA): Protocols such as those used over the ethernet.
Medium access control (MAC): Protocols arbitrate channel access between all nodes on a wireless LAN.
Mobile host (MH) nodes: The nodes of wireless LAN.
Supervisor host (SH): The node that takes care of flow-control and other protocol processing for all connections.

References

1. Bantz, D.F. and Bauchot, F.J., Wireless LAN design alternatives. *IEEE Network,* 8(2), 43–53, 1994.
2. Barke, A. and Badrinath, B.R., I-TCP: indirect TCP for mobile hosts. Tech. Rept. DCS-TR-314, Dept. Computer Science, Rutgers University, Piscataway, NJ, 1994.
3. Chen, K.-C., Medium access control of wireless LANs for mobile computing. *IEEE Network,* 8(5), 50–63, 1994.
4. Chen, K.-C. and Lee, C.H., RAP: a novel medium access control protocol for wireless data networks. *Proc. IEEE GLOBECOM'93,* IEEE Press, Piscataway, NJ, 08854. 1713–1717, 1993.
5. Ghai, R. and Singh, S., An architecture and communication protocol for picocellular networks. *IEEE Personal Comm. Mag.,* 1(3), 36–46, 1994.
6. Glisic, S.G., 1-Persistent carrier sense multiple access in radio channel with imperfect carrier sensing. *IEEE Trans. on Comm.,* 39(3), 458–464, 1991.
7. Goodman, D.J., Cellular packet communications. *IEEE Trans. on Comm.,* 38(8), 1272–1280, 1990.
8. Ioanidis, J., Duchamp, D., and Maguire, G.Q., IP-based protocols for mobile internetworking. *Proc. of ACM SIGCOMM'91,* ACM Press, New York, NY, 10036 (Sept.), 235–245, 1991.
9. Seal, K. and Singh, S., Loss profiles: a quality of service measure in mobile computing. *J. Wireless Networks,* 2, 45–61, 1996.
10. Singh, S., Quality of service guarantees in mobile computing. *J. of Computer Comm.,* 19, 359–371, 1996.
11. Teraoka, F. and Tokoro, M., Host migration transparency in IP networks: the VIP approach. *Proc. of ACM SIGCOMM,* ACM Press, New York, NY, 10036 (Jan.), 45–65, 1993.
12. Tobagi, F. and Kleinrock, L., Packet switching in radio channels: Part I carrier sense multiple access models and their throughput delay characteristic. *IEEE Trans. on Comm.,* 23(12), 1400–1416, 1975a.
13. Tobagi, F. and Kleinrock, L., Packet switching in radio channels: Part II the hidden terminal problem in CSMA and busy-one solution. *IEEE Trans. on Comm.,* 23(12), 1417–1433, 1975b.

Further Information

A good introduction to physical layer issues is presented in Bantz and Bauchot [1] and MAC layer issues are discussed in Chen [3]. For a discussion of network and transport layer issues, see Singh [10] and Ghai and Singh [5].

86

Wireless Data

Allen H. Levesque
Worcester Polytechnic Institute

Kaveh Pahlavan
Worcester Polytechnic Institute

86.1 Introduction

Wireless data services and systems represent a steadily growing and increasingly important segment of the communications industry. While the wireless data industry is becoming increasingly diverse, one can identify two mainstreams that relate directly to users' requirement for data services. On one hand, there are requirements for relatively low-speed data services provided to mobile users over wide geographical areas, as provided by private mobile data networks and by data services implemented on common-carrier cellular telephone networks. On the other hand, there are requirements for high-speed data services in local areas, as provided by cordless private branch exchange (PBX) systems and wireless local area networks (LANs), as well as by the emerging personal communications services (PCS). Wireless LANs are treated in Chapter 85. In this chapter we mainly address wide-area wireless data systems, commonly called *mobile data systems*, and touch upon data services to be incorporated into the emerging digital cellular systems. We also briefly address paging and messaging services.

Mobile data systems provide a wide variety of services for both business users and public safety organizations. Basic services supporting most businesses include electronic mail, enhanced paging, modem and facsimile transmission, remote access to host computers and office LANs, information broadcast services and, increasingly, Internet access. Public safety organizations, particularly law-enforcement agencies, are making increasing use of wireless data communications over traditional VHF and UHF radio dispatch networks, over commercial mobile data networks, and over public cellular telephone networks. In addition, there are wireless services supporting vertical applications that are more or less tailored to the needs of specific companies or industries, such as transaction processing, computer-aided delivery dispatch, customer service, fleet management, and emergency medical services. Work currently in progress to develop the national Intelligent Transportation System (ITS) includes the definition of a wide array of new traveler services, many of which will be supported by standardized mobile data networks.

Much of the growth in use of wireless data services has been spurred by the rapid growth of the paging service industry and increasing customer demand for more advanced paging services, as well as the desire to increase work productivity by extending to the mobile environment the suite of digital communications services readily available in the office environment. There is also a desire to make more cost-efficient use of the mobile radio and cellular networks already in common use for mobile voice communications by incorporating efficient data transmission services into these networks. The services and networks that have evolved to date represent a variety of specialized solutions and, in general, they are not interoperable with each other. As the wireless data industry expands, there is an increasing demand for an array of attractively priced standardized services and equipment accessible to mobile users over wide geographic areas. Thus, we see the growth of nationwide privately operated service networks as well as new data services built upon the first and second generation cellular telephone networks. The implementation of PCS networks in the 2-GHz bands as well as the eventual implementation of third generation (3G) wireless networks will further extend this evolution.

In this chapter we describe the principal existing and evolving wireless data networks and the related standards activities now in progress. We begin with a discussion of the technical characteristics of wireless data networks.

86.2 Characteristics of Wireless Data Networks

From the perspective of the data user, the basic requirement for wireless data service is convenient, reliable, low-speed access to data services over a geographical area appropriate to the user's pattern of daily business operation. By low speed we mean data rates comparable to those provided by standard data modems operating over the public switched telephone network (PSTN). This form of service will support a wide variety of short-message applications, such as notice of electronic mail or voice mail, as well as short file transfers or even facsimile transmissions that are not overly lengthy. The user's requirements and expectations for these types of services are different in several ways from the requirements placed on voice communication over wireless networks. In a wireless voice service, the user usually understands the general characteristics and limitations of radio transmission and is tolerant of occasional *signal fades* and brief dropouts. An overall level of acceptable voice quality is what the user expects. In a data service, the user is instead concerned with the accuracy of delivered messages and data, the time-delay characteristics of the service network, the ability to maintain service while traveling about, and, of course, the cost of the service. All of these factors are dependent on the technical characteristics of wireless data networks, which we discuss next.

Radio Propagation Characteristics

The chief factor affecting the design and performance of wireless data networks is the nature of radio propagation over wide geographic areas. The most important mobile data systems operate in various land-mobile radio bands from roughly 100 to 200 MHz, the specialized mobile radio (SMR) band around 800 MHz, and the cellular telephone bands at 824–894 MHz. In these frequency bands, radio transmission is characterized by distance-dependent field strength, as well as the well-known effects of *multipath fading*, signal shadowing, and signal blockage. The signal coverage provided by a radio transmitter, which in turn determines the area over which a mobile data receiving terminal can receive a usable signal, is governed primarily by the *power–distance relationship*, which gives signal power as a function of distance between transmitter and receiver. For the ideal case of single-path transmission in free space, the relationship between transmitted power P_t and received power P_r is given by

$$P_r/P_t = G_t G_r (\lambda/4\pi d)^2 \tag{86.1}$$

where G_t and G_r are the transmitter and receiver antenna gains, respectively, d is the distance between the transmitter and the receiver, and λ is the wavelength of the transmitted signal. In the mobile radio

environment, the power-distance relationship is in general different from the free-space case just given. For propagation over an Earth plane at distances much greater than either the signal wavelength or the antenna heights, the relationship between P_t and P_r is given by

$$P_r/P_t = G_t G_r (h_1^2 h_2^2/d^4) \tag{86.2}$$

where h_1 and h_2 are the transmitting and receiving antenna heights. Note here that the received power decreases as the fourth power of the distance rather than the square of distance seen in the ideal free-space case. This relationship comes from a propagation model in which there is a single signal reflection with phase reversal at the Earth's surface, and the resulting received signal is the vector sum of the direct line-of-sight signal and the reflected signal. When user terminals are deployed in mobile situations, the received signal is generally characterized by rapid fading of the signal strength, caused by the vector summation of reflected signal components, the vector summation changing constantly as the mobile terminal moves from one place to another in the service area. Measurements made by many researchers show that when the fast fading is averaged out, the signal strength is described by a Rayleigh distribution having a log-normal mean. In general, the power-distance relationship for mobile radio systems is a more complicated relationship that depends on the nature of the terrain between transmitter and receiver.

Various propagation models are used in the mobile radio industry for network planning purposes, and a number of these models are described in [1]. Propagation models for mobile communications networks must take account of the terrain irregularities existing over the intended service area. Most of the models used in the industry have been developed from measurement data collected over various geographic areas. A very popular model is the *Longley–Rice model* [8,14]. Many wireless networks are concentrated in urban areas. A widely used model for propagation prediction in urban areas is one usually referred to as the *Okumura–Hata model* [4,9].

By using appropriate propagation prediction models, one can determine the range of signal coverage for a base station of given transmitted power. In a wireless data system, if one knows the level of received signal needed for satisfactory performance, the area of acceptable performance can, in turn, be determined. Cellular telephone networks utilize base stations that are typically spaced 1–5 mi apart, though in some mid-town areas, spacings of 1/2 mi or less are now being used. In packet-switched data networks, higher power transmitters are used, spaced about 5–15 mi apart.

An important additional factor that must be considered in planning a wireless data system is the in-building penetration of signals. Many applications for wireless data services involve the use of mobile data terminals inside buildings, for example, for trouble-shooting and servicing computers on customers' premises. Another example is wireless communications into hospital buildings in support of emergency medical services. It is usually estimated that in-building signal penetration losses will be in the range of 15–30 dB. Thus, received signal strengths can be satisfactory in the outside areas around a building but totally unusable inside the building. This becomes an important issue when a service provider intends to support customers using mobile terminals inside buildings.

One important consequence of the rapid fading experienced on mobile channels is that errors tend to occur in bursts, causing the transmission to be very unreliable for short intervals of time. Another problem is signal dropouts that occur, for example, when a data call is handed over from one base station to another, or when the mobile user moves into a location where the signal is severely attenuated. Because of this, mobile data systems employ various error-correction and error-recovery techniques to ensure accurate and reliable delivery of data messages.

86.3 Market Issues

There are two important trends that are tending to propel growth in the use of wireless data services. The first is the rapidly increasing use of portable devices such as laptop computers, pen-pads, notebook computers, and other similar devices. Increasingly, the laptop or notebook computer is becoming a standard item of equipment for traveling professional or business person, along with the cellular telephone

and pager. This trend has been aided by the steady decrease in prices, increases in reliability, and improvements in capability and design for such devices. The second important trend tending to drive growth in wireless data services is the explosive growth in the use of the Internet. As organizations become increasingly reliant upon the Internet for their everyday operations, they will correspondingly want their employees to have convenient access to the Internet while travelling, just as they do in the office environment. Wireless data services can provide the traveler with the required network access in many situations where wired access to the public network is impractical or inconvenient. Mobile data communication services and newer messaging services discussed here provide a solution for wireless access over wide areas. Recent estimates of traffic composition indicate that data traffic now accounts for less than 1% of the traffic on wireless networks, compared to 50% for wireline networks. Therefore, the potential for growth in the wireless data market continues to be strong.

86.4 Modem Services Over Cellular Networks

A simple form of wireless data communication now in common use is data transmission using modems or facsimile terminals over analog cellular telephone links. In this form of communication, the mobile user simply accesses a cellular channel just as he would in making a standard voice call over the cellular network. The user then operates the modem or facsimile terminal just as would be done from office to office over the PSTN. A typical connection is shown in Fig. 86.1, where the mobile user has a laptop computer and portable modem in the vehicle, communicating with another modem and computer in the office. Typical users of this mode of communication include service technicians, real estate agents, and traveling sales people. In this form of communication, the network is not actually providing a data service but simply a voice link over which the data modem or fax terminal can interoperate with a corresponding data modem or fax terminal in the office or service center. The connection from the mobile telephone switching office (MTSO) is a standard landline connection, exactly the same as is provided for an ordinary cellular telephone call. Many portable modems and fax devices are now available in the market and are sold as elements of the so-called "mobile office" for the traveling business person. Law enforcement personnel are also making increasing use of data communication over cellular telephone and dispatch radio networks to gain rapid access to databases for verification of automobile registrations and drivers' licenses. Portable devices are currently available that operate at transmission rates up to 9.6 or 14.4 kb/s. Error-correction modem protocols such as MNP-10, V.34, and V.42 are used to provide reliable delivery of data in the error-prone wireless transmission environment.

In another form of mobile data service, the mobile subscriber uses a portable modem or fax terminal as already described but now accesses a modem provided by the cellular service operator as part of a *modem pool*, which is connected to the MTSO. This form of service is shown in Fig. 86.2. The modem pool might provide the user with a choice of several standard modem types. The call connection from the modem pool to the office is a digital data connection, which might be supported by any of a number of public packet data networks, such as those providing X.25 service. Here, the cellular operator is providing a special service in the form of modem pool access, and this service in general carries a higher tariff than does standard cellular telephone service, due to the operator's added investment in the modem pools.

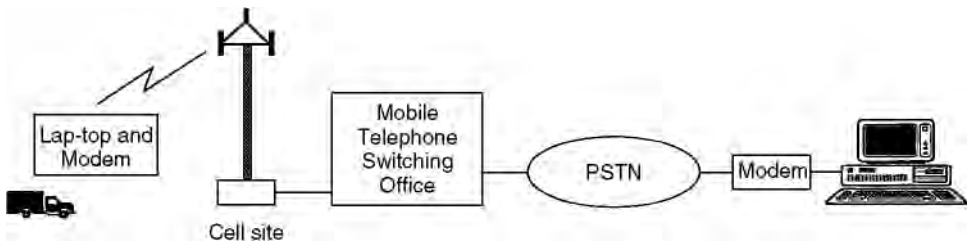


FIGURE 86.1 Modem operation over an analog cellular voice connection.

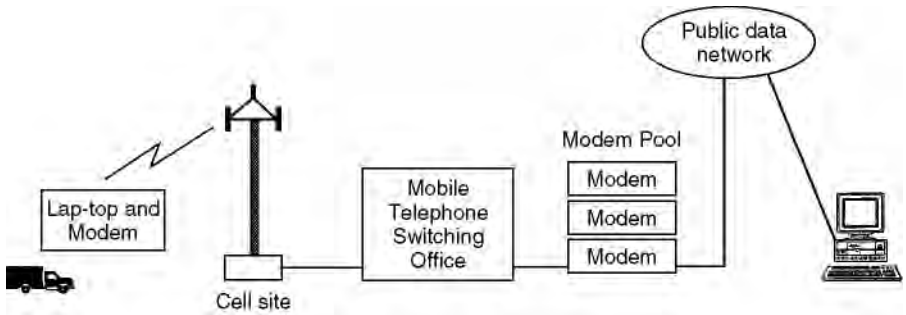


FIGURE 86.2 Cellular data service supported by modem pools in the network.

In this form of service, however, the user in the office or service center does not require a modem but instead has a direct digital data connection to the desk-top or host computer.

Each of the types of wireless data transmission just described is in effect an appliqué onto an underlying cellular telephone service and, therefore, has limitations imposed by the characteristics of the underlying voice-circuit connection. That is, the cellular segment of the call connection is a circuit-mode service, which might be cost effective if the user needs to send long file transfers or fax transmissions but might be relatively costly if only short messages are to be transmitted and received. This is because the subscriber is being charged for a circuit-mode connection, which stays in place throughout the duration of the communication session, even if only intermittent short message exchanges are needed. The need for systems capable of providing cost-effective communication of relatively short message exchanges led to the development of wireless packet data networks, which we describe next.

86.5 Packet Data and Paging/Messaging Networks

Here we describe two packet data networks that provide mobile data services to users in major metropolitan areas of the United States and briefly discuss paging/messaging networks.

ARDIS (Motient Data Network)

ARDIS is a two-way radio service developed as a joint venture between IBM and Motorola and first implemented in 1983. In mid-1994, IBM sold its interest in ARDIS to Motorola and early in 1998 ARDIS was acquired by the American Mobile Satellite (now Motient Corporation). The ARDIS network consists of four network control centers with 32 network controllers distributed through 1250 base stations in 400 cities in the U.S. The service is suitable for two-way transfers of data files of size less than 10 kilobytes, and much of its use is in support of computer-aided dispatching, such as is used by field service personnel, often while they are on customers' premises. Remote users access the system from laptop radio terminals, which communicate with the base stations. Each of the ARDIS base stations is tied to one of the 32 radio network controllers, as shown in Fig. 86.3. The backbone of the network is implemented with leased telephone lines. The four ARDIS hosts, located in Chicago, New York, Los Angeles, and Lexington, KY, serve as access points for a customer's mainframe computer, which can be linked to an ARDIS host using async, bisync, SNA, or X.25 dedicated circuits.

The operating frequency band is 800 MHz, and the RF links use separate transmit and receive frequencies spaced by 45 MHz. The system was initially implemented with an RF channel data rate 4800 b/s per 25-kHz channel, using the MDC-4800 protocol. In 1993 the access data rate was upgraded to 19.2 kb/s, using the RD-LAP protocol, which provides a user data rate of about 8000 b/s. In the same year, ARDIS implemented a nationwide roaming capability, allowing users to travel between widely separated regions without having to preregister their portable terminals in each new region. The ARDIS system architecture is cellular, with cells overlapped to increase the probability that the signal transmission from a portable

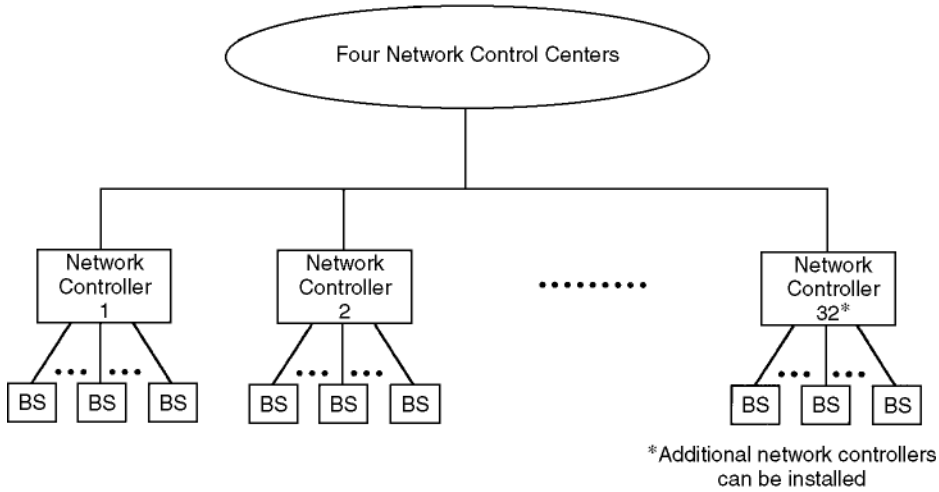


FIGURE 86.3 ARDIS network architecture.

transmitter will reach at least one base station. The base station power is 40 W, which provides line-of-sight coverage up to a radius of 10–15 miles. The portable units operate with 4 W of radiated power. The overlapping coverage, combined with designed power levels, and error-correction coding in the transmission format, ensures that the ARDIS can support portable communications from inside buildings, as well as on the street. This capability for in-building coverage is an important characteristic of the ARDIS service. The modulation technique is frequency-shift keying (FSK), the access method is frequency division multiple access (FDMA), and the transmission packet length is 256 bytes.

Although the use of overlapping coverage, almost always on the same frequency, provides reliable radio connectivity, it poses the problem of interference when signals are transmitted simultaneously from two adjacent base stations. The ARDIS network deals with this by turning off neighboring transmitters, for 0.5–1 s, when an outbound transmission occurs. This scheme has the effect of constraining overall network capacity.

The laptop portable terminals access the network using a random access method called data sense multiple access (DSMA) [11]. A remote terminal listens to the base station transmitter to determine if a "busy bit" is on or off. When the busy bit is off, the remote terminal is allowed to transmit. If two remote terminals begin to transmit at the same time, however, the signal packets may collide, and retransmission will be attempted, as in other contention-based multiple access protocols. The busy bit lets a remote user know when other terminals are transmitting and, thus, reduces the probability of packet collision.

MOBITEX (Cingular Wireless)

The MOBITEX system is a nationwide, interconnected trunked radio network developed by Ericsson and Swedish Telecom. The first MOBITEX network went into operation in Sweden in 1986, and networks have either been implemented or are being deployed in 22 countries. A MOBITEX operations association oversees the open technical specifications and coordinates software and hardware developments [6]. In the U.S., MOBITEX service was introduced by RAM Mobile Data in 1991. In 1992 Bell South Enterprises became a partner with RAM. Currently, MOBITEX data service in the U.S. is operated by Cingular Wireless, formerly known as Bell South Wireless Data, LP. The Cingular MOBITEX service now covers over 90% of the U.S. urban business population with about 2000 base stations, and it provides automatic "roaming" across all service areas. By locating its base stations close to major business centers, the system provides a degree of in-building signal coverage. Although the MOBITEX system was designed to carry both voice and data service, the U.S. and Canadian networks are used to provide data service only.

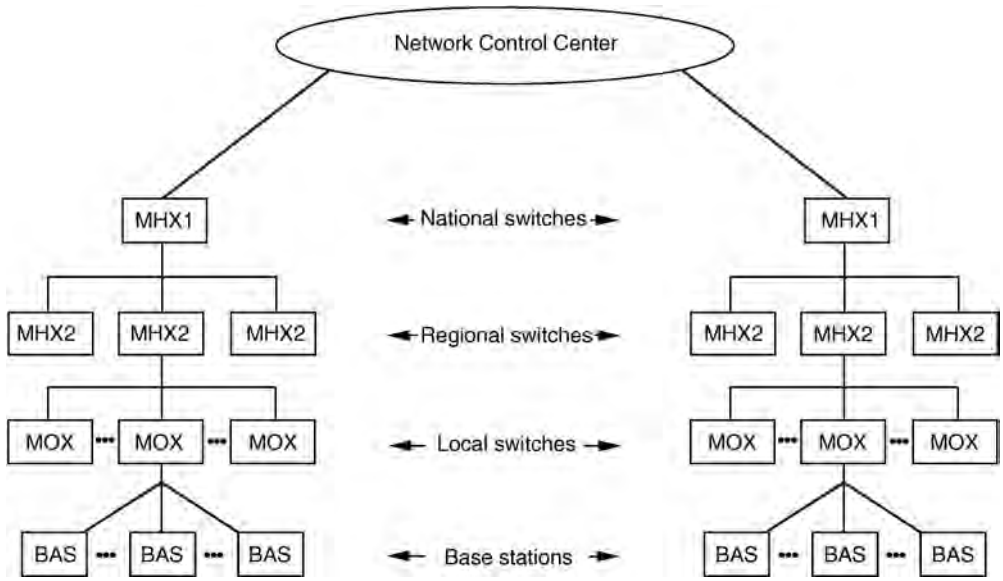


FIGURE 86.4 MOBITEX network architecture.

MOBITEX is an intelligent network with an open architecture that allows establishing virtual networks. This feature facilitates the mobility and expandability of the network [7,12].

The MOBITEX network architecture is hierarchical, as shown in Fig. 86.4. At the top of the hierarchy is the network control center (NCC), from which the entire network is managed. The top level of switching is a national switch (MHX1) that routes traffic between service regions. The next level comprises regional switches (MHX2s), and below that are local switches (MOXs), each of which handles traffic within a given service area. At the lowest level in the network, multichannel trunked-radio base stations communicate with the mobile and portable data sets. MOBITEX uses packet-switching techniques, as does ARDIS, to allow multiple users to access the same channel at the same time. Message packets are switched at the lowest possible network level. If two mobile users in the same service area need to communicate with each other, their messages are relayed through the local base station, and only billing information is sent up to the network control center.

The base stations are laid out in a grid pattern using the same frequency reuse rules as are used for cellular telephone networks. In fact, the MOBITEX system operates in much the same way as a cellular telephone system, except that handoffs are not managed by the network. That is, when a radio connection is to be changed from one base station to another, the decision is made by the mobile terminal, not by a network computer as in cellular telephone systems.

To access the network, a mobile terminal finds the base station with the strongest signal and then registers with that base station. When the mobile terminal enters an adjacent service area, it automatically re-registers with a new base station, and the user's whereabouts are relayed to the higher level network nodes. This provides automatic routing of messages bound for the mobile user, a capability known as *roaming*. The MOBITEX network also has a store-and-forward capability.

The mobile units transmit at 896 to 901 MHz and the base stations at 935 to 940 MHz. The base stations use a trunked radio design employing 2 to 30 radio channels in each service area. The system uses dynamic power setting, in the range of 100 mW–10 W for mobile units and 100 mW–4 W for portable units. The Gaussian minimum shift keying (GMSK) modulation technique is used, with $BT = 0.3$ and noncoherent demodulation. The transmission rate is 8000 b/s half-duplex in 12.5-kHz channels, and the service is suitable for file transfers up to 20 kilobytes. The MOBITEX system uses a proprietary network-layer protocol called MPAK, which provides a maximum packet size of 512 bytes and a 24-b address field. Forward-error correction, as well as retransmissions, are used to ensure the bit-error-rate

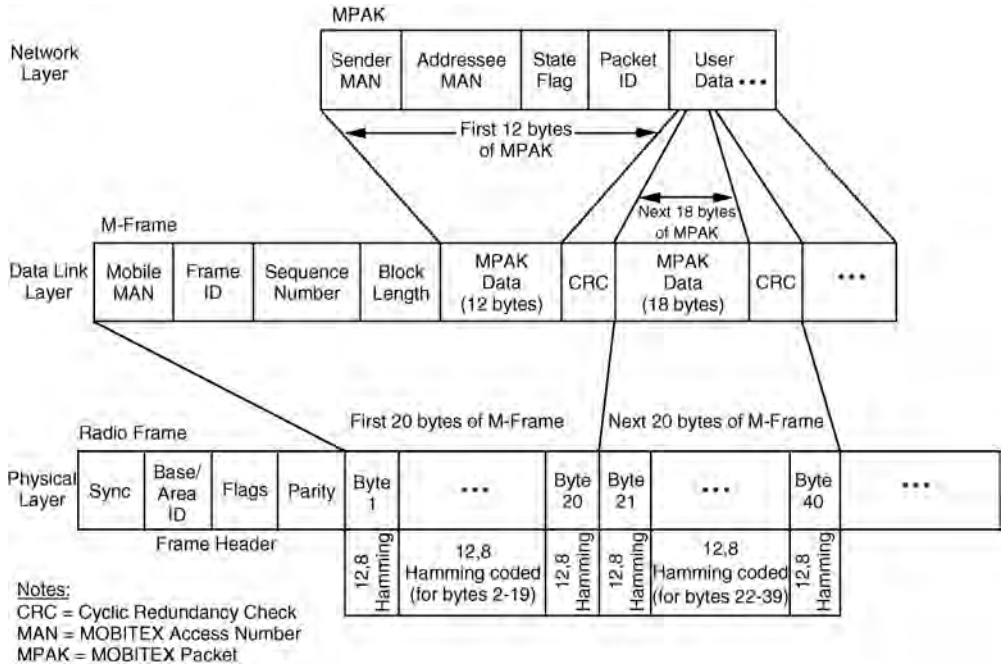


FIGURE 86.5 MOBITEX packet and frame structure at three layers of the protocol stack.

quality of delivered data packets. Figure 86.5 shows the packet structure at various layers of the MOBITEX protocol stack. The system uses the reservation-slotted ALOHA (R-S-ALOHA) random access method.

Paging and Messaging Networks

The earliest paging networks, first launched in the early 1960s, supported simple one-way tone pagers. Tone paging was soon replaced by special paging codes, beginning with numeric codes introduced in North American networks in the 1970s [19]. In 1976, European and U.S. manufacturers and paging industry representatives formed the Post Office Code Advisory Group (POCSAG), and they subsequently defined the POCSAG code. The POCSAG code supported tone, numeric, and alpha-text services and was widely adopted for paging networks. In 1985 the organization of European Post, Telephone and Telegraph administrations, known as CEPT, began work to develop a unified European paging code standard. In 1993 their recommendation for an enhanced protocol was approved by the European Telecommunications Standards Institute (ETSI) as the Enhanced Radio Message System (ERMES). The ERMES code was based upon the POCSAG code, and incorporated enhancements for multicasting and multifrequency roaming as well as for transmission of longer text messages. However, ERMES protocol did not receive widespread support, and the market demands for simpler short numeric message paging tended to prevail.

At present, a suite of alphanumeric paging protocols called FLEX™, developed by Motorola in 1993, has become the *de facto* world standard for paging services. While the FLEX protocols bear some similarity to the POCSAG and ERMES protocols, they overcome many of the weaknesses of those protocols. FLEX carries less transmission overhead than does ERMES, providing improved alphanumeric messaging capacity. The FLEX protocol is a synchronous time-slotted protocol referenced to a GPS time base. Each pager is assigned to a base frame in a set of 128 frames transmitted during a 4-minute cycle (32 frames per minute, 1.875 seconds per frame.) Fifteen FLEX cycles occur each hour and are synchronized to the start of the GPS hour. Each FLEX frame consists of a synchronization portion and a data portion. The synchronization portion consists of a sync signal and an 11-bit frame information word allowing the pager to identify the frame and cycle in which it resides. A second sync signal indicates the rate at which the data portion is transmitted.

Three data rates are provided, 1600, 3200, or 6400 bps, and the modulation used is either 2-level or 4-level frequency-shift keying (FSK). For two-way paging service, which is becoming increasingly popular in the U.S., the FLEX forward channel (in the 931 MHz band) protocol is combined with a return channel (901 MHz band) protocol called ReFLEX™.

86.6 Cellular Data Networks and Services

Cellular Digital Packet Data (CDPD)

The cellular digital packet data (CDPD) system was designed to provide packet data services as an overlay onto the existing analog cellular telephone network, which is called advanced mobile phone service (AMPS). CDPD was developed by IBM in collaboration with the major cellular carriers. Any cellular carrier owning a license for AMPS service is free to offer its customers CDPD service without any need for further licensing. A basic concept of the CDPD system is to provide data services on a noninterfering basis with the existing analog cellular telephone services using the same 30-kHz channels. This is accomplished in either of two ways. First, one or a few AMPS channels in each cell site can be devoted to CDPD service. Second, CDPD is designed to make use of a cellular channel that is temporarily not being used for voice traffic and to move to another channel when the current channel is assigned to voice service. In most of the CDPD networks deployed to date, the fixed-channel implementation is being used.

The compatibility of CDPD with the existing cellular telephone system allows it to be installed in any AMPS cellular system in North America, providing data services that are not dependent on support of a digital cellular standard in the service area. The participating companies issued release 1.0 of the CDPD specification in July 1993, and release 1.1 was issued in late 1994 [2]. At this writing (early 2002), CDPD service is implemented in many of the major market areas in the U.S. Typical applications for CDPD service include: electronic mail, field support servicing, package delivery tracking, inventory control, credit card verification, security reporting, vehicle theft recovery, traffic and weather advisory services, and a wide range of information retrieval services. Some CDPD networks support Palm Handheld devices.

Although CDPD cannot increase the number of channels usable in a cell, it can provide an overall increase in user capacity if data customers use CDPD instead of voice channels. This capacity increase results from the inherently greater efficiency of a connectionless packet data service relative to a connection-oriented service, given bursty data traffic. That is, a packet data service does not require the overhead associated with setup of a voice traffic channel in order to send one or a few data packets. In the following paragraphs we briefly describe the CDPD network architecture and the principles of operation of the system. Our discussion follows [13], closely.

The basic structure of a CDPD network (Fig. 86.6) is similar to that of the cellular network with which it shares transmission channels. Each mobile end system (M-ES) communicates with a mobile data base station (MDBS) using the protocols defined by the air-interface specification, to be described subsequently. The MDBSs are typically collocated with the cell equipment providing cellular telephone service to facilitate the channel-sharing procedures. All of the MDBSs in a service area are linked to a mobile data intermediate system (MD-IS) by microwave or wireline links. The MD-IS provides a function analogous to that of the mobile switching center (MSC) in a cellular telephone system. The MD-IS may be linked to other MD-ISs and to various services provided by end systems outside the CDPD network. The MD-IS also provides a connection to a network management system and supports protocols for network management access to the MDBSs and M-ESs in the network.

Service endpoints can be local to the MD-IS or remote, connected through external networks. A MD-IS can be connected to any external network supporting standard routing and data exchange protocols. A MD-IS can also provide connections to standard modems in the PSTN by way of appropriate modem interworking functions (modem emulators). Connections between MD-ISs allow routing of data to and from M-ESs that are roaming, that is, operating in areas outside their home service areas. These connections also allow MD-ISs to exchange information required for mobile terminal authentication, service authorization, and billing.

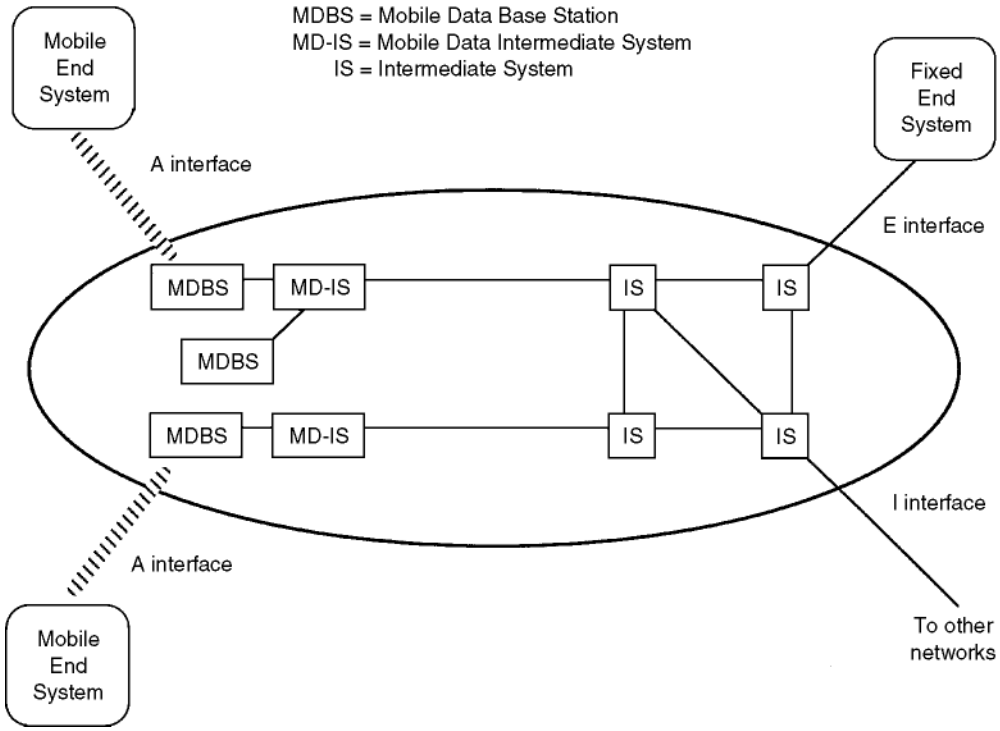


FIGURE 86.6 Cellular digital packet data network architecture.

CDPD employs the same 30-kHz channelization as used in existing AMPS cellular systems throughout North America. Each 30-kHz CDPD channel supports channel transmission rates up to 19.2 kb/s. Degraded radio channel conditions, however, will limit the actual information payload throughput rate to lower levels, and will introduce additional time delay due to the error-detection and retransmission protocols.

The CDPD radio link physical layer uses GMSK modulation at the standard cellular carrier frequencies, on both forward (base-to-mobile) and reverse (mobile-to-base) links. The Gaussian pulse-shaping filter is specified to have bandwidth-time product $B_bT = 0.5$. The specified B_bT product ensures a transmitted waveform with bandwidth narrow enough to meet adjacent-channel interference requirements, while keeping the intersymbol interference small enough to allow simple demodulation techniques. The choice of 19.2 kb/s as the channel bit rate yields an average power spectrum that satisfies the emission requirements for analog cellular systems and for dual-mode digital cellular systems.

The forward channel carries data packets transmitted by the MDBS, whereas the reverse channel carries packets transmitted by the M-ESs. In the forward channel, the MDBS forms data frames by adding standard high level data link control (HDLC) terminating flags and inserted zero bits, and then segments each frame into blocks of 274 b. These 274 b, together with an 8-b *color code* for MDBS and MD-IS identification, are encoded into a 378-b coded block using a (63, 47) Reed–Solomon code over a 64-ary alphabet. A 6-b synchronization and flag word is inserted after every 9 code symbols. The flag words are used for reverse link access control. The forward link block structure is shown in Fig. 86.7.

In the reverse channel, when an M-ES has data frames to send, it formats the data with flags and inserted zeros in the same manner as in the forward link. That is, the reverse link frames are segmented and encoded into 378-b blocks using the same Reed–Solomon code as in the forward channel. The M-ES may form up to 64 encoded blocks for transmission in a single reverse channel transmission burst. During the transmission, a 7-b transmit continuity indicator is interleaved into each coded block and is set to

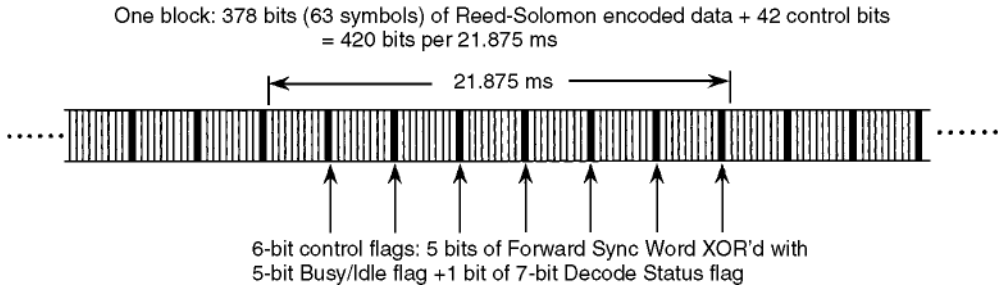


FIGURE 86.7 Cellular digital packet data forward link block structure.

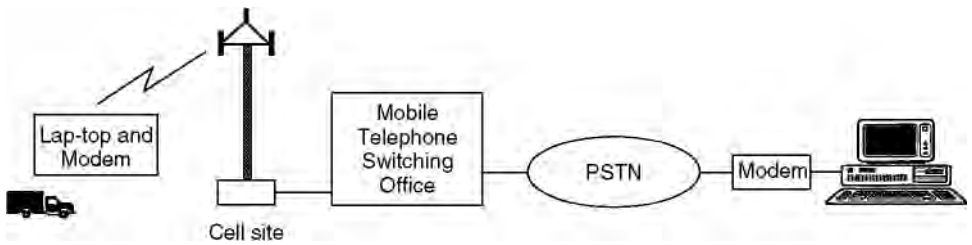


FIGURE 86.8 Cellular digital packet data reverse link block structure.

all ones to indicate that more blocks follow, or all zeros to indicate that this is the last block of the burst. The reverse channel block structure is shown in Fig. 86.8.

The media access control (MAC) layer in the forward channel is relatively simple. The receiving M-ES removes the inserted zeros and HDLC flags and reassembles data frames that were segmented into multiple blocks. Frames are discarded if any of their constituent blocks are received with uncorrectable errors.

On the reverse channel (M-ES to MDBS), access control is more complex, since several M-ESs must share the channel. CDPD uses a multiple access technique called digital sense multiple access (DSMA), which is closely related to the carrier sense multiple access/collision detection (CSMA/CD) access technique.

The network layer and higher layers of the CDPD protocol stack are based on standard ISO and Internet protocols. The CDPD specification stipulates that there be no changes to protocols above the network layer of the seven-layer ISO model, thus ensuring the compatibility of applications software used by CDPD subscribers.

The selection of a channel for CDPD service is accomplished by the radio resource management entity in the MDBS. Through the network management system, the MDBS is informed of the channels in its cell or sector that are available either as dedicated data channels or as potential CDPD channels when they are not being used for analog cellular service, depending on which channel allocation method is implemented. For the implementation in which CDPD service is to use "channels of opportunity," there are two ways in which the MDBS can determine whether the channels are in use. If a communication link is provided between the analog system and the CDPD system, the analog system can inform the CDPD system directly about channel usage. If such a link is not available, the CDPD system can use a forward power monitor ("sniffer" antenna) to detect channel usage on the analog system. Circuitry to implement this function can be built into the cell sector interface.

Another version of CDPD called circuit-switched CDPD (C-SCDPD) is designed to provide service to subscribers traveling in areas where the local cellular service provider has not implemented the CDPD service. With C-SCDPD, the subscriber establishes a standard analog cellular circuit connection to a

prescribed number, and then transmits and receives CDPD data packets over that circuit connection. The called number is a gateway that provides connection to the CDPD backbone packet network.

Digital Cellular Data Services

In response to the rapid growth in demand for cellular telephone service throughout the U.S and Canada, the Cellular Telecommunications Industry Association (CTIA) and the Telecommunications Industry Association (TIA) have developed standards for new digital cellular systems to replace the existing analog cellular system (advanced mobile phone system or AMPS). Two air-interface standards have now been published. The IS-54 standard specifies a three-slot TDMA system, and the I-95 standard specifies a CDMA spread spectrum system. In both systems, a variety of data services are provided or planned.

Following the development of the IS-95 standard for CDMA voice service, the cellular industry has worked on defining various data services to operate in the same networks. The general approach taken in the definition of IS-95 data services has been to base the services on standard data protocols, to the greatest extent possible [17]. The previously-specified physical layer of the IS-95 protocol stack was adopted for the physical layer of the data services, with an appropriate radio link protocol (RLP) overlaid. The first CDMA data service to be defined was asynchronous ("start-stop" interface) data and Group-3 facsimile. This service provides for interoperability with many standard PSTN data modems as well as standard office fax machines. This service is in the category of *circuit-mode service*, since a circuit connection is first established, just as with a voice call, and the circuit connection is maintained until the user disconnects.

Following the standardization of the asynchronous data service, the industry defined a service that carries packet-formatted data over a CDMA circuit connection. It is important to note that this is not a true packet-data service over the radio link, since the full circuit connection is maintained regardless of how little packet data is transmitted. One potential application for this type of service is to provide subscribers with CDPD access from a CDMA network.

It is recognized that in order to make use more efficient, it will be highly desirable to provide a contention-based packet data service in CDMA cellular networks. This is currently a subject of study in CDMA data services standardization groups.

In parallel with the CDMA data services efforts, another TIA task group, TR45.3.2.5, has defined standards for digital data services for the TDMA digital cellular standard IS-54 [15,18]. As with the IS-95 data services effort, initial priority was given to standardizing circuit-mode asynchronous data and Group-3 facsimile services [16].

86.7 Other Developing Standards

Terrestrial Trunked Radio (TETRA)

As has been the case in North America, there is interest in Europe in establishing fixed wide-area standards for mobile data communications. Whereas the Pan-European standard for digital cellular, termed Global Systems for Mobile communications (GSM), will provide an array of data services, data will be handled as a circuit-switched service, consistent with the primary purpose of GSM as a voice service system. Therefore, the European Telecommunications Standards Institute (ETSI) began developing a public standard in 1988 for trunked radio and mobile data systems, and this standardization process continues today. The standards, which are now known generically as Terrestrial Trunked Radio (TETRA) (formerly Trans-European Trunked Radio), were made the responsibility of the ETSI RES 6 subtechnical committee [3]. In 1996, the TETRA standardization activity was elevated within RES-6 with the creation of project TETRA.

TETRA is being developed as a family of standards. One branch of the family is a set of radio and network interface standards for trunked voice (and data) services. The other branch is an air-interface standard optimized for wide-area packet data services for both fixed and mobile subscribers and supporting standard network access protocols. Both versions of the standard will use a common physical layer, based

TABLE 86.1 Characteristics and Parameters of Five Mobile Data Services

System:	ARDIS	MOBITEX	CDPD	IS-95[b]	TETRA[b]
Frequency band					
Base to mobile, (MHz).	(800 band,	935–940[a]	869–894	869–894	(400 and
Mobile to base, (MHz).	45-kHz sep.)	896–901	824–849	824–849	900 Bands)
RF channel spacing	25 kHz (U.S.)	12.5 kHz	30 kHz	1.25 MHz	25 kHz
Channel access/ multiuser access	FDMA/DSMA	FDMA/ dynamic-R-S-ALOHA	FDMA/DSMA	FDMA/ CDMA-SS	FDMA/ DSMA & SAPR[c]
Modulation method	FSK, 4-FSK	GMSK	GMSK	4-PSK/DSSS	π/4-QDPSK
Channel bit rate, kb/s	19.2	8.0	19.2	9.6	36
Packet length	Up to 256 bytes (HDLC)	Up to 512 bytes	24–928 b	(Packet service-TBD)	192 b (short) 384 b (long)
Open architecture	No	Yes	Yes	Yes	Yes
Private or Public Carrier	Private	Private	Public	Public	Public
Service Coverage	Major metro. areas in U.S.	Major metro. areas in U.S.	All AMPS areas	All CDMA cellular areas	European trunked radio
Type of coverage	In-building & mobile	In-building & mobile	Mobile	Mobile	Mobile

[a] Frequency allocation in the U.S. In the U.K., 380–450 MHz band is used.
[b] IS-95 and TETRA data services standardization in progress.
[c] Slotted-ALOHA packet reservation.

on π/4 differential quadrature phase shift keying (π/4-DQPSK) modulation operating at a channel rate of 36 kb/s in each 25-kHz channel. The composite data rate of 36 kb/s comprises four 9 kb/s user channels multiplexed in a TDMA format. The TETRA standard provides both connection-oriented and connectionless data services, as well as mixed voice and data services.

TETRA has been designed to operate in the frequency range from VHF (150 MHz) to UHF (900 MHz). The RF carrier spacing in TETRA is 25 kHz. In Europe, harmonized bands have been designated in the frequency range 380–400 MHz for public safety users. It is expected that commercial users will adopt the 410–430 MHz band. The Conference of European Posts and Telecommunications Administrations (CEPT) has made additional recommendations for use in the 450–470 MHz and 870–876/915–921 MHz frequency bands.

Table 86.1 compares the chief characteristics and parameters of the wireless data services described.

86.8 Conclusions

Mobile data radio systems have grown out of the success of the paging-service industry and the increasing customer demand for more advanced services. The growing use of portable, laptop, and palmtop computers and other data services will propel a steadily increasing demand for wireless data services. Today, mobile data services provide length-limited wireless connections with in-building penetration to portable users in metropolitan areas. The future direction is toward wider coverage, higher data rates, and capability for wireless Internet access.

References

1. Bodson, D., McClure, G.F., and McConoughey, S. R., Eds., *Land-Mobile Communications Engineering,* Selected Reprint Ser., IEEE Press, New York, 1984.
2. CDPD Industry Coordinator. Cellular Digital Packet Data Specification, Release 1.1, November 1994, Pub. CDPD, Kirkland, WA, 1994.
3. Haine, J.L., Martin, P.M., and Goodings, R.L.A., A European standard for packet-mode mobile data, *Proceedings of Personal, Indoor, and Mobile Radio Conference (PIMRC'92),* Boston, MA. Pub. IEEE, New York, 1992.

4. Hata, M., Empirical formula for propagation loss in land-mobile radio services. *IEEE Trans. on Vehicular Tech.,* 29(3), 317–325, 1980.
5. International Standards Organization (ISO), Protocol for providing the connectionless-mode network service. Pub. ISO 8473, 1987.
6. Khan, M. and Kilpatrick, J., MOBITEX and mobile data standards. *IEEE Comm. Maga.,* 33(3), 96– 101, 1995.
7. Kilpatrick, J.A., Update of RAM Mobile Data's packet data radio service. *Proceedings of the 42nd IEEE Vehicular Technology Conference (VTC'92),* Denver, CO, 898–901, Pub. IEEE, New York, 1992.
8. Longley, A. G. and Rice, P. L., Prediction of tropospheric radio transmission over irregular terrain. A computer method—1968, Environmental Sciences and Services Administration Tech. Rep. ERL 79-ITS 67, U.S. Government Printing Office, Washington, D.C., 1968.
9. Okumura, Y., Ohmori, E., Kawano, T., and Fukuda, K., Field strength and its variability in VHF and UHF land-mobile service. *Review of the Electronic Communication Laboratory,* 16, 825–873, 1968.
10. Pahlavan, K. and Levesque, A.H., Wireless data communications. *Proceedings of the IEEE,* 82(9), 1398–1430, 1994.
11. Pahlavan, K. and Levesque, A.H., *Wireless Information Networks,* J. Wiley & Sons, New York, 1995.
12. Parsa, K., The MOBITEX packet-switched radio data system. *Proceedings of the Personal, Indoor and Mobile Radio Conference,* (*PIMRC'92*), Boston, MA, 534–538. Pub. IEEE, New York, 1992.
13. Quick, R.R., Jr. and Balachandran, K., Overview of the cellular packet data (CDPD) system. *Proceedings of the Personal, Indoor and Mobile Radio Conference,* (*PIMRC'93*), Yokohama, Japan, 338–343. Pub. IEEE, New York, 1993.
14. Rice, P.L., Longley, A.G., Norton, K.A., and Barsis, A.P., Transmission loss predictions for tropospheric communication circuits. National Bureau of Standards, Tech. Note 101, Boulder, CO, 1967.
15. Sacuta, A., Data standards for cellular telecommunications—a service-based approach. *Proceedings of the 42nd IEEE Vehicular Technology Conference,* Denver CO, 263–266. Pub. IEEE, New York, 1992.
16. Telecommunications Industry Association. Async data and fax. Project No. PN-3123, and Radio link protocol 1. Project No. PN-3306, Nov. 14. Issued by TIA, Washington, D.C., 1994.
17. Tiedemann, E., Data services for the IS-95 CDMA standard. Presented at Personal, Indoor and Mobile Radio Conf. PIMRC'93. Yokohama, Japan, 1993.
18. Weissman, D., Levesque, A.H., and Dean, R.A., Interoperable wireless data. *IEEE Comm. Mag.,* 31(2), 68–77, 1993.
19. Beaulieu, M., *Wireless Internet Applications and Architecture,* Addison-Wesley, Boston, MA, 2002.

Further Information

Reference [11] provides a comprehensive survey of the wireless data field as of mid-1994. Reference [19] covers the wireless data field up to 2001, with emphasis on current and emerging applications for wireless data networking technology. The monthly journals *IEEE Communications Magazine* and *IEEE Personal Communications Magazine,* and the bimonthly journal *IEEE Transactions on Vehicular Technology* report advances in many areas of mobile communications, including wireless data. For subscription information contact: IEEE Service Center, 445 Hoes Lane, P.O. Box 1331, Piscataway, NJ, 08855-1131. Phone (800) 678-IEEE.

87

Wireless ATM: Interworking Aspects

Melbourne Barton
Telcordia Technologies

Matthew Cheng
Mobilink Telecom

Li Fung Chang
Mobilink Telecom

87.1 Introduction

The ATM Forum's **wireless asynchronous transfer mode (WATM)** Working Group (WG) is developing specifications intended to facilitate the use of ATM technology for a broad range of wireless network access and interworking scenarios, both public and private. These specifications are intended to cover the following two broad WATM application scenarios:

- *End-to-End WATM*—This provides seamless extension of ATM capabilities to mobile terminals, thus providing ATM virtual channel connections (VCCs) to wireless hosts. For this application, high data rates are envisaged, with limited coverage, and transmission of one or more ATM cells over the air.
- *WATM Interworking*—Here the fixed ATM network is used primarily for high-speed transport by adding mobility control in the ATM infrastructure network, without changing the non-ATM air interface protocol. This application will facilitate the use of ATM as an efficient and cost-effective infrastructure network for next generation non-ATM wireless access systems, while providing a smooth migration path to seamless end-to-end WATM.

This chapter focuses on the ATM interworking application scenario. It describes various interworking and non-ATM wireless access options and their requirements. A generic **personal communications services (PCS)**[1]-to-ATM interworking scenario is described which enumerates the architectural features, protocol reference models, and signalling issues that are being addressed for mobility support in the ATM

[1]The term PCS is being used in a generic sense to mean emerging digital wireless systems, which support mobility in microcellular and other environments. It is currently defined in ANSI T1.702-1995 as "a set of capabilities that allows some combination of terminal mobility, personal mobility, and service profile management."

infrastructure network. Evolution strategies intended to eventually provide end-to-end WATM capabilities and a methodology to consistently support a range of quality of service (QoS) levels on the radio link are also described.

87.2 Background and Issues

ATM is the switching and multiplexing standard for **broadband integrated services digital network (BISDN)**, which will ultimately be capable of supporting a broad range of applications over a set of high capacity multiservice interfaces. ATM holds out the promise of a single network platform that can simultaneously support multiple bandwidths and latency requirements for fixed access and wireless access services without being dedicated to any one of them. In today's wireline ATM network environment, the **user network interface (UNI)** is fixed and remains stationary throughout the connection lifetime of a call. The technology to provide fixed access to ATM networks has matured. Integration of fixed and wireless access to ATM will present a cost-effective and efficient way to provide future tetherless multimedia services, with common features and capabilities across both wireline and wireless network environments. Early technical results [19,20,25] have shown that standard ATM protocols can be used to support such integration and extend mobility control to the subscriber terminal by incorporating wireless specific layers into the ATM user and control planes.

Integration of wireless access features into wireline ATM networks will place additional demands on the fixed network infrastructure due primarily to the additional user data and signalling traffic that will be generated to meet future demands for wireless multimedia services. This additional traffic will allow for new signalling features including registration, call delivery, and handoff during the connection lifetime of a call. Registration keeps track of a wireless user's location, even though the user's communication link might not be active. Call delivery, establishes a connection link to/from a wireless user with the help of location information obtained from registration. The registration and call delivery functions are referred to as **location management**. Handoff is the process of switching (rerouting) the communication link from the old coverage area to the new coverage area when a wireless user moves during active communication. This function is also referred to as **mobility management**.

In June 1996 the ATM Forum established a WATM WG to develop requirements and specifications for WATM. The WATM standards are to be compatible with ATM equipment adhering to the (then) current ATM Forum specifications. The technical scope of the WATM WG includes development of: (1) **radio access layer (RAL)** protocols for the physical (PHY), medium access control (MAC), and data link control (DLC) layers; (2) wireless control protocols for radio resource management; (3) mobile protocol extensions for ATM (mobile ATM) including handoff control, routing considerations, location management, traffic and QoS control, and wireless network management; and (4) wireless interworking functions for mapping between non-ATM wireless access and ATM signalling and control entities. Phase-1 WATM specifications are being developed for short-range, high-speed, end-to-end WATM devices using wireless terminals that operate in the 5 GHz frequency band. Operating speeds will be up to 25 Mb/s, with a range of 30 m–50 m indoor and 200 m–300 m outdoor. The European Telecommunications Standards Institute (ETSI) Broadband Radio Access Networks (BRAN) project is developing the RAL for Phase-1 WATM specifications using the **High Performance Radio LAN (HIPERLAN)** functional requirements. The ATM Forum plans to release the Phase-1 WATM specifications by the second quarter of 1999.

There are a number of emerging wireless access systems (including digital cellular, PCS, legacy LANs based on the IEEE 802.11 standards, satellite, and IMT-2000 systems), that could benefit from access to, and interworking with, the fixed ATM network. These wireless systems are based on different access technologies and require development of different **interworking functions (IWFs)** at the wireless access network and fixed ATM network boundaries to support WATM interworking. For example, a set of network interfaces have already been identified to support PCS access to the fixed ATM network infrastructure, without necessarily modifying the PCS air interface protocol to provide end-to-end ATM capabilities [4,6,28]. The WATM WG might consider forming sub-working groups which could work in parallel to identify other network interfaces and develop IWF specifications for each of (or each subset of)

the wireless access options that are identified. These WATM interworking specifications would be available in the Phase-2 and later releases of the WATM standards.

Some service providers and network operators see end-to-end WATM as a somewhat limited service option at this time because it is being targeted to small enterprise networks requiring high-speed data applications, with limited coverage and low mobility. On the other hand, WATM interworking can potentially support a wider range of services and applications, including low-speed voice and data access, without mandatory requirements to provide over-the-air transmission of ATM cells. It will allow for wider coverage, possibly extending to macrocells, while supporting higher mobility. WATM interworking will provide potential business opportunities, especially for public switched telephone network (PSTN) operators and service providers, who are deploying emerging digital wireless technologies such as PCS. Existing wireless service providers (WSPs) with core network infrastructures in place can continue to use them while upgrading specific network elements to provide ATM transport. On the other hand, a new WSP entrant without such a network infrastructure can utilize the public (or private) ATM transport network to quickly deploy the WATM interworking service, and not be burdened with the cost of developing an overlay network. If the final goal is to provide end-to-end WATM services and applications, then WATM interworking can provide an incremental development path.

87.3 Wireless Interworking With Transit ATM Networks

Figure 87.1 shows one view of the architectural interworking that will be required between public/private wireless access networks and the fixed ATM network infrastructure. It identifies the network interfaces where modifications will be required to allow interworking between both systems. A desirable objective in formulating WATM specifications for this type of wireless access scenario should be to minimize modifications to the transit ATM network and existing/emerging wireless access system specifications. This objective can be largely met by limiting major modifications to the network interfaces between the boundaries of the transit ATM network and public/private wireless networks, and where possible, adopting existing network standard processes (i.e., SS7, IS-41, MAP, AIN.3, Q.931, Q.932, Q.2931, etc.) to minimize development costs and upgrades to existing service providers' network infrastructure. Development of standard network interfaces that allow interworking of a reasonable subset of non-ATM digital wireless access systems with the fixed ATM network infrastructure ensure that:

- Large-scale revisions and modifications are not necessary to comply with later versions of the WATM specifications to accommodate other emerging digital wireless access systems that do not require end-to-end ATM connectivity
- WATM systems are supported by open interfaces with a rich set of functionality to provide access to both ATM and non-ATM wireless access terminal devices

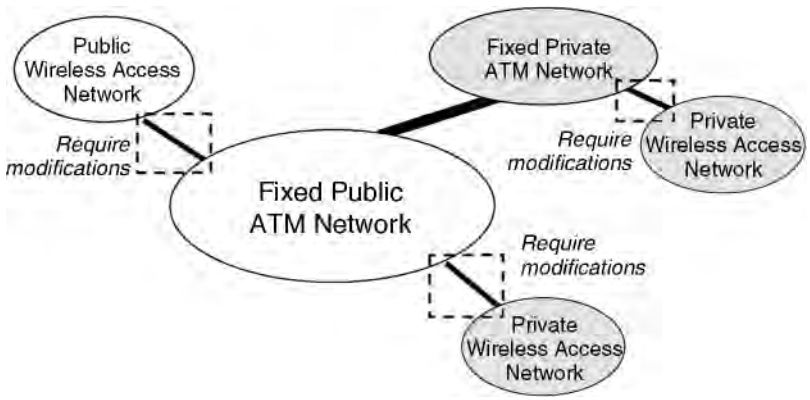


FIGURE 87.1 Wireless ATM interworking architecture.

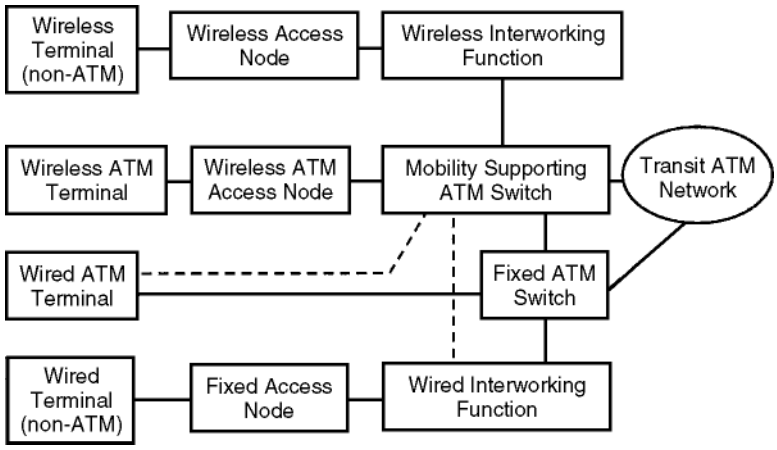


FIGURE 87.2 Wireless and wireline system integration with transit ATM Network.

- WATM services can reach a much larger potential market including those markets providing traditional large-scale support for existing voice services and vertical voice features

Integrated Wireless–Wireline ATM Network Architecture

Figure 87.2 shows one example of a mature, multifunctional ATM transport network platform, which provides access to fixed and mobile terminals for wide-area coverage. Four distinct network interfaces are shown supporting: (1) fixed access with non-ATM terminal, (2) fixed access with ATM terminal, (3) wireless access with non-ATM terminal (WATM interworking), and (4) wireless access with ATM terminal (end-to-end WATM).

International Telecommunications Union (ITU) and ATM Forum standard interfaces either exist today or are being developed to support fixed access to ATM networks through various network interfaces. These include frame relay service (FRS) UNI, cell relay service (CRS) UNI, circuit emulation service (CES) UNI, and switched multimegabit data service (SMDS) subscriber NI (SNI). The **BISDN intercarrier interface (B-ICI)** specification [1] provides examples of wired IWFs that have been developed for implementation above the ATM layer to support intercarrier service-specific functions developed at the network nodes, and distributed in the public ATM/BISDN network. These distributed, service-specific functions are defined by B-ICI for FRS, CRS, CES, and SMDS. Examples of such functions include ATM cell conversion, clock recovery, loss of signal and alarm indication detection, virtual channel identifier (VCI) mapping, access class selection, encapsulation/mapping, and QoS selection. In addition to the B-ICI, the **private network–network interface (PNNI)** specification [2] defines the basic call control signalling procedures (e.g., connection setup and release) in private ATM networks. It also has capabilities for autoconfiguration, scalable network hierarchy formation, topology information exchange, and dynamic routing.

On the wireless access side, existing ITU specifications provide for the transport of wireless services on public ATM networks (see, e.g., [10,11]). For example, if the ATM UNI is at the mobile switching center (MSC), then message transfer part (MTP) 1 and 2 would be replaced by the PHY and ATM layers, respectively, the broadband ISDN user part (BISUP) replaced by MTP 3, and a common channel signalling (CCS) interface deployed in the ATM node. BISUP is used for ATM connection setup and any required feature control. If the ATM UNI is at the base station controller (BSC), then significant modifications are likely to be required. Equipment manufacturers have not implemented, to any large degree, the features that are available with the ITU specifications. In any case, these features are not sufficient to support the WATM scenarios postulated in this chapter.

The two sets of WATM scenarios postulated in this chapter are shown logically interfaced to the ATM network through a **mobility-enabled ATM (ME-ATM)** switch. This enhanced ATM access switch will have capabilities to support mobility management and location management. In addition to supporting handoff and rerouting of ATM connections, it will be capable of locating a mobile user anywhere in the network.

It might be desirable that functions related to mobility not be implemented in standard ATM switches so that network operators and service providers are not required to modify their ATM switches in order to accommodate WATM and related services. A feasible strategy that has been proposed is to implement mobility functions in servers (i.e., service control modules or SCMs) that are logically separated from the ATM switch. In these servers, all mobility features, service creation logic, and service management functions will be implemented to allow logical separation of ATM switching and service control from mobility support, service creation, and management. The open network interface between the ATM access switch and SCM will be standardized to enable multivendor operation. It would be left up to the switch manufacturer to physically integrate the SCM into a new ATM switching fabric, or implement the SCM as a separate entity.

Wireless Access Technology Options

Figure 87.3 identifies various digital wireless systems that could be deployed to connect mobile terminals to the transit ATM network through IWFs that have to be specified. The main emphasis is on developing mobility support in the transit ATM infrastructure network to support a range of wireless access technologies. Significant interests have been expressed, through technical contributions and related activities in the WATM WG, in developing specifications for the IWFs shown in Fig. 87.3 to allow access to the fixed ATM network through standard network interfaces. The main standardization issues that need to be addressed for each wireless access system are described below.

PCS

This class of digital wireless access technologies include the low-tier PCS systems as described in Cox [8]. Digital cellular systems which are becoming known in the U.S. as high-tier PCS, especially when implemented in the 1.8-1.9 GHz PCS frequency bands, are addressed in the Digital Cellular section below. The PCS market has been projected to capture a significant share of the huge potential revenues to be

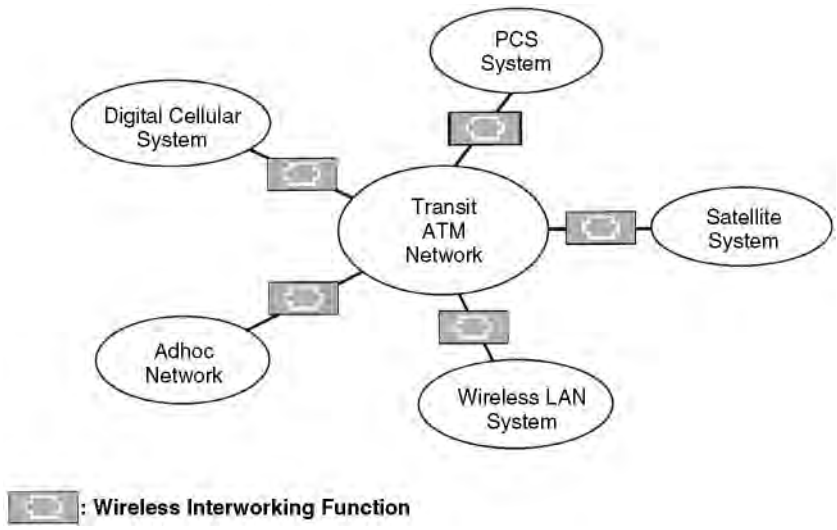


FIGURE 87.3 Various wireless access technologies supported by wireless ATM interworking.

generated by business and residential customers. In order to provide more flexible, widespread, tetherless portable communications than can be provided by today's limited portable communications approaches, low-power exchange access radio needs to be integrated with network intelligence provided by the wireline network.

Today, the network for supporting PCS is narrowband ISDN, along with network intelligence based on advanced intelligent network (AIN) concepts, and a signalling system for mobility/location management and call control based on the signalling system 7 (SS7) network architecture. It is expected that the core network will evolve to BISDN/ATM over time with the capability to potentially integrate a wide range of network services, both wired and wireless, onto a single network platform. Furthermore, there will be no need for an overlay network for PCS. In anticipation of these developments, the WATM WG included the development of specifications and requirements for PCS access to, and interworking with, the ATM transit network in its charter and work plan. To date, several contributions have been presented at WATM WG meetings that have identified some of the key technical issues relating to PCS-to-ATM interworking. These include (1) architectures, (2) mobility and location management signalling, (3) network evolution strategies, and (4) PCS service scenarios.

Wireless LAN

Today, wireless LAN (WLAN) is a mature technology. WLAN products are frequently used as LAN extensions to access areas of buildings with wiring difficulties and for cross-building interconnect and nomadic access. Coverage ranges from tens to a few hundreds of meters, with data rates ranging from hundreds of kb/s to more than 10 Mb/s. Several products provide 1 or 2 Mb/s. ATM LAN products provide LAN emulation services over the connection-oriented (CO) ATM network using various architectural alternatives [24]. In this case, the ATM network provides services that permit reuse of existing LAN applications by stations attached directly to an ATM switch, and allow interworking with legacy LANs. Furthermore, the increasing importance of mobility in data access networks and the availability of more usable spectrum are expected to speed up the evolution and adoption of mobile access to WLANs. Hence, it is of interest to develop wireless LAN products that have LAN emulation capabilities similar to wireline ATM LANs.

The ETSI BRAN project is developing the HIPERLAN RAL technology for wireless ATM access and interconnection. It will provide short-range wireless access to ATM networks at approximately 25 Mb/s in the 5 GHz frequency band. HIPERLAN is an ATM-based wireless LAN technology that will have end-to-end ATM capabilities. It does not require the development of an IWF to provide access to ATM. A previous version of HIPERLAN (called HIPERLAN I), which has been standardized, supports data rates from 1–23 Mb/s in the 5 GHz band using a non-ATM RAL [27]. This (and other non-ATM HIPERLAN standards being developed) could benefit from interworking with the backbone ATM as a means of extending the marketability of these products in the public domain such as areas of mass transit and commuter terminals.

Several proposals have been submitted to IEEE 802.11 to provide higher speed extensions of current IEEE 802.11 systems operating in the 2.4 GHz region and the development of specifications for new systems operating in the 5 GHz frequency band. The proposed 2.4 GHz extensions support different modulation schemes, but are interoperable with the current IEEE 802.11 low rate PHY and are fully compliant with the IEEE 802.11 defined MAC. The 5 GHz proposals are not interoperable with the current 2.4 GHz IEEE 802.11 systems. One of the final three 5 GHz proposals being considered is based on orthogonal frequency division multiplexing (OFDM), or multicarrier modulation. The other two are single carrier proposals using offset QPSK (OQPSK)/offset QAM (OQAM), and differential pulse position modulation (DPPM). The OFDM proposal has been selected. With 16-QAM modulation on each subcarrier and rate-3/4 convolutional coding, the OFDM system has a peak data rate capability of 30 Mb/s.

It is clear that a whole range of WLAN systems either exist today, or are emerging, that are not based on ATM technology, and therefore cannot provide seamless access to the ATM infrastructure network. The development of IWF specifications that allow these WLANs to provide such access through standard network interfaces will extend the range of applications and service features provided by WLANs. In Pahlavan [17],

a number of architectural alternatives to interconnect WLAN and WATM to the ATM and/or LAN backbone are discussed, along with service scenarios and market and product issues.

Digital Cellular

Digital cellular mobile radio systems include the 1.8–1.9 GHz (high-tier PCS) and the 800–900 MHz systems that provide high-mobility, wide-area coverage over macrocells. Cellular radio systems at 800–900 MHz have evolved to digital in the form of Global System for Mobile Communications (GSM) in Europe, Personal Digital Cellular (PDC) in Japan, and IS-54 Time Division Multiple Access (TDMA) and IS-95 Code Division Multiple Access (CDMA) in the U.S. The capabilities in place today for roaming between cellular networks provide for even wider coverage. Cellular networks have become widespread, with coverage extending beyond some national boundaries. These systems integrate wireless access with large-scale mobile networks having sophisticated intelligence to manage mobility of users.

Cellular networks (e.g., GSM) and ATM networks are evolving somewhat independently. The development of IWFs that allow digital cellular systems to utilize ATM transport will help to bridge the gap. Cellular networks have sophisticated mechanisms for authentication and handoff, and support for rerouting through the home network. In order to facilitate the migration of cellular systems to ATM transport, one of the first issues that should be addressed is that of finding ways to enhance the basic mobility functions already performed by them for implementation in WATM. Contributions presented at WATM WG meetings have identified some basic mobility functions of cellular mobile networks (and cordless terminal mobility) which might be adopted and enhanced for WATM interworking. These basic mobility functions include rerouting scenarios (including path extension), location update, call control, and authentication.

Satellite

Satellite systems are among the primary means of establishing connectivity to untethered nodes for long-haul radio links. This class of applications has been recognized as an essential component of the National Information Infrastructure (NII) [16]. Several compelling reasons have been presented in the ATM Forum's WATM WG for developing standard network interfaces for **satellite ATM (SATATM)** networks. These include (1) ubiquitous wide area coverage, (2) topology flexibility, (3) inherent point-to-multipoint and broadcast capability, and (4) heavy reliance by the military on this mode of communications. Although the geostationary satellite link represents only a fraction of satellite systems today, WATM WG contributions that have addressed this interworking option have focused primarily on geostationary satellites. Some of these contributions have also proposed the development of WATM specifications for SATATM systems having end-to-end ATM capabilities.

Interoperability problems between satellite systems and ATM networks could manifest themselves in at least four ways.

1. Satellite links operate at much higher bit error rates (BERs) with variable error rates and bursty errors.
2. The approximately 540 ms round trip delay for geosynchronous satellite communications can potentially have adverse impacts on ATM traffic and congestion control procedures.
3. Satellite communications bandwidth is a limited resource, and might be incompatible with less bandwidth efficient ATM protocols.
4. The high availability rates (at required BERs) for delivery of ATM (e.g., 99.95%) is costly, hence the need to compromise between performance relating to availability levels and cost.

A number of experiments have been performed to gain insights into these challenges [21]. The results can be used to guide the development of WATM specifications for the satellite interworking scenario. Among the work items that have been proposed for SATATM access using geostationary satellites links are (1) identification of requirements for RAL and mobile ATM functions; (2) study of the impact of satellite delay on traffic management and congestion control procedures; (3) development of requirements and specifications for bandwidth efficient operation of ATM speech over satellite links; (4) investigation of various WATM access scenarios; and (5) investigation of frequency spectrum availability issues.

One interesting SATATM application scenario has been proposed to provide ATM services to multiuser airborne platforms via satellite links for military and commercial applications. In this scenario, a satellite constellation is assumed to provide contiguous overlapping coverage regions along the flight path of the airborne platforms. A set of interworked ground stations form the mobile enhanced ATM network that provides connectivity between airborne platforms and the fixed terrestrial ATM network via bent-pipe satellite links. Key WATM requirements for this scenario have been proposed, which envisage among other things modifications to existing PNNI signalling and routing mechanisms to allow for mobility of (ATM) switches.

In related work, the Telecommunications Industry Association (TIA) TR34.1 WG has also proposed to develop technical specifications for SATATM networks. Three ATM network architectures are proposed for bent-pipe satellites and three others for satellites with onboard ATM switches [23]. Among the technical issues that are likely to be addressed by TR34.1 are protocol reference models and architecture specifications, RAL specifications for SATATM, and support for routing, rerouting, and handoff of active connections. A liaison has been established between the ATM Forum and TR34.1, which is likely to lead to the WATM WG working closely with TR34.1 to develop certain aspects of the TR34.1 SATATM specifications.

Ad Hoc Networks

The term ad hoc network is used to characterize wireless networks that do not have dedicated terminals to perform traditional BS and/or wireless resource control functions. Instead, any mobile terminal (or a subset of mobile terminals) can be configured to perform these functions at any time. Ad hoc networking topologies have been investigated by wireless LAN designers, and are part of the HIPERLAN and IEEE 802.11 wireless LAN specifications [13]. As far as its application to WATM is concerned, low-cost, plug-and-play, and flexibility of system architecture are essential requirements. Potential application service categories include rapidly deployable networks for government use (e.g., military tactical networks, rescue missions in times of natural disasters, law enforcement operations, etc.), ad hoc business conferencing devoid of any dedicated coordinating device, and ad hoc residential network for transfer of information between compatible home appliances.

There are some unique interworking features inherent in ad hoc networks. For example, there is a need for location management functions not only to identify the location of terminals but also to identify the current mode of operation of such terminals. Hence, the WATM WG is considering proposals to develop separate requirements for ad hoc networks independent of the underlying wireless access technology. It is likely that requirements will be developed for an ad hoc RAL, mobility management signalling functions, and location management functions for supporting interworking of ad hoc networks that provide access to ATM infrastructure networks.

The range of potential wireless access service features and wireless interworking scenarios presented in the five wireless access technologies discussed above is quite large. For example, unlicensed satellite systems could provide 32 kb/s voice, and perhaps up to 10 Mb/s for wireless data services. The main problem centers around the feasibility of developing specifications for IWFs to accommodate the range of applications and service features associated with the wireless access options shown in Fig. 87.3. The WATM WG might consider forming sub-working groups which would work in parallel to develop the network interface specifications for each (or a subset of each) of the above non-ATM wireless access options.

87.4 The PCS-to-ATM Interworking Scenario

This section presents a more detailed view of potential near-term and longer-term architectures and reference models for PCS-to-ATM interworking. A signalling link evolution strategy to support mobility is also described. Mobility and location management signalling starts with the current CCS network, which is based on the SS7 protocol and 56 kb/s signalling links, and eventually migrates to ATM signalling. The system level architectures, and the mobility and location management signalling issues addressed in

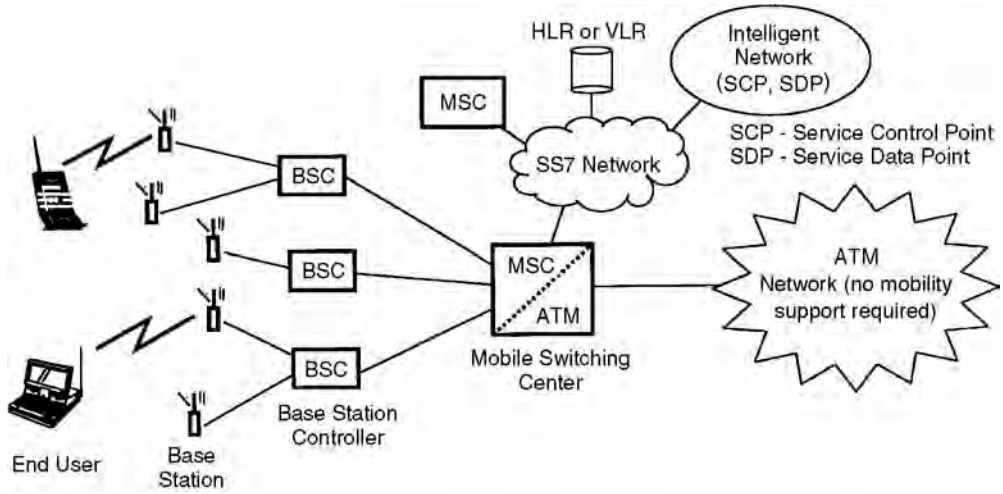


FIGURE 87.4 A near-term architecture for PCS-to-ATM interworking.

this section serve to illustrate the range of technical issues that need to be addressed for the other WATM interworking options.

Architecture and Reference Model

The near-term approach for the PCS-to-ATM interworking scenario is shown in Fig. 87.4. The near-term approach targets existing PCS providers with network infrastructures in place, who wish to continue to use them while upgrading specific network elements (e.g., MSCs) to provide ATM transport for user data. The existing MSCs in the PCS network are upgraded to include fixed ATM interfaces. ATM is used for transport and switching of user data, while mobility/location management and call control signalling is carried by the SS7 network. No mobility support is required in the ATM network. Synchronization problems might develop because different traffic types relating to the same call may traverse the narrowband SS7 network and the broadband ATM network and arrive at the destination in an uncoordinated manner. Upgrading the SS7 network to broadband SS7 (e.g., T1 speeds or higher) should partially alleviate this potential problem.

A longer-term approach for PCS-to-ATM interworking has been proposed in some technical contributions to the WATM WG. This is illustrated in Fig. 87.5, together with the protocol stacks for both data and signalling. The ATM UNI is placed at the BSC, which acts as the PCS-to-ATM gateway. The ATM network carries both data and signalling. ATM cells are not transmitted over the PCS link. Communications between the BS and BSC are specific to the PCS access network, and could be a proprietary interface. Compared with existing/emerging BSCs, additional protocol layer functionality is required in the BSC to provide (1) transfer/translation and/or encapsulation of PCS **protocol data units (PDUs)**, (2) ATM to wireless PDU conversion, and (3) a limited amount of ATM multiplexing/demultiplexing capabilities.

The BSC is connected to the ATM network through a ME-ATM access switch instead of a MSC. The ME-ATM switch provides switching and signalling protocol functions to support ATM connections together with mobility and location management. These functions could be implemented in servers that are physically separate from, but logically connected to, the ME-ATM switch. The WATM WG is expected to formulate requirements and specifications for these mobile-specific functions. On the other hand, an IWF can be introduced between the BSC and the ME-ATM switch shown in Fig. 87.5. In this case, the UNI is between the IWF and the ME-ATM switch, and another standard interface (not necessarily ATM) can be used to connect the BSC to the IWF. The BSC then requires no modification, but a new entity (i.e., the IWF) is required. The IWF will perform protocol conversion and it may serve multiple BSCs.

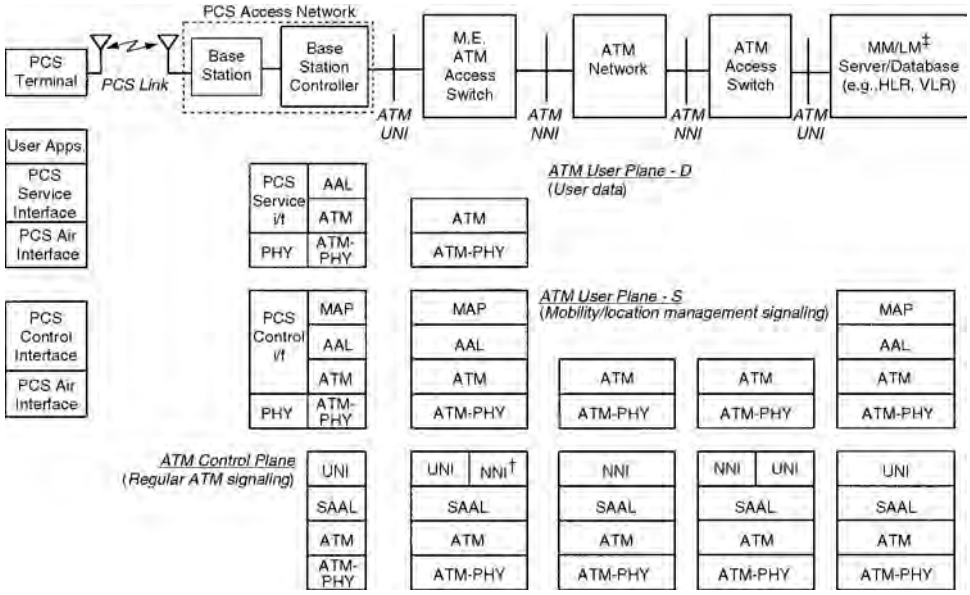


FIGURE 87.5 A longer-term architecture and reference model PCS-to-ATM interworking.

† NNI here is used in a generic sense, to refer to both public and private networks. In public networks, there will be an additional MTP3 layer between the NNI and SAAL layers.

‡ MM—Mobility Management, LM—Location Management.

A unique feature of this architecture is that modifications to network entities to allow for interworking of PCS with the transit ATM network are only required at the edges of the ATM and wireless access networks, i.e., to the BSC and the ME-ATM access switch. In order to minimize the technical impact of mobility on existing/emerging transit ATM networks and PCS specifications, an initial interworking scenario is envisaged in which there are no (or only very limited) interactions between PCS and ATM signalling entities. PCS signalling would be maintained over the air interface, traditional ATM signalling would be carried in the control plane (C-Plane), and PCS signalling would be carried over the ATM network as user traffic in the user plane (U-Plane). In the long term, this architecture is expected to eventually evolve to an end-to-end ATM capability.

On the ATM network side, mobility and location management signalling is implemented in the user plane as the **mobile application part** (**MAP**) layer above the ATM adaptation layer (AAL), e.g., AAL5. The MAP can be based on the MAP defined in the existing PCS standards (e.g., IS41-MAP or GSM-MAP), or based on a new set of mobile ATM protocols. The U-plane is logically divided into two parts, one for handling mobility and location management signalling messages and the other for handling traditional user data. This obviates the need to make modifications to the ATM UNI and NNI signalling protocols, currently being standardized. The MAP layer also provides the necessary end-to-end reliability management for the signalling because it is implemented in the U-Plane, where reliable communication is not provided in the lower layers of the ATM protocol stack as is done in the signalling AAL (SAAL) layer. The MAP functionality can be distributed at individual BSCs or ME-ATM access switches or centralized in a separate server to further reduce modifications to existing implementations. The setup and release of ATM connections are still handled by the existing ATM signalling layer. This approach allows easy evolution to future wireless ATM architectures, which will eventually integrate the ATM and PCS network segments to form a homogeneous network to carry both mobility/location management and traditional ATM signalling in the C-Plane. Issues relating to mobility support for PCS interworking with ATM networks are addressed in more detail in Cheng [7].

Signalling Link Evolution

Signalling will play a crucial role in supporting end-to-end connections (without location restrictions) in an integrated WATM network infrastructure. The signalling and control message exchanges required to support mobility will occur more frequently than in wireline ATM networks. Today's CCS network, which is largely based on the SS7 protocol and 56 kb/s signalling links, will not be able to support the long-term stringent service and reliability requirements of WATM. Two broad deployment alternatives have been proposed for migrating the CCS/SS7 network (for wireline services) to using the broadband signalling platform [22].

- Migration to high-speed signalling links using the narrowband signalling platform supported by current digital transmission (e.g., DS1) facilities. The intent of this approach is to support migration to high-speed signalling links using the existing CCS infrastructure with modified signalling network elements. This would allow the introduction of high-speed (e.g., 1.5 Mb/s) links with possible minimal changes in the CCS network. One option calls for modifications of existing MTP2 procedures while maintaining the current protocol layer structure. Another option replaces MTP2 with some functionality of the ATM/SAAL link layer, while continuing to use the same transport infrastructure as DS1 and transport messages over the signalling links in variable length signal units delimited by flags.
- Alternative 2 supports migration of signalling links to a broadband/ATM signalling network architecture. Signaling links use both ATM cells and the ATM SAAL link layer, with signalling message transported over synchronous optical network (SONET) or existing DS1/DS3 facilities at rates of 1.5 Mb/s or higher. This alternative is intended primarily to upgrade the current CCS network elements to support an ATM-based interface, but could also allow for the inclusion of signalling transfer point (STP) functions in the ATM network elements to allow for internetworking between ATM and existing CCS networks.

The second alternative provides a good vehicle for the long-term goal of providing high-speed signalling links on a broadband signalling platform supported by the ATM technology. One signalling network configuration and protocol option is the extended Q.93B signalling protocol over ATM in associated mode [22]. The PSTN's CCS networks are currently quasi-associated signalling networks. Q.93B is primarily intended for point-to-point bearer control in user-to-network access, but can also be extended for (link-to-link) switch-to-switch and some network-to-network applications. Standards activities to define high-speed signalling link characteristics in the SS7 protocol have been largely finalized. Standards to support SS7 over ATM are at various stages of completion. These activities provide a good basis for further evolving the SS7 protocol to provide the mobility and location management features that will be required to support PCS (and other wireless systems) access to the ATM network.

The functionality required to support mobility in cellular networks is currently defined as part of the MAP. Both IS-41 and GSM MAPs are being evolved to support PCS services with SS7 as the signalling transport protocol [14]. Quite similar to the two alternatives described above, three architectural alternatives have been proposed for evolving today's IS-41 MAP on SS7 to a future modified (or new) IS-41 on ATM signalling transport platform [28]. They are illustrated in Fig. 87.6. In the first (or near-term) approach, user data is carried over the ATM network, while signalling is carried over existing SS7 links. The existing SS7 network can also be upgraded to broadband SS7 network (e.g., using T1 links) to alleviate some of the capacity and delay constraints in the narrowband SS7 network. This signalling approach can support the near-term PCS-to-ATM interworking described in the previous subsection. In the second (or midterm) approach, a hybrid-mode operation is envisaged, with the introduction of broadband SS7 network elements into the ATM network. This results in an SS7-over-ATM signalling transport platform. Taking advantage of the ATM's switching and routing capabilities, the MTP3 layer could also be modified to utilize these capabilities and eliminate the B-STP functionality from the network. In the third phase (or long-term approach), ATM replaces the SS7 network with a unified network for both signalling and user data. No SS7 functionality exists in this approach. Here, routing of signalling messages is completely determined by the ATM layer, and the MAP may be implemented in a

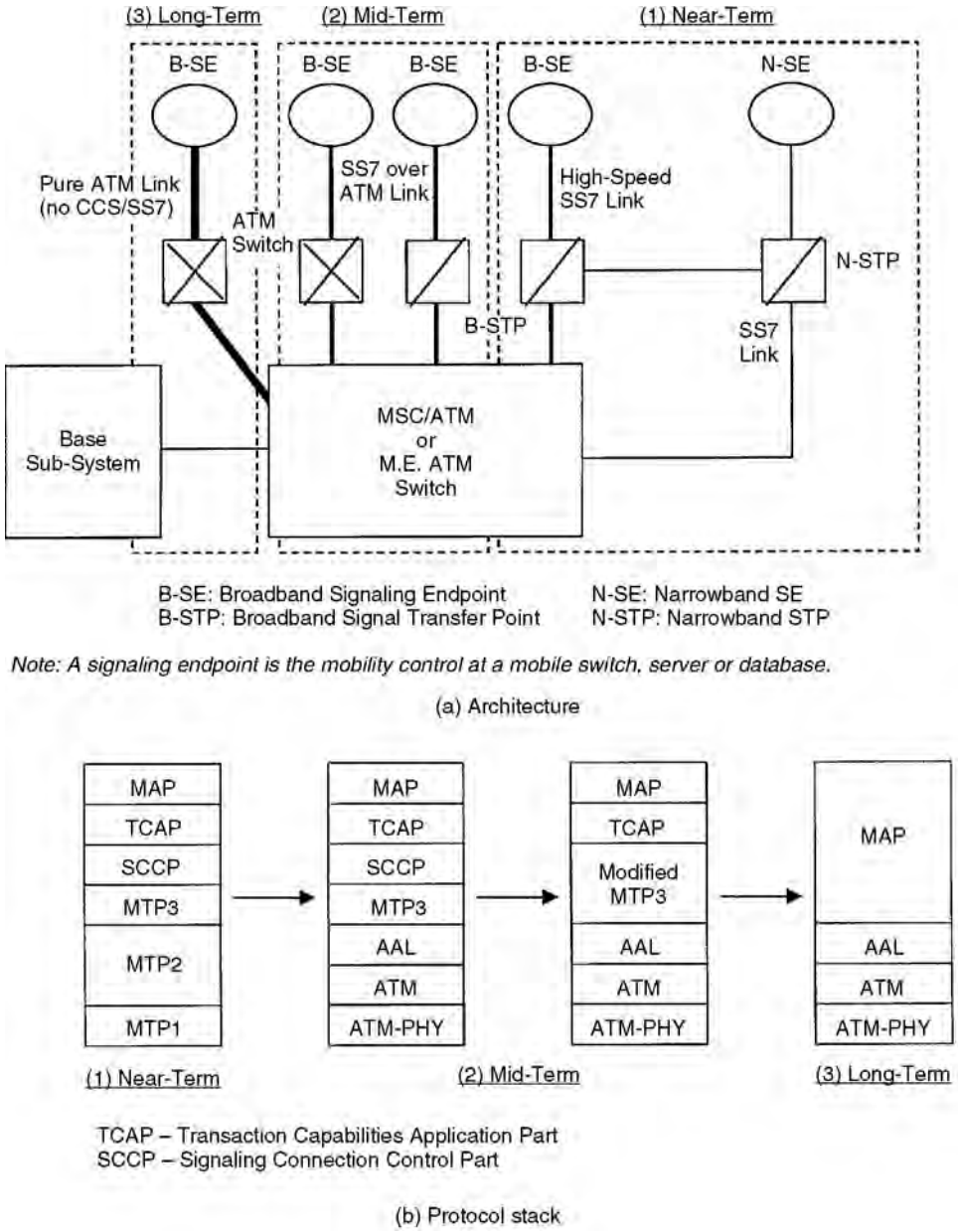


FIGURE 87.6 Signalling link evolution for mobility/location management over a public ATM network.

format other than the transaction capability application part (TCAP), so that the unique features of ATM are best utilized. The longer-term PCS-to-ATM interworking approach is best supported for this signalling approach. The performance of several signalling protocols for PCS mobility support for this long-term architecture is presented in Cheng [6,7]. The above discussion mainly focuses on public networks. In private networks, there are no existing standard signalling networks. Therefore, the third approach can be deployed immediately in private networks.

There are several ways to achieve the third signalling approach in ATM networks. The first approach is to overlay another "network layer" protocol (e.g., IP) over the ATM network, but this requires the management of an extra network. The second approach is to enhance the current ATM network-network interface

(e.g., B-ICI for public networks and PNNI for private networks) to handle the new mobility/location management signalling messages and information elements, but this requires modifications in the existing ATM network. The third approach is to use dedicated channels (PVCs or SVCs) between the mobility control signalling points. This does not require any modifications in existing ATM specifications. However, signalling latency may be high in the case of SVCs, and a full mesh of PVCs between all mobility control signalling points is difficult to manage. The fourth approach is to use the generic functional protocols (e.g., connection oriented-bearer independent (CO-BI) or connectionless-BI (CL-BI) transport mechanism) defined in ITU-T's Q.2931.2 [12]. This cannot be done in existing ATM networks without modifications, but these functions are being included in the next version of PNNI (PNNI 2.0) to provide a generic support for supplementary services. There are also technical contributions to the ATM Forum proposing "connectionless ATM" [26], which attempt to route ATM cells in a connectionless manner by using the routing information obtained through the PNNI protocol. However, the "connectionless ATM" concept is still being debated.

87.5 QoS Support

One immediate impact of adding mobility to an otherwise fixed ATM infrastructure network is the need to manage the changing QoS levels that are inherent in a mobile environment due to the vagaries of the wireless link, the need for rerouting of traffic due to handoff, and available bandwidth, etc. Dynamic QoS negotiation and flow control will be required to flexibly support QoS guarantees for multimedia service applications that are likely to be encountered in this environment. QoS provisioning is based on the notion that the wireless channel is likely to demand more stringent measures than the fixed ATM network to support end-to-end QoS. QoS metrics include (1) throughput, (2) delay sensitivity, (3) loss sensitivity, and (4) BER performance. Issues relating to end-to-end QoS provisioning in multimedia wireless networks are discussed in Naghshineh [15] and articles therein. Here, the focus is on BER maintenance in the context of PCS-to-ATM interworking using forward error correction (FEC) at the radio PHY layer. This is the first step toward developing a hybrid automatic repeat request (ARQ)/FEC protocol for error control of the wireless link, with FEC at the PHY supplemented by ARQ at the DLC layer. A comparison of commonly used FEC and ARQ techniques and their potential application to WATM is presented in Ayanoglu [3].

One adaptive FEC coding scheme that has been proposed for PCS-to-ATM interworking is based on the use of **rate-compatible punctured convolution** (**RCPC**), punctured Bose-Chaudhuri-Hocquenghem (BCH) or Reed-Solomon (RS) coding at the wireless PHY layer to provide unequal error protection of the wireless PDU [5]. These coding schemes can support a broad range of QoS levels consistent with the requirements of multimedia services, minimize the loss of information on the wireless access segment, and prevent misrouting of cells on the fixed ATM network. Code rate puncturing is a procedure used to periodically discard a set of predetermined coded bits from the sequence generated by an encoder for the purposes of constructing a higher rate code. With the rate-compatibility restriction, higher rate codes are embedded in the lower rate codes, allowing for continuous code rate variation within a data frame.

An example set of three wireless PDU formats that might be appropriate for a PCS system that provides access to the ATM network is shown in Table 87.1. It is desirable to establish a tight relationship between the wireless PDU and wireline ATM cell to minimize incompatibilities between them. This will reduce

TABLE 87.1 Examples of Wireless Protocol Data Unit (PDU) Formats

PDU type	PDU header (bits)	Information payload (bits)	PDU trailer (bits)	PDU size (bits)
PDU-1	24	128	—	152
PDU-2	40	256	8	304
PDU-3	56	384	16	456

Note: Information payloads are limited to submultiples of a 48-byte ATM cell information payload.

the complexity of the IWF at the PCS-to-ATM gateway by limiting the amount of processing required for protocol conversion. The wireless PDU can be tightly coupled to the wireline ATM cell in two ways.

1. Encapsulate the equivalent of a full 48-byte ATM information payload, along with wireless-specific overhead (and a full/compressed ATM header, if required by the mobile terminal), in the wireless PDU (e.g., PDU-3).
2. Transmit a submultiple of 48-byte ATM information payload, along with wireless-specific overhead (and a compressed ATM header, if required by the mobile terminal), in the wireless PDU (e.g., PDU-1 and PDU-2).

If the second option is used, then the network has to decide whether to send partially filled ATM cells over the ATM infrastructure network, or wait until enough wireless PDUs arrive to fill an ATM cell.

The performance of RCPC and punctured BCH (and RS) codes have been evaluated in terms of the decoded bit BER [5] on a Rayleigh fading channel. For RCPC coding, the Viterbi upper bound on the decoded BER is given by:

$$P_b \leq \frac{1}{P}\sum_{d=d_f}^{\infty} b_d P_d \tag{87.1}$$

where P is the puncturing period, β_d is the weight coefficient of paths having distance d, P_d is the probability of selecting the wrong path, and d_f is the free distance of the RCPC code. Closed-form expressions for P_d are presented in Hagenauer [9] for a flat Rayleigh fading channel, along with tables for determining β_d. Since the code weight structure of BCH codes is known only for a small subset of these codes, an upper BER performance bound is derived in the literature independent of the structure. Assume that for a t-error correcting BCH(n,k) code, a pattern of i channel errors ($i > t$) will cause the decoded word to differ from the correct word in $i + t$ bit positions. For flat Rayleigh fading channel conditions and differential quadrature phase-shift keying (DQPSK) modulation, the decoded BER for the BCH(n, k) code is [18]:

$$P_b = \frac{2}{3}\sum_{i=t+1}^{n} \frac{i+t}{n}(1-P_s)^{n-1}P_s^i \tag{87.2}$$

where P_s is the raw symbol error rate (SER) on the channel. An upper bound on the decoded BER for a t-error correcting RS(n,k) code with DQPSK signalling on a flat Rayleigh fading channel is similar in form to Eq. (87.2). However, P_s should be replaced with the RS-coded digit error rate P_e, and the unit of measure for the data changed from symbols do digits. A transmission system that uses the RS code [from GF(2^m)][2] with M-ary signalling generates $r = m/\log_2 M$ symbols per m-bit digit. For statistically independent symbol errors, $P_e = 1 - (1 - P_s)^r$. For the binary BCH(n,k) code, $n = 2^m - 1$.

Table 87.2 shows performance results for RCPC and punctured BCH codes that are used to provide adaptive FEC on the PHY layer of a simulated TDMA-based PCS system that accesses the fixed ATM network. The FEC codes provide different levels of error protection for the header, information payload, and trailer of the wireless PDU. This is particularly useful when the wireless PDU header contains information required to route cells in the fixed ATM network, for example. Using a higher level of protection for the wireless PDU header increases the likelihood that the PDU will reach its destination, and not be misrouted in the ATM network.

The numerical results show achievable average code rates for the three example PDU formats in Table 87.1. The PCS system operates in a microcellular environment at 2 GHz, with a transmission bit rate of 384 kb/s, using DQPSK modulation. The channel is modeled as a time-correlated Rayleigh fading channel. The PCS transmission model assumes perfect symbol and frame synchronization, as well as

[2] GF(2^m) denotes the Galois Field of real numbers from which the RS code is constructed. Multiplication and addition of elements in this field are based on modulo-2 arithmetic, and each RS code word consists of m bits.

TABLE 87.2 Average Code Rates for RCPC Coding[a] and Punctured BCH Coding[b] for a 2-GHz TDMA-Based PCS System with Access to an ATM Infrastructure Network

	Average code rate (10^{-3} BER for information payload)			Average code rate (10^{-6} BER for information payload)		
Coding scheme	PDU-1	PDU-2	PDU-3	PDU-1	PDU-2	PDU-3
RCPC: no diversity	0.76	0.77	0.78	0.61	0.62	0.63
RCPC: 2-branch diversity	0.91	0.93	0.93	0.83	0.84	0.85
Punctured BCH: no diversity	0.42	0.44	0.51	0.34	0.39	0.46
Punctured BCH: 2-branch diversity	0.83	0.81	0.87	0.75	0.76	0.82

[a] with soft decision decoding and no channel state information at the receiver.
[b] with and without 2-branch frequency diversity.

perfect frequency tracking. Computed code rates are shown with and without the use of diversity combining. All overhead blocks in the wireless PDUs are assumed to require a target BER of 10^{-9}. On the other hand, the information payload has target BERs of 10^{-3} and 10^{-6}, which might be typical for voice and data, respectively. Associated with this target BERs is a design goal of 20 dB for the SNR. The mother code rate for the RCPC code is $R = 1/3$, the puncturing period $P = 8$, and the memory length $M = 6$. For BCH coding, the parameter $m \geq 8$.

The numerical results in Table 87.2 show the utility of using code rate puncturing to improve the QoS performance of the wireless access segment. The results for punctured RS coding are quite similar to those for punctured BCH coding. Adaptive PHY layer FEC coding can be further enhanced by implementing an ARQ scheme at the DLC sublayer, which is combined with FEC to form a hybrid ARQ/FEC protocol [9] to supplement FEC at the PHY layer. This approach allows adaptive FEC to be distributed between the wireless PHY and DLC layers.

87.6 Conclusions

The ATM Forum is developing specifications intended to facilitate the use of ATM technology for a broad range of wireless network access and interworking scenarios, both public and private. These specifications are intended to cover requirements for seamless extension of ATM to mobile devices and mobility control in ATM infrastructure networks to allow interworking of non-ATM wireless terminals with the fixed ATM network. A mobility-enhanced ATM network that is developed from specifications for WATM interworking may be used in near-term cellular/PCS/satellite/wireless LAN deployments, while providing a smooth migration path to the longer-term end-to-end WATM application scenario. It is likely to be cost-competitive with other approaches that adopt non-ATM overlay transport network topologies.

This chapter describes various WATM interworking scenarios where the ATM infrastructure might be a public (or private) transit ATM network, designed primarily to support broadband wireline services. A detailed description of a generic PCS-to-ATM architectural interworking scenario is presented, along with an evolution strategy to eventually provide end-to-end WATM capabilities in the long term. One approach is described for providing QoS support using code rate puncturing at the wireless PHY layer, along with numerical results. The network architectures, protocol reference models, signalling protocols, and QoS management strategies described for PCS-to-ATM interworking can be applied, in varying degrees, to the other WATM interworking scenarios described in this chapter.

Defining Terms

BISDN intercarrier interface (B-ICI): A carrier-to-carrier public interface that supports multiplexing of different services such as SMDS, frame relay, circuit emulation, and cell relay services.

Broadband integrated services digital network (BISDN): A cell-relay-based information transfer technology upon which the next-generation telecommunications infrastructure is to be based.

High-Performance Radio LAN (HIPERLAN): Family of standards being developed by the European Telecommunications Standards Institute (ETSI) for high-speed wireless LANs, to provide short-range and remote wireless access to ATM networks and for wireless ATM interconnection.

Interworking functions (IWFs): A set of network functional entities that provide interaction between dissimilar subnetworks, end systems, or parts thereof, to support end-to-end communications.

Location management: A set of registration and call delivery functions.

Mobile application part (MAP): Application layer protocols and processes that are defined to support mobility services such as intersystem roaming and handoffs.

Mobility-enabled ATM (ME-ATM): An ATM switch with additional capabilities and features to support location and mobility management.

Mobility management: The handoff process associated with switching (rerouting) of the communication link from the old coverage area to the new coverage area when a wireless user moves during active communication.

Personal communications services (PCS): Emerging digital wireless systems which support mobility in microcellular and other environments, which have a set of capabilities that allow some combination of terminal mobility, personal mobility, and service profile management.

Protocol data units (PDUs): The physical layer message structure used to carry information across the communications link.

Private network–network interface (PNNI): The interface between two private networks.

Radio access layer (RAL): A reference to the physical, medium access control, and data link control layers of the radio link.

Rate-compatible punctured convolution (RCPC): Periodic discarding of predetermined coded bits from the sequence generated by a convolutional encoder for the purposes of constructing a higher rate code. The rate-compatibility restriction ensures that the higher rate codes are embedded in the lower rate codes, allowing for continuous code rate variation to change from low to high error protection within a data frame.

Satellite ATM (SATATM): A satellite network that provides ATM network access to fixed or mobile terminals, high-speed links to interconnect fixed or mobile ATM networks, or form an ATM network in the sky to provide user access and network interconnection services.

User network interface (UNI): A standardized interface providing basic call control functions for subscriber access to the telecommunications network.

Wireless asynchronous transfer mode (WATM): An emerging wireless networking technology that extends ATM over the wireless access segment, and/or uses the ATM infrastructure as a transport network for a broad range of wireless network access scenarios, both public and private.

References

1. ATM. B-ICI Specification, V 2.0. ATM Forum, 1995.
2. ATM. Private Network-Network Interface (PNNI) Specification, Version 1.0. ATM Forum, 1996.
3. Ayanoglu, E., et al., Wireless ATM: Limits, challenges, and protocols. *IEEE Personal Commun.,* 3(4), 18–34, Aug. 1996.
4. Barton, M., Architecture for wireless ATM networks. *PIMRC'95.* 778–782, Sep. 1995.
5. Barton, M., Unequal error protection for wireless ATM applications. *GLOBECOM'96,* 1911–1915, Nov. 1996.
6. Cheng, M., Performance comparison of mobile assisted network controlled, and mobile-controlled hand-off signalling for TDMA PCS with an ATM backbone. *ICC'97.* Jun. 1997a.
7. Cheng, M., Rajagopalan, S., Chang, L.F., Pollini, G.P., and Barton, M., PCS mobility support over fixed ATM networks. *IEEE Commun. Mag.,* 35(11), 82–92, Nov. 1997b.
8. Cox, D.C., Wireless personal communications: what is it? *IEEE Personal Commun. Mag.,* 2(2), 20–35, Apr. 1995.
9. Hagenauer, J., Rate-compatible punctured convolution codes (RCPC codes) and their applications. *IEEE Trans. Commun.,* COM-36(4), 389–400, Apr. 1988.

10. ITU. Message Transfer Part Level 3 Functions and Messages Using the Service of ITU-T Recommendations, Q.2140. International Telecommunications Union-Telecommunications Standardization Sector, Geneva, Switzerland. TD PL/11–97. 1995a.
11. ITU. B-ISDN Adaptation Layer—Service Specific Coordination Function for Signaling at the Network Node Interface (SCCF) at NNI. ITU-T Q.2140. International Telecommunications Union, Telecommunications Standardization Sector, Geneva, Switzerland. 1995b.
12. ITU. Digital Subscriber Signaling System No. 2—Generic Functional Protocol: Core Functions. ITU-T Recommendation Q.2931.2. International Telecommunications Union, Telecommunications Standardization Sector, Geneva, Switzerland. 1996.
13. LaMaire, R.O., Krishna, A., Bhagwat, P., and Panian, J., Wireless LANs and mobile networking: standards and future directions. *IEEE Commun. Mag.,* 34(8), 86–94, Aug. 1996.
14. Lin, Y.B. and Devries, S.K., PCS network signalling using SS7. *IEEE Commun. Mag.,* 2(3), 44–55, Jun. 1995.
15. Naghshineh, M. and Willebeek-LeMair, M., End-to-end QoS provisioning in multimedia wireless/mobile networks using an adaptive framework. *IEEE Commun. Mag.,* 72–81, Nov. 1997.
16. NSTC. *Strategic Planning Document—Information and Communications,* National Science and Technology Council. 10, Mar. 1995.
17. Pahlavan, K., Zahedi, A., and Krishnamurthy, P., Wideband local wireless: wireless LAN and wireless ATM. *IEEE Commun. Mag.,* 34–40, Nov. 1997.
18. Proakis, J.G., *Digital Communications.* McGraw-Hill, New York. 1989.
19. Raychaudhuri, D. and Wilson, N.D., ATM-based transport architecture for multiservices wireless personal communication networks. *IEEE J. Select. Areas Commun.,* 12(8), 1401–1414, Oct. 1994.
20. Raychaudhuri, D., Wireless ATM networks: architecture, system design, and prototyping. *IEEE Personal Commun.,* 3(4), 42–49, Aug. 1996.
21. Schmidt, W.R., et al., Optimization of ATM and legacy LAN for high speed satellite communications. Transport Protocols for High-Speed Broadband Networks Workshop. *GLOBECOM'96.* Nov. 1996.
22. SR. Alternatives for Signaling Link Evolution. Bellcore Special Report. *Bellcore SR-NWT-002897.* (1), Feb. 1994.
23. TIA. TIA/EIA Telecommunications Systems Bulletin (TSB)—91. Satellite ATM Networks: Architectures and Guidelines. Telecommunications Industry Association. TIA/EIA/TSB-91. Apr. 1998.
24. Truong, H.L. et al., LAN emulation on an ATM network. *IEEE Commun. Mag.,* 70–85, May 1995.
25. Umehira, M. et al., An ATM wireless system for tetherless multimedia services. *ICUPC'95.* Nov. 1995.
26. Veeraraghavan, M., Pancha, P., and Eng, K.Y., Connectionless Transport in ATM Networks. ATM Forum Contribution. ATMF/97-0141, 9–14, Feb. 1997.
27. Wilkinson, T. et al., A report on HIPERLAN standardization. *Intl. J. Wireless Inform. Networks.* 2(2), 99–120, 1995.
28. Wu, T.H. and Chang, L.F., Architecture for PCS mobility management on ATM transport networks. *ICUPC'95.* 763–768, Nov. 1995.

Further Information

Information supplementing the wireless ATM standards work may be found in the ATM Forum documents relating to the Wireless ATM Working Group's activities (web page http://www.atmforum.com). Special issues on wireless ATM have appeared in the August 1996 issue of *IEEE Personal Communications,* the January 1997 issue of the *IEEE Journal on Selected Areas in Communications,* and the November 1997 issue of *IEEE Communications Magazine*.

Reports on proposals for higher speed wireless LAN extensions in the 2.4 GHz and 5 GHz bands can be found at the IEEE 802.11 web site (http://grouper.ieee.org/groups/802/11/Reports). Additional information on HIPERLAN and related activities in ETSI BRAN can be obtained from their web site (http://www.etsi.fr/bran).

88

Wireless ATM: QoS and Mobility Management

Bala Rajagopalan
Tellium, Inc.

Daniel Reininger
Semandex Networks, Inc.

88.1 Introduction

Wireless ATM (WATM) refers to the technology that enables ATM end-system mobility as well as tetherless access to ATM core networks. Wireless ATM has two distinct components: the radio access technology, and enhancements to existing ATM technology to support end-system mobility. The latter component is referred to as "MATM" (mobility-enhanced ATM) and it is independent of the radio access technology. The rationale for wireless ATM has been discussed at length elsewhere [1,2]. In this chapter, we restrict our discussion to two challenging issues in wireless ATM: the provisioning of ATM quality of service (QoS) for connections that terminate on mobile end systems over a radio link, and the protocols for mobility management in the MATM infrastructure.

Figure 88.1 illustrates the WATM reference model considered in this chapter [3]. "W" UNI in this figure indicates the ATM user-network interface established over the wireless link. "M" NNI refers to the ATM network-node interface supporting mobility management protocols. The figure depicts the scenario where an MATM network has both end system mobility-supporting ATM switches (EMAS) and traditional ATM switches with no mobility support. Thus, one of the features of mobility management protocols in an MATM network is the ability to work transparently over switches that do not implement mobility support.

88.2 QoS in Wireless ATM

The type of QoS guarantees to be provided in wireless ATM systems is debatable [4]. On the one hand, the QoS model for traditional ATM networks is based on fixed terminals and high quality links. Terminal mobility and error-prone wireless links introduce numerous problems [5]. On the other hand, maintaining the existing QoS model allows the transparent extension of fixed ATM applications into the domain of mobile networking. Existing prototype implementations have chosen the latter approach [6]–[8]. This is also the decision of the ATM Forum wireless ATM working group [9]. Our discussion,

TABLE 88.1 ATM Traffic and QoS Parameters

Attribute	ATM Service Category				
	CBR	rt-VBR	nrt-VBR	UBR	ABR
Traffic Parameters					
PCR and CDVT	Yes	Yes	Yes	Yes	Yes
SCR and MBS	N/A	Yes	Yes	N/A	N/A
MCR	N/A	N/A	N/A	N/A	Yes
QoS Parameters					
CDV	Yes	Yes	No	No	No
Maximum CTD	Yes	Yes	No	No	No
CLR	Yes	Yes	Yes	No	No

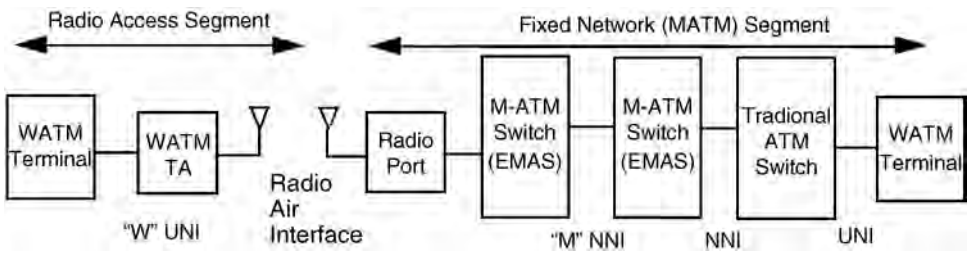


FIGURE 88.1 WATM reference model.

therefore, is oriented in the same direction, and to this end we first briefly summarize the existing ATM QoS model.

ATM QoS Model

Five service categories have been defined under ATM [10]. These categories are differentiated according to whether they support constant or variable rate traffic, and real-time or non-real-time constraints. The service parameters include a characterization of the traffic and a reservation specification in the form of QoS parameters. Also, traffic is policed to ensure that it conforms to the traffic characterization, and rules are specified for how to treat nonconforming traffic. ATM provides the ability to tag nonconforming cells and specify whether tagged cells are policed (and dropped) or provided with best-effort service.

Under UNI 4.0, the service categories are constant bit rate (CBR), real-time variable bit rate (rt-VBR), non-real-time variable bit rate (nrt-VBR), unspecified bit rate (UBR) and available bit rate (ABR). The definition of these services can be found in [10]. Table 88.1 summarizes the traffic descriptor parameters and QoS parameters relevant to each service category in ATM traffic management specifications version 4.0 [11]. Here, the traffic parameters are peak cell rate (PCR), cell delay variation tolerance (CDVT), sustainable cell rate (SCR), maximum burst size (MBS) and minimum cell rate (MCR). The QoS parameters are cell loss ratio (CLR), maximum cell transfer delay (max CTD) and cell delay variation (CDV). The explanation of these parameters can be found in [11].

Functions related to the implementation of QoS in ATM networks are usage parameter control (UPC) and connection admission control (CAC). In essence, the UPC function (implemented at the network edge) ensures that the traffic generated over a connection conforms to the declared traffic parameters. Excess traffic may be dropped or carried on a best-effort basis (i.e., QoS guarantees do not apply). The CAC function is implemented by each switch in an ATM network to determine whether the QoS requirements of a connection can be satisfied with the available resources. Finally, ATM connections can be either point-to-point or point-to-multipoint. In the former case, the connection is bidirectional, with

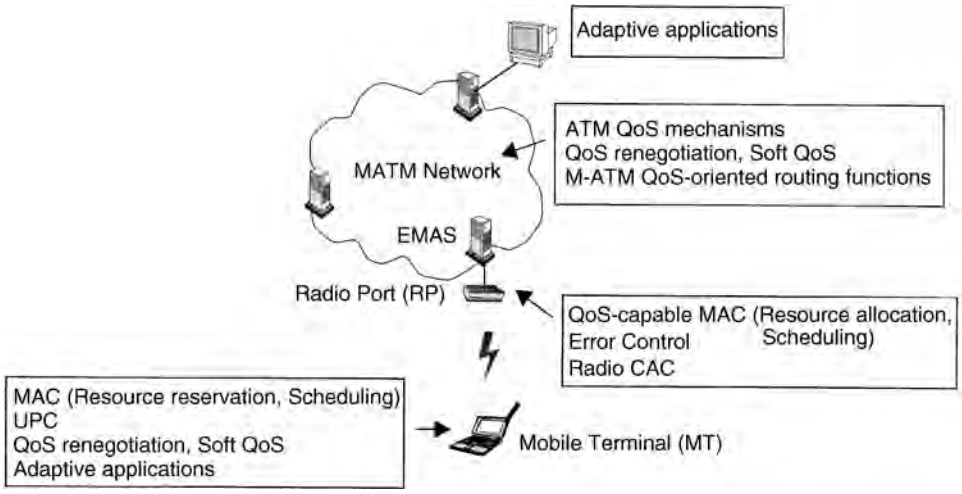


FIGURE 88.2 QoS mechanisms in wireless ATM.

separate traffic and QoS parameters for each direction, while in the latter case it is unidirectional. In this chapter, we consider only point-to-point connections for the sake of simplicity.

QoS Approach in Wireless ATM

QoS in wireless ATM requires a combination of several mechanisms acting in concert. Figure 88.2 illustrates the various points in the system where QoS mechanisms are needed:

- *At the radio interface:* A QoS-capable medium access control (MAC) layer is required. The mechanisms here are resource reservation and allocation for ATM virtual circuits under various service categories, and scheduling to meet delay requirements. Furthermore, an error control function is needed to cope with radio link errors that can otherwise degrade the link quality. Finally, a CAC mechanism is required to limit access to the multiple access radio link in order to maintain QoS for existing connections.
- *In the network:* ATM QoS mechanisms are assumed in the network. In addition, a capability for QoS renegotiation will be useful. This allows the network or the mobile terminal (MT) to renegotiate the connection QoS when the existing connection QoS cannot be maintained during handover. Renegotiation may also be combined with *soft* QoS mechanisms, as described later. Finally, mobility management protocols must include mechanisms to maintain QoS of connections rerouted within the network during handover.
- *At the MT:* The MT implements the complementary functions related to QoS provisioning in the MAC and network layers. In addition, application layer functions may be implemented to deal with variations in the available QoS due to radio link degradation and/or terminal mobility. Similar functions may be implemented in fixed terminals communicating with MTs.

In the following, we consider the QoS mechanisms in some detail. We focus on the MAC layer and the new network layer functions such as QoS renegotiation and soft QoS. The implementation of existing ATM QoS mechanisms have been described in much detail by others (for example, see [12]).

MAC Layer Functions

The radio link in a wireless ATM system is typically a broadcast multiple access channel shared by a number of MTs. Different multiple access technologies are possible, for instance frequency, time, or code division multiple access (FDMA, TDMA, and CDMA, respectively). A combination of FDMA and

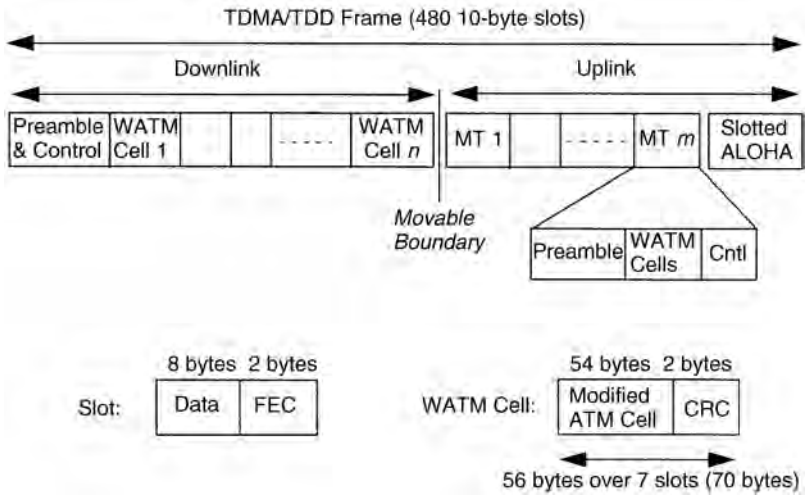


FIGURE 88.3 TDMA logical frame format.

dynamic TDMA is popular in wireless ATM implementations. That is, each radio port (RP) operates on a certain frequency band and this bandwidth is shared dynamically among ATM connections terminating on multiple MTs using a TDMA scheme [13]. ATM QoS is achieved under dynamic TDMA using a combination of a resource reservation/allocation scheme and a scheduling mechanism. This is further explained using the example of two wireless ATM implementations: NEC's WATMnet 2.0 prototype [13] and the European Union's Magic WAND (Wireless ATM Network Demonstrator) project [8].

Resource Reservation and Allocation Mechanisms (WATMnet 2.0)

WATMnet utilizes a TDMA/TDD (time division duplexing) scheme for medium access. The logical transmission frame structure under this scheme is shown in Fig. 88.3. As shown, this scheme allows the flexibility to partition the frame dynamically for downlink (from EMAS to MTs) and uplink (from MTs to EMAS) traffic, depending on the traffic load in each direction. Other notable features of this scheme are:

- A significant portion of each slot is used for forward error control (FEC)
- A separate contention region in the frame is used for MTs to communicate with the EMAS
- 8-byte control packets are used for bandwidth request and allocation announcements. An MT can tag request packets along with the WATM cells it sends or in the contention slots
- WATM cells are modified ATM cells with data link control (DLC) and cyclic redundancy check (CRC) information added

In the downlink direction, the WATM cells transported belong to various ATM connections terminating on different MTs. After such cells arrive at the EMAS from the fixed network, the allocation of TDMA slots for specific connections is done at the EMAS based on the connections' traffic and QoS parameters. This procedure is described in the next section. In the uplink direction, the allocation is based on requests from MTs. For bursty traffic, an MT makes a request only after each burst is generated. Once uplink slots are allocated to specific MTs, the transmission of cells from multiple active connections at an MT is again subject to the scheduling scheme. Both the request for uplink slot allocation from the MTs and the results from the EMAS are carried in control packets whose format is shown in Fig. 88.4. Here, the numbers indicate the bits allocated for various fields. The sequence number is used to recover from transmission losses. Request and allocation types indicate one of four types: CBR, VBR, ABR or UBR. The allocation packet has a start slot field which indicates where in the frame the MT should start transmission. The number of allocated slots is also indicated.

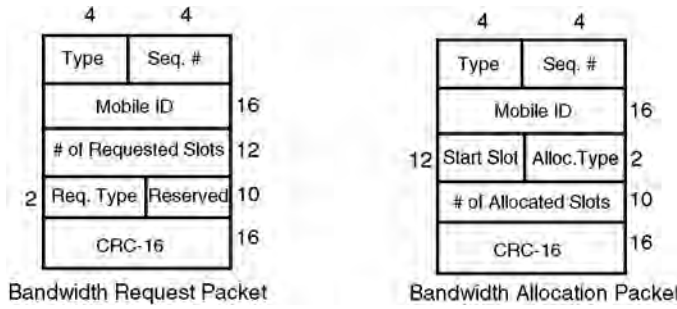


FIGURE 88.4 Bandwidth control packet formats.

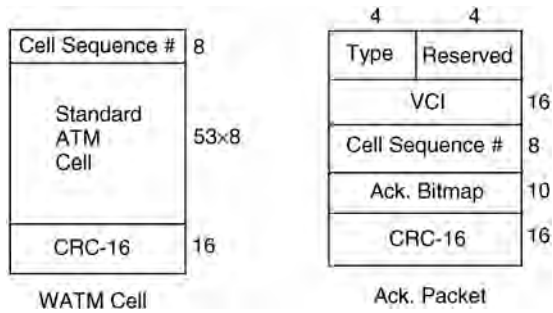


FIGURE 88.5 WATM formats at the DLC layer.

The DLC layer implementation in WATMnet is used to reduce the impact of errors that cannot be corrected using the FEC information. The DLC layer is responsible for selective retransmission of cells with uncorrectable errors or lost cells. Furthermore, the DLC layer provides request/reply control interface to the ATM layer to manage the access to the wireless bandwidth, based on the instantaneous amount of traffic to be transmitted. The wireless ATM cell sent over the air interface is a modified version of the standard ATM cell with DLC and CRC information, as shown in Fig. 88.5. The same figure also shows the acknowledgment packet format for implementing selective retransmission. In this packet, the VCI field specifies the ATM connection for which the acknowledgment is being sent. The sequence number field indicates the beginning sequence number from which the 16-bit acknowledgment bitmap indicates the cells correctly received (a "1" in the bit map indicates correct reception).

The TDMA/TDD scheme thus provides a mechanism for dynamic bandwidth allocation to multiple ATM connections. How active connections are serviced at the EMAS and the MTs to maintain their QoS needs is another matter. This is described next using Magic WAND as the example.

Scheduling (Magic WAND)

The Magic WAND system utilizes a TDMA scheme similar to that used by the WATM-net. This is shown in Fig. 88.6. Here, each MAC Protocol Data Unit (MPDU) consists of a header and a sequence of WATM cells from the same MT (or the EMAS) referred to as a *cell train*. In the Magic WAND system, the scheduling of both uplink and downlink transmissions is done at the EMAS. Furthermore, the scheduling is based on the fact that the frame length is variable. A simplified description of the scheduling scheme is presented below. More details can be found in [14,15].

At the beginning of each TDMA frame, the scheduling function at the EMAS considers pending transmission requests, uplink and downlink, from active connections. The scheduler addresses two issues:

1. The determination of the number of cells to be transmitted from each connection in the frame, and
2. The transmission sequence of the selected cells within the frame

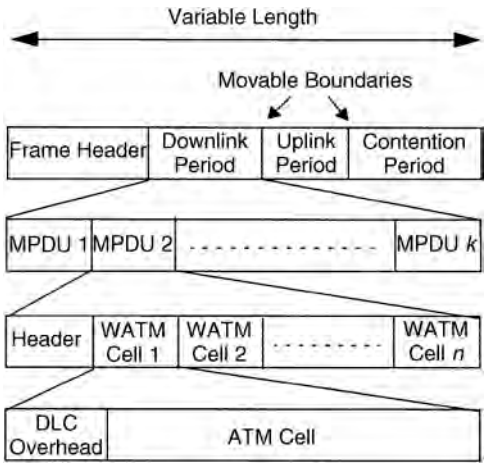


FIGURE 88.6 Magic WAND TDMA frame structure.

The objective of the scheduler is to regulate the traffic over the radio interface as per the declared ATM traffic parameters of various connections and to ensure that the delay constraints (if any) are met for these connections over this interface.

The selection of cells for transmission is done based on the service categories of active connections as well as their traffic characteristics. First, for each connection, a priority based on its service category is assigned. CBR connections are assigned the highest priority, followed by rt-VBR, nrt-VBR, ABR. In addition, for each active connection that is not of type UBR, a token pool is implemented. Tokens for a connection are generated at the declared SCR of the connection and tokens may be accumulated in the pool as long as their number does not exceed the declared MBS for the connection. The scheduler services active connections in two passes: in the first pass, connections are considered in priority order from CBR to ABR (UBR is omitted) and within each priority class only connections with a positive number of tokens in their pools are considered. Such connections are serviced in the decreasing order of the number of tokens in their pools. Whenever a cell belonging to a connection is selected for transmission, a token is removed from its pool. At the end of the first pass, either all the slots in the downlink portion of the frame are used up or there are still some slots available. In the latter case, the second pass is started. In this pass, the scheduler services remaining excess traffic in each of CBR, rt and nrt-VBR and ABR classes, and UBR traffic in the priority order. It is clear that in order to adequately service active connections, the mean bandwidth requirement of the connections cannot exceed the number of downlink slots available in each frame. The CAC function is used to block the setting up of new connections over a radio interface when there is a danger of overloading. Another factor that can result in overloading is the handover of connections. The CAC must have some knowledge of the expected load due to handovers so that it can limit new connection admissions. Preserving the QoS for handed over connections while not degrading existing connections at a radio interface requires good network engineering. In addition, mechanisms such as QoS renegotiation and soft QoS (described later) may be helpful.

Now, at the end of the selection phase, the scheduler has determined the number of cells to be transmitted from each active connection. Some of these cells are to be transmitted uplink while the others are to be transmitted downlink. The scheduler attempts to place a cell for transmission within the frame such that the cell falls in the appropriate portion of the frame (uplink or downlink, Fig. 88.6) and the delay constraint (CDT) of the corresponding connection is met. To do this, first the delay allowed over the radio segment is determined for each connection with a delay constraint (it is assumed that this value can be obtained during the connection routing phase by decomposing the path delay into delays for each hop). Then, for downlink cells, the arrival time for the cell from the fixed network is marked. For uplink cells, the arrival time is estimated from the time at which the request was received from the MT.

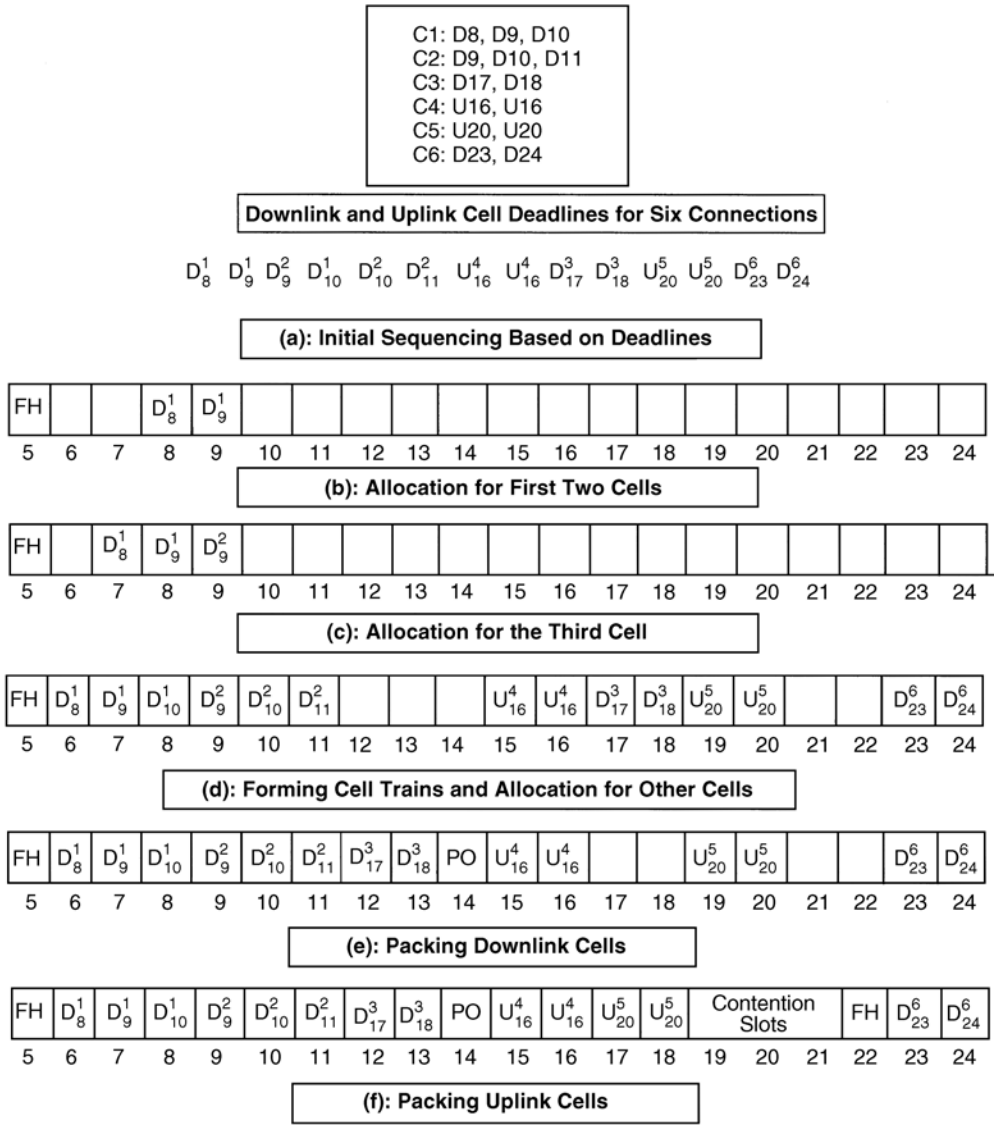


FIGURE 88.7 Scheduling example.

The deadline for the transmission of a cell (uplink or downlink) is computed as the arrival time plus the delay allowed over the radio link.

The final placement of the cells in the frame is based on a three-step process, as illustrated by an example with six connections (Fig. 88.7). Here, D*n* and U*n* indicate downlink and uplink cells with deadline = slot *n*, respectively, and D_j^i and U_j^i indicate downlink and uplink cells of the *i*th connection with deadline = slot *j*, respectively. In the first step, the cells are ordered based on their deadlines [Fig. 88.7(a)]. Several cells belonging to the same connection may have been selected for transmission in a frame. When assigning a slot for the first cell of such a "cell train" the scheduler positions the cell such that its transmission will be before and as close to its deadline as possible. Some cells may have to be shifted from their previously allocated slots to make room for the new allocation. This is done only if the action does not violate the deadline of any cell. When assigning a slot for another cell in the train, the scheduler attempts to place it in the slot next to the one allocated to the previous cell. This may

require shifting of existing allocations as before. This is shown in Figs. 88.7(b)–88.7(d). Here, the transmission frame is assumed to begin at slot 5.

At the end of the first step, the cell sequencing may be such that uplink and downlink cells may be interleaved. The second step builds the downlink portion of the frame by first shifting all downlink cells occurring before the first uplink cell as close to the beginning of the frame as possible. In the space between the last such downlink cell and the first uplink cell, as many downlink cells as possible are packed. This is illustrated in Fig. 88.7(e). A slot for *period overhead* (PO) is added between the downlink and uplink portions. Finally, in the last step, the uplink cells are packed, by moving all uplink cells occurring before the next downlink cell as shown in Fig. 88.7(f). The contention slots are added after the last uplink cell and the remaining cells are left for the next frame.

Thus, scheduling can be a rather complicated function. The specific scheduling scheme used in the Magic WAND system relies on the fact that the frame length is variable. Scheduling schemes for other frame structures could be different.

Network and Application Layer Functions

Wireless broadband access is subject to sudden variations in bandwidth availability due to the dynamic nature of the service demand (e.g., terminals moving in and out of RPs coverage area, variable bit-rate interactive multimedia connections) and the natural constraints of the physical channel (e.g., fading and other propagation conditions). QoS control mechanisms should be able to handle efficiently both the mobility and the heterogeneous and dynamic bandwidth needs of multimedia applications. In addition, multimedia applications themselves should be able to adapt to terminal heterogeneity, computing limitations, and varying availability of network resources [16].

In this section, the network and application layer QoS control functions are examined in the context of NECs WATMnet system [6]. In this system, a concept called *soft-QoS* is used to effectively support terminal mobility, maintain acceptable application performance and high network capacity utilization. Soft-QoS relies on a QoS control framework that permits the allocation of network resources to dynamically match the varying demands of mobile multimedia applications. Network mechanisms under this framework for connection admission, QoS renegotiation, and handoff control are described. The soft-QoS concept and the realization of the soft-QoS controller based on this concept are described next. Finally, experimental results on the impact of network utilization and soft-QoS provisioning for video applications are discussed.

A Dynamic Framework for QoS Control

The bit-rates of multimedia applications vary significantly among sessions and within a session due to user interactivity and traffic characteristics. Contributing factors include the presence of heterogeneous media (e.g., video, audio, and images) compression schemes (e.g., MPEG, JPEG), presentation quality requirements (e.g., quantization, display size), and session interactivity (e.g., image scaling, VCR-like control). Consider, for example, a multimedia application using several media components or media objects (such as a multiwindow multimedia user interface or future MPEG-4 encoded video) which allows users to vary the relative importance-of-presence (IoP) of a given media object to match the current viewing priorities. In this case there would be a strong dependency of user/application interaction on the bandwidth requirements of individual media components. Figure 88.8 shows the bit-rate when the user changes the video level-of-detail (LoD) during a session. A suitable network service for these applications should support bandwidth renegotiation to simultaneously achieve high network utilization and maintain acceptable performance. For this purpose, an efficient network service model should support traffic contract renegotiation during a session. It has been experimentally verified that bandwidth renegotiation is key for efficient QoS support of network-aware adaptive multimedia applications, and that the added implementation complexity is reasonable [17,18].

In the mobile multimedia communication scenario, bandwidth renegotiation is particularly important. Conventional network services use static bandwidth allocation models that lack the flexibility needed to

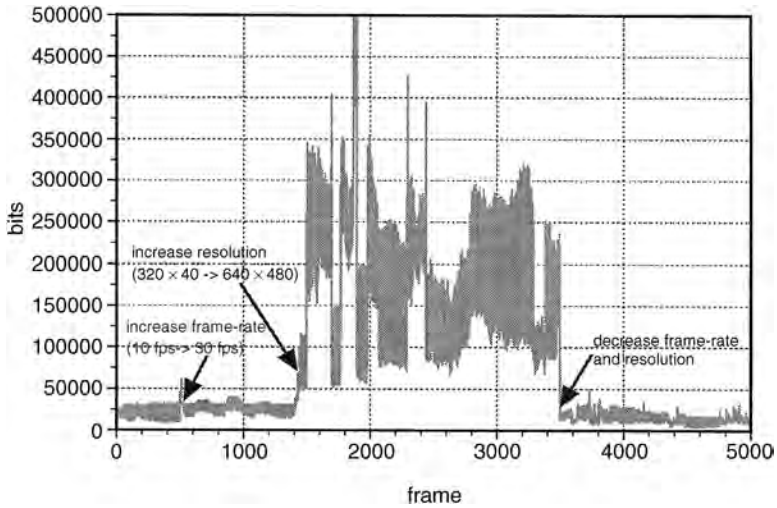


FIGURE 88.8 Video bit-rate changes on an interactive multimedia session.

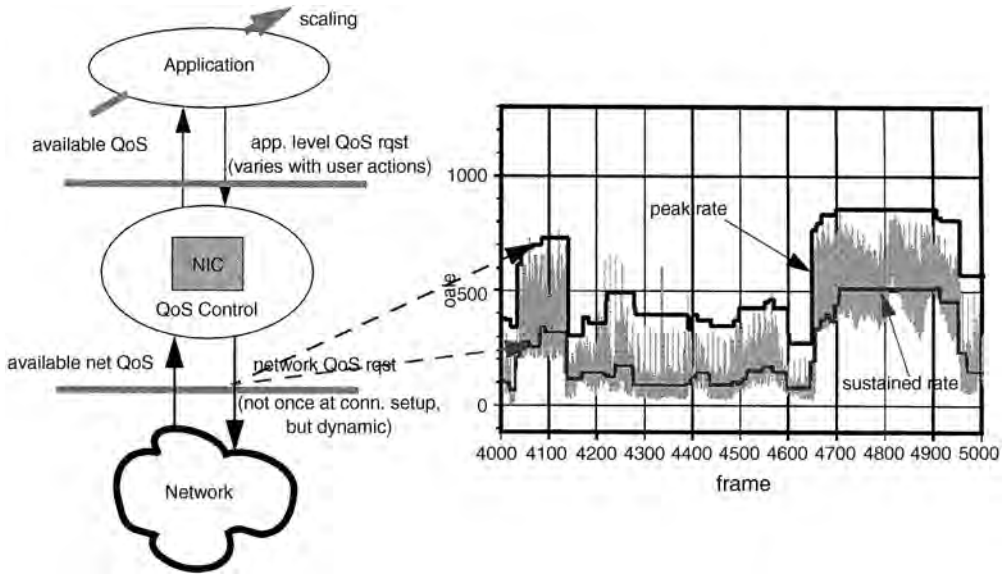


FIGURE 88.9 System and API model for QoS control with bandwidth renegotiation.

cope with multimedia interactivity and session mobility. These session properties enlarge the dynamic range of bandwidth requirements and make dynamic bandwidth management protocols a requirement for effective end-to-end QoS support. Renegotiation may be required during handover, as well as when resource allocation changes are warranted due to instantaneous application needs and sudden changes in network resource availability.

Figure 88.9 shows the system and API model for QoS control with bandwidth renegotiation. The application programming interface (API) between the adaptive application and the QoS control module is dynamic, i.e., its parameters can be modified during the session. For example, the Winsock 2 API under Microsoft Windows [19] allows the dynamic specification of QoS parameters suitable for the application. In addition, the API between the QoS controller and the network allows the traffic descriptor

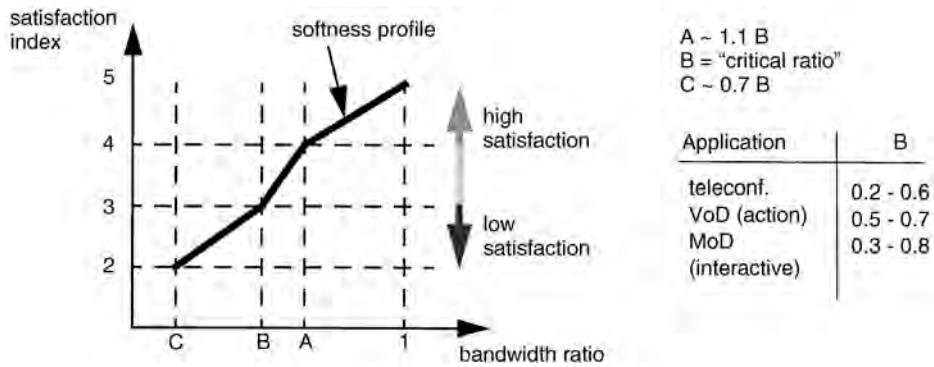


FIGURE 88.10 Example softness profile.

to be varied to track the bit-rate requirements of the bitstream. A new network service, called VBR^+ allows renegotiation of traffice descriptors between the network elements and the terminals [20]. VBR^+ allows multimedia applications to request "bandwidth-on-demand" suitable for their needs.

Soft-QoS Model

Although multimedia applications have a wide range of bandwidth requirements, most can gracefully adapt to sporadic network congestion while still providing acceptable performance. This graceful adaptation can be quantified by a softness profile [17]. Figure 88.10 shows the characteristics of a softness profile. The softness profile is a function defined on the scales of two parameters: satisfaction index and bandwidth ratio. The satisfaction index is based on the subjective mean-opinion-score (MOS), graded from 1 to 5; a minimum satisfaction divides the scale in two operational regions: the acceptable satisfaction region and the low satisfaction region. The bandwidth ratio is defined by dividing the current bandwidth allocated by the network to the bandwidth requested to maintain the desired application performance. Thus, the bandwidth ratio is graded from 0 to 1; a value of 1 means that the allocated bandwidth is sufficient to achieve the desired application performance. The point indicated as *B* is called the critical bandwidth ratio since it is the value that results in minimum acceptable satisfaction. As shown in Fig. 88.10, the softness profile is approximated by piecewise linear "S-shaped" function consisting of three linear segments. The slope of each linear segment represents the rate at which applications performance degrades (satisfaction index decreases) when the network allocates only a portion of the requested bandwidth: the steeper the slope is, the "harder" the corresponding profile is.

The softness profile allows efficient match of application requirements to network resource availability. With the knowledge of the softness profile, network elements can perform soft-QoS control—QoS-fair allocation of resources among contending applications when congestion arises. Applications can define a softness profile that best represents their needs. For example, the softness profile for digital compressed video is based on the nonlinear relationship between coding bit-rate and quality, and the satisfaction index is correlated to the user perception of quality [21,22]. While video-on-demand (VoD) applications may, in general, tolerate bit-rate regulations within a small dynamic range, applications such as surveillance or teleconference may have a larger dynamic range for bit-rate control. Other multimedia applications may allow a larger range of bit-rate control by resolution scaling [18]. In these examples, VoD applications are matched to a "harder" profile than the other, more adaptive multimedia applications. Users on wireless mobile terminals may select a "softer" profile for an application in order to reduce the connection's cost, while a "harder" profile may be selected when the application is used on wired desktop terminal. Thus, adaptive multimedia applications able to scale their video quality could specify their soft-QoS requirements dynamically to control the session's cost.

Figure 88.11 conceptually illustrates the role of application QoS/bandwidth renegotiation, service contract, and session cost in the service model. The soft-QoS service model is suitable for adaptive

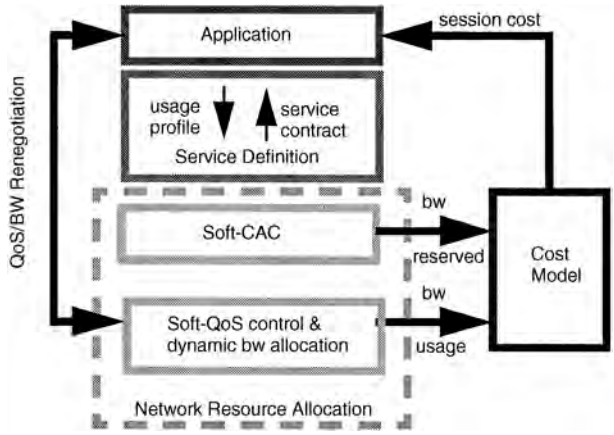


FIGURE 88.11 Service model for dynamic bandwidth allocation with soft-QoS.

multimedia applications capable of gracefully adjusting their performance to variable network conditions. The service definition is needed to match the requirements of the application with the capabilities of the network. The service definition consists of two parts: a usage profile that specifies the target regime of operation and the service contract that statistically quantifies the soft-QoS service to be provided by the network. The usage profile, for example, can describe the media type (e.g., MPEG video), interactivity model (e.g., multimedia browsing, video conference), mobility model (indoors, urban semi-mobile, metropolitan coverage area), traffic, and softness profiles. The service contract quantifies soft-QoS in terms of the probability that the satisfaction of a connection will fall outside the acceptable range (given in its softness profile), the expected duration of "satisfaction outage," and the new connection blocking probability.

Network resource allocation is done in two phases. First, a connection admission control procedure, called soft-CAC, checks the availability of resources on the terminals coverage area at connection set-up time. The necessary resources are estimated based on the service definition. The new connection is accepted if sufficient resources are estimated to be available for the connection to operate within the service contract without affecting the service of other ongoing connections. Otherwise the connection is blocked. Second, while the connection is in progress, dynamic bandwidth allocation is performed to match the requirements of interactive variable bit-rate traffic. When congestion occurs, the soft-QoS control mechanism (re)-allocates bandwidth among connections to maintain the service of all ongoing connections within their service contracts. The resulting allocation improves the satisfaction of undersatisfied connections while maintaining the overall satisfaction of other connections as high as possible [17]. Under this model, connections compete for bandwidth in a "socially responsible" manner based on their softness profiles. Clearly, if a cost model is not in place, users would request the maximum QoS possible. The cost model provides feedback on session cost to the applications; the user can adjust the long-term QoS requirements to maintain the session cost within budget.

Soft-QoS Control in the WATMnet System

In the WATMnet system, soft-QoS control allows effective support of mobile multimedia applications with high network capacity utilization. When congestion occurs, the soft-QoS controller at the EMASs allocates bandwidth to connections based on their relative robustness to congestion given by the applications softness profiles. This allocation improves the satisfaction of undersatisfied connections while maintaining the overall satisfaction of other connections as high as possible. Within each EMAS, connections compete for bandwidth in a "socially responsible" manner based on their softness profiles.

ATM UNI signalling extensions are used in the WATMnet system to support dynamic bandwidth management. These extensions follow ITU-T recommendations for ATM traffic parameter modification

while the connection is active [23]. Although these procedures are not finalized at the time of this writing, an overview of the current state of the recommendation is given next with an emphasis on its use to support soft-QoS in the mobile WATM scenario.

ITU-T Q.2963 allows all three ATM traffic parameters, PCR, SCR, and MBS, to be modified during a call. All traffic parameters must be increased or decreased; it is not possible to increase a subset of the parameters while decreasing the others. The user who initiates the modification request expects to receive from the network a new set of traffic parameters that are greater than or equal to (or less than) the existing traffic parameters if the modification request is an increase (or decrease). Traffic parameter modification is applicable only to point-to-point connections and may be requested only by the terminal that initiated the connection while in the active state.

The following messages are added to the UNI:

- MODIFY REQUEST message is sent by the connection owner to request modification of the traffic descriptor; its information element (IE) is the ATM traffic descriptor.
- MODIFY ACKNOWLEDGE message is sent by the called user or network to indicate that the modify request is accepted. The broadband report type IE is included in the message when the called user requires confirmation of the success of modification.
- CONNECTION AVAILABLE is an optional message issued by the connection owner to confirm the connection modification performed in the addressed user to requesting user direction. The need for explicit confirmation of modification is indicated by the "modification confirmation" field in the MODIFY ACKNOWLEDGE broadband report IE.
- MODIFY REJECT message is sent by the called user or network to indicate that the modify connection request is rejected. The cause of the rejection is informed through the cause IE.

Figures 88.12, 88.13, and 88.14 show the use of these messages for successful, addressed user rejection and network rejection of modification request, respectively.

Additionally, the soft-QoS control framework of the WATMnet system uses the following modifications to the Q2963 signalling mechanisms:

- BANDWIDTH CHANGE INDICATION (BCI) message supports network-initiated and called user-initiated modification. The message is issued by the network or called user to initiate a modification procedure. The traffic descriptor to be used by the connection owner when issuing the corresponding MODIFY REQUEST message is specified in the BCIs ATM traffic descriptor IE. Figure 88.15 illustrates the use of BCI for called user–initiated modification. Timer T362 is set when issuing the BCI message and cleared when the corresponding MODIFY REQUEST message

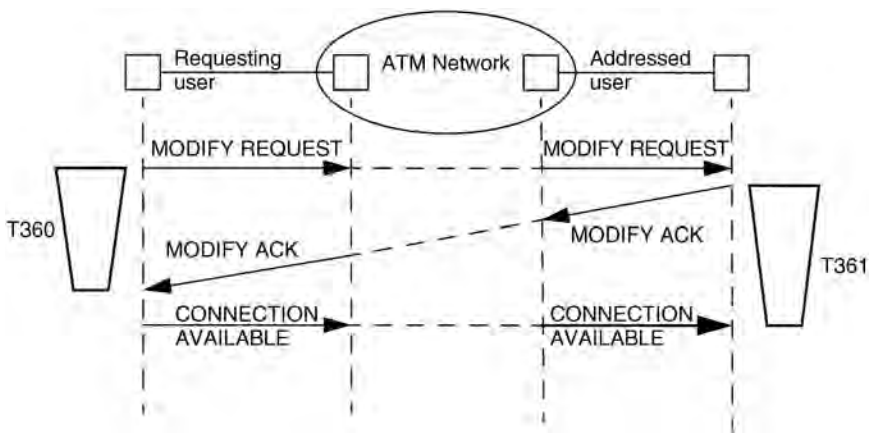


FIGURE 88.12 Successful Q2963 modification of ATM traffic parameters with (optional) confirmation.

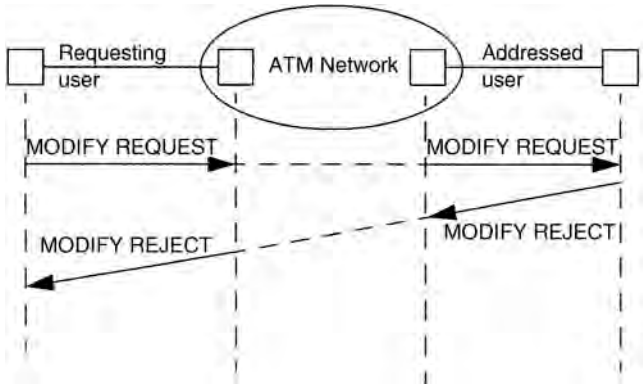


FIGURE 88.13 Addressed user rejection of modification.

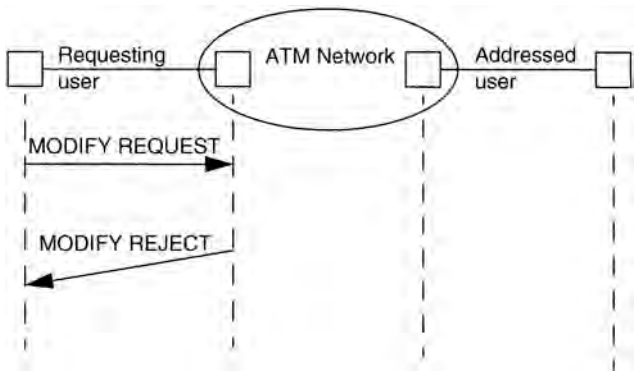


FIGURE 88.14 Network rejection of modification request.

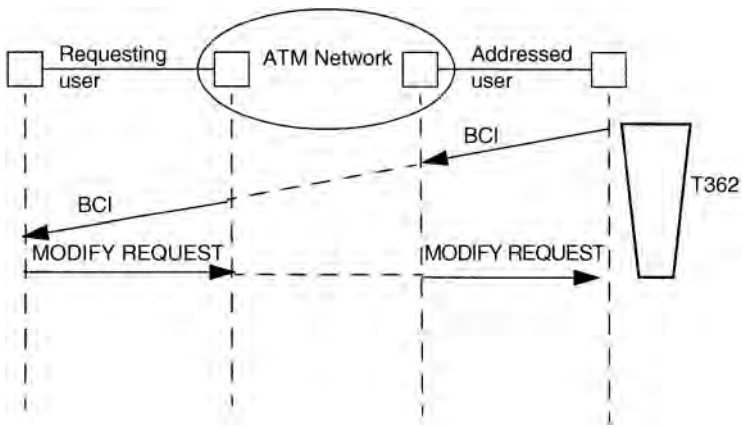


FIGURE 88.15 Use of BCI message for addressed user–initiated modification.

is received; if T362 expires, the terminal and/or network element can modify the traffic policers to use the ATM traffic descriptor issued in the BCI message.

- Specification of softness profile and associated minimum acceptable satisfaction level (sat_{min}) in the MODIFY REQUEST message. The softness profile and sat_{min} are used for QoS-fair allocation within the soft-QoS control algorithm.

- Specification of available bandwidth fraction (ABF) for each ATM traffic descriptor parameter. ABF is defined as the ratio of the available to requested traffic descriptor parameter. This results in ABF-PCR, ABF-SCR, and ABF-MBS for the peak, sustained, and maximum burst size, respectively. These parameters are included in the MODIFY REJECT message. Using the ABF information, the connection owner may recompute the requested ATM traffic descriptor and reissue an appropriate MODIFY REQUEST message.

Two additional call states are defined to support modification. An entity enters the modify request state when it issues a MODIFY REQUEST of BCI message to the other side of the interface. An entity enters the modify received state when it receives a MODIFY REQUEST of BCI message from the other side of the interface.

Soft-QoS control is particularly useful during the handover procedure as a new MT moves into a cell and places demands on resources presently allocated to connections from other MTs. A flexible way of prioritizing the bandwidth allocation to various session VCs is through their softness profiles. If a mobile terminal faces significant drop in bandwidth availability as it moves from one cell to another, rather than dropping the handover connections, the EMAS might be able to reallocate bandwidth among selected active connections in the new cell.

Within the soft-QoS framework, the soft-QoS controller selects a set of connections, called donors, and changes their bandwidth reservation so as to ensure satisfactory service for all [17]. This process is called network-initiated renegotiation. Network-initiated renegotiation improves the session handover success probability since multiple connections within and among sessions can share the available resources at the new EMAS, maintaining the satisfaction of individual connections above the minimum required. This mechanism allows multimedia sessions to transparently migrate the relative priority of connections as the MT moves across cells without a need to further specify details of the media session's content.

Figure 88.16 shows an MT moving into the coverage area of a new RP under a new EMAS and issuing a MODIFY REQUEST message (MRE). As a result, the EMAS might have to reallocate bandwidth of other connections under the RP to successfully complete the handover. This is accomplished by issuing BCI messages to a selected set of connections, called donors. At the time of receiving the first MRE message for a connection being handed over, no state exists for that connection within the new EMAS. This event differentiates the MRE messages from ongoing connections and connections being handed over. Different algorithmic provisions can be made to expedite bandwidth allocation to MRE messages of connections being handed over, reducing the probability of handover drop. For example, upon identifying a MRE from such connection, the soft QoS controller can use cached bandwidth reserves to maintain the satisfaction of the connection above the minimum. The size of the bandwidth cache is made adaptive to the ratio of handed-over to local bandwidth demand. The bandwidth cache for each RP can

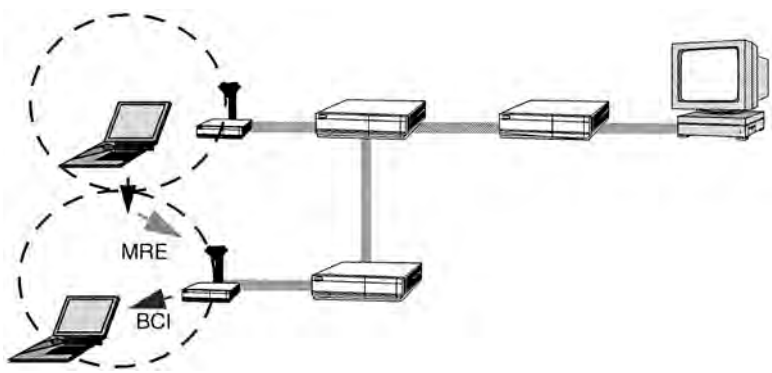


FIGURE 88.16 Handover procedure with soft-QoS in the WATMnet system.

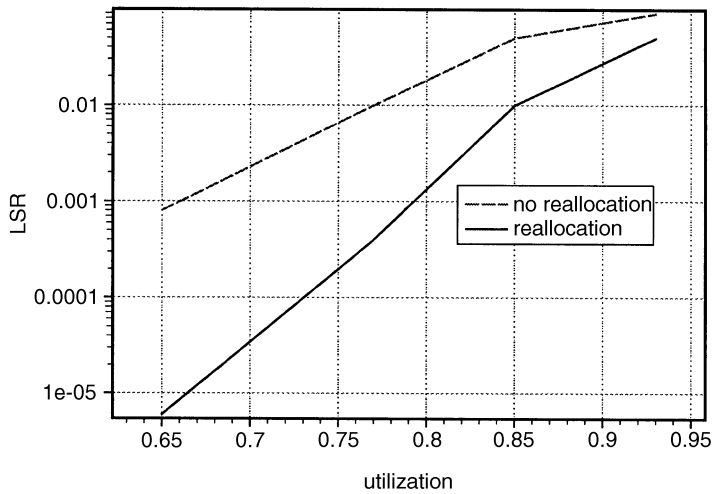


FIGURE 88.17 Effect of soft-QoS control with and without network-initiated modification.

be replenished off-line using the network-initiated modification procedure. In this way, a handed-over connection need not wait for the network-initiated modification procedure to end before being able to use the bandwidth. The outcome of the reallocation enables most connections to sustain a better than minimum application performance while resources become available. Short-term congestion may occur due to statistical multiplexing. If long-term congestion arises due to the creation of a hot spot, dynamic channel allocation (DCA) may be used to provide additional resources. It is also possible that if connection rerouting is required inside the network for handover, the required resources to support the original QoS request may not be available within the network along the new path. Renegotiation is a useful feature in this case also.

An important performance metric for the soft-QoS service model is the low satisfaction rate (LSR). LSR measures the probability of failing to obtain link capacity necessary to maintain acceptable application performance. Figure 88.17 compares the LSR with and without network-initiated modification over a wide range of link utilization for an MPEG-based interactive video application. The softness profiles for MPEG video were derived from empirical results reported in [21,22]. The figure shows that network-initiated modification has an important contribution to soft-QoS control performance: robust operation (LSR $< 10^{-3}$) is achievable while maintaining 70 to 80% utilization. In the WATM scenario, the handoff success probability with soft-QoS control is related to the LSR by Prob(handoff success) = $P(sat^{+} > sat_{min}) > 1 - LSR \gg 1 - P_b$, where sat^{+} represents the satisfaction after handover to the new AP completes. Since it is better to block a new connection than to drop an existing connection for lack of capacity, the condition $LSR \ll P_b$ is used, where P_b is the connection blocking probability. The operating goal for the system is to maintain utilization >70%, LSR ~ 10^{-5}, $P_b \sim 10^{-3}$.

Although the results presented are based on softness profiles for video, the definition of soft-QoS is appropriate for adaptive multimedia applications in general. Various representative softness profiles can be defined and refined as users' experience with distributed multimedia applications grows. New profiles can easily be incorporated within the framework as they become available.

88.3 Mobility Management in Wireless ATM

Allowing end system mobility in ATM networks gives rise to the problem of mobility management, i.e., maintaining service to end systems regardless of their location or movement. A fundamental design choice here is whether mobility management deals with user mobility or terminal mobility. When a network supports user mobility, it recognizes the user as the subscriber with an associated service profile.

The user can then utilize any MT for access to the subscribed services. This results in flexibility in service provisioning and usage, but some extra complexity is introduced in the system implementation, as exemplified by the GSM system [24]. Support for user mobility implies support for terminal mobility. A network may support only terminal mobility and not recognize the user of the terminal, resulting in a simpler implementation. In either case, the mobility management tasks include:

- *Location Management:* Keeping track of the current location of an MT in order to permit correspondent systems to set up connections to it. A key requirement here is that the correspondent systems need not be aware of the mobility or the current location of the MT.
- *Connection Handover:* Maintaining active connections to an MT as it moves between different points of attachment in the network. The handover function requires protocols at both the radio layer and at the network layer. The issue of preserving QoS during handovers was described earlier and this introduces some complexity in handover implementations.
- *Security Management:* Authentication of mobile users (or terminals) and establishing cryptographic procedures for secure communications based on the user (or terminal) profile [25].
- *Service Management:* Maintaining service features as a user (or terminal) roams among networks managed by different administrative entities. Security and service management can be incorporated as part of location management procedures.

Early wireless ATM implementations have considered only terminal mobility [6,7]. This has been to focus the initial effort on addressing the core technical problems of mobility management, i.e., location management and handover [26]. Flexible service management in wide-area settings, in the flavor of GSM, has not been an initial concern in these systems. The mobility management protocol standards being developed by the ATM Forum WATM working group may include support for user mobility. In the following, therefore, we concentrate on the location management and handover functions required to support terminal mobility in wireless ATM.

Location Management in Wireless ATM

Location management (LM) in WATM networks is based on the notions of permanent and temporary ATM addresses. A permanent ATM address is a location-invariant, unique address assigned to each MT. As the MT attaches to different points in a WATM network, it may be assigned different temporary ATM addresses. As all ATM end system addresses, both permanent and temporary addresses are derived from the addresses of switches in the network, in this case EMASs. This allows connection set-up messages to be routed toward the MT, as described later. The EMAS whose address is used to derive the permanent address of an MT is referred to as the home EMAS of that MT. The LM function in essence keeps track of the current temporary address corresponding to the permanent address of each MT. Using this function, it becomes possible for correspondent systems to establish connections to an MT using only its permanent address and without knowledge of its location.

Network Entities Involved in LM

The LM functions are distributed across four entities:

- *The Location Server (LS):* This is a logical entity maintaining the database of associations between the permanent and temporary addresses of mobile terminals. The LS responds to query and update requests from EMASs to retrieve and modify database entries. The LS may also keep track of service-specific information for each MT.
- *The AUthentication Server (AUS):* This is a logical entity maintaining a secure database of authentication and privacy related information for each MT. The authentication protocol may be implemented between EMASs and the AUS, or directly between MTs and the AUS.

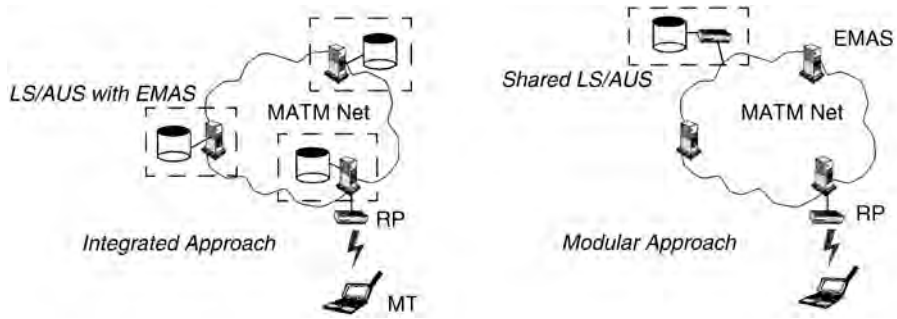


FIGURE 88.18 Server organizations.

- *The Mobile Terminal:* The MT is required to execute certain functions to initiate location updates and participate in authentication and privacy protocols.
- *The EMAS:* Certain EMASs are required to identify connection set-up messages destined to MTs and invoke location resolution functions. These can be home EMASs or certain intermediate EMASs in the connection path. All EMASs in direct contact with MTs (via their RPs) may be required to execute location update functions. Home EMASs require the ability to redirect a connection set-up message. In addition, all EMASs may be required to participate in the redirection of a connection set-up message to the current location of an MT.

There could be multiple LSs and AUSs in a WATM network. Specifically, an LS and an AUS may be incorporated with each home EMAS, containing information on all the MTs that the EMAS is home to. On the other hand, an LS or an AUS may be shared between several EMASs, by virtue of being separate entities. These choices are illustrated in Fig. 88.18, where the terms "integrated" and "modular" are used to indicate built-in and separated LS and AUS. In either case, protocols must be implemented for reliably querying and updating the LS, and mechanisms to maintain the integrity and security of the AUS. NEC's WATMnet [6] and BAHAMA [7] are examples of systems implementing integrated servers. The modular approach is illustrated by GSM [24] and next-generation wireless network proposals [27].

Location Management Functions and Control Flow

At a high level, the LM functions are registration, location update, connection routing to home or gateway EMAS, location query, and connection redirect.

Registration and Location Update

When an MT connects to a WATM network, a number of resources must be instantiated for that mobile. This instantiation is handled by two radio layer functions: association, which establishes a channel for the MT to communicate with the edge EMAS, and registration, which binds the permanent address of the MT to a temporary address. In addition, the routing information pertaining to the mobile at one or more location servers must be updated whenever a new temporary address is assigned. This is done using location updates.

The authentication of a mobile terminal and the establishment of encryption parameters for further communication can be done during the location updating procedure. This is illustrated in Fig. 88.19 which shows one possible control flow when an MT changes location from one EMAS to another. Here, the Broadcast_ID indicates the identity of the network, the location area, and the current radio port. Based on this information, e.g., by comparing its access rights and the network ID, the MT can decide to access the network. After an association phase, which includes the setting up of the signalling channel to the EMAS, the MT sends a registration message to the switch. This message includes the MT's home address and authentication information. The location update is initiated by the visited EMAS and the further progression is as shown. The LS/AUS are shown logically separate from the home EMAS for generality.

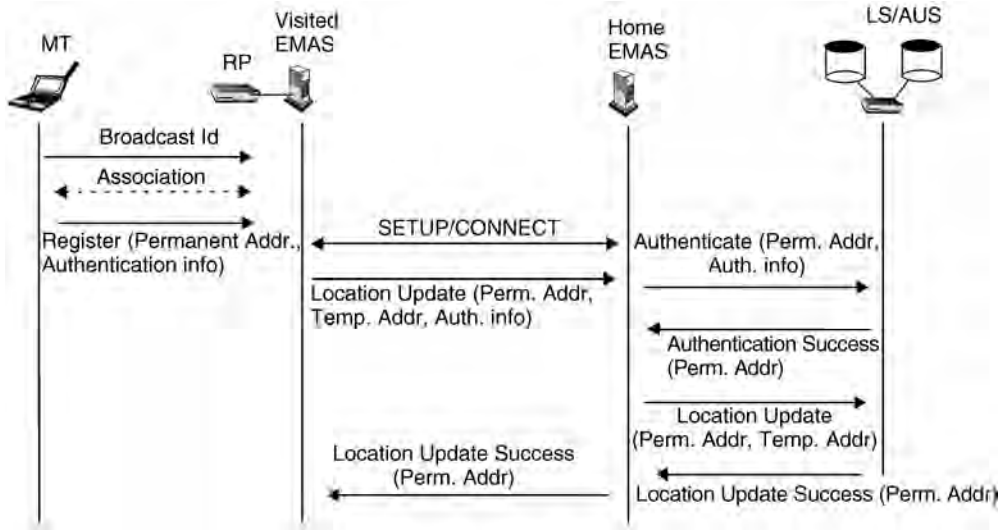


FIGURE 88.19 Location update control flow.

They can be integrated with the home EMAS. Details on the implementation of a similar location update scheme can be found in [28].

Now, there are other possible configurations of LSs that give rise to different control message flow. For example, a two-level hierarchical LS arrangement can be used. Under this organization, the LS in the visiting network is updated as long as the MT remains in this network, and the home LS is updated only when the MT moves to a different WATM network. The information kept in the home LS must, therefore, point to a gateway EMAS in the visited network, since precise location in the visited network will not be available at the home LS. GSM location management is an example of this scheme [24].

Connection Forwarding, Location Query, and Connection Redirect

After a location update, a location server handling the MT has the correct association between its permanent and temporary ATM addresses. When a new connection to the MT is established, the set-up message must be routed to some EMAS that can query the LS to determine the current address of the MT. This is the connection forwarding function. Depending on how MT addresses are assigned, the location query can occur very close to the origination of the connection or it must progress to the home EMAS of the MT. For instance, if MT addresses are assigned from a separately reserved ATM address space within a network, a gateway EMAS in the network can invoke location query when it processes a set-up message with a destination address known to be an MT address. To reach some EMAS that can interpret an MT address, it is sufficient to always forward connection set-up messages toward the home EMAS. This ensures that at least the home EMAS can invoke the query if no other EMAS enroute can do this. The location query is simply a reliable control message exchange between an EMAS and an LS. If the LS is integrated with the EMAS, this is a trivial operation. Otherwise, it requires a protocol to execute this transaction.

The control flow for connection establishment when the MT is visiting a foreign network is shown in Fig. 88.20. The addresses of various entities shown have been simplified for illustration purposes. Here, a fixed ATM terminal (A.1.1.0) issues a SETUP toward the MT whose permanent address is C.2.1.1. The SETUP message is routed toward the home EMAS whose address is C.2.1. It is assumed that no other EMAS in the path to the home EMAS can detect MT addresses. Thus, the message reaches the home EMAS which determines that the end system whose address is C.2.1.1 is an MT. It then invokes a location query to the LS which returns the temporary address for the MT (B.3.3). The home EMAS issues a redirected SETUP toward the temporary address. In this message, the MT is identified by its permanent

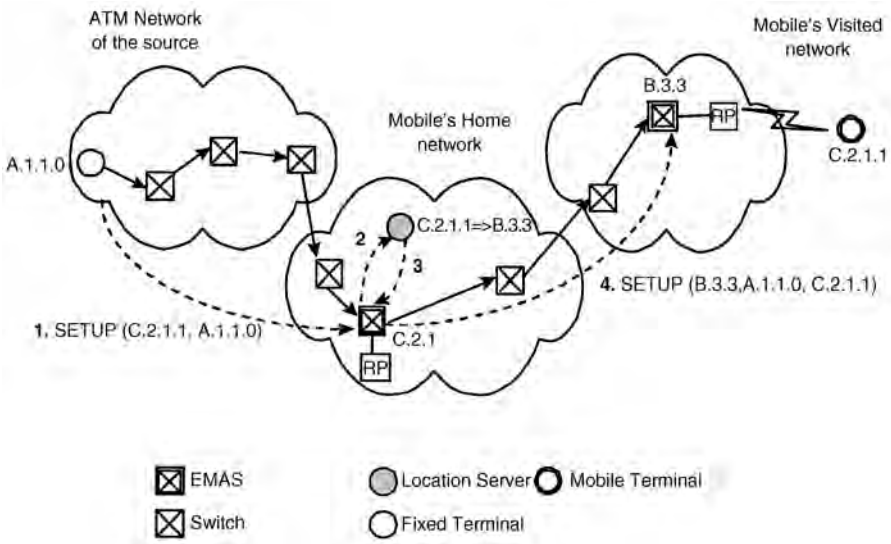


FIGURE 88.20 MT in foreign network.

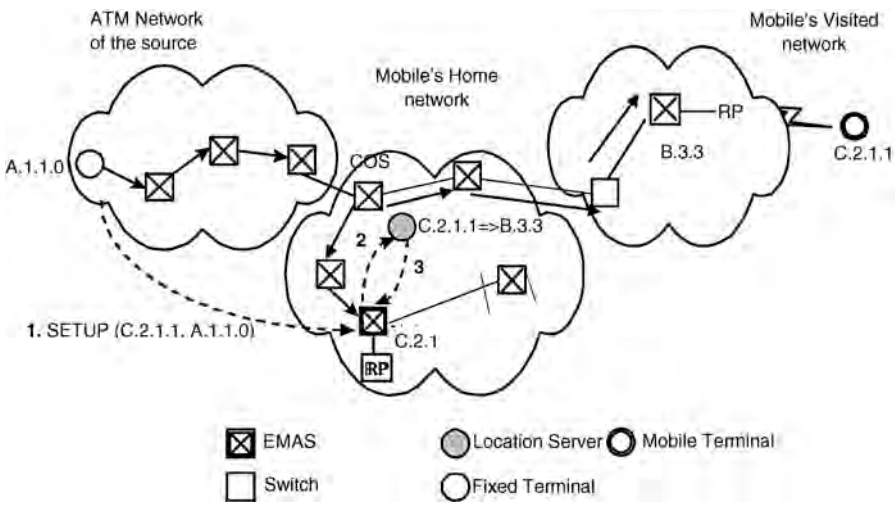


FIGURE 88.21 Connection redirect.

address thereby enabling the visited EMAS to identify the MT and proceed with the connection SETUP signalling.

It should be noted that in the topology shown in Fig. 88.20, the redirection of the connection set-up does not result in a nonoptimal path. But, in general, this may not be the case. To improve the overall end-to-end path, redirection can be done with partial teardown in which case a part of the established path is released and the connection set-up is redirected from an EMAS that occurs further upstream of the home EMAS. This is shown in Fig. 88.21. Here, the EMAS labelled COS (Cross Over Switch) occurs in the original connection path upstream of the home EMAS. To redirect the set-up to B.3.3, the connection already established to the home EMAS is torn down up to COS, and new segment is established from the COS to B.3.3. This requires additional signalling procedures.

Finally, in Fig. 88.22, the case of hierarchical location servers is illustrated. Here, as long as the MT is in the visited network, the address of the gateway EMAS (B.1.1) is registered in its home LS. The connection

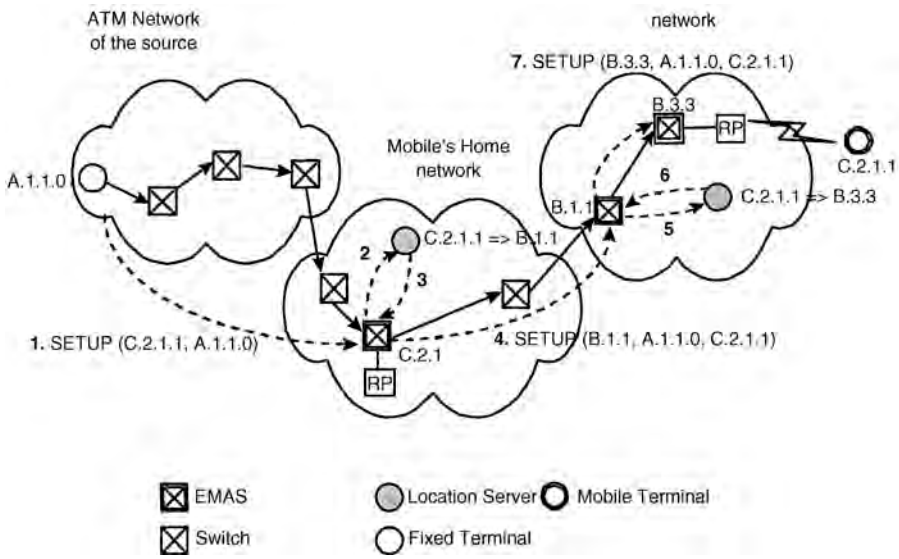


FIGURE 88.22 Hierarchical LS configuration.

set-up is sent via the home EMAS to the gateway EMAS. The gateway then queries its local LS to obtain the exact location (B.3.3) of the MT. It is assumed that the gateway can distinguish the MT address (C.2.1.1) in a SETUP message from the fact that this address has a different network prefix (C) than the one used in the visited network (B).

Signalling and Control Messages for LM

The signalling and control messages required can be derived from the scenarios above. Specifically, interactions between the EMAS and the LS require control message exchange over a VC established for this purpose. The ATM signalling support needed for connection set-up and redirection is described in [28].

Connection Handover in Wireless ATM

Wireless ATM implementations rely on mobile-initiated handovers whereby the MT is responsible for monitoring the radio link quality and decides when to initiate a handover [26]. A handover process typically involves the following steps:

1. *Link quality monitoring:* When there are active connections, the MT constantly monitors the strength of the signal it receives from each RP within range.
2. *Handover trigger:* At a given instance, all the connections from/to the MT are routed through the same RP, but deterioration in the quality of the link to this RP triggers the handover procedure.
3. *Handover initiation:* Once a handover is triggered, the MT initiates the procedure by sending a signal to the edge EMAS with which it is in direct contact. This signal indicates to the EMAS the list of candidate RPs to which active connections can be handed over.
4. *Target RP selection:* The edge EMAS selects one RP as the handover target from the list of candidates sent by the MT. This step may make use of network-specific criteria for spreading the traffic load among various RPs and interaction between the edge EMAS and other EMASs housing the candidate RPs.
5. *Connection rerouting:* Once the target RP is selected, the edge EMAS initiates the rerouting of all connections from/to the MT within the MATM network to the target RP. The complexity of

this step depends on the specific procedures chosen for rerouting connections, as described next. Due to constraints on the network or radio resources, it is possible that not all connections are successfully rerouted at the end of this step.

6. *Handover completion:* The MT is notified of the completion of handover for one or more active connections. The MT may then associate with the new RP and begin sending/receiving data over the connections successfully handed over.

Specific implementations may differ in the precise sequence of events during handover. Furthermore, the handover complexity and capabilities may be different. For instance, some systems may implement lossless handover whereby cell loss and missequencing of cells are avoided during handover by buffering cells inside the network [29]. The handover control flow is described in detail below for two types of handovers:

- *Backward handover:* The MT initiates handover through the current RP it is connected to. This is the normal scenario.
- *Forward handover:* The MT loses connectivity to the current RP due to a sudden degeneration of the radio link. It then chooses a new RP and initiates the handover of active connections.

The following description allows only hard handovers, i.e., active connections are routed via exactly one RP at a given instance, as opposed to soft handovers in which the MT can receive data for active connections simultaneously from more than one RP during handover.

Backward Handover Control Flow

Figure 88.23 depicts the control sequence for backward handover when the handover involves two different EMASs. Here, "old" and "new" EMAS refer to the current EMAS and the target EMAS, respectively. The figure does not show handover steps (1) and (2), which are radio layer functions, but starts with step (3). The following actions take place:

1. The MT initiates handover by sending an *HO_REQUEST* message to the old EMAS. With this message, the MT identifies a set of candidate RPs. Upon receiving the message, the old EMAS identifies a set of candidate EMASs that house the indicated RPs. It then sends an *HO_REQUEST_QUERY* to each candidate EMAS, identifying the candidate RP as well as the set of connections (including the traffic and QoS parameters) to be handed over. The connection identifiers are assumed to be unique within the network [28].
2. After receiving the *HO_REQUEST_QUERY* message, a candidate EMAS checks the radio resources available on all the candidate RPs it houses and selects the one that can accommodate the most number of connections listed. It then sends an *HO_REQUEST_RESPONSE* message to the old EMAS identifying the target RP chosen and a set of connections that can be accommodated (this may be a subset of connections indicated in the *QUERY* message).
3. After receiving an *HO_REQUEST_RESPONSE* message from all candidate EMASs, the old EMAS selects one target RP, based on some local criteria (e.g., traffic load spreading). It then sends an *HO_RESPONSE* message to the MT, indicating the target RP. At the same time, it also sends an *HO_COMMAND* to the new EMAS. This message identifies the target RP and all the connections to be handed over along with their ATM traffic and QoS parameters. This message may also indicate the connection rerouting method. Rerouting involves first the selection of a cross-over switch (COS) which is an EMAS in the existing connection path. A new connection segment is created from the new EMAS to the COS and the existing segment from the COS to the old EMAS is deleted. Some COS selection options are:

   VC Extension: The old EMAS-E itself serves as the COS.
   Anchor-Based Rerouting: The COS is determined apriori (e.g., a designated EMAS in the network) or during connection set-up (e.g., the EMAS that first served the MT when the

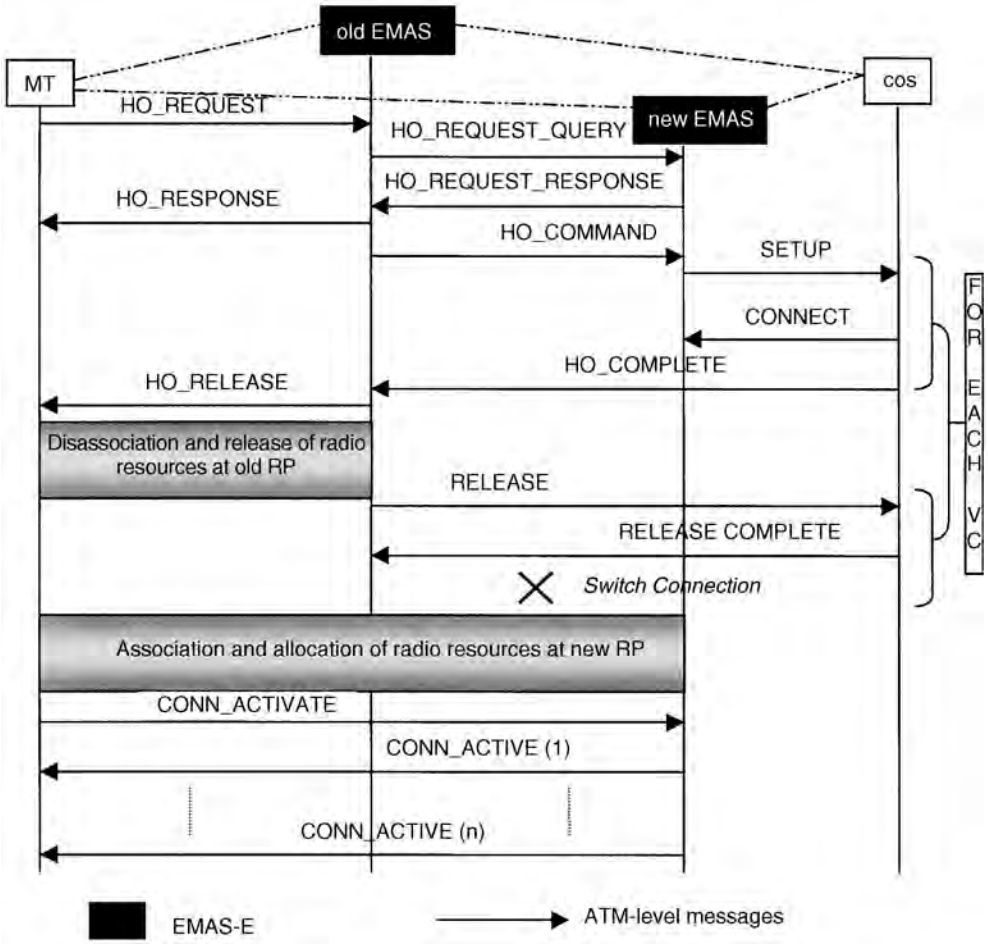


FIGURE 88.23 Backward handover control flow.

connection was set up). The selected COS is used for all handovers during the lifetime of the connection.

Dynamic COS Discovery: The COS is dynamically determined during each handover.

These procedures are illustrated in Fig. 88.24. While anchor-based rerouting and VC extension result in the same COS being used for all the connections being handed over, dynamic COS selection may result in different COSs for different connections. The *HO_COMMAND* message indicates which COS selection method is used, and if VC extension or anchor-based rerouting is used, it also includes the identity of the COS. If dynamic COS selection is used, the message includes the identity of the first EMAS in the connection path from the source (this information is collected during connection set-up [28]). For illustrative purposes, we assume that the dynamic COS selection procedure is used.

4. Upon receiving the *HO_COMMAND* message, the new EMAS allocates radio resources in the target RP for as many connections in the list as possible. It then sends a *SETUP* message toward the COS of each such connection. An EMAS in the existing connection path that first processes this message becomes the actual COS. This is illustrated in Fig. 88.25. This action, if successful, establishes a new segment for each connection from the new EMAS to the COS.

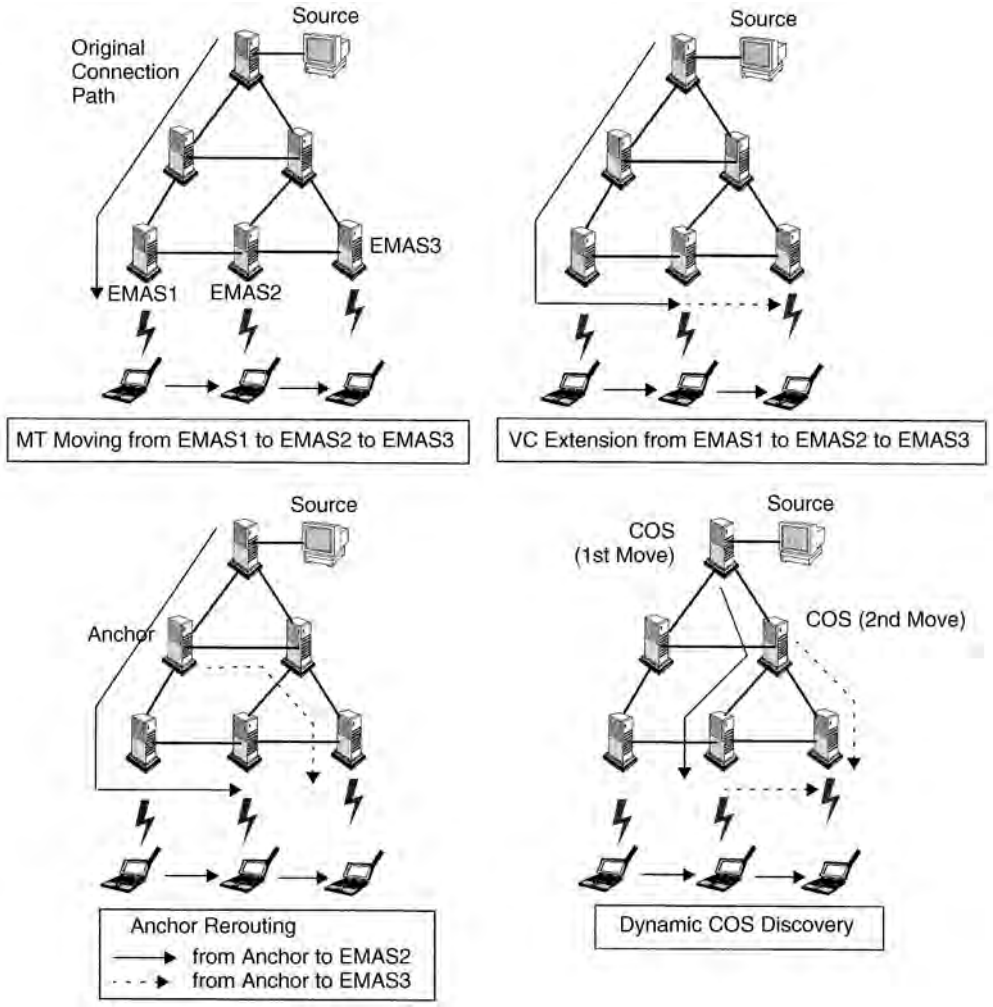


FIGURE 88.24 COS discovery.

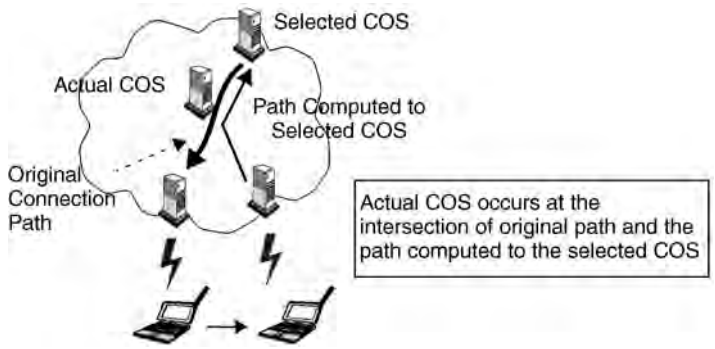


FIGURE 88.25 COS selection when MT moves.

5. If the set-up attempt is not successful, new EMAS sends an *HO_FAILURE* message to the old EMAS, after releasing all the local resources reserved in step (4). This message identifies the connection in question and it is forwarded by the old EMAS to the MT. What the MT does in response to an *HO_FAILURE* message is not part of the backward handover specification.
6. If the COS successfully receives the *SETUP* message, it sends a *CONNECT* message in reply, thereby completing the partial connection establishment procedure. It then sends an *HO_COMPLETE* message to old EMAS-E. The *HO_COMPLETE* message is necessary to deal with the situation when handover is simultaneously initiated by both ends of the connection when two MTs communicate (for the sake of simplicity, we omit further description of this situation, but the reader may refer to [30] for further details).
7. The old EMAS-E waits to receive *HO_COMPLETE* messages for all the connections being handed over. However, the waiting period is limited by the expiry of a timer. Upon receiving the *HO_COMPLETE* message for the last connection, or if the timer expires, the old EMAS sends an *HO_RELEASE* message to MT. Waiting for the *HO_RELEASE* message allows the MT to utilize the existing connection segment as long as possible thereby minimizing data loss. However, if the radio link deteriorates rapidly, the MT can switch over to the new RP without receiving the *HO_RELEASE* message.
8. The old EMAS initiates the release of each connection for which an *HO_COMPLETE* was received by sending a *RELEASE* message to the corresponding COS.
9. Upon receiving the *RELEASE* message, the COS sends a *RELEASE COMPLETE* to the previous switch in the path (as per regular ATM signalling) and switches the data flow from the old to the new connection segment.
10. Meanwhile, after receiving the *HO_RELEASE* message from the old EMAS or after link deterioration, the MT dissociates from the old RP and associates with the new RP. This action triggers the assignment of radio resources for the signalling channel and user data connections for which resources were reserved in step (4).
11. Finally, the MT communicates to the new EMAS its readiness to send and receive data on all connections that have been handed over by sending a *CONN_ACTIVATE* message.
12. Upon receiving the *CONN_ACTIVATE* message from the MT, new EMAS responds with a *CONN_ACTIVE* message. This message contains the identity of the connections that have been handed over, including their new ATM VC identifiers. Multiple *CONN_ACTIVE* messages may be generated, if all the connections have not been handed over when the *CONN_ACTIVATE* message was received. However, handover of remaining connections and the subsequent generation of *CONN_ACTIVE* signals are timer-bound: if the MT does not receive information about a connection in a *CONN_ACTIVE* message before the corresponding timer expires, it assumes that the connection was not successfully handed over. The recovery in this case is left up to the MT.

The description above has left open some questions. Among them: what mechanisms are used to reliably exchange control messages between the various entities that take part in handover? What actions are taken when network or radio link failures occur during handover? How can lossless handover be included in the control flow? What effect do transient disruptions in service during handover have on application behavior?, and what are the performance impacts of signalling for handover? The short answers to these questions are: reliability can be incorporated by implementing a reliable transfer protocol for those control messages that do not already use such a transport (*SETUP, RELEASE*, etc., do, but *HO_REQUEST_QUERY* for example, requires attention). Actions taken during network failures require further analysis, but forward handover can be used to recover from radio link failures during handover. Lossless handover requires inband signalling within each connection and buffering in the network. Details on this can be found in [29]. The effect of transient disruptions on applications can be minimal, depending on how rerouting is implemented during handover. This is described in detail in [31]. Finally, some of the performance issues related to mobility management are investigated in [32].

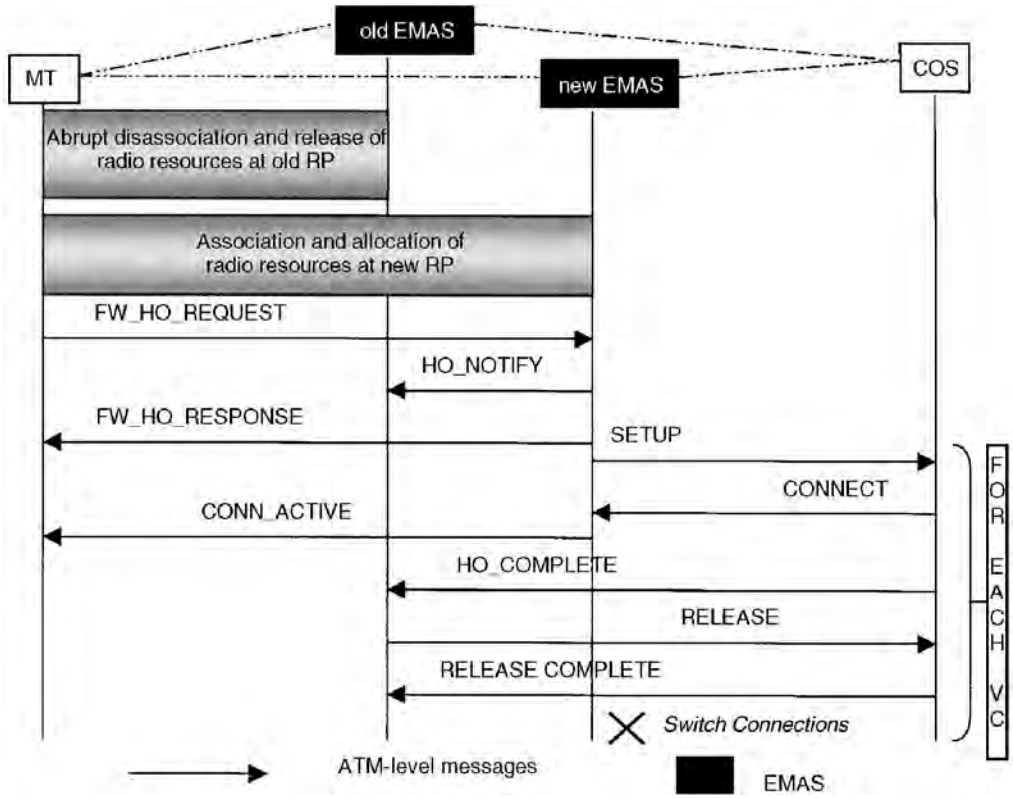


FIGURE 88.26 Forward handover control flow.

Forward Handover

Forward handover is considered as the measure of last resort, to be invoked when the current radio link deteriorates suddenly. The forward handover procedure is simpler than backward handover, as illustrated in Fig. 88.26. Here,

1. The radio link with the old EMAS degrades abruptly. This results in the MT being dissociated with the old RP. The MT then chooses a new RP and associates with it. In the example shown, the new RP is housed by a different EMAS (the new EMAS).
2. The MT initiates forward handover by sending a *FW_HO_REQUEST* message to the new EMAS. This message indicates the active connections by their network-wide unique identifiers, along with their ATM traffic and QoS parameters, the identity of the previous EMAS (the "old" EMAS), and the COS information (this information may be obtained when the connection is initially set up).
3. The new EMAS sends an *HO_NOTIFY* message to the old EMAS indicating the initiation of handover. This serves to keep the old EMAS from prematurely releasing the existing connections.
4. The new EMAS reserves radio resources for as many listed connections as possible on the radio port to which the MT is associated. It then sends a *FW_HO_RESPONSE* message to the MT identifying the connections that can be handed over.
5. For each such connection, the new EMAS generates a SETUP message toward the COS to establish the new connection segment. This message includes the identity of the new EMAS. An EMAS in the existing connection path that first processes this message becomes the COS (Fig. 88.25).
6. Upon receiving the *SETUP* message, the COS completes the establishment of the new connection segment by sending a *CONNECT* message to the new EMAS.

7. After receiving the *CONNECT* message, new EMAS sends a *CONN_ACTIVE* message to the MT, indicating the connection has become active. Reception of *CONN_ACTIVE* by the MT is subject to a timer expiry: if it does not receive information about a connection in any *CONN_ACTIVE* message before the corresponding timer expires, it may initiate any locally defined recovery procedure.
8. If the new connection segment cannot be setup, the new EMAS sends an *HO_FAILURE* message to the old EMAS and the MT, after releasing all the local resources reserved for the connection. Recovery in this case is left up to the MT.
9. If the COS did send a *CONNECT* in step 7, it switches the connection data to the new segment and sends an *HO_COMPLETE* message to old EMAS. As in the case of backward handover, the *HO_COMPLETE* message is necessary to resolve conflicts in COS selection when handover is simultaneously initiated by both ends of the connection when two MTs communicate.
10. Upon receiving the *HO_COMPLETE* message, the old EMAS releases the existing connection segment by sending a *RELEASE* message to the COS. In response, the COS sends a *RELEASE COMPLETE* to the previous switch.

88.4 Summary and Conclusions

In this chapter, the QoS and mobility management aspects of wireless ATM were described. WATM implementations, as well as the WATM standards being developed, allow the same ATM service categories in WATM networks as found in fixed ATM networks. The support for these classes of service in wireless ATM requires a variety of QoS control mechanisms acting in concert. The implementation of QoS in the wireless MAC layer, as well as the new QoS control mechanisms in the network and application layers, were described. The role of QoS renegotiation during handover was also described.

While mobility management in wide area involves various service-related aspects, the wireless ATM implementations and the standards efforts have so far focused on the core technical problems of location management and connection handover. These were described in some detail, along the lines of the specifications being developed by the ATM Forum with examples from WATM implementations. In conclusion, we note that wireless ATM is still an evolving area and much work is needed to fully understand how to efficiently support QoS and mobility management.

References

1. Raychaudhuri, D. and Wilson, N., ATM based transport architecture for multiservices wireless personal communication network, *IEEE J. Selected Areas in Commun.,* 1401–1414, Oct. 1994.
2. Raychaudhuri, D., Wireless ATM: An enabling technology for personal multimedia communications, in *Proc. Mobile Multimedia Commun. Workshop,* Bristol, U.K., Apr. 1995.
3. Raychaudhuri, D., Wireless ATM networks: Architecture, system design and prototyping, *IEEE Personal Commun.,* 42–49, Aug. 1996.
4. Singh, S., Quality of service guarantees in mobile computing, *Computer Communications,* 19, 359–371, 1996.
5. Naghshineh, M., Schwartz, M., and Acampora, A.S., Issues in wireless access broadband networks, in *Wireless Information Networks, Architecture, Resource Management, and Mobile Data,* Holtzman, J.M., Ed., Kluwer, 1996.
6. Raychaudhuri, D., French, L.J., Siracusa, R.J., Biswas, S.K., Yuan, R., Narasimhan, P., and Johnston, C.A., WATMnet: A prototype wireless ATM system for multimedia personal communication, *IEEE J. Selected Areas in Commun.,* 83–95, Jan. 1997.
7. Veeraraghavan, M., Karol, M.J., and Eng, K.Y., Mobility and connection management in a wireless ATM LAN, *IEEE J. Selected Areas in Commun.,* 50–68, Jan. 1997.
8. Ala-Laurila, J. and Awater, G., The magic WAND: Wireless ATM network demonstrator, in *Proc. ACTS Mobile Summit 97,* Denmark, Oct. 1997.

9. The ATM Forum, *Wireless ATM Capability Set O Spec*, 2002.
10. The ATM Forum Technical Committee, *ATM User-Network Signalling Specification,* Version 4.0, AF-95-1434R9, Jan. 1996.
11. The ATM Forum Technical Committee, *Traffic Management Specification,* Version 4.0, AF-95-0013R11, Mar. 1996.
12. Liu, K., Petr, D.W., Frost, V.S., Zhu, H., Braun, C., and Edwards, W.L., A bandwidth management framework for ATM-based broadband ISDN, *IEEE Communications Mag.,* 138–145, May 1997.
13. Johnston, C.A., Narasimhan, P., and Kokudo, J., Architecture and implementation of radio access protocols in wireless ATM networks, in *Proc. IEEE ICC 98,* Atlanta, Jun. 1998.
14. Passas, N., Paskalis, S., Vali, D., and Merakos, L., Quality-of-service-oriented medium access control for wireless ATM networks, *IEEE Communications Mag.,* 42–50, Nov. 1997.
15. Passas, N., Merakos, L., and Skyrianoglou, D., Traffic scheduling in wireless ATM networks, in *Proc. IEEE ATM 97 Workshop,* Lisbon, May 1997.
16. Raychaudhuri, D., Reininger, D., Ott, M., and Welling, G., Multimedia processing and transport for the wireless personal terminal scenario, *Proceedings SPIE Visual Communications and Image Processing Conference,* VCIP95, May 1995.
17. Reininger, D. and Izmailov, R., Soft Quality-of-Service with VBR$^+$ video, *Proceedings of 8th International Workshop on Packet Video (AVSPN97),* Aberdeen, Scotland, Sept. 1997.
18. Ott, M., Michelitsch, G., Reininger, D., and Welling, G., An architecture for adaptive QoS and its application to multimedia systems design *Computers and Communications,* 1997.
19. Microsoft Corporation., Windows Quality of Service Technology, White paper available on-line at http://www.microsoft.com/ntserver/
20. Reininger, D., Raychaudhuri, D., and Hui, J., Dynamic bandwidth allocation for VBR video over ATM networks, *IEEE Journal on Selected Areas in Communications,* 14(6), 1076–1086, Aug. 1996.
21. Lourens, J.G., Malleson, H.H., and Theron, C.C., Optimization of bit-rates, for digitally compressed television services as a function of acceptable picture quality and picture complexity, *Proceedings IEE Colloquium on Digitally Compressed TV by Satellite,* 1995.
22. Nakasu, E., Aoi, K., Yajima, R., Kanatsugu, Y., and Kubota, K., A statistical analysis of MPEG-2 picture quality for television broadcasting, *SMPTE Journal,* 702–711, Nov. 1996.
23. ITU-T, ATM Traffic Descriptor Modification by the connection owner, ITU-T Q.2963.2, Sept. 1997.
24. Mouly, M. and Pautet, M.-B., *The GSM System for Mobile Communications,* Cell & Sys, Palaiseau, France, 1992.
25. Brown, D., Techniques for privacy and authentication in personal communication systems, *IEEE Personal Comm.,* Aug. 1985.
26. Acharya, A., Li, J., Rajagopalan, B., and Raychaudhuri, D., Mobility management in wireless ATM networks, *IEEE Communications Mag.,* 100–109, Nov. 1997.
27. Tabbane, S., Location management methods for third-generation mobile systems, *IEEE Communications Mag.,* Aug. 1997.
28. Acharya, A., Li, J., Bakre, A., and Raychaudhuri, D., Design and prototyping of location management and handoff protocols for wireless ATM networks, in *Proc. ICUPC,* San Diego, Nov. 1997.
29. Mitts, H., Hansen, H., Immonen, J., and Veikkolainen, Lossless handover in wireless ATM, *Mobile Networks and Applications,* 299–312, Dec. 1996.
30. Rajagopalan, B., Mobility management in integrated wireless ATM networks, *Mobile Networks and Applications,* 273–286, Dec. 1996.
31. Mishra, P. and Srivastava, M., Effect of connection rerouting on application performance in mobile networks, in *Proc. IEEE Conf. on Distributed Computing Syst.,* May 1997.
32. Pollini, G.P., Meier-Hellstern, K.S., and Goodman, D.J., Signalling traffic volume generated by mobile and personal communications, *IEEE Communications Mag.,* Jun. 1995.

89

An Overview of cdma2000, WCDMA, and EDGE

Tero Ojanperä
Nokia Group

Steven D. Gray
Nokia Research Center

In response to the International Telecommunications Union's (ITU) call for proposals, third generation cellular technologies are evolving at a rapid pace where different proposals are vying for the future market place in digital wireless multimedia communications. While the original intent for third generation was to have a convergence of cellular based technologies, this appears to be an unrealistic expectation. As such, three technologies key for the North American and European markets are the third generation extension of TIA/EIA-95B based Code Division Multiple Access (CDMA) called cdma2000, the European third generation CDMA called WCDMA, and the third generation Time Division Multiple Access (TDMA) system based on EDGE. For packet data, EDGE is one case where second generation technologies converged to a single third generation proposal with convergence of the US TDMA system called TIA/EIA-136 and the European system GSM. This chapter provides an overview of the air interfaces of these key technologies. Particular attention is given to the channel structure, modulation, and offered data rates of each technology. A comparison is also made between cdma2000 and WCDMA to help the reader understand the similarities and differences of these two CDMA approaches for third generation.

89.1 Introduction

The promise of third generation is a world where the subscriber can access the World Wide Web (WWW) or perform file transfers over packet data connections capable of providing 144 kbps for high mobility, 384 kbps with restricted mobility, and 2 Mbps in an indoor office environment [1]. With these guidelines on rate from the ITU, standards bodies started the task of developing an air interface for their third generation system. In North America, the Telecommunications Industry Association (TIA) evaluated proposals from TIA members pertaining to the evolution of TIA/EIA-95B and TIA/EIA-136. In Europe, the European Telecommunications Standards Institute (ETSI) evaluated proposals from ETSI members pertaining to the evolution of GSM.

While TIA and ETSI were still discussing various targets for third generation systems, Japan began to roll out their contributions for third generation technology and develop proof-of-concept prototypes. In the beginning of 1997, the Association for Radio Industry and Business (ARIB), a body responsible for standardization of the Japanese air interface, decided to proceed with the detailed standardization of a wideband CDMA system. The technology push from Japan accelerated standardization in Europe and the U.S. During 1997, joint parameters for Japanese and European wideband CDMA proposals were agreed. The air interface is commonly referred to as WCDMA. In January 1998, the strong support behind wideband CDMA led to the selection of WCDMA as the UMTS terrestrial air interface scheme for FDD (Frequency Division Duplex) frequency bands in ETSI. In the U.S., third generation CDMA came through a detailed proposal process from vendors interested in the evolution of TIA/EIA-95B. In February 1998, the TIA committee TR45.5 responsible for TIA/EIA-95B standardization adopted a framework that combined the different vendors' proposals and later became known as cdma2000.

For TDMA, the focus has been to offer IS-136 and GSM operators a competitive third generation evolution. WCDMA is targeted toward GSM evolution; however, Enhanced Data Rates for Global TDMA Evolution (EDGE) allows the operators to supply IMT-2000 data rates without the spectral allocation requirements of WCDMA. Thus, EDGE will be deployed by those operators who wish to maintain either IS-136 or GSM for voice services and augment these systems with a TDMA-based high rate packet service. TDMA convergence occurred late in 1997 when ETSI approved standardization of the EDGE concept and in February 1998 when TIA committee TR45.3 approved the UWC-136 EDGE-based proposal.

The push to third generation was initially focused on submission of an IMT-2000 radio transmission techniques (RTT) proposal. To date, the evaluation process has recently started in ITU [2] where Fig. 89.1 depicts the time schedule of the ITU RTT development. Since at the same time regional standards have started the standards writing process, it is not yet clear what is the relationship between the ITU and regional standards. Based upon actions in TIA and ETSI, it is reasonable to assume that standards will exist for cdma2000, WCDMA, and EDGE and all will be deployed based upon market demands.

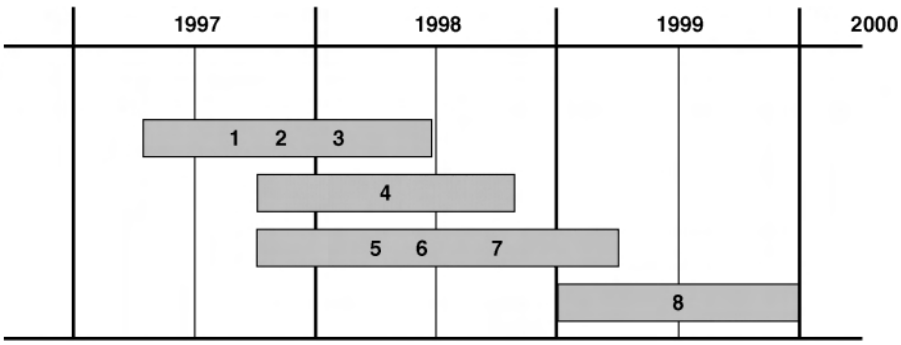


FIGURE 89.1 ITU timelines: 1, 2, 3—RTTs request, development, and submission; 4—RTT evaluation; 5—review outside evaluation; 6—assess compliance with performance parameters; 7—consideration of evaluation results and consensus on key characteristics; 8—development of detailed radio interface specifications.

The chapter is organized as follows: issues effecting third generation CDMA are discussed followed by a brief introduction of cdma2000, WCDMA, and EDGE. A table comparing cdma2000 and WCDMA is given at the end of the CDMA section. For TDMA, an overview of the IS-136-based evolution is given including the role played by EDGE.

89.2 CDMA-Based Schemes

Third generation CDMA system descriptions in TIA and ETSI have similarities and differences. Some of the similarities between cdma2000 and WCDMA are variable spreading, convolutional coding, and QPSK data modulation. The major differences between cdma2000 and WCDMA occur with the channel structure, including the structure of the pilot used on the forward link. To aid in comparison of the two CDMA techniques, a brief overview is given to some important third generation CDMA issues, the dedicated channel structure of cdma2000 and WCDMA, and a table comparing air interface characteristics.

89.3 CDMA System Design Issues

Bandwidth

An important design goal for all third generation proposals is to limit spectral emissions to a 5 MHz dual-sided passband. There are several reasons for choosing this bandwidth. First, data rates of 144 and 384 kbps, the main targets of third generation systems, are achievable within 5 MHz bandwidth with reasonable coverage. Second, lack of spectrum calls for limited spectrum allocation, especially if the system has to be deployed within the existing frequency bands already occupied by the second generation systems. Third, the 5 MHz bandwidth improves the receiver's ability to resolve multipath when compared to narrower bandwidths, increasing diversity and improving performance. Larger bandwidths of 10, 15, and 20 MHz have been proposed to support highest data rates more effectively.

Chip Rate

Given the bandwidth, the choice of chip rate depends on spectrum deployment scenarios, pulse shaping, desired maximum data rate and dual-mode terminal implementation. Figure 89.2 shows the relation between chip rate (CR), pulse shaping filter roll-off factor (α) and channel separation (Δf). If raised cosine filtering is used, spectrum is zero (in theory) after $CR/2(1 + \alpha)$. In Fig. 89.2, channel separation is selected such that two adjacent channel spectra do not overlap. Channel separation should be selected this way, if there can be high power level differences between the adjacent carriers. For example, for WCDMA parameters minimum channel separation (Δf_{min}) for nonoverlapping carriers is $\Delta f_{min} = 4.096(1 + 0.22) = 4.99712$ MHz. If channel separation is selected in such a way that the spectrum of two adjacent channel signals overlap,

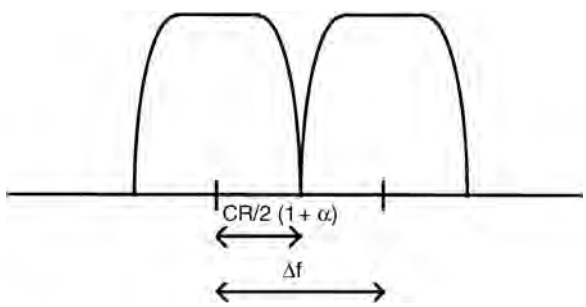


FIGURE 89.2 Relationship between chip rate (CR), roll-off factor (α), and channel separation (Δf).

some power leaks from one carrier to another. Partly overlapping carrier spacing can be used, for example, in micro cells where the same antenna masts are used for both carriers.

A designer of dual-mode terminals needs to consider the relation between the different clock frequencies of different modes. Especially important are the transmitter and receiver sampling rates and the carrier raster. A proper selection of these frequencies for the standard would ease the dual mode terminal implementation. The different clock frequencies in a terminal are normally derived from a common reference oscillator by either direct division or synthesis by the use of a PLL. The use of a PLL will add some complexity. The WCDMA chip rate has been selected based on consideration of backward compatibility with GSM and PDC. cdma2000 chip rate is a direct derivation of the TIA/EIA-95B chip rate.

Multirate

Multirate design means multiplexing different connections with different quality of service requirements in a flexible and spectrum efficient way. The provision for flexible data rates with different quality of service requirements can be divided into three subproblems: how to map different bit rates into the allocated bandwidth, how to provide the desired quality of service, and how to inform the receiver about the characteristics of the received signal. The first problem concerns issues like multicode transmission and variable spreading. The second problem concerns coding schemes. The third problem concerns control channel multiplexing and coding.

Multiple services belonging to the same session can be either time- or code-multiplexed as depicted in Fig. 89.3. The time multiplexing avoids multicode transmissions thus reducing peak-to-average power of the transmission. A second alternative for service multiplexing is to treat parallel services completely separate with separate channel coding/interleaving. Services are then mapped to separate physical data channels in a multicode fashion as illustrated in the lower part of Fig. 89.3. With this alternative scheme, the power, and consequently the quality, of each service can be controlled independently.

Spreading and Modulation Solutions

A complex spreading circuit as shown in Fig. 89.4 helps to reduce the peak-to-average power and thus improves power efficiency.

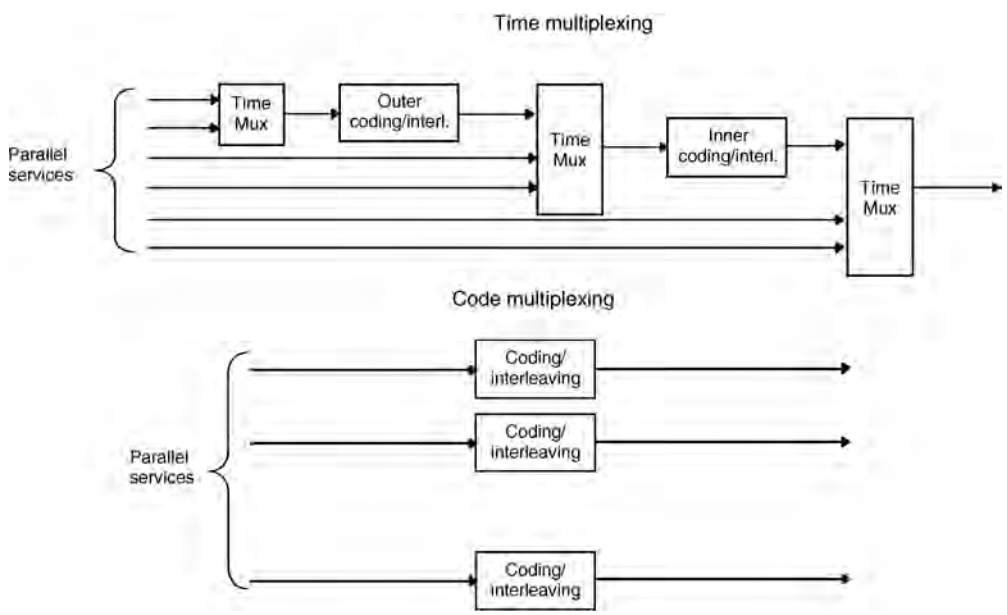


FIGURE 89.3 Time and code multiplexing principles.

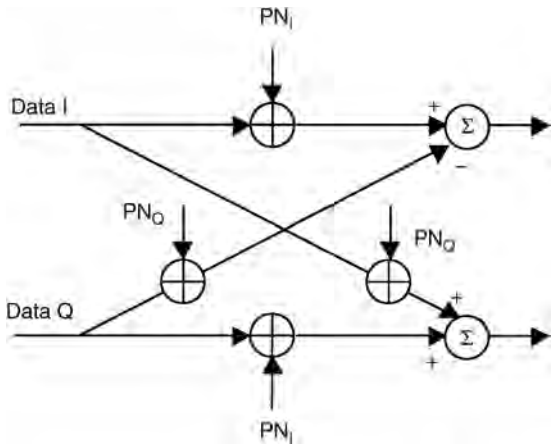


FIGURE 89.4 Complex spreading.

The spreading modulation can be either balanced- or dual-channel QPSK. In the balanced QPSK spreading the same data signal is split into I and Q channels. In dual-channel QPSK spreading the symbol streams on the I and Q channels are independent of each other. In the forward link, QPSK data modulation is used in order to save code channels and allow the use of the same orthogonal sequence for I and Q channels. In the reverse link, each mobile station uses the same orthogonal codes; this allows for efficient use of BPSK data modulation and balanced QPSK spreading.

Coherent Detection in the Reverse Link

Coherent detection can improve the performance of the reverse link up to 3 dB compared to noncoherent reception used by the second generation CDMA system. To facilitate coherent detection a pilot signal is required. The actual performance improvement depends on the proportion of the pilot signal power to the data signal power and the fading environment.

Fast Power Control in Forward Link

To improve the forward link performance fast power control is used. The impact of the fast power control in the forward link is twofold. First, it improves the performance in a fading multipath channel. Second, it increases the multiuser interference variance within the cell since orthogonality between users is not perfect due to multipath channel. The net effect, however, is improved performance at low speeds.

Additional Pilot Channel in the Forward Link for Beamforming

An additional pilot channel on the forward link that can be assigned to a single mobile or to a group of mobiles enables deployment of adaptive antennas for beamforming since the pilot signal used for channel estimation needs to go through the same path as the data signal. Therefore, a pilot signal transmitted through an omnicell antenna cannot be used for the channel estimation of a data signal transmitted through an adaptive antenna.

Seamless Interfrequency Handover

For third generation systems hierarchical cell structures (HCS), constructed by overlaying macro cells on top of smaller micro or pico cells, have been proposed to achieve high capacity. The cells belonging to different cell layers will be in different frequencies, and thus an interfrequency handover is required. A key requirement for the support of seamless interfrequency handover is the ability of the mobile station

to carry out cell search on a carrier frequency different from the current one, without affecting the ordinary data flow. Different methods have been proposed to obtain multiple carrier frequency measurements. For mobile stations with receiver diversity, there is a possibility for one of the receiver branches to be temporarily reallocated from diversity reception and instead carry out reception on a different carrier. For single-receiver mobile stations, slotted forward link transmission could allow interfrequency measurements. In the slotted mode, the information normally transmitted during a certain time, e.g., a 10 ms frame, is transmitted in less than that time, leaving an idle time that the mobile can use to measure on other frequencies.

Multiuser Detection

Multiuser detection (MUD) has been the subject of extensive research since 1986 when Verdu formulated an optimum multiuser detector for AWGN channel, maximum likelihood sequence estimation (MLSE) [3]. In general, it is easier to apply MUD in a system with short spreading codes since cross-correlations do not change every symbol as with long spreading codes. However, it seems that the proposed CDMA schemes would all use long spreading codes. Therefore, the most feasible approach seems to be interference cancellation algorithms that carry out the interference cancellation at the chip level, thereby avoiding explicit calculation of the cross-correlation between spreading codes from different users [4]. Due to complexity, MUD is best suited for the reverse link. In addition, the mobile station is interested in detecting its own signal in contrast to the base station, which needs to demodulate the signals of all users. Therefore, a simpler interference suppression scheme could be applied in the mobile station. Furthermore, if short spreading codes are used, the receiver could exploit the cyclostationarity, i.e., the periodic properties of the signal, to suppress interference without knowing the interfering codes.

Transmit Diversity

The forward link performance can be improved in many cases by using transmit diversity. For direct spread CDMA schemes, this can be performed by splitting the data stream and spreading the two streams using orthogonal sequences or switching the entire data stream between two antennas. For multicarrier CDMA, the different carriers can be mapped into different antennas.

89.4 WCDMA

To aid in the comparison of cdma2000 and WCDMA, the dedicated frame structure of WCDMA is illustrated in Figs. 89.5 and 89.6. The approach follows a time multiplex philosophy where the Dedicated Physical Control Channel (DPCCH) provides the pilot, power control, and rate information and the Dedicated Physical Data Channel (DPDCH) is the portion used for data transport. The forward and

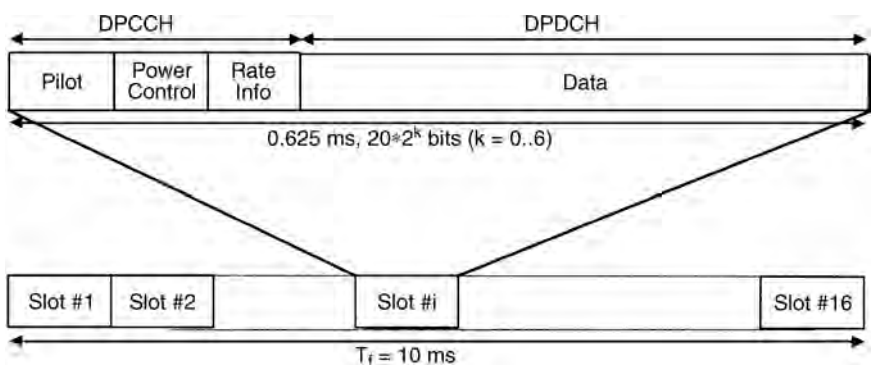


FIGURE 89.5 Forward link dedicated channel structure in WCDMA.

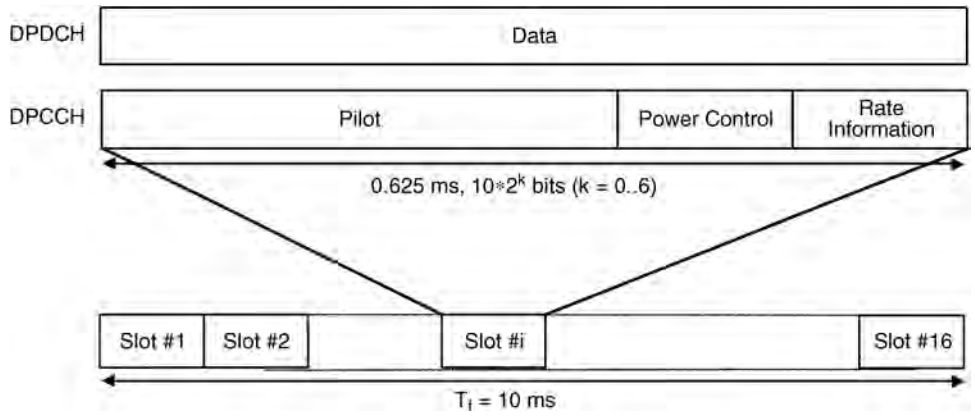


FIGURE 89.6 Reverse link dedicated channel structure in WCDMA.

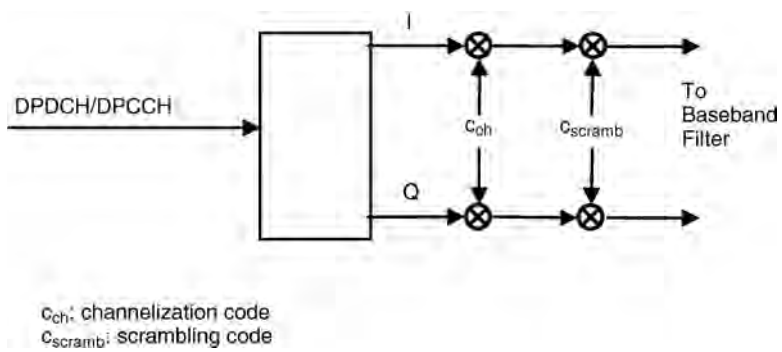


FIGURE 89.7 Forward link spreading of DPDCH and DPCCH.

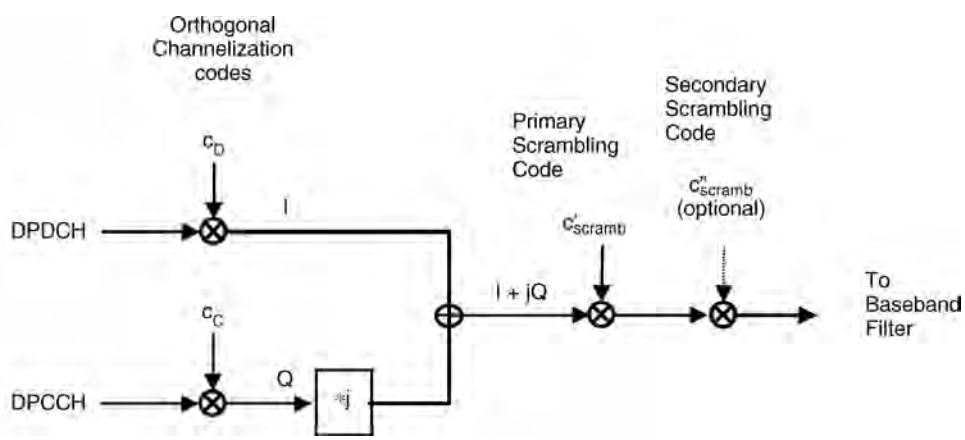


FIGURE 89.8 Reverse link spreading for the DPDCH and DPCCH.

reverse DPDCH channels have been convolutional encoded and interleaved prior to framing. The major difference between the forward and reverse links is that the reverse channel structure of the DPCCH is a separate code channel from the DPDCH.

After framing, the forward and reverse link channels are spread as shown in Figs. 89.7 and 89.8. On the forward link orthogonal, variable rate codes, c_{ch}, are used to separate channels and pseudo random

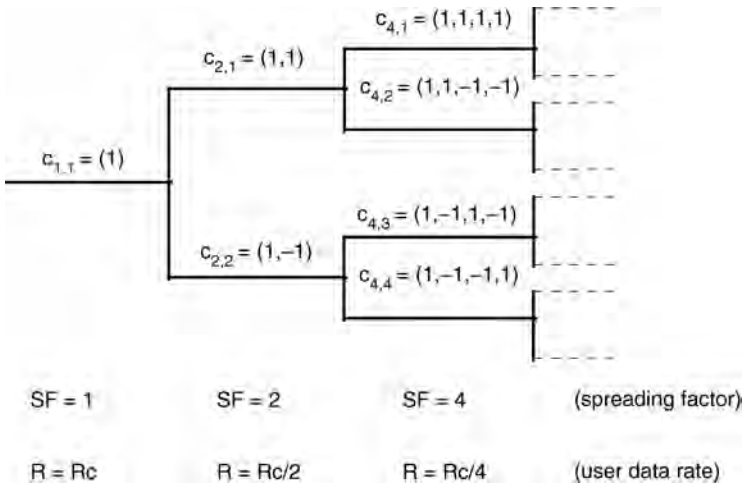


FIGURE 89.9 Construction of orthogonal spreading codes for different spreading factors.

scrambling sequences, c_{scramb}, are used to spread the signal evenly across the spectrum and separate different base stations. On the reverse link, the orthogonal channelization codes are used as in the forward link to separate CDMA channels. The scrambling codes, c'_{scramb} and c''_{scramb}, are used to identify mobile stations and to spread the signal evenly across the band. The optional scrambling code is used as a means to group mobiles under a common scrambling sequence.

Spreading Codes

WCDMA employs long spreading codes. Different spreading codes are used for cell separation in the forward link and user separation in the reverse link. In the forward link Gold codes of length 2^{18} are truncated to form cycles of 2^{16} times 10 ms frames. In order to minimize the cell search time, a special short code mask is used. The synchronization channel of WCDMA is masked with an orthogonal short Gold code of length 256 chips spanning one symbol. The mask symbols carry information about the BS long code group. Thus, the mobile station first acquires the short mask code and then searches the corresponding long code. A short VL-Kasami code has been proposed for the reverse link to ease the implementation of multiuser detection. In this case, code planning would also be negligible because the number of VL-Kasami sequences is more than one million. However, in certain cases, the use of short codes may lead to bad correlation properties, especially with very small spreading factors. If multiuser detection were not used, adaptive code allocation could be used to restore the cross-correlation properties. The use of short codes to ease the implementation of advanced detection techniques is more beneficial in the forward link since the cyclostationarity of the signal could be utilized for adaptive implementation of the receiver.

Orthogonality between the different spreading factors can be achieved by tree-structured orthogonal codes whose construction is illustrated in Fig. 89.9 [5]. The tree-structured codes are generated recursively according to the following equation,

$$c_{2n} = \begin{pmatrix} c_{2n,1} \\ c_{2n,2} \\ \vdots \\ c_{2n,2n} \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} c_{n,1} & c_{n,1} \\ c_{n,1} & -c_{n,1} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} c_{n,n} & c_{n,n} \\ c_{n,n} & -c_{n,n} \end{pmatrix} \end{pmatrix}$$

where C_{2n} is the orthogonal code set of size $2n$. The generated codes within the same layer constitute a set of orthogonal functions and are thus orthogonal. Furthermore, any two codes of different layers are also orthogonal except for the case that one of the two codes is a mother code of the other. For example code $c_{4,4}$ is not orthogonal with codes $c_{1,1}$ and $c_{2,2}$.

Coherent Detection and Beamforming

In the forward link, time-multiplexed pilot symbols are used for coherent detection. Because the pilot symbols are user dedicated, they can be used for channel estimation with adaptive antennas as well. In the reverse link, WCDMA employs pilot symbols multiplexed with power control and rate information for coherent detection.

Multirate

WCDMA traffic channel structure is based on a single code transmission for small data rates and multicode for higher data rates. Multiple services belonging to the same connection are, in normal cases, time multiplexed as was depicted in the upper part of Fig. 89.3. After service multiplexing and channel coding, the multiservice data stream is mapped to one or more dedicated physical data channels. In the case of multicode transmission, every other data channel is mapped into Q and every other into I channel. The channel coding of WCDMA is based on convolutional and concatenated codes. For services with BER = 10^{-3}, convolutional code with constraint length of 9 and different code rates (between 1/2-1/4) is used. For services with BER = 10^{-6}, a concatenated coding with an outer Reed-Solomon code has been proposed. Typically, block interleaving over one frame is used. WCDMA is also capable of interframe interleaving, which improves the performance for services allowing longer delay. Turbo codes for data services are under study. Rate matching is performed by puncturing or symbol repetition.

Packet Data

WCDMA has two different types of packet data transmission possibilities. Short data packets can be appended directly to a random access burst. The WCDMA random-access burst is 10 ms long, it is transmitted with fixed power, and the access principle is based on the slotted Aloha scheme. This method, called common channel packet transmission, is used for short infrequent packets, where the link maintenance needed for a dedicated channel would lead to an unacceptable overhead. Larger or more frequent packets are transmitted on a dedicated channel. A large single packet is transmitted using a single-packet scheme where the dedicated channel is released immediately after the packet has been transmitted. In a multipacket scheme the dedicated channel is maintained by transmitting power control and synchronization information between subsequent packets.

89.5 cdma2000

The dedicated channels used in cdma2000 system are the fundamental, supplemental, pilot,[1] and dedicated control channels. Shown for the forward link in Fig. 89.10 and for the reverse in Fig. 89.11, the fundamental channel provides for the communication of voice, low rate data, and signalling where power control information for the reverse channels is punctured on the forward fundamental channel. For high rate data services, the supplemental channel is used where one important difference between the supplemental and the fundamental channel is the addition of parallel-concatenated turbo codes. For different service options, multiple supplemental channels can be used. The code multiplex pilot channel allows for phase coherent detection. In addition, the pilot channel on the forward link is used for determining soft handoff and the pilot channel on the reverse is used for carrying power control information for the

[1]Dedicated for the reverse link and common for the forward link.

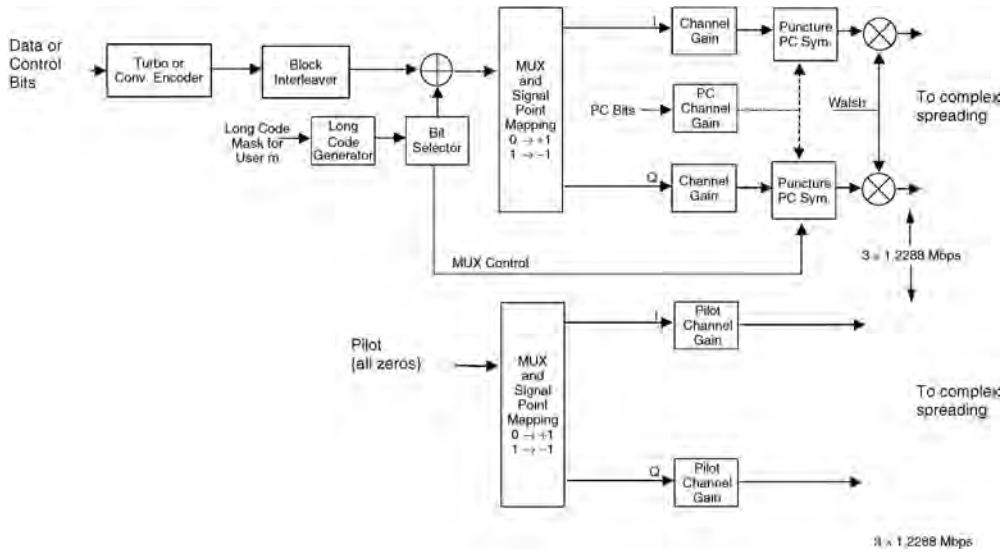


FIGURE 89.10 Forward link channel structure in cdma2000 for direct spread. (*Note:* dashed line indicates that it is only used for the fundamental channel).

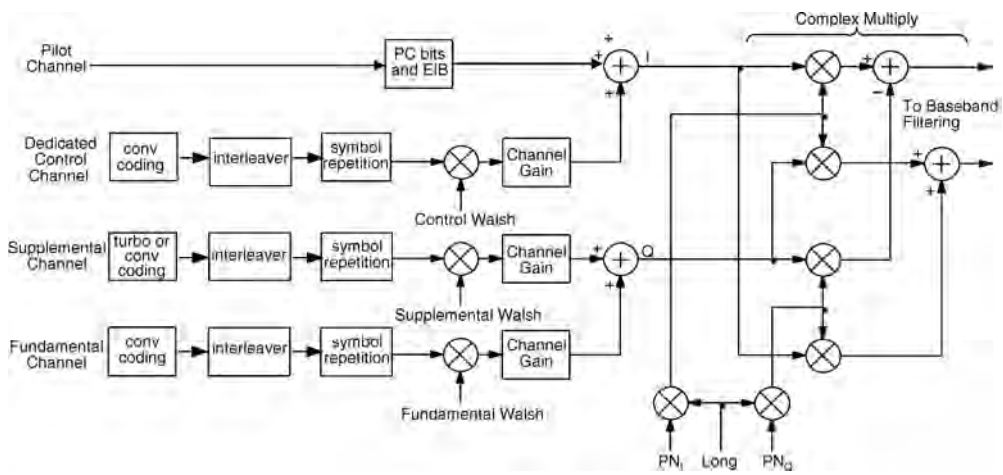


FIGURE 89.11 Reverse link channel structure in cdma2000.

forward channels. Finally, the dedicated control channel, also shown in Fig. 89.10 for the forward link and in Fig. 89.11, for the reverse, is used primarily for exchange of high rate Media Access Control (MAC) layer signalling.

Multicarrier

In addition to direct spread, a multicarrier approach has been proposed for the cdma2000 forward link since it would maintain orthogonality between the cdma2000 and TIA/EIA-95B carriers [6]. The multicarrier variant is achieved by using three 1.25 MHz carriers for a 5 MHz bandwidth where all carriers have separate channel coding and are power controlled in unison.

Spreading Codes

On the forward link, the cell separation for cdma2000 is performed by two M-sequences of length 3×2^{15}, one for I and one for Q channel, which are phase shifted by PN-offset for different cells. Thus, during the cell search process only these sequences are searched. Because there are a limited number of PN-offsets, they need to be planned in order to avoid PN-confusion [7]. In the reverse link, user separation is performed by different phase shifts of M-sequence of length 2^{41}. The channel separation is performed using variable spreading factor Walsh sequences, which are orthogonal to each other.

Coherent Detection

In the forward link, cdma2000 has a common pilot channel, which is used as a reference signal for coherent detection when adaptive antennas are not employed. When adaptive antennas are used, an auxiliary pilot is used as a reference signal for coherent detection. Code multiplexed auxiliary pilots are generated by assigning a different orthogonal code to each auxiliary pilot. This approach reduces the number of orthogonal codes available for the traffic channels. This limitation is alleviated by expanding the size of the orthogonal code set used for the auxiliary pilots. Since a pilot signal is not modulated by data, the pilot orthogonal code length can be extended, thereby yielding an increased number of available codes, which can be used as additional pilots. In the reverse link, the pilot signal is time multiplexed with power control and erasure indicator bit (EIB).

Multirate Scheme

cdma2000 has two traffic channel types, the fundamental and the supplemental channel, which are code multiplexed. The fundamental channel is a variable rate channel which supports basic rates of 9.6 kbps and 14.4 kbps and their corresponding subrates, i.e., Rate Set 1 and Rate Set 2 of TIA/EIA-95B. It conveys voice, signalling, and low rate data. The supplemental channel provides high data rates. Services with different QoS requirements are code multiplexed into supplemental channels. The user data frame length of cdma2000 is 20 ms. For the transmission of control information, 5 and 20 ms frames can be used on the fundamental channel or dedicated control channel. On the fundamental channel a convolutional code with constraint length 9 is used. On supplemental channels convolutional coding is used up to 14.4 kbps. For higher rates Turbo codes with constraint length 4 and rate 1/4 are preferred. Rate matching is performed by puncturing, symbol repetition, and sequence repetition.

Packet Data

cdma2000 also allows short data burst using the slotted Aloha principle. However, instead of fixed transmission power it increases the transmission power for the random access burst after an unsuccessful access attempt. When the mobile station has been allocated a traffic channel, it can transmit without scheduling up to a predefined bit rate. If the transmission rate exceeds the defined rate, a new access request has to be made. When the mobile station stops transmitting, it releases the traffic channel but not the dedicated control channel. After a while it also releases the dedicated control channel as well but maintains the link layer and network layer connections in order to shorten the channel set-up time when new data needs to be transmitted.

Parametric Comparison

For comparison, Table 89.1 lists the parameters of cdma2000 and WCDMA. cdma2000 uses a chip rate of 3.6864 Mcps for the 5 MHz band allocation with the direct spread forward link option and a 1.2288 Mcps chip rate with three carriers for the multicarrier option. WCDMA uses direct spread with a chip rate of 4.096 Mcps. The multicarrier approach is motivated by a spectrum overlay of cdma2000 carriers with existing TIA/EIA-95B carriers [6]. Similar to EIA/TIA-95B, the spreading codes of cdma2000 are generated using different phase shifts of the same M-sequence. This is possible due to the synchronous

TABLE 89.1 Parameters of WCDMA and cdma2000

	WCDMA	cdma2000
Channel bandwidth	5, 10, 20 MHz	1.25, 5, 10, 15, 20 MHz
Forward link RF channel structure	Direct spread	Direct spread or multicarrier
Chip rate	4.096/8.192/16.384 Mcps	1.2288/3.6864/7.3728/11.0593/14.7456 Mcps for direct spread $n \times 1.2288$ Mcps (n = 1,3,6,9,12) for multicarrier
Roll-off factor	0.22	Similar to TIA/EIA-95B
Frame length	10 ms/20 ms (optional)	20 ms for data and control/5 ms for control information on the fundamental and dedicated control channel
Spreading modulation	Balanced QPSK (forward link) Dual channel QPSK (reverse link) Complex spreading circuit	Balanced QPSK (forward link) Dual channel QPSK (reverse link) Complex spreading circuit
Data modulation	QPSK (forward link) BPSK (reverse link)	QPSK (forward link) BPSK (reverse link)
Coherent detection	User dedicated time multiplexed pilot (forward link and reverse link), common pilot in forward link	Pilot time multiplexed with PC and EIB (reverse link) Common continuous pilot channel and auxiliary pilot (forward link)
Channel multiplexing in reverse link	Control and pilot channel time multiplexed I&Q multiplexing for data and control channel	Control, pilot fundamental, and supplemental code multiplexed I&Q multiplexing for data and control channels
Multirate	Variable spreading and multicode	Variable spreading and multicode
Spreading factors	4-256 (4.096 Mcps)	4-256 (3.6864 Mcps)
Power control	Open and fast closed loop (1.6 kHz)	Open loop and fast closed loop (800 Hz)
Spreading (forward link)	Variable length orthogonal sequences for channel separation. Gold sequences for cell and user separation	Variable length Walsh sequences for channel separation, M-sequence 3×2^{15} (same sequence with time shift utilized in different cells different sequence in I&Q channel)
Spreading (reverse link)	Variable length orthogonal sequences for channel separation. Gold sequence 2^{41} for user separation (different time shifts in I and Q channel, cycle 2^{16} 10 ms radio frames)	Variable length orthogonal sequences for channel separation, M-sequence 2^{15} (same for all users different sequences in I&Q channels), M-sequence 2^{41} for user separation (different time shifts for different users)
Handover	Soft handover Interfrequency handover	Soft handover Interfrequency handover

network operation. Since WCDMA has an asynchronous network, different long codes rather than different phase shifts of the same code are used for the cell and user separation. The code structure determines how code synchronization, cell acquisition, and handover synchronization are performed.

89.6 TDMA-Based Schemes

As discussed, TIA/EIA-136 and GSM evolution have similar paths in the form of EDGE. The UWC-136 IMT 2000 proposal contains, in addition to the TIA/EIA-136 30 kHz carriers, the high rate capability provided by the 200 kHz and 1.6 MHz carriers shown in Table 89.2. The targets for the IS-136 evolution were to meet IMT-2000 requirements and an initial deployment within 1 MHz spectrum allocation. UWC-136 meets these targets via modulation enhancement to the existing 30 kHz channel (136+) and by defining complementary wider band TDMA carriers with bandwidths of 200 kHz for vehicular/outdoor environments and 1.6 MHz for indoor environments. The 200 kHz carrier, 136 HS (vehicular/outdoor)

TABLE 89.2 Parameters of 136 HS

	136 HS (Vehicular/Outdoor)	136 HS (Indoor)
Duplex method	FDD	FDD and TDD
Carrier spacing	200 kHz	1.6 MHz
Modulation	Q-O-QAM	Q-O-QAM
	B-O-QAM	B-O-QAM
	8 PSK	
	GMSK	
Modulation bit rate	722.2 kbps (Q-O-QAM)	5200 kbps (Q-O-QAM)
	361.1 kbps (B-O-QAM)	2600 kbps (B-O-QAM)
	812.5 kbps (8 PSK)	
	270.8 kbps (GMSK)	
Payload	521.6 kbps (Q-O-QAM)	4750 kbps (Q-O-QAM)
	259.2 kbps (B-O-QAM)	2375 kbps (B-O-QAM)
	547.2 kbps (8 PSK)	
	182.4 kbps (GMSK)	
Frame length	4.615 ms	4.615 ms
Number of slots	8	64 (72 μs)
		16 (288 μs)
Coding	Convolutional	Convolutional
	1/2, 1/4, 1/3, 1/1	1/2, 1/4, 1/3, 1/1
	ARQ	Hybrid Type II ARQ
Frequency hopping	Optional	Optional
Dynamic channel allocation	Optional	Optional

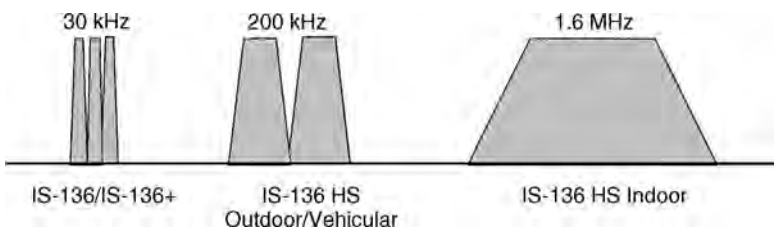


FIGURE 89.12 UWC-136 carrier types.

with the same parameters as EDGE provides medium bit rates up to 384 kbps and the 1.6 MHz carrier, 136 HS (indoor), highest bit rates up to 2 Mbps. The parameters of the 136 HS proposal submitted to ITU are listed in Table 89.2 and the different carrier types of UWC-136 are shown in Fig. 89.12.

Carrier Spacing and Symbol Rate

The motivation for the 200 kHz carrier is twofold. First, the adoption of the same physical layer for 136 HS (Vehicular/Outdoor) and GSM data carriers provides economics of scale and therefore cheaper equipment and faster time to market. Second, the 200 kHz carrier with higher order modulation can provide bit rates of 144 and 384 kbps with reasonable range and capacity fulfilling IMT-2000 requirements for pedestrian and vehicular environments. The 136 HS (Indoor) carrier can provide 2 Mbit/s user data rate with a reasonably strong channel coding.

Modulation

First proposed modulation methods were Quaternary Offset QAM (Q-O-QAM) and Binary Offset QAM (B-O-QAM). Q-O-QAM could provide higher data rates and good spectral efficiency. For each symbol two bits are transmitted and consecutive symbols are shifted by $\pi/2$. An offset modulation was proposed,

because it causes smaller amplitude variations than 16QAM, which can be beneficial when using amplifiers that are not completely linear. The second modulation B-O-QAM has been introduced, which has the same symbol rate of 361.111 ksps, but where only the outer signal points of the Q-O-QAM modulation are used. For each symbol one bit is transmitted and consecutive symbols are shifted by $\pi/2$. A second modulation scheme with the characteristic of being a subset of the first modulation scheme and having the same symbol rate as the first modulation allows seamless switching between the two modulation types between bursts. Both modulation types can be used in the same burst. From a complexity point of view the addition of a modulation, which is subset of the first modulation, adds no new requirements for the transmitter or receiver.

In addition to the originally proposed modulation schemes, Quaternary Offset QAM (Q-O-QAM) and Binary Offset QAM (B-O-QAM), other modulation schemes, CPM (Continuous Phase Modulation) and 8-PSK, have been evaluated in order to select the modulation best suited for EDGE. The outcome of this evaluation is that 8 PSK was considered to have implementation advantages over Q-O-QAM. Parties working on EDGE are in the process of revising the proposals so that 8 PSK would replace the Q-O-QAM and GMSK can be used as the lower level modulation instead of B-O-QAM. The symbol rate of the 8 PSK will be the same as for GMSK and the detailed bit rates will be specified early in 1999.

Frame Structures

The 136 HS (Vehicular/Outdoor) data frame length is 4.615 ms and one frame consists of eight slots. The burst structure is suitable for transmission in a high delay spread environment. The frame and slot structures of the 136 HS (Indoor) carrier were selected for cell coverage for high bit rates. The HS-136 Indoor supports both FDD and TDD duplex methods. Figure 89.13 illustrates the frame and slot structure. The frame length is 4.615 ms and it can consist of

- 64 1/64 time slots of length 72 μs
- 16 1/16 time slots of length 288 μs

In the TDD mode, the same burst types as defined for the FDD mode are used. The 1/64 slot can be used for every service from low rate speech and data to high rate data services. The 1/16 slot is to be used for medium to high rate data services. Figure 89.13 also illustrates the dynamic allocation of resources between the reverse link and the forward link in the TDD mode.

The physical contents of the time slots are bursts of corresponding length. Three types of traffic bursts are defined. Each burst consists of a training sequence, two data blocks, and a guard period. The bursts differ in the length of the burst (72 μs and 288 μs) and in the length of the training sequence

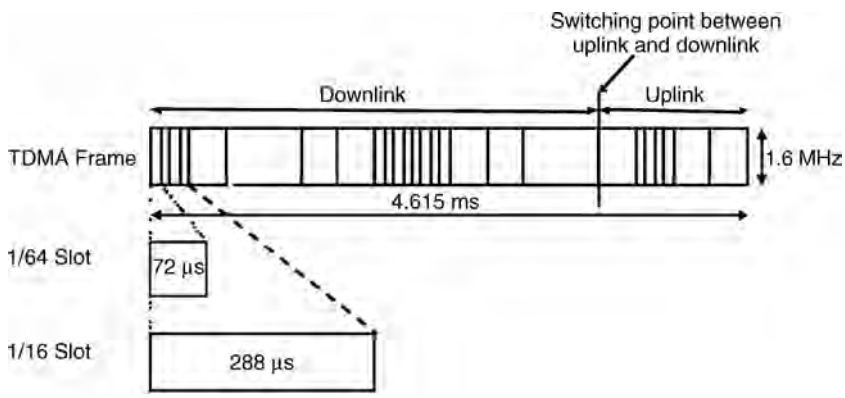


FIGURE 89.13 Wideband TDMA frame and slot structure.

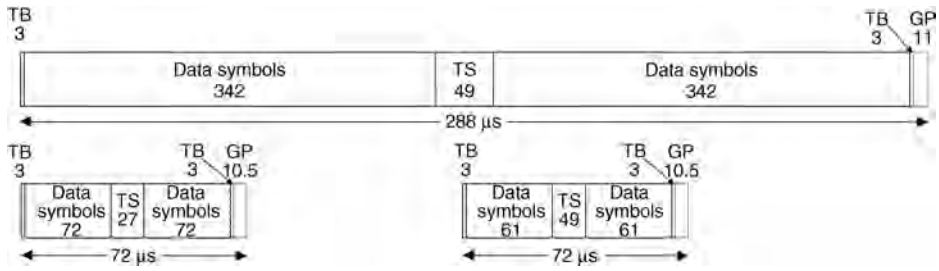


FIGURE 89.14 Burst structure.

(27 symbols and 49 symbols) leading to different numbers of payload symbols and different multipath delay performances (Fig. 89.14). The number of required reference symbols in the training sequence depends on the length of the channel's impulse response, the required signal-to-noise ratio, the expected maximum Doppler frequency shift, and the number of modulation levels. The number of reference symbols should be matched to the channel characteristics, remain practically stable within the correlation window, and have good correlation properties. All 136 based schemes can use interference cancellation as a means to improve performance [8]. For 136 HS (Indoor), the longer sequence can handle about 7 μs of time dispersion and the shorter one 2.7 μs. It should be noted that if the time dispersion is larger, the drop in performance is slow and depends on the power delay profile.

Multirate Scheme

The UWC-136 multirate scheme is based on a variable slot, code, and modulation structure. Data rates up to 43.2 kbps can be offered using the 136+ 30 kHz carrier and multislot transmission. Depending on the user requirements and channel conditions a suitable combination of modulation, coding, and number of data slots is selected. 136 HS can offer packet switched, and both transparent and nontransparent circuit switched data services. Asymmetrical data rates are provided by allocating a different number of time slots in the reverse and forward links. For packet switched services the RLC/MAC protocol provides fast medium access via a reservation based medium access scheme, supplemented by selective ARQ for efficient retransmission.

Similar to 136 HS (Outdoor/Vehicular), the 136 HS (Indoor) uses two modulation schemes and different coding schemes to provide variable data rates. In addition, two different slot sizes can be used. For delay tolerant packet data services, error control is based on a Type II hybrid ARQ (automatic repeat request) scheme [5]. The basic idea is to first send all data blocks using a simple error control coding scheme. If decoding at the receiver fails, a retransmission is requested using a stronger code. After the second retransmission, diversity combining can be performed between the first and second transmissions prior to hard decisions. This kind of ARQ procedure can be used due to the ability of the RLC/MAC protocol to allocate resources fast and to send transmission requests reliably in the feedback channel [5].

Radio Resource Management

The radio resource management schemes of UWC-136 include link adaptation, frequency hopping, power control, and dynamic channel allocation. Link adaptation offers a mechanism for choosing the best modulation and coding alternative according to channel and interference conditions. Frequency hopping averages interference and improves link performance against fast fading. For 136 HS (Indoor) fast power control (frame-by-frame) could be used to improve the performance in cases where frequency hopping cannot be applied, for example, when only one carrier is available. Dynamic channel allocation can be used for channel assignments. However, when deployment with minimum spectrum is desired, reuse 1/3 and fractional loading with fixed channel allocation is used.

89.7 Time Division Duplex (TDD)

The main discussion about the IMT-2000 air interface has been concerned with technologies for FDD. However, there are several reasons why TDD would be desirable. First, there will likely be dedicated frequency bands for TDD within the identified UMTS frequency bands. Furthermore, FDD requires exclusive paired bands and spectrum is, therefore, hard to find. With a proper design including powerful FEC, TDD can be used even in outdoor cells. The second reason for using TDD is flexibility in radio resource allocation, i.e., bandwidth can be allocated by changing the number of time slots for the reverse link and forward link. However, the asymmetric allocation of radio resources leads to two interference scenarios that will impact the overall spectrum efficiency of a TDD scheme:

- asymmetric usage of TDD slots will impact the radio resource in neighboring cells, and
- asymmetric usage of TDD slots will lead to blocking of slots in adjacent carriers within their own cells.

Figure 89.15 depicts the first scenario. MS2 is transmitting at full power at the cell border. Since MS1 has a different asymmetric slot allocation than MS2, its forward link slots received at the sensitivity limit are interfered by MS1, which causes blocking. On the other hand, since the BS1 can have much higher EIRP (effective isotropically radiated power) than MS2, it will interfere BS2's ability to receive MS2. Hence, the radio resource algorithm needs to avoid this situation.

In the second scenario, two mobiles would be connected into the same cell but using different frequencies. The base station receives MS1 on the frequency $f1$ using the same time slot it uses on the frequency $f2$ to transmit into MS2. As shown in Table 89.3, the transmission will block the reception due to the irreducible noise floor of the transmitter regardless of the frequency separation between $f1$ and $f2$.

Both TDMA- and CDMA-based schemes have been proposed for TDD. Most of the TDD aspects are common to TDMA- and CDMA-based air interfaces. However, in CDMA-based TDD systems the slot duration on the forward and reverse links must be equal to enable the use of soft handoff and prevent the interference situation described in the first scenario. Because TDMA systems do not have soft handoff on a common frequency, slot imbalances from one BS to the next are easier to accommodate. Thus, TDMA-based solutions have higher flexibility. The frame structure for the wide band TDMA for the TDD system was briefly discussed in the previous section. WCDMA has been proposed for TDD in Japan and Europe. The frame structure is the same as for the FDD component, i.e., a 10 ms frame split into 16 slots of 0.625 ms each. Each slot can be used either for reverse link or forward link. For cdma2000, the TDD frame structure is based on a 20 ms frame split into 16 slots of 1.25 ms each.

TABLE 89.3 Adjacent Channel Interference Calculation

BTS transmission power for MS2 in forward link 1W	30 dBm
Received power for MS1	-100 dBm
Adjacent channel attenuation due to irreducible noise floor	50 to 70 dB
Signal to adjacent channel interference ratio	-60 to -80 dB

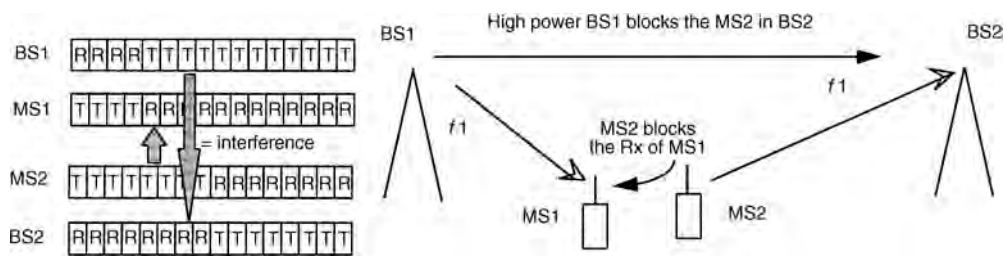


FIGURE 89.15 TDD interference scenario.

89.8 Conclusions

Third generation cellular systems are a mechanism for evolving the telecommunications business based primarily on voice telephony to mobile wireless datacomm. In light of events in TIA, ETSI, and ARIB, cdma2000, WCDMA, and EDGE will be important technologies used to achieve the datacomm goal. Standardization related to radio access technologies discussed in this chapter were under way at the time of writing and will offer the European, U.S., and Japanese markets both CDMA and TDMA third-generation options. In comparing CDMA evolution, the European, U.S., and Japanese based systems have some similarities, but differ in the chip rate and channel structure. In the best circumstances, some harmonization will occur between cdma2000 and WCDMA making deployment of hardware capable of supporting both systems easier. In TDMA, the third generation paths of GSM and TIA/EIA-136 are through a common solution. This alignment will offer TDMA systems an advantage in possible global roaming for data services. In spite of the regional standards differences, third generation will be the mechanism for achieving wireless multimedia enabling services beyond the comprehension of second generation systems.

Acknowledgments

The authors would like to thank Harri Holma, Pertti Lukander, and Antti Toskala from Nokia Telecommunications, George Fry, Kari Kalliojarvi, Riku Pirhonen, Rauno Ruismaki, and Zhigang Rong from Nokia Research Center, Kari Pehkonen from Nokia Mobile Phones, and Kari Pulli from the University of Stanford for helpful comments. In addition, contributions related to spectrum and modulation aspects from Harri Lilja from Nokia Mobile Phones are acknowledged.

References

1. ITU-R M.1225, Guidelines for Evaluation of Radio Transmission Technologies for IMT-2000, 1998.
2. Special Issue on IMT-2000: Standards Efforts of the ITU, *IEEE Pers. Commun.,* 4(4), Aug. 1997.
3. Verdu, S., Minimum probability of error for asynchronous gaussian multiple access, *IEEE Trans. on IT.,* IT-32(1), 85–96, Jan. 1986.
4. Monk, A.M. et al., A noise-whitening approach to multiple access noise rejection—Pt I: Theory and background, *IEEE J. Select. Areas Commun.,* 12(5), 817–827, Jun. 1997.
5. Nikula, E., Toskala, A., Dahlman, E., Girard, L., and Klein, A., FRAMES multiple access for UMTS and IMT-2000, *IEEE Pers. Commun.,* Apr. 1998.
6. Tiedemann, E.G., Jr., Jou, Y.-C., and Odenwalder, J.P., The evolution of IS-95 to a third generation system and to the IMT-2000 era, *Proc. ACTS Summit,* Aalborg, Denmark, 924–929, Oct. 1997.
7. Chang, C.R., Van, J.Z., and Yee, M.F., PN offset planning strategies for nonuniform CDMA networks, *Proc. VTC'97,* 3, Phoenix, Arizona, 1543–1547, May 4–7, 1997.
8. Ranta, P., Lappetelainen, A., and Honkasalo, Z.-C., Interference cancellation by joint detection in random frequency hopping TDMA networks, *Proc. ICUPC96,* 1, Cambridge, MA, 428–432, Sept./Oct. 1996.

90

Multiple-Input Multiple-Output (MIMO) Wireless Systems[1]

Helmut Bölcskei
ETH Zurich

Arogyaswami J. Paulraj
Stanford University

90.1 Introduction

Wireless transmission is impaired by signal **fading** and interference. The increasing requirements on data rate and quality of service for wireless communications systems call for new techniques to increase spectrum efficiency and to improve link reliability. The use of multiple antennas at both ends of a wireless link promises significant improvements in terms of spectral efficiency and link reliability. This technology is known as multiple-input multiple-output (MIMO) wireless (see Fig. 90.1). MIMO systems offer **diversity gain** and **multiplexing gain**.

Diversity Gain

Diversity is used in wireless systems to combat small scale fading caused by multipath effects. The basic principle of diversity is that if several replicas of the information signal are received through independently fading links (branches), then with high probability at least one or more of these links will not be in a fade at any given instant. Clearly, this probability will increase if the number of diversity branches increases. In a noise limited link, without adequate diversity, the transmit power will need to be higher or the range

[1]This work was supported in part by FWF-grant J1868-TEC.

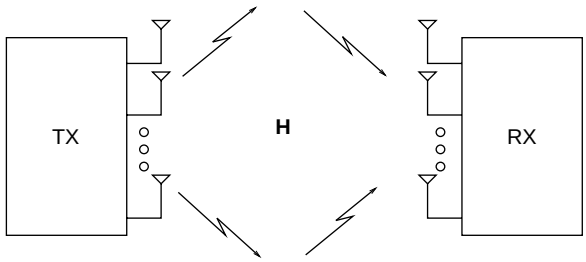


FIGURE 90.1 Schematic representation of a MIMO wireless system.

smaller to protect the link against fading. Likewise in co-channel interference limited links, without adequate diversity, the reuse factor will have to be increased to make co-channel interference well below the signal average level. Diversity processing that reduces fading is a powerful tool to increase capacity and coverage in radio networks. The three main forms of diversity traditionally exploited in wireless communications systems are *temporal diversity*, *frequency diversity*, and *spatial (or antenna) diversity*.

Temporal diversity is applicable in a channel that has time selective fading. The information is transmitted with spreading over a time span that is larger than the coherence time of the channel. The coherence time is the minimum time separation between independent channel fades. Time diversity is usually exploited via **interleaving**, **forward error correction coding** (**FEC**), and **automatic repeat request** (**ARQ**). One drawback of time diversity is the inherent delay incurred in time spreading.

Frequency diversity is effective when the fading is frequency-selective. This diversity can be exploited by spreading the information over a frequency span larger than the coherence bandwidth of the channel. The coherence bandwidth is the minimum frequency separation between independent channel fades and is inversely dependent on the delay spread in the channel. Frequency diversity can be exploited through spread spectrum techniques or through interleaving and FEC in conjunction with multicarrier modulation [1].

Spatial diversity. In this chapter, we are mostly concerned with spatial (or antenna) diversity. In space diversity we receive or transmit information signals from antennas that are spaced by more than the coherence distance apart. The coherence distance is the minimum spatial separation of antennas for independent fading and depends on the angle spread of the multipaths arriving at or departing from an antenna array. For example, if the multipath signals arrive from all directions in the azimuth, antenna spacing on the order of 0.4λ-0.6λ is adequate [2] for independent fading. On the other hand, if the multipath angle spread is smaller, the coherence distance is larger. Empirical measurements show a strong coupling between antenna height and coherence distance for base station antennas. Higher antenna heights imply larger coherence distances. At the terminal end, which is usually low and buried in scatterers, a 0.4λ-0.6λ separation will be adequate. Receive diversity, i.e., the use of multiple antennas only at the receive side, is a well-studied subject [3]. Driven mostly by mobile wireless applications, where it is difficult to employ multiple antennas in the handset, the use of multiple antennas at the base station in combination with **transmit diversity** has become an active area of research in the past few years [4,5,6,7,8,9]. Transmit diversity in the case where the channel is known at the transmitter involves transmission such that the signals launched from the individual antennas arrive in phase at the receiver antenna. In the case where the channel is not known at the transmitter, transmit diversity requires more sophisticated methods such as **space-time coding** which use coding across antennas (space) and time. Here, the basic idea is to send the information bitstream with different preprocessing (coding, modulation, delay, etc.) from different antennas such that the receiver can combine these signals to obtain diversity.

In a MIMO system both transmit and receive antennas combine to give a large diversity order. Assuming M_T transmit and M_R receive antennas, a maximum of $M_T M_R$ links are available; if all of these links fade independently, we get $M_T M_R$th order diversity. The value of spatial MIMO diversity is demonstrated in Figs. 90.2(a) and (b), where the signal level at the output of a 1-input 1-output (SISO) and

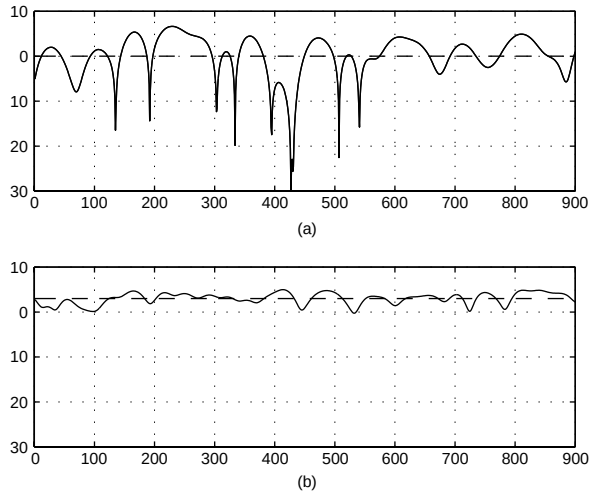


FIGURE 90.2 The value of spatial diversity. Signal level as a function of time for a (a) 1-input 1-output, and a (b) 2-input 2-output system.

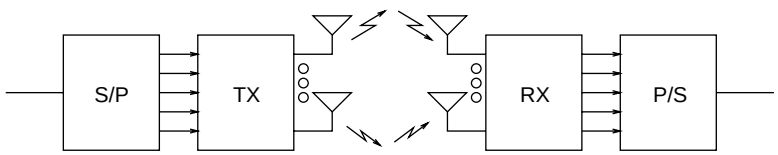


FIGURE 90.3 Schematic of a spatial multiplexing system. S/P and P/S denote serial-to-parallel and parallel-to-serial conversion, respectively.

of a 2-input 2-output system, respectively, is shown, assuming that the channel is known at the receiver and proper diversity combining is employed. Clearly, the use of multiple antennas reduces the signal fluctuations and eliminates deep fades.

Transmit and receive diversity are both similar and different in many ways. While receive diversity needs merely multiple antennas which fade independently, and is independent of coding/modulation schemes, transmit diversity needs special modulation/coding schemes in order to be effective. Also, receive diversity provides **array gain**, whereas transmit diversity does not provide array gain when the channel is unknown in the transmitter (the usual case).

Multiplexing Gain

While spatial diversity gain can be obtained when multiple antennas are present at either the transmit or the receive side, **spatial multiplexing** requires multiple antennas at both ends of the link [10,11]. The idea of spatial multiplexing is that the use of multiple antennas at the transmitter and the receiver in conjunction with rich scattering in the propagation environment opens up multiple data pipes within the same frequency band to yield a linear (in the number of antennas) increase in capacity. This increase in capacity comes at no extra bandwidth or power consumption and is therefore very attractive. The basic principle of spatial multiplexing is illustrated in Fig. 90.3. The symbol stream to be transmitted is broken up into several parallel symbol streams which are then transmitted simultaneously and within the same frequency band from the antennas. Due to multipath propagation, each transmit antenna induces a different spatial signature at the receiver. The receiver exploits these signature differences to

separate the individual data streams. The price to be paid for multiplexing gain is increased hardware cost due to the use of multiple antennas.

We conclude by noting that MIMO systems can be used for transmit–receive diversity, spatial multiplexing, and combinations of diversity and multiplexing.

90.2 The MIMO Fading Channel Model

We focus on a single-user communication model and consider a point-to-point link where the transmitter is equipped with M_T antennas and the receiver employs M_R antennas (see Fig. 90.1). We restrict our discussion to the case where the channel is frequency-flat (narrowband assumption). A commonly used channel model in MIMO wireless communications is the block fading model, where the channel matrix entries are i.i.d. complex Gaussian (Rayleigh fading), constant during a block of symbols, and change in an independent fashion from block to block. We furthermore assume that the channel is unknown at the transmitter and perfectly known (and tracked) at the receiver. Channel knowledge in the receiver can be obtained by sending training data and estimating the channel. We do not discuss the case where the channel is known at the transmitter. Acquiring channel knowledge at the transmitter requires the use of feedback from the receiver or the use of transmit-receive duplexing based channel mapping methods and is therefore often hard to realize in practice.

The input-output relation of the $M_R \times M_T$ matrix channel can be written as

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{90.1}$$

where $\mathbf{r} = [r_0\ r_1 \ldots r_{M_R-1}]^T$ is the $M_R \times 1$ receive signal vector, $\mathbf{H}$ is the $M_R \times M_T$ channel transfer matrix, $\mathbf{s} = [s_0\ s_1 \ldots s_{M_T-1}]^T$ is the $M_T \times 1$ transmit signal vector, and $\mathbf{n} = [n_0\ n_1 \ldots n_{M_R-1}]^T$ is the $M_R \times 1$ noise vector. The elements of the channel transfer matrix $\mathbf{H}$ are i.i.d. circularly symmetric complex Gaussian with zero mean and unit variance. A circularly symmetric complex Gaussian random variable is a random variable $z = (x + jy) \sim \mathcal{CN}(0, \sigma^2)$, in which x and y are i.i.d. real Gaussian distributed with $\mathcal{N}(0, \sigma^2/2)$. Equivalently, each entry of $\mathbf{H}$ has uniformly distributed phase and Rayleigh distributed magnitude. This model is typical of an environment with rich scattering and enough separation between the antennas at the transmitter and the receiver. The noise components n_i ($i = 0, 1, \ldots, M_R - 1$) are assumed to be i.i.d. $\mathcal{CN}(0, 1)$. The average transmitted power across all antennas, which is equal to the average SNR at each receive antenna with this normalization of noise power and channel loss, is limited to be no greater than ρ regardless of M_T. The transmit power is assumed to be equally distributed across the transmit antennas.

90.3 Capacity of MIMO Channels

As discussed earlier, MIMO systems offer significant capacity gains over SISO channels. Before analyzing the capacity of random i.i.d. MIMO fading channels, let us briefly consider deterministic SISO and MIMO channels.

The Deterministic Case

The formula for Shannon capacity for a deterministic SISO channel with input-output relation $r = Hs + n$, where ρ is the average transmit power and $\mathcal{E}\{|n|^2\} = 1$, is given by [12]

$$C = \log_2(1 + \rho|H|^2) \text{ bps/Hz}. \tag{90.2}$$

It is evident that for high SNRs, a 3 dB increase in ρ gives an increase in capacity of one bps/Hz. Assuming that the transmitted signal vector is composed of M_T statistically independent equal power components, each with a circularly symmetric complex Gaussian distribution, the capacity of a deterministic MIMO

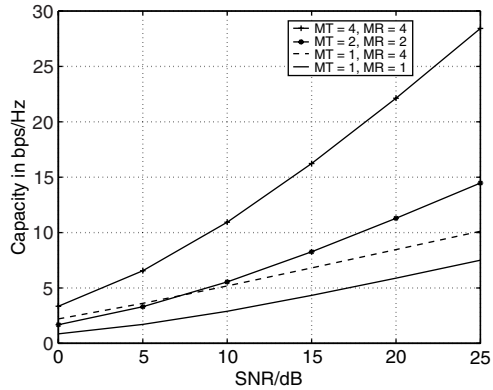


FIGURE 90.4 Ergodic capacity of MIMO fading channels.

channel **H** is given by* [12]

$$C = \log_2\left[\det\left(\mathbf{I}_{M_R} + \frac{\rho}{M_T}\mathbf{H}\mathbf{H}^H\right)\right] \text{ bps/Hz.} \tag{90.3}$$

Let us next specialize Eq. (90.3) to $M_T = M_R$ and $\mathbf{H} = \mathbf{I}_{M_T}$. We get [12]

$$C = M_T \log_2\left(1 + \frac{\rho}{M_T}\right) \to \rho/\ln(2) \quad \text{as} \quad M_T \to \infty. \tag{90.4}$$

Unlike in Eq. (90.2), capacity scales linearly, rather than logarithmically, with increasing SNR. This result hints at a significant advantage of using multiple antennas.

The Random Case

In the following the channel **H** is assumed to be random according to the definitions in Section 90.2. The ergodic capacity of this channel is given by [13]

$$C = \mathcal{E}_H\left\{\log_2\left[\det\left(\mathbf{I}_{M_R} + \frac{\rho}{M_T}\mathbf{H}\mathbf{H}^H\right)\right]\right\} \text{ bps/Hz,} \tag{90.5}$$

where $\mathcal{E}_H$ denotes expectation with respect to the random channel. Note that for fixed M_R as M_T gets large, $\frac{1}{M_T}\mathbf{H}\mathbf{H}^H \to \mathbf{I}_{M_R}$ and, hence, the ergodic capacity in the limit of large M_T equals

$$C = M_R \log_2(1 + \rho) \text{ bps/Hz.}$$

Thus, the ergodic capacity grows linearly in the number of receive antennas, which hints at significant capacity gains of MIMO fading channels. Figure 90.4 shows the ergodic capacity as a function of SNR for various multi-antenna systems in a fading environment. We can see that the use of multiple antennas on one side of the link only leads to smaller capacity gains than the use of multiple antennas on both sides of the link. For example, in the high SNR regime $M_T = M_R = 2$ yields much higher capacity than $M_T = 1$ and $M_R = 4$.

*The superscripts T and H denote transposition and conjugate transposition, respectively.

90.4 Spatial Multiplexing

We recall from Section 90.1 that multiplexing gain is available if multiple antennas are used at the transmit and receive side of the wireless link. In this section, we describe various signal processing techniques which allow realization of this gain.

In a spatial multiplexing system [10,11] the data stream to be transmitted is demultiplexed into M_T lower rate streams which are then simultaneously sent from the M_T transmit antennas after coding and modulation. Note that these streams occupy the same frequency band (i.e., they are co-channel signals). Each receive antenna observes a superposition of the transmitted signals. The receiver then separates them into constituent data streams and remultiplexes them to recover the original data stream (see Fig. 90.3). Clearly, the separation step determines the computational complexity of the receiver. Competing receiver architectures have different performance-complexity tradeoffs. In the following, we discuss various receiver structures.

Zero Forcing Receiver

A simple linear receiver is the zero forcing (ZF) receiver which basically inverts the channel transfer matrix, i.e., assuming that $\mathbf{H}$ is square and invertible. An estimate of the $M_T \times 1$ transmitted data symbol vector $\mathbf{s}$ is obtained as

$$\hat{\mathbf{s}} = \mathbf{H}^{-1}\mathbf{r} \tag{90.6}$$

The ZF receiver hence perfectly separates the co-channel signals $s_i (i = 0, 1, \ldots, M_T - 1)$. For ill-conditioned $\mathbf{H}$, the ZF receiver incurs significant noise enhancement. In practice, if $\mathbf{H}$ is rank-deficient or non square the receiver left-multiplies $\mathbf{r}$ by the pseudo-inverse of $\mathbf{H}$ to recover rank($\mathbf{H}$) symbol streams. If FEC is employed, spatial and temporal redundancy can be exploited to recover lost data.

Minimum Mean-Square Error Receiver

The ZF receiver yields perfect separation of the co-channel signals at the cost of noise enhancement. An alternative linear receiver is the minimum mean-square error (MMSE) receiver, which minimizes the overall error due to noise and mutual interference between the co-channel signals. In this case, an estimate of $\mathbf{s}$ is obtained according to

$$\hat{\mathbf{s}} = \frac{\rho}{M_T}\mathbf{H}^H\left(\sigma_n^2\mathbf{I}_{M_R} + \frac{\rho}{M_T}\mathbf{H}\mathbf{H}^H\right)^{-1}\mathbf{r}, \tag{90.7}$$

where it was assumed that $\mathcal{E}\{\mathbf{s}\mathbf{s}^H\} = \frac{\rho}{M_T}\mathbf{I}_{M_T}$ and $\mathcal{E}\{\mathbf{n}\mathbf{n}^H\} = \sigma_n^2\mathbf{I}_{M_R}$. The MMSE receiver is less sensitive to noise at the cost of reduced signal separation quality. In other words, the co-channel signals are in general not perfectly separated. In the high SNR case ($\sigma_n^2 \approx 0$), the MMSE receiver converges to the ZF receiver.

V-BLAST Receiver

An attractive alternative to ZF and MMSE receivers which, in general, yields improved performance at the cost of increased computational complexity is the so-called V-BLAST algorithm [14,15]. In V-BLAST rather than jointly decoding all the transmit signals, we first decode the "strongest" signal, then subtract this strongest signal from the received signal, proceed to decode the strongest signal of the remaining transmit signals, and so on. The optimum detection order in such a *nulling and cancellation* strategy is

from the strongest to the weakest signal [14]. Assuming that the channel **H** is known, the main steps of the V-BLAST algorithm can be summarized as follows:

- *Nulling:* An estimate of the strongest transmit signal is obtained by nulling out all the weaker transmit signals (e.g., using the zero forcing criterion).
- *Slicing:* The estimated signal is detected to obtain the data bits.
- *Cancellation:* These data bits are remodulated and the channel is applied to estimate its vector signal contribution at the receiver. The resulting vector is then subtracted from the received signal vector and the algorithm returns to the nulling step until all transmit signals are decoded.

For a more in-depth treatment of the V-BLAST algorithm the interested reader is referred to [14,15].

Maximum Likelihood Receiver

The receiver which yields the best performance in terms of error rate is the maximum likelihood (ML) receiver. However, this receiver also has the highest computational complexity which moreover exhibits exponential growth in the number of transmit antennas. Assuming that channel state information has been acquired, the ML receiver computes the estimate $\hat{\mathbf{s}}$ according to

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} \|\mathbf{r} - \mathbf{H}\mathbf{s}\|^2,$$

where the minimization is performed over all possible codeword vectors **s**. Note that if the size of the scalar constellation used is Q (e.g., $Q = 4$ for QPSK), the receiver has to perform an enumeration over a set of size Q^{M_T}. For higher-order modulation such as 64-QAM, this complexity can become prohibitive even for a small number of transmit antennas. For example, for 64-QAM and $M_T = 3$, the receiver has to enumerate over 262,144 different vectors on the symbol rate.

90.5 Transmit Diversity

Transmit diversity is a technique which realizes spatial diversity gain in systems with multiple transmit antennas without requiring channel knowledge in the transmitter [4,5,6,7,8,16]. In this section, we shall first discuss indirect transmit diversity schemes and then proceed with a discussion of direct transmit diversity schemes such as space-time block and space-time Trellis codes.

Indirect Transmit Diversity

In indirect transmit diversity schemes we convert spatial diversity into time or frequency diversity, which can then readily be exploited by the receiver (often using FEC and interleaving). In the following, we shall discuss two techniques, namely *delay diversity*, which converts spatial diversity into frequency diversity, and *intentional frequency offset diversity*, which converts spatial diversity into time diversity.

Delay diversity. In order to simplify the presentation let us assume that $M_T = 2$ and $M_R = 1$. The principle underlying delay diversity is to convert the available spatial diversity into frequency diversity by transmitting the data bearing signal from the first antenna and a delayed replica thereof from the second antenna (see Fig. 90.5) [4,5]. Assuming that the delay is one symbol interval, the effective SISO channel seen by the receiver is given by

$$H_e(e^{j2\pi\theta}) = h_0 + h_1 e^{-j2\pi\theta}, \tag{90.8}$$

where h_0 and h_1 denote the channel gains between transmit antennas 1 and 2, and the receive antenna, respectively. We assume that h_0 and h_1 are i.i.d. $\mathcal{CN}(0, 1)$. Therefore, the channel transfer function $H_e(e^{j2\pi\theta})$

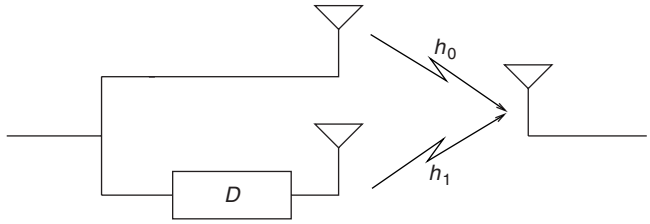


FIGURE 90.5 Schematic representation of delay diversity. A delayed version of the data bearing signal is transmitted from the second antenna.

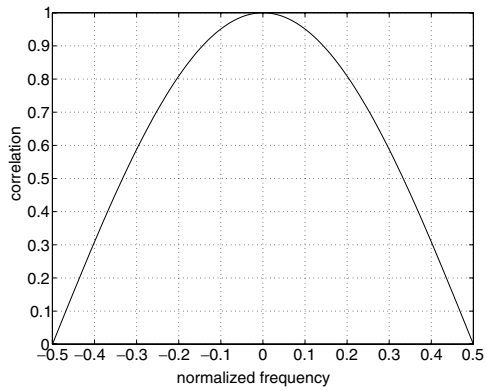


FIGURE 90.6 The correlation function $|R(e^{j2\pi\nu})|$ for a delay of one symbol interval.

is random as well. In order to see how the spatial diversity is converted into frequency diversity we compute the frequency correlation function $R(e^{j2\pi\nu}) = \frac{1}{2}\,\mathcal{E}\{H_e(e^{j2\pi\theta})\,H_e^*(e^{j2\pi(\theta-\nu)})\}$. Using Eq. (90.8) we get

$$R(e^{j2\pi\nu}) \;=\; \frac{1}{2}(1 + e^{-j2\pi\nu}) \tag{90.9}$$

The function $|R(e^{j2\pi\nu})|$ in Eq. (90.9) is depicted in Fig. 90.6. We can see that frequencies separated by $\nu = \frac{1}{2}$ are fully decorrelated. Such a channel looks exactly like a 2-paths channel with independent path fades and same average path energy. Therefore a Viterbi (ML) sequence detector will capture the diversity in the system. The extension of delay diversity to more than two transmit antennas is described in [5].

Intentional frequency offset diversity. An alternative indirect transmit diversity scheme which converts spatial diversity into temporal diversity was first described in [17]. Let us again assume that $M_T = 2$ and $M_R = 1$. After coding and modulation, we transmit the data bearing signal from the first antenna and a frequency shifted (phase rotated) version thereof from the second antenna (see Fig. 90.7). The effective SISO channel seen by the receiver is given by

$$h_e[n] = h_0 + h_1 e^{j2\pi n\theta_1}, \quad n \in \mathbf{Z}, \tag{90.10}$$

where θ_1 with $|\theta_1| < 1/2$ is the intentional frequency offset introduced at the second antenna. It is again assumed that h_0 and h_1 are i.i.d. $\mathcal{CN}(0, 1)$. In order to see how spatial diversity is converted into temporal diversity, we compute the temporal correlation function of the stochastic channel $h_e[n]$ given

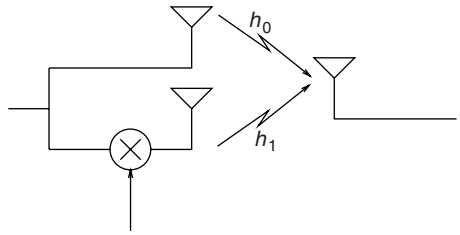


FIGURE 90.7 Conversion of spatial diversity into time diversity through intentional frequency offset. A modulated version of the data bearing signal is transmitted from the second antenna.

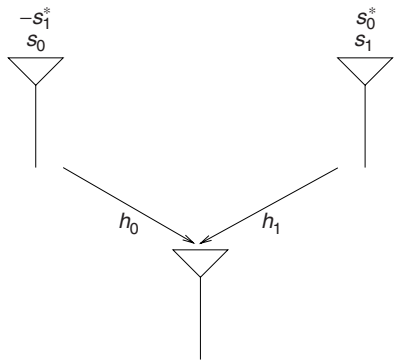


FIGURE 90.8 Schematic of the Alamouti scheme.

by $R[k] = \frac{1}{2}\mathcal{E}\{h_e[n]h_e^*[n-k]\}$. Using Eq. (90.10), we get

$$R[k] = \frac{1}{2}(1 + e^{j2\pi k\theta_1})$$

If $|R[k]|$ is small data symbols spaced k symbol intervals apart undergo close to independent fading. The resulting temporal diversity can be exploited by using FEC in combination with time interleaving just as we may do in naturally time fading channels.

Direct Transmit Diversity

Space-time block coding. Space-time block coding has attracted much attention for practical applications [8,18]. A simple space-time block code known as Alamouti scheme [18] performs very similar to maximum-ratio combining (MRC), a technique which realizes spatial diversity gain by employing multiple receive or transmit antennas (but needs channel knowledge in the transmitter for the latter). We briefly review receive MRC for $M_T = 1$ and $M_R = 2$. The receive signals are given by $r_0 = h_0 s + n_0$ and $r_1 = h_1 s + n_1$, respectively, where h_0 and h_1 denote the $\mathcal{CN}(0, 1)$ i.i.d. channel gains between the transmit antenna and the two receive antennas and n_0 and n_1 are $\mathcal{CN}(0, 1)$ i.i.d. noise samples. The receiver estimates the transmitted data symbols by forming the decision variable

$$y = h_0^* r_0 + h_1^* r_1 = (|h_0|^2 + |h_1|^2)s + h_0^* n_0 + h_1^* n_1.$$

Clearly, if either h_0 or h_1 is not faded, we have a good channel. Thus, we get second order diversity. We shall next describe Alamouti's scheme where $M_T = 2$ and $M_R = 1$. Note that now the two antennas are used at the transmitter rather than at the receiver. Figure 90.8 shows a schematic of Alamouti's scheme,

which consists of three steps:

- *Encoding and transmitting*. At a given symbol period, two signals are simultaneously transmitted from the two antennas. In the first time instant, the signal transmitted from antenna 1 is s_0 and the signal transmitted from antenna 2 is s_1. In the next time instant, $-s_1^*$ is transmitted from antenna 1 and s_0^* is transmitted from antenna 2. The received signals r_0 and r_1 are, hence, given by

$$\begin{aligned} r_0 &= h_0 s_0 + h_1 s_1 + n_0 \\ r_1 &= -h_0 s_1^* + h_1 s_0^* + n_1, \end{aligned} \tag{90.11}$$

  where h_0 and h_1 denote the $\mathcal{CN}(0,1)$ i.i.d. channel gains between transmit antenna 1 and the receive antenna and transmit antenna 2 and the receive antenna, respectively. Furthermore, the noise samples n_0 and n_1 are $\mathcal{CN}(0,1)$ i.i.d. The equations in Eq. (90.11) can be rewritten in vector-matrix form as

$$\mathbf{r} = \mathbf{H}_a \mathbf{s} + \mathbf{n},$$

  where $\mathbf{r} = [r_0 \quad r_1^*]^T$ is the received signal vector (note that r_1 was conjugated),

$$\mathbf{H}_a = \begin{bmatrix} h_0 & h_1 \\ h_1^* & -h_0^* \end{bmatrix} \tag{90.12}$$

  is the equivalent channel matrix, $\mathbf{s} = [s_0 \quad s_1]^T$, and $\mathbf{n} = [n_0 \quad n_1^*]$ is the noise vector. Note that the columns of the channel matrix $\mathbf{H}_a$ are orthogonal irrespective of the specific values of h_0 and h_1.
- *The combining step*. In the receiver, assuming that perfect channel knowledge has been acquired, the vector $\mathbf{r}$ is left-multiplied by $\mathbf{H}_a^H$ which results in

$$\hat{\mathbf{s}} = \mathbf{H}_a^H \mathbf{r} = \begin{bmatrix} |h_0|^2 + |h_1|^2 & 0 \\ 0 & |h_0|^2 + |h_1|^2 \end{bmatrix} \mathbf{s} + \underbrace{\mathbf{H}_a^H \mathbf{n}}_{\tilde{\mathbf{n}}}.$$

  The estimates of the symbols s_0 and s_1 are now given by

$$\begin{aligned} \hat{s}_0 &= (|h_0|^2 + |h_1|^2) s_0 + \tilde{n}_0 \\ \hat{s}_1 &= (|h_0|^2 + |h_1|^2) s_1 + \tilde{n}_1, \end{aligned}$$

  where $\tilde{n}_0 = h_0^* n_0 + h_1 n_1^*$ and $\tilde{n}_1 = h_1^* n_0 - h_0 n_1^*$.
- *ML detection*. The symbols $\hat{s}_0$ and $\hat{s}_1$ are independently sent to an ML detector.

We can see that the Alamouti scheme yields the same diversity order as MRC. Note, however, that if the overall transmit power in the Alamouti scheme is kept the same as in MRC, the Alamouti scheme incurs a 3 dB SNR loss compared to MRC [18]. This 3 dB difference is due to the fact that the Alamouti scheme does not offer array gain since it does not know the channel in the transmitter. If more than one receive antenna is employed, the Alamouti scheme realizes $2M_R$-fold diversity, i.e., full spatial diversity gain. Figure 90.9 shows the bit error rate obtained for the Alamouti scheme in comparison to the SISO fading channel using BPSK modulation.

The Alamouti scheme is a special case of so-called space-time block codes [8]. Space-time block codes achieve full diversity gain and drastically simplify the receiver structure since the vector detection problem decouples into simpler scalar detection problems.

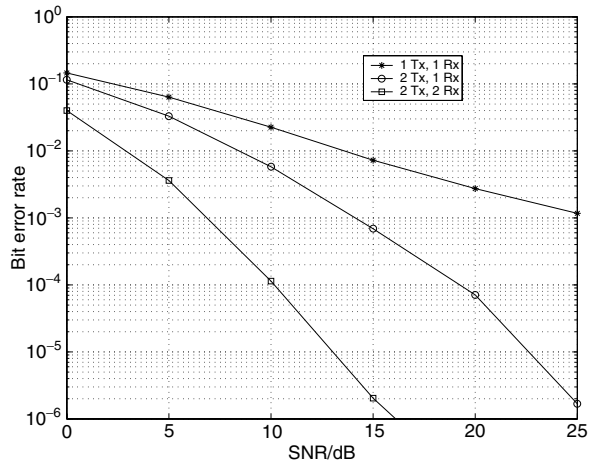


FIGURE 90.9 Bit error rate for Alamouti scheme in comparison to the SISO fading channel.

Space-time trellis codes. Space-time Trellis codes are an extension of Trellis codes [19] to the case of multiple transmit and receive antennas. In space-time Trellis coding, the bit stream to be transmitted is encoded by the space-time encoder into blocks of size $M_T \times T$, where T is the size of the burst over which the channel is assumed to be constant. One data burst therefore consists of T vectors $\mathbf{c}_k (k = 0, 1, \ldots, T - 1)$ of size $M_T \times 1$ with the data symbols taken from a finite complex alphabet chosen such that the average energy of the constellation elements is 1. The kth receive symbol vector is given by $\mathbf{r}_k = \sqrt{E_s}\,\mathbf{H}\mathbf{c}_k + \mathbf{n}_k (k = 0, 1, \ldots, T - 1)$ where $\mathbf{n}_k$ is complex-valued Gaussian noise satisfying $\mathcal{E}\{\mathbf{n}_k \mathbf{n}_l^H\} = \mathbf{I}_{M_R}\,\delta[k - l]$. Assuming perfect channel state information, the ML decoder computes the vector sequence $\hat{\mathbf{c}}_k\ (k = 0, 1, \ldots, T - 1)$ according to

$$\hat{\mathbf{c}}_k = \arg \min_{\mathbf{C}} \sum_{k=0}^{T-1} \left\| \mathbf{r}_k - \sqrt{E_s}\mathbf{H}\mathbf{c}_k \right\|^2,$$

where $\mathbf{C} = [\mathbf{c}_0\ \mathbf{c}_1 \ldots \mathbf{c}_{T-1}]$ and the minimization is over all possible codeword matrices $\mathbf{C}$.

Let us next briefly review the design criteria for space-time Trellis codes assuming that the receiver has perfect channel state information. We consider the pairwise error probability. Let $\mathbf{C} = [\mathbf{c}_0\ \mathbf{c}_1 \ldots \mathbf{c}_{T-1}]$ and $\mathbf{E} = [\mathbf{e}_0\ \mathbf{e}_1 \ldots \mathbf{e}_{T-1}]$ be two different space-time codewords of size $M_T \times T$ and assume that $\mathbf{C}$ was transmitted. In the high SNR case, the average probability (averaged over all channel realizations) that the receiver decides erroneously in favor of the signal $\mathbf{E}$ is upper bounded by [7]

$$P(\mathbf{C} \to \mathbf{E}) \le \left(\frac{E_s}{4}\right)^{-r(\mathbf{B}_{c,e})M_R} \prod_{i=0}^{r(\mathbf{B}_{c,e})-1} \lambda_i(\mathbf{B}_{c,e})^{-M_R} \tag{90.13}$$

where*

$$\mathbf{B}_{c,e} = (\mathbf{C} - \mathbf{E})^T (\mathbf{C} - \mathbf{E})^*,$$

and $r(\mathbf{B}_{c,e})$ denotes the rank of the matrix $\mathbf{B}_{c,e}$ and $\lambda_i(\mathbf{B}_{c,e})$ denotes the nonzero eigenvalues of $\mathbf{B}_{c,e}$.

The design criteria for space-time Trellis codes can now be summarized as follows [7]:

- *The rank criterion:* In order to achieve the maximum diversity $M_T M_R$, the matrix $\mathbf{B}_{c,e}$ has to be full rank for every pair of distinct codewords $\mathbf{C}$ and $\mathbf{E}$.

*The superscript $*$ stands for elementwise conjugation.

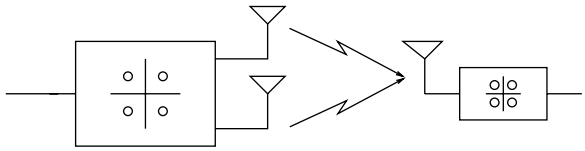


FIGURE 90.10 Schematic of space-time Trellis coding.

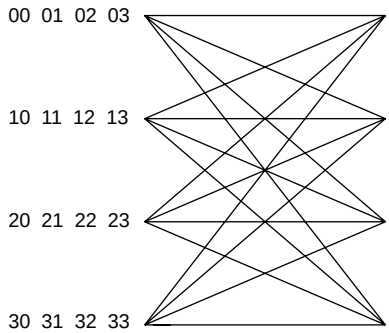


FIGURE 90.11 Trellis diagram for 4-PSK, 4 state Trellis code for $M_T = 2$ with rate 2 bps/Hz.

- *The determinant criterion:* If a diversity advantage of $M_T M_R$ is the design target, the minimum of the determinant of $\mathbf{B}_{c,e}$ taken over all pairs of distinct codewords **C** and **E** must be maximized.

The design criteria for arbitrary diversity order can be found in [7]. Space-time Trellis codes generally offer better performance than space-time block codes at the cost of increased decoding complexity [7]. Figure 90.10 shows a schematic of a space-time Trellis coding system, and Fig. 90.11 shows the Trellis diagram for a simple 4-PSK, 4-state space-time Trellis code for $M_T = 2$ with rate 2 bps/Hz. The Trellis diagram is to be read as follows. The symbols to the left of the nodes correspond to the encoder output for the two transmit antennas. For example, in state 0, if the incoming two bits are 10, the encoder outputs a zero on antenna 1 and a 2 on antenna 2 and changes to state 2. The encoder outputs 0, 1, 2, 3 are mapped to the symbols 1, j, -1, -j, respectively.

90.6 Summary and Conclusion

The use of multiple antennas at both ends of a wireless radio link provides significant gains in terms of spectral efficiency and link reliability. Spatial multiplexing is a technique which requires multiple antennas at both sides of the link and is capable of increasing the spectral efficiency. Rich scattering in the propagation environment is, however, needed in order to obtain multiplexing gain. MIMO diversity techniques realize spatial diversity gain from the transmit and receive antennas. Space-time block codes can realize full diversity gain and decouple the vector-ML decoding problem into scalar problems, which dramatically reduces receiver complexity. Space-time Trellis codes yield better performance than space-time block codes at the cost of increased receiver complexity.

The area of MIMO communication theory is new and full of challenges. Some promising MIMO research areas are MIMO in combination with OFDM and CDMA, new coding, modulation and receivers, combinations of space-time coding and spatial multiplexing, MIMO technology for cellular communications, and adaptive modulation and link adaptation in the context of MIMO.

Defining Terms

Array gain: Improvement in SNR obtained by coherently combining the signals on multiple transmit or multiple receive antennas.

Automatic request for repeat: An error control mechanism in which received packets that cannot be corrected are retransmitted.

Diversity gain: Improvement in link reliability obtained by transmitting the same data on independently fading branches.

Fading: Fluctuation in the signal level due to shadowing and multipath effects.

Forward error correction (FEC): A technique that inserts redundant bits during transmission to help detect and correct bit errors during reception.

Interleaving: A form of data scrambling that spreads bursts of bit errors evenly over the received data allowing efficient forward error correction.

Multiplexing gain: Capacity gain at no additional power or bandwidth consumption obtained through the use of multiple antennas at both sides of a wireless radio link.

Space-time coding: Coding technique that realizes spatial diversity gain without knowing the channel in the transmitter by spreading information across antennas (space) and time.

Spatial multiplexing: Technique to realize multiplexing gain.

Transmit diversity: Simple technique to realize spatial diversity gain without knowing the channel in the transmitter by sending modified versions of the data bearing signal from multiple transmit antennas.

References

1. B. LeFloch, M. Alard, and C. Berrou, Coded orthogonal frequency division multiplex, *Proc. IEEE,* 83, 982, 1995.
2. W. C. Y. Lee, *Mobile Communications Engineering,* McGraw-Hill, New York, 1982.
3. W. C. Jakes, *Microwave Mobile Communications,* Wiley, New York, 1974.
4. A. Wittneben, Base station modulation diversity for digital SIMULCAST, *Proc. IEEE VTC,* May 1993, 505–511.
5. N. Seshadri and J. Winters, Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity, *Int. J. Wireless Inf. Networks,* 1(1), 49–60 1994.
6. J. Guey, M. Fitz, M. Bell, and W. Kuo, Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels, *Proc. IEEE VTC,* 1996, 136–140.
7. V. Tarokh, N. Seshadri, and A. R. Calderbank, Space-time codes for high data rate wireless communication: Performance criterion and code construction, *IEEE Trans. Inf. Theory,* 44, 744–765, 1998.
8. V. Tarokh, H. Jafarkhani, and A. R. Calderbank, Space-time block codes from orthogonal designs, *IEEE Trans. Inf. Theory,* 45, 1456–1467, 1999.
9. T. L. Marzetta and B. M. Hochwald, Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading, *IEEE Trans. Inf. Theory,* 45, 139–157, 1999.
10. A. J. Paulraj and T. Kailath, Increasing capacity in wireless broadcast systems using distributed transmission/directional reception, U.S. Patent no. 5,345,599, 1994.
11. G. J. Foschini, Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas, *Bell Labs Tech. J.,* 41–59, Autumn 1996.
12. G. J. Foschini and M. J. Gans, On limits of wireless communications in a fading environment when using multiple antennas, *Wireless Personal Communications,* 6, 311–335, 1998.
13. I. E. Telatar, Capacity of multi-antenna Gaussian channels, *European Trans. Telecommunications,* 10(6), 585–595, 1999.

14. G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, Detection algorithm and initial laboratory results using the V-BLAST space-time communication architecture, *Electron. Lett.,* 35(1), 14–15, 1999.
15. G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, Simplified processing for high spectral efficiency wireless communication employing multi-antenna arrays, *IEEE J. Sel. Areas Comm.,* 17(11), 1841–1852, 1999.
16. B. M. Hochwald and T. L. Marzetta, Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading, *IEEE Trans. Inf. Theory,* 46(2), 543–564, 2000.
17. A. Hiroike, F. Adachi, and N. Nakajima, Combined effects of phase sweeping transmitter diversity and channel coding, *IEEE Trans. Veh. Technol.,* 41, 170–176, 1992.
18. S. M. Alamouti, A simple transmit diversity technique for wireless communications, *IEEE J. Sel. Areas Comm.,* 16, 1451–1458, 1998.
19. E. Biglieri, D. Divsalar, P. J. McLane, and M. K. Simon, *Introduction to Trellis-Coded Modulation with Applications.* Macmillan, New York, 1991.

91

Near-Instantaneously Adaptive Wireless Transceivers of the Near Future*

Lie-Liang Yang
University of Southampton

Lajos Hanzo
University of Southampton

91.1 Introduction

There is a range of activities in various parts of the globe concerning the standardization, research, and development of the third-generation (3G) mobile systems known as the Universal Mobile Telecommunications System (UMTS) in Europe, which was termed as the IMT-2000 system by the International Telecommunications Union (ITU) [1,2]. This is mainly due to the explosive expansion of the Internet and the continued dramatic increase in demand for all types of advanced wireless multimedia services, including video telephony as well as the more conventional voice and data services. However, advanced high-rate services such as high-resolution interactive video and "telepresence" services require data rates in excess of 2Mb/s, which are unlikely to be supported by the 3G systems [3–7]. These challenges remain to be solved by future mobile broadband systems (MBS).

The most recent version of the IMT-2000 standard is, in fact, constituted by a range of five independent standards. These are the UTRA Frequency Division Duplex (FDD) Wideband Code Division Multiple Access (W-CDMA) mode [8], the UTRA Time Division Duplex (TDD) CDMA mode, the Pan-American multi-carrier CDMA configuration mode known as cdma2000 [8], the Pan-American Time Division Multiple Access (TDMA) mode known as UWT-136, and the Digital European Cordless Telecommunications (DECT) [8] mode. It would be desirable for future systems to become part of this standard

[1]This chapter is based on L.-L. Yang and L. Hanzo, Software Defined Radio Assisted Adaptive Broadband Frequency Hopping Multicarrier DS-CDMA, ©IEEE [11].

framework without having to define new standards. The framework proposed in this contribution is capable of satisfying this requirement.

More specifically, these future wireless systems are expected to cater for a range of requirements. Firstly, MBSs are expected to support extremely high bitrate services while exhibiting different traffic characteristics and satisfying the required quality of service guarantees [3]. The objective is that mobile users become capable of accessing the range of broadband services available for fixed users at data rates up to 155 Mb/s. Multi-standard operation is also an essential requirement. Furthermore, these systems are expected to be highly flexible, supporting multimode and multiband operation as well as global roaming, while achieving the highest possible spectral efficiency. These features have to be sustained under adverse operating conditions, i.e., for high-speed users, for dynamically fluctuating traffic loads, and over hostile propagation channels. These requirements can be conveniently satisfied with the aid of broadband mobile wireless systems based on the concept of adaptive software defined radio (SDR) architectures [9,10].

In the first part of this chapter, a broadband multiple access candidate scheme meeting the above requirements is presented, which is constituted by frequency-hopping (FH) based multicarrier DS-CDMA (FH/MC DS-CDMA) [16–19]. Recent investigations demonstrated that channel-quality controlled rate adaptation is an efficient strategy for attaining the highest possible spectral efficiency in terms of b/s/Hz [20–24], while maintaining a certain target integrity. Hence, in the second part of the chapter we consider adaptive rate transmission (ART) schemes associated with supporting both time-variant rate and multirate services. These ART techniques are discussed in the context of the proposed FH/MC DS-CDMA system, arguing that SDRs constitue a convenient framework for their implementation. Therefore, in the final part of this contribution, the concept of SDR-assisted broadband FH/MC DS-CDMA is presented and the range of reconfigurable parameters are described with the aim of outlining a set of promising research topics. Let us now commence our detailed discourse concerning the proposed FH/MC DS-CDMA system.

91.2 FH/MC DS-CDMA

The transmitter schematic of the proposed FH/MC DS-CDMA arrangement is depicted in Fig. 91.1. Each subcarrier of a user is assigned a pseudo-noise (PN) spreading sequence. These PN sequences can be simultaneously assigned to a number of users, provided that only one user activates the same PN sequence on the same subcarrier. These PN sequences produce narrow-band DS-CDMA signals. In Fig. 91.1, C(Q, U) represents a constant-weight code having U number of 1s and (Q - U) number of 0s. Hence, the weight of C(Q, U) is U. This code is read from a so-called constant-weight code book, which represents the frequency-hopping patterns. The constant-weight code C(Q, U) plays two different roles. Its first role is that its weight—namely U—determines the number of subcarriers invoked, while its second function is that the positions of

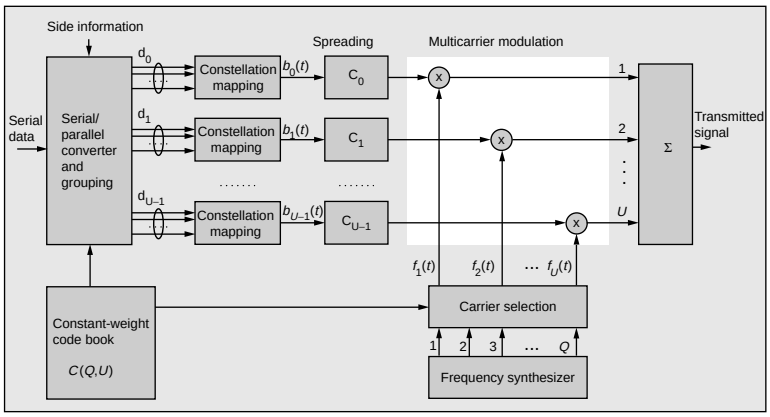


FIGURE 91.1 Transmitter diagram of the frequency-hopping multicarrier DS-CDMA system using adaptive transmission, ©IEEE [11].

the U number of binary 1s determines the selection of a set of U number of subcarrier frequencies from the Q outputs of the frequency synthesizer. Furthermore, in the transmitter, side-information reflecting the channel's instantaneous quality might be employed in order to control its transmission and coding mode so that the required target throughput and transmission integrity requirements are met.

As shown in Fig. 91.1, the original bit stream having a bit duration of T_b is first serial-to-parallel (S-P) converted. Then, these parallel bit streams are grouped and mapped to the potentially time-variant modulation constellations of the U active subcarriers. Let us assume that the number of bits transmitted by an FH/MC DS-CDMA symbol is M, and let us denote the symbol duration of the FH/MC DS-CDMA signal by T_s. Then, if the system is designed for achieving a high processing gain and for mitigating the inter-symbol-interference (ISI) in a constant-rate transmission scheme, the symbol duration can be extended to a multiple of the bit duration, i.e., $T_s = MT_b$. In contrast, if the design aims to support multiple transmission rates or channel-quality matched variable information rates, then a constant bit duration of $T_0 = T_s$ can be employed. Both multirate and variable rate transmissions can be implemented by employing a different number of subcarriers associated with different modulation constellations as well as different spreading gains. As seen in Fig. 91.1, after the constellation mapping stage, each branch is DS spread using the assigned PN sequence, and then this spread signal is carrier modulated using one of the active subcarrier frequencies derived from the constant-weight code C(Q, U). Finally, all U active branch signals are multiplexed in order to form the transmitted signal.

In the FH/MC DS-CDMA receiver of Fig. 91.2, the received signal associated with each active subcarrier is detected using, for example, a RAKE combiner. Alternatively, multiuser detection (MUD) can be invoked in order to approach the single-user bound. In contrast to the transmitter side, where only U out of Q subcarriers are transmitted by a user, at the receiver different detector structures might be implemented based on the availability [19] or lack [25] of the FH pattern information. During the FH pattern acquisition stage, which usually happens at the beginning of transmission or during handover, tentatively all Q subcarriers can be demodulated. The transmitted information can be detected and the FH patterns can be acquired simultaneously by using blind joint detection algorithms exploiting the characteristics of the constant-weight codes [17,18]. If, however, the receiver has the explicit knowledge of the FH patterns, then only U subcarriers have to be demodulated. However, if fast Fourier transform (FFT) techniques are employed for demodulation, as often is the case in multicarrier CDMA [26] or OFDM [27] systems, then all Q subcarriers might be demodulated, where the inactive subcarriers only output noise. In the decision unit of Fig. 91.2, these noise-output-only branches can be eliminated by exploiting the knowledge of the FH patterns [19,25]. Hence, the decision unit only outputs the information transmitted

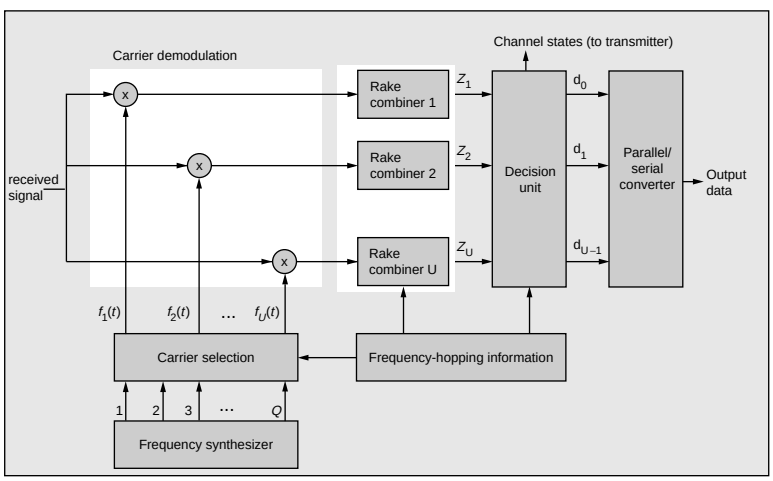


FIGURE 91.2 Receiver block diagram of the frequency-hopping multicarrier DS-CDMA system using conventional RAKE receiver, ©IEEE [11].

by the active subcarriers. Finally, the decision unit's output information is parallel-to-serial converted to form the output data.

At the receiver, the channel states associated with all the subchannels might be estimated or predicted using pilot signals [22,24]. This channel state information can be utilized for coherent demodulation. It can also be fed back to the transmitter as highly protected side-information in order to invoke a range of adaptive transmission schemes including power control and adaptive-rate transmission.

91.3 Characteristics of the FH/MC DS-CDMA Systems

In the proposed FH/MC DS-CDMA system, the entire bandwidth of future broadband systems can be divided into a number of subbands, and each subband can be assigned a subcarrier. According to the prevalent service requirements, the set of legitimate subcarriers can be distributed in line with the users' instantaneous information rate requirements. FH techniques are employed for each user in order to evenly occupy the whole system bandwidth available and to efficiently utilize the available frequency resources. In this respect, FH/MC DS-CDMA systems exhibit compatibility with the existing 2nd- and 3rd-generation CDMA systems and, hence, constitute a highly flexible air interface.

Broadband Wireless Mobile System—To elaborate a little further, our advocated FH/MC DS-CDMA system is a broadband wireless mobile system constituted by multiple narrow-band DS-CDMA subsystems. Again, FH techniques are employed for each user in order to evenly occupy the whole system bandwidth and to efficiently utilize the available frequency resources. The constant-weight code based FH patterns used in the schematic of Fig. 91.1 are invoked in order to control the number of subcarriers invoked which is kept constant during the FH procedure. In contrast to single-carrier broadband DS-CDMA systems such as wideband CDMA (W-CDMA) [2] exhibiting a bandwidth in excess of 5 MHz—which inevitably results in extremely high-chip-rate spreading sequences and high-complexity—the proposed FH/MC DS-CDMA system does not have to use high chip-rate DS spreading sequences since each subcarrier conveys a narrow-band DS-CDMA signal. In contrast to broadband OFDM systems [6]—which have to use a high number of subcarriers and usually result in a high peak-to-average power fluctuation—due to the associated DS spreading, the number of subcarriers of the advocated broadband wireless FH/MC DS-CDMA system may be significantly lower. This potentially mitigates the crest-factor problem. Additionally, with the aid of FH, the peak-to-average power fluctuation of the FH/MC DS-CDMA system might be further decreased. In other words, the FH/MC DS-CDMA system is capable of combining the best features of single-carrier DS-CDMA and OFDM systems while avoiding many of their individual shortcomings. Finally, in comparison to the FH/MC DS-CDMA system, both broadband single-carrier DS-CDMA systems and broadband OFDM systems are less amenable to interworking with the existing 2nd- and 3rd-generation wireless communication systems. Let us now characterize some of the features of the proposed system in more depth.

Compatibility—The broadband FH/MC DS-CDMA system can be rolled out over the bands of the 2nd- and 3rd-generation mobile wireless systems and/or in the band licensed for future broadband wireless communication systems. In FH/MC DS-CDMA systems, the subbands associated with different subcarriers are not required to be of equal bandwidth. Hence, existing 2nd- and 3rd-generation CDMA systems can be supported using one or more subcarriers. For example, Fig. 91.3 shows the spectrum of a frequency hopping, orthogonal multicarrier DS-CDMA signal using a subchannel bandwidth of 1.25 MHz, which constitutes the bandwidth of a DS-CDMA signal in the IS-95 standard [1]. In Fig. 91.3, we also show that seven subchannels, each having a bandwidth of 1.25 MHz, can be replaced by one subchannel with a bandwidth of 5 MHz ($= 8 \times 1.25/2$ MHz). Hence, the narrow-band IS-95 CDMA system can be supported by a single subcarrier, while the UMTS and IMT-2000 W-CDMA systems might be supported using seven subchannels' bandwidth amalgamated into one W-CDMA channel. Moreover, with the aid of SDRs, FH/MC DS-CDMA is also capable of embracing other existing standards, such as the time-division multiple-access (TDMA) based global system of mobile communications known as GSM [8].

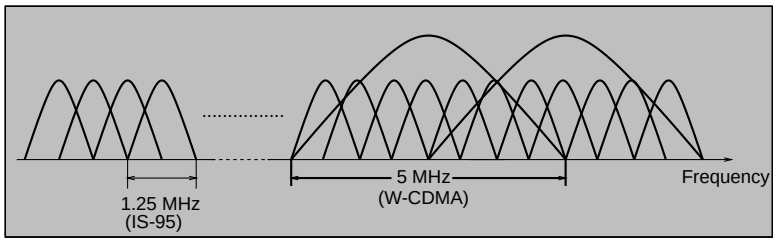


FIGURE 91.3 Spectrum of FH/MC DS-CDMA signal using subchannel bandwidth of 1.25 MHz and/or 5 MHz, ©IEEE [11].

FH Strategy—In FH/MC DS-CDMA systems, both slow FH and fast FH techniques can be invoked, depending on the system's design and the state-of-the-art. In slow FH, several symbols are transmitted after each frequency hopping, while in fast FH, several frequency hops take place in a symbol duration, i.e., each symbol is transmitted using several subcarriers. Moreover, from a networking point of view, random FH, uniform FH, and adaptive FH [19] schemes can be utilized in order to maximize the efficiency of the network. In the context of *random* FH [19], the subcarriers associated with each transmission of a user are determined by the set of pre-assigned FH patterns constituting a group of constant-weight codewords [25]. The active subcarriers are switched from one group of frequencies to another without the knowledge of the FH patterns of the other users. In contrast, for the FH/MC DS-CDMA system using *uniform* FH [19], the FH patterns of all users are determined jointly under the control of the base station (BS), so that each subcarrier is activated by a similar number of users. It can be shown that for the downlink (DL), uniform FH can be readily implemented since the BS has the knowledge of the FH patterns of all users. However, for implementing uplink (UL) transmissions, the FH patterns to be used must be signalled by the BS to each mobile station (MS) in order to be able to implement uniform FH. Finally, if the near-instantaneous channel quality information is available at the transmitter, advanced adaptive FH can be invoked, where information is only transmitted over a group of subchannels exhibiting a satisfactory signal-to-interference ratio (SIR).

Implementation of Multicarrier Modulation—The multicarrier modulation block in Fig. 91.1 and the multicarrier demodulation block in Fig. 91.2 can be implemented using FFT techniques, provided that each of the subchannels occupies the same bandwidth. Since not all of the subcarriers are activated at each transmission in the proposed FH/MC DS-CDMA system, the deactivated subcarriers can be set to zero in the FFT or IFFT algorithm. However, if an unequal bandwidth is associated with the subchannels, multicarrier modulation/demodulation can only be implemented using less efficient conventional, rather than FFT based carrier modulation/demodulation, schemes.

Access Strategy—When a new user attempts to access the channel and commences his/her transmission, a range of different access strategies might be offered by the FH/MC DS-CDMA system in order to minimize the interference inflicted by the new user to the already active users. Specifically, if there are subchannels which are not occupied by any other users, or if there are subchannels that exhibit a sufficiently high SIR, then the new user can access the network using these passive subchannels or the subchannels exhibiting a high SIR. However, if all the subchannels have been occupied and the SIR of each of the subchannels is insufficiently high, then the new user accesses the network by spreading its transmitted energy evenly across the subchannels. This access scheme imposes the minimum possible impact on the QoS of the users already actively communicating. However, the simplest strategy for a new user to access the network is by randomly selecting one or several subchannels.

Multirate and Variable Rate Services—In FH/MC DS-CDMA systems, multirate and variable rate services can be implemented using a variety of approaches. Specifically, the existing techniques, such as employing a variable spreading gain, multiple spreading codes, a variable constellation size, variable-rate forward error correction (FEC) coding, etc., can be invoked to provide multirate and variable rate services. Furthermore, since the proposed FH/MC DS-CDMA systems use constant-weight code based FH patterns,

multirate and variable rate services can also be supported by using constant-weight codes having different weights, i.e., by activating a different number of subcarriers. Note that the above-mentioned techniques can be implemented either separately or jointly in a system.

Diversity—The FH/MC DS-CDMA system includes frequency hopping, multicarrier modulation, as well as direct-sequence spreading, hence a variety of diversity schemes and their combinations can be implemented. The possible diversity schemes include the following arrangements.

- If the chip-duration of the spreading sequences is lower than the maximum delay spread of the fading channels, then frequency diversity can be achieved on each of the subcarrier signals.
- Frequency diversity can also be achieved by transmitting the same signal using a number of different subcarriers.
- Time diversity can be achieved by using slow frequency hopping in conjunction with error control coding as well as interleaving.
- Time-frequency diversity can be achieved by using fast frequency hopping techniques, where the same symbol is transmitted using several time slots assigned to different frequencies.
- Spatial diversity can be achieved by using multiple transmit antennas, multiple receiver antennas, and polarization.

Initial Synchronization—In our FH/MC DS-CDMA system, initial synchronization can be implemented by first accomplishing DS code acquisition. The fixed-weight code book index of the FH pattern used can be readily acquired once DS code acquisition is achieved. During DS code acquisition, the training signal supporting the initial synchronization, which is usually the carrier modulated signal without data modulation, can be transmitted using a group of subcarriers. These subcarrier signals can be combined at the receiver using, for example, equal gain combining (EGC) [28]. Hence, frequency diversity can be employed during the DS code acquisition stage of the FH/MC DS-CDMA system's operation, and, consequently, the initial synchronization performance can be significantly enhanced. Following the DS code acquisition phase, data transmission can be activated and the index of the FH pattern used can be signalled to the receiver using a given set of fixed subchannels. Alternatively, the index of the FH pattern can be acquired blindly from the received signal with the aid of a group of constant-weight codes having a given minimum distance [18].

Interference Resistance—The FH/MC DS-CDMA system concerned can mitigate the effects of intersymbol interference encountered during high-speed transmissions, and it readily supports partial-band and multitone interference suppression. Moreover, the multiuser interference can be suppressed by using multiuser detection techniques [20], potentially approaching the single-user performance.

Advanced Technologies—The FH/MC DS-CDMA system can efficiently amalgamate the techniques of FH, OFDM, and DS-CDMA. Simultaneously, a variety of performance enhancement techniques, such as multiuser detection [29], turbo coding [30], adaptive antennas [31], space-time coding and transmitter diversity [32], near-instantaneously adaptive modulation [21], etc., might be introduced.

Flexibility—The future generation broadband mobile wireless systems will aim to support a wide range of services and bit rates. The transmission rates may vary from voice and low-rate messages to very high-rate multimedia services requiring rates in excess of 100 Mb/s [3]. The communications environments vary in terms of their grade of mobility, the cellular infrastructure, the required symmetric and asymmetric transmission capacity, and whether indoor, outdoor, urban, or rural area propagation scenarios are encountered. Hence, flexible air interfaces are required which are capable of maximizing the area spectrum efficiency expressed in terms of bits/s/Hz/km^2 in a variety of communication environments. Future systems are also expected to support various types of services based on ATM and IP, which require various levels of quality of service (QoS). As argued before, FH/MC DS-CDMA systems exhibit a high grade of compatibility with existing systems. These systems also benefit from the employment of FH, MC, and DS spreading-based diversity assisted adaptive modulation [27]. In short, FH/MC DS-CDMA systems constitute a high-flexibility air interface.

91.4 Adaptive Rate Transmission

Why Adaptive Rate Transmission?

There are a range of issues which motivate the application of adaptive rate transmissions (ART) in the broadband mobile wireless communication systems of the near future. The explosive growth of the Internet and the continued dramatic increase in demand for all types of wireless services are fueling the demand for increasing the user capacity, data rates, and the variety of services supported. Typical low-data-rate applications include audio conferencing, voice mail, messaging, email, facsimile, and so on. Medium- to high-data-rate applications encompass file transfer, Internet access, high-speed packet- and circuit-based network access, as well as high-quality video conferencing. Furthermore, the broadband wireless systems in the future are also expected to support real-time multimedia services, which provide concurrent video, audio, and data services to support advanced interactive applications. Hence, in the future generation mobile wireless communication systems, a wide range of information rates must be provided in order to support different services which demand different data rates and different QoS. In short, an important motivation for using ART is to support a variety of services, which we refer to as service-motivated ART (S-ART). However, there is a range of other motivating factors which are addressed below.

The performance of wireless systems is affected by a number of propagation phenomena: (1) path-loss variation vs. distance; (2) random slow shadowing; (3) random multipath fading; (4) inter-symbol interference (ISI), co-channel interference, and multiuser interference; and (5) background noise. For example, mobile radio links typically exhibit severe multipath fading, which leads to serious degradation in the link's signal-to-noise ratio (SNR) and, consequently, a higher bit error rate (BER). Fading compensation techniques such as an increased link budget margin or interleaving with channel coding are typically required to improve the link's performance. However, today's cellular systems are designed for the worst-case channel conditions, typically achieving adequate voice quality over 90–95% of the coverage area for voice users, where the signal-to-interference plus noise ratio (SINR) is above the designed target [21]. Consequently, the systems designed for the worst-case channel conditions result in poor exploitation of the available channel capacity a good percentage of the time. Adapting the transmitter's certain parameters to the time-varying channel conditions leads to better exploitation of the channel capacity available. This ultimately increases the area spectral efficiency expressed in terms of b/s/Hz/km^2. Hence, the second reason for the application of ART is constituted by the time-varying nature of the channel, which we refer to as channel quality motivated ART (C-ART).

What Is Adaptive Rate Transmission?

Broadly speaking, ART in mobile wireless communications implies that the transmission rates at both the base stations and the mobile terminals can be adaptively adjusted according to the instantaneous operational conditions, including the communication environment and service requirements. With the expected employment of SDR-based wireless systems, the concept of ART might also be extended to adaptively controlling the multiple access schemes—including FDMA, TDMA, narrow-band CDMA, W-CDMA, and OFDM—as well as the supporting network structures—such as local area networks and wide area networks. In this contribution, only C-ART and S-ART are concerned in the context of the proposed FH/MC DS-CDMA scheme. Employing ART in response to different service requests indicates that the transmission rate of the base station and the mobile station can be adapted according to the requirements of the services concerned, as well as to meet their different QoS targets. In contrast, employing ART in response to the time-varying channel quality implies that for a given service supported, the transmission rate of the base station and that of the mobile station can be adaptively controlled in response to their near-instantaneous channel conditions. The main philosophy behind C-ART is the real-time balancing of the link budget through adaptive variation of the symbol rate, modulation constellation size and format, spreading factor, coding rate/scheme, etc., or, in fact, any combination of these parameters. Thus, by taking advantage of the time-varying nature of the wireless channel and interference conditions, the C-ART schemes can provide

a significantly higher average spectral efficiency than their fixed-mode counterparts. This takes place without wasting power, without increasing the co-channel interference, or without increasing the BER. We achieve these desirable properties by transmitting at high speeds under favorable interference/channel conditions and by responding to degrading interference and/or channel conditions through a smooth reduction of the associated data throughput. Procedures that exploit the time-varying nature of the mobile channel are already in place for all the major cellular standards worldwide [21], including IS-95 CDMA, cdma2000, and UMTS W-CDMA [8], IS-136 TDMA, the General Packet Radio Service of GSM, and the Enhanced Data rates for Global Evolution (EDGE) schemes. The rudimentary portrayal of a range of existing and future ART schemes is given below. Note that a system may employ a combination of several ART schemes, listed below, in order to achieve the desired data rate, BER, or the highest possible area spectrum efficiency.

- **Multiple Spreading Codes**—In terms of S-ART, higher rate services can be supported in CDMA based systems by assigning a number of codes. For example, in IS-95B, each high-speed user can be assigned one to eight Walsh codes, each of which supports a data rate of 9.6 kb/s. In contrast, multiple codes cannot be employed in the context of C-ART in order to compensate for channel fading, path-loss, and shadowing unless they convey the same information, and, hence, achieve diversity gain. However, if the co-channel interference is low—which usually implies in CDMA based systems that the number of simultaneously transmitting users is low—then multiple codes can be transmitted by an active user in order to increase the user's transmission rate.
- **Variable Spreading Factors**—In the context of S-ART, higher rate services are supported by using lower spreading factors without increasing the bandwidth required. For example, in UMTS W-CDMA [1], the spreading factors of 4/8/16/32/64/128/256 may be invoked to achieve the corresponding data rates of 1024/512/256/128/64/32/16 kb/s. In terms of C-ART, when the SINR experienced is increased, reduced spreading factors are assigned to users for the sake of achieving higher data rates.
- **Variable Rate FEC Codes**—In a S-ART scenario, higher rate services can be supported by assigning less powerful, higher rate FEC codes associated with reduced redundancy. In a C-ART scenario, when the SINR improves, a higher-rate FEC code associated with reduced redundancy is assigned in an effort to achieve a higher data rate.
- **Different FEC Schemes**—The range of coding schemes might entail different classes of FEC codes, code structures, encoding/decoding schemes, puncturing patterns, interleaving depths and patterns, and so on. In the context of S-ART, higher rate services can be supported by coding schemes having a higher coding rate. In the context of C-ART, usually an appropriate coding scheme is selected in order to maximize the spectrum efficiency. The FEC schemes concerned may entail block or convolutional codes, block or convolutional constituent code based turbo codes, trellis codes, turbo trellis codes, etc. The implementational complexity and error correction capability of these codes can be varied as a function of the coding rate, code constraint length, the number of turbo decoding iterations, the puncturing pattern, etc. A rule of thumb is that the coding rate is increased toward unity, as the channel quality improves, in order to increase the system's effective throughput.
- **Variable Constellation Size**—In S-ART schemes, higher rate services can be supported by transmitting symbols having higher constellation sizes. For example, an adaptive modem may employ BPSK, QPSK, 16QAM, and 64QAM constellations [27], which corresponds to 1, 2, 4, and 6 bits per symbol. The highest data rate provided by the 64QAM constellation is a factor six higher than that provided by employing the BPSK constellation. In C-ART scenarios, when the SINR increases, a higher number of bits per symbol associated with a higher order constellation is transmitted for increasing the system's throughput.
- **Multiple Time Slots**—In a S-ART scenario, higher rate services can also be supported by assigning a corresponding number of time slots. A multiple time slot based adaptive rate scheme is used in GPRS-136 (1–3 time slots/20 ms) and in enhanced GPRS (EGPRS) (1–8 time slots/4.615GSM frame) in order to achieve increased data rates. In the context of C-ART, multiple time slots associated

with interleaving or frequency hopping can be implemented for achieving time diversity gain. Hence, C-ART can be supported by assigning a high number of time slots for the compensation of severe channel fading at the cost of tolerating a low data throughput. In contrast, assigning a low number of time slots over benign non-fading channels allows us to achieve a high throughput.

- **Multiple Bands**—In the context of S-ART, higher rate services can also be supported by assigning a higher number of frequency bands. For example, in UMTS W-CDMA [1], two 5 MHz bands may be assigned to a user in order to support the highest data rate of 2 Mb/s (=2 × 1024 kb/s), which is obtained by using a spreading factor of 4 on each subband signal. In the context of C-ART associated with multiple bands, frequency-hopping associated with time-variant redundancy and/or variable rate FEC coding schemes or frequency diversity techniques might be invoked in order to increase the spectrum efficiency of the system. For example, in C-ART schemes associated with double-band assisted frequency diversity, if the channel quality is low, the same signal can be transmitted in two frequency bands for the sake of maintaining a diversity order of two. However, if the channel quality is sufficiently high, two independent streams of information can be transmitted in these bands and, consequently, the throughput can be doubled.
- **Multiple Transmit Antennas**—Employing multiple transmit antennas based on space-time coding [32] is a novel method of communicating over wireless channels which was also adapted for use in the 3rd-generation mobile wireless systems. ART can also be implemented using multiple transmit antennas associated with different space-time codes. In S-ART schemes, higher rate services can be supported by a higher number of transmit antennas associated with appropriate space-time codes. In terms of C-ART schemes, multiple transmit antennas can be invoked for achieving a high diversity gain. Therefore, when the channel quality expressed in terms of the SINR is sufficiently high, the diversity gain can be decreased. Consequently, two or more symbols can be transmitted in each signalling interval, and each stream is transmitted by only a fraction of the transmit antennas associated with the appropriate space-time codes. Hence, the throughput is increased. However, when the channel quality is poor, all the transmit antennas can be invoked for transmitting one stream of data, hence achieving the highest possible transmit diversity gain of the system while decreasing the throughput.

Above we have summarized the philosophy of a number of ART schemes which can be employed in wireless communication systems. An S-ART scheme requires a certain level of transmitted power in order to achieve the required QoS. Specifically, a relatively high transmitted power is necessitated for supporting high-data rate services, and a relatively low transmitted power is required for offering low-data rate services. Hence, a side effect of an S-ART scheme supporting high data rate services is the concomitant reduction of the number of users supported due to the increased interference or/and increased bandwidth. By contrast, a cardinal requirement of a C-ART scheme is the accurate channel quality estimation or prediction at the receiver as well as the provision of a reliable side-information feedback between the channel quality estimator or predictor of the receiver and the remote transmitter [22,24] where the modulation/coding mode requested by the receiver is activated. The parameters capable of reflecting the channel quality may include BER, SINR, transmission frame error rate, received power, path loss, automatic repeat request (ARQ) status, etc. A C-ART scheme typically operates under the constraint of constant transmit power and constant bandwidth. Hence, without wasting power and bandwidth, or without increasing the co-channel interference compromising the BER performance, C-ART schemes are capable of providing a significantly increased average spectral efficiency by taking advantage of the time-varying nature of the wireless channel when compared to fixed-mode transmissions.

Adaptive Rate Transmission in FH/MC DS-CDMA Systems

Above we have discussed a range of ART schemes which were controlled by the prevalent service requirements and the channel quality. Future broadband mobile wireless systems are expected to provide a wide range of services characterized by highly different data rates while achieving the highest possible spectrum efficiency. The proposed FH/MC DS-CDMA based broadband mobile wireless system

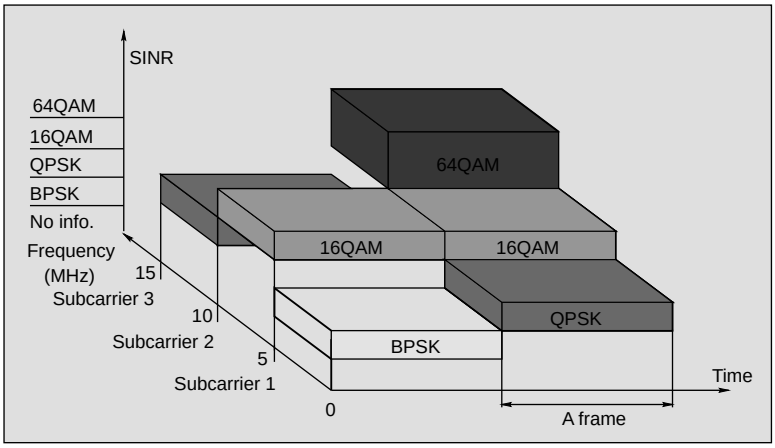


FIGURE 91.4 A frame structure of burst-by-burst adaptive modulation in multicarrier DS-CDMA systems, ©IEEE [11].

constitutes an attractive candidate system since it is capable of efficiently utilizing the available frequency resources, as discussed previously, and simultaneously achieving a high grade of flexibility by employing ART techniques. More explicitly, the FH/MC DS-CDMA system can provide a wide range of data rates by combining the various ART schemes discussed above. At the same time, for any given service, the FH/MC DS-CDMA system may also invoke a range of adaptation schemes in order to achieve the highest possible spectrum efficiency in various propagation environments such as indoor, outdoor, urban, rural scenarios at low to high speeds. Again, the system is expected to support different services at a variety of QoS, including voice mail, messaging, email, file transfer, Internet access, high-speed packet- and circuit-based network access, real-time multimedia services, and so on. As an example, a channel-quality motivated burst-by-burst ART assisted FH/MC DS-CDMA system is shown in Fig. 91.4, where we assume that the number of subcarriers is three, the bandwidth of each subchannel is 5 MHz and the channel quality metric is the SINR. In response to the SINR experienced, the transmitter may transmit a frame of symbols selected from the set of BPSK, QPSK, 16QAM, or 64QAM constellations or may simply curtail transmitting information if the SINR is too low.

91.5 Software Defined Radio Assisted FH/MC DS-CDMA

The range of existing wireless communication systems is shown in Fig. 91.5. Different legacy systems will continue to coexist, unless ITU, by some miracle, succeeds in harmonizing all the forthcoming standards under a joint framework while at the same time ensuring compatibility with the existing standards. In the absence of the perfect standard, the only solution is employing multiband, multimode, multistandard transceivers based on the concept of software defined radios (SDR) [9,10].

In SDRs, the digitization of the received signal is performed at some stage downstream from the antenna, typically after wideband filtering, low-noise amplification, and down-conversion to a lower frequency. The reverse processes are invoked by the transmitter. In contrast to most wireless communication systems which employ digital signal processing (DSP) only at baseband, SDRs are expected to implement the DSP functions at an intermediate frequency (IF) band. An SDR defines all aspects of the air interface, including RF channel access and waveform synthesis in software. In SDRs, wide-band analog-to-digital and digital-to-analog converters (ADC and DAC) transform each RF service band from digital and analog forms at IF. The wideband digitized receiver stream of bandwidth W_s accommodates all subscriber channels, each of which has a bandwidth of $W_c (W_c \ll W_s)$. Because they use programmable DSP chips at both the intermediate frequency as well as at baseband, SDRs can efficiently support

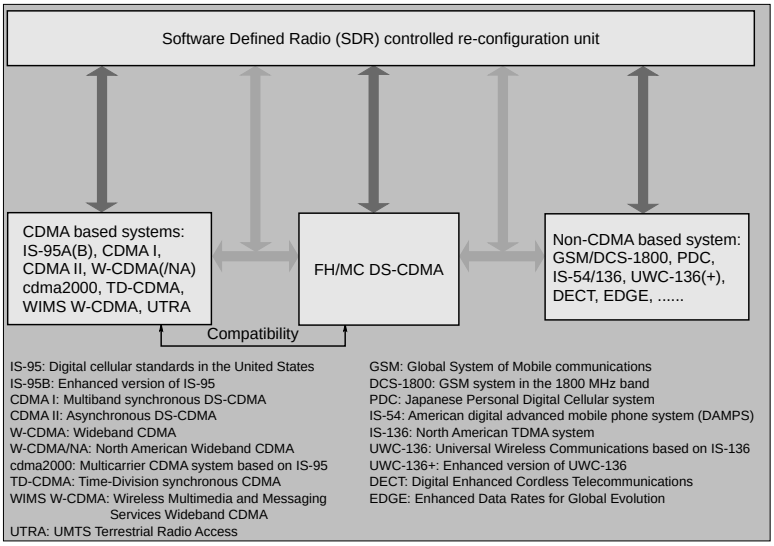


FIGURE 91.5 Software defined radio assisted FH/MC DS-CDMA and its reconfiguration modes, ©IEEE [11].

multiband and multistandard communications. An SDR employs one or more reconfigurable processors embedded in a real-time multiprocessing fabric, permitting flexible reprogramming and reconfiguration using software downloaded, for example, with the aid of a signalling channel from the base station. Hence, SDRs offer an elegant solution to accommodating various modulation formats, coding, and radio access schemes. They also have the potential of reducing the cost of introducing new technology superseding legacy systems and are amenable to future software upgrades, potentially supporting sophisticated future signal processing functions such as array processing, multiuser detection, and as yet unknown coding techniques.

FH/MC DS-CDMA systems will be designed in order to provide the maximum grade of compatibility with the existing CDMA based systems, such as IS-95 and W-CDMA based systems. For example, the frequency bands of the IS-95 CDMA system in North America are 824–849 MHz (uplink) and 869–894 MHz (downlink), respectively. The corresponding frequency bands for the UMTS-FDD wideband CDMA system are 1850–1910 MHz (uplink) and 1930–1990 MHz (downlink) in North America, and 1920–1980 MHz (uplink) and 2110–2170 MHz (downlink) in Europe. In order to ensure compatibility with these systems, the proposed FH/MC DS-CDMA system's spectrum can be assigned according to Fig. 91.6. Specially, in the frequency band of IS-95, 39 orthogonal subcarriers are assigned, each having a bandwidth of 1.25 MHz, while in the frequency band of UMTS-FDD W-CDMA, 23 orthogonal subcarriers are allocated, each with a bandwidth of 5 MHz. The multicarrier modulation used in the FH/MC DS-CDMA system obeying the above spectrum allocation can be readily implemented using two IFFT subsystems at the transmitter and two FFT subsystems at the receiver, where a pair of IFFT-FFT subsystems carries out modulation/demodulation in the IS-95 band, while another pair of IFFT-FFT subsystems transmits and receives in the UMTS-FDD band. If the chip rate for the 1.25 MHz bandwidth subcarriers is 1.2288M chips per second (cps) and for the 5 MHz bandwidth subcarriers is 3.84 Mcps, then the FH/MC DS-CDMA system will be compatible with both the IS-95 and the UMTS-FDD W-CDMA systems.

However, the terminals of future broadband mobile wireless systems are expected not only to support multimode and multistandard communications, but also to possess the highest possible grade of flexibility while achieving a high spectrum efficiency. Hence, these systems are expected to be capable of software reconfiguration both between different standards as well as within a specific standard. In contrast to reconfiguration between different standards invoked, mainly for the sake of compatibility, the objective of reconfiguration within a specific standard is to support a variety of services at the highest possible

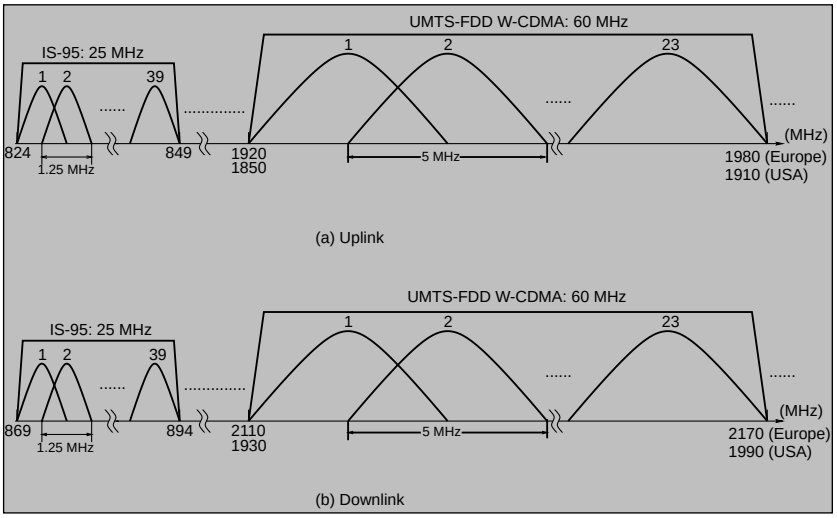


FIGURE 91.6 Exhibition of spectrum compatibility of the broadband FH/MC DS-CDMA system with IS-95 and UMTS-FDD wideband CDMA systems, ©IEEE [11].

spectrum efficiency. The SDR assisted broadband FH/MC DS-CDMA system is operated under the control of the software reconfiguration unit shown in Fig. 91.5. The set of reconfigured parameters of the broadband FH/MC DS-CDMA system may include:

- **Services:** data rate, QoS, real-time or non-real-time transmission, and encryption/decryption schemes and parameters
- **Error Control:** CRC, FEC codes, coding/decoding schemes, coding rate, number of turbo decoding steps, and interleaving depth and pattern
- **Modulation:** modulation schemes, signal constellation, and partial response filtering
- **PN Sequence:** spreading sequences (codes), chip rate, chip waveform, spreading factor, and PN acquisition and tracking schemes
- **Frequency Hopping:** FH schemes (slow, fast, random, uniform, and adaptive), FH patterns, and weight of constant-weight codes
- **Detection:** detection schemes (coherent or non-coherent, etc.) and algorithms (maximum likelihood sequence detection (MLSD) or minimum mean square estimation (MMSE), etc.), parameters associated with space/time as well as frequency diversity, beam-forming, diversity combining schemes, equalization schemes as well as the related parameters (such as the number of turbo equalization iterations, etc.), and channel quality estimation algorithms
- **Others:** subchannel bandwidth and power control parameters

In the context of different standards—in addition to the parameters listed above—the transceiver parameters that must be reconfigurable have to include the clock rate, the radio frequency (RF) bands, and air interface modes.

91.6 Final Remarks

We have presented a flexible broadband mobile wireless communication system based on FH/MC DS-CDMA and reviewed the existing as well as a range of forthcoming techniques which might be required to develop broadband mobile wireless systems exhibiting high flexibility and high efficiency. We argued that this broadband FH/MC DS-CDMA system exhibits a high grade of compatibility with the existing CDMA

based systems since it is capable of incorporating a wide range of techniques developed for the 2nd- and 3rd-generation mobile systems. At the time of this writing, research is well under way toward the SDR-based implementation of a range of existing systems [33]. It is expected that these efforts will soon encompass a generic scheme not too different from the FH/MC DS-CDMA scheme advocated in this chapter. For a range of related topics, the interested reader is referred to [34–36].

References

1. M. Zeng, A. Annamalai, and V. K. Bhargava, Recent advances in cellular wireless communications, *IEEE Commun. Mag.,* 128–138, 1999.
2. P. Chaudhury, W. Mohr, and S. Onoe, The 3GPP proposal for imt-2000, *IEEE Commun. Mag.,* 72–81, 1999.
3. M. Progler, C. Evci, and M. Umehira, Air interface access schemes for broadband mobile systems, *IEEE Commun. Mag.,* 106–115, 1999.
4. L. Kleinrock, On some principles of nomadic computing and multi-access communications, *IEEE Commun. Mag.,* 46–50, 2000.
5. P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushayana, and A. Viterbi, CDMA/HDR: A bandwidth-efficient high-speed wireless data service for nomadic users, *IEEE Commun. Mag.,* 70–77, 2000.
6. J. Chuang and N. Sollenberger, Beyond 3G: wideband wireless data access based on OFDM and dynamic packet assignment, *IEEE Commun. Mag.,* 78–87, 2000.
7. N. Dimitriou, R. Tafazolli, and G. Sfikas, Quality of service for multimedia CDMA, *IEEE Commun. Mag.,* 88–94, 2000.
8. R. Steele and L. Hanzo, Eds., *Mobile Radio Communications,* 2nd ed., IEEE Press–John Wiley, New York, 1999.
9. J. Mitola III, The software radio architecture, *IEEE Commun. Mag.,* 26–38, 1995.
10. J. Mitola III, Technical challenges in the globalization of software radio, *IEEE Commun. Mag.,* 84–89, 1999.
11. L. L. Yang and L. Hanzo, Software defined radio assisted adaptive broadband frequency hopping multicarrier DS-CDMA, *IEEE Commun. Mag.,* to appear March 2002.
12. H. Tsurumi and Y. Suzuki, Broadband RF stage architecture for software-defined radio in handheld terminal applications, *IEEE Commun. Mag.,* 90–95, 1999.
13. R. H. Walden, Performance trends for analogy-to-digital converters, *IEEE Commun. Mag.,* 96–101, 1999.
14. S. Srikanteswara, J. H. Reed, P. Athanas, and R. Boyle, A soft radio architecture for reconfigurable platforms, *IEEE Commun. Mag.,* 140–147, 2000.
15. D. Murotake, J. Oates, and A. Fuchs, Real-time implementation of a reconfiguable imt-2000 base station channel modem, *IEEE Commun. Mag.,* 148–152, 2000.
16. E. A. Sourour and M. Nakagawa, Performance of orthogonal multicarrier CDMA in a multipath fading channel, *IEEE Trans. Commun.,* 44, 356–367, 1996.
17. L. L. Yang and L. Hanzo, Blind soft-detection assisted frequency-hopping multicarrier DS-CDMA systems, Proceedings of IEEE GLOBECOM'99, Rio de Janeiro, Brazil, December 5–9, 1999, 842–846.
18. S. B. Slimane, MC-CDMA with quadrature spreading for wireless communication systems, *Eur. Trans. Telecommun.,* 9, 371–378, 1998.
19. L. L. Yang and L. Hanzo, Slow frequency-hopping multicarrier ds-cdma over Nakagami multipath fading channels, submitted to *IEEE J. Sel. Areas Commun.,* 19(7), 1211–1222, 1999.
20. S. Verdu, Wireless bandwidth in the making, *IEEE Commun. Mag.,* 53–58, 2000.
21. S. Nanda, K. Balachandran, and S. Kumar, Adaptation techniques in wireless packet data services, *IEEE Commun. Mag.,* 54–64, 2000.
22. A. J. Goldsmith and S. G. Chua, Variable-rate variable-power MQAM for fading channels, *IEEE Trans. Commun.,* 45, 1218–1230, 1997.

23. M. S. Alouini and A. J. Goldsmith, Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques, *IEEE Trans. Vehicular Technol.,* 48, 1165–1181, 1999.
24. A. Duel-Hallen, S. Hu, and H. Hallen, Long-range prediction of fading signals, *IEEE Signal Proc. Mag.,* 62–75, 2000.
25. L. L. Yang and L. Hanzo, Blind joint soft-detection assisted slow frequency-hopping multicarrier DS-CDMA, *IEEE Trans. Commun.,* 48, 1520–1529, 2000.
26. R. Prasad and S. Hara, Overview of multicarrier CDMA, *IEEE Commun. Mag.,* 126–133, 1997.
27. L. Hanzo, W. T. Webb, and T. Keller, *Single- and Multi-carrier Quadrature Amplitude Modulation: Principles and Applications for Personal Communications, WLANs and Broadcasting,* 2nd ed., IEEE Press and John Wiley & Sons, London, 1999.
28. L. L. Yang and L. Hanzo, Parallel code-acquisition for multicarrier DS-CDMA systems communicating over multipath Nakagami fading channels, Proceedings of IEEE GLOBECOM'99, San Francisco, California, November 27–December 1, 2000, 890–894.
29. S. Verdu, *Multiuser Detection*. Cambridge University Press, London, 1998.
30. C. Berrou and A. Glavieux, Near optimum error correcting coding and decoding: turbo codes, *IEEE Trans. Commun.,* 44, 1261–1271, 1996.
31. Y. Li and N. R. Sollenberger, Adaptive antenna arrays for OFDM systems with cochannel interference, *IEEE Trans. Commun.,* 47, 217–229, 1999.
32. V. Tarokh, N. Seshadri, and A. R. Calderbank, Space-time codes for high data rate wireless communication: performance criterion and code construction, *IEEE Trans. Inf. Theory,* 44, 744–765, 1998.
33. W. Tuttlebee, Ed., *Software Defined Radio*, vol. I and II, John Wiley & Sons, New York, 2002.
34. L. Hanzo, C. Wong, and M. Yee, *Adaptive Wireless Transceivers*, John Wiley, IEEE Press, 2002.
35 L. Hanzo, P. Cherriman, and J. Streit, *Wireless Video Communications: From Second to Third Generation Systems*, WLANs and beyond, IEEE Press.
36. L. Hanzo, T. Liew, and B. Yeap, *Turbo Coding, Turbo Equalisation and Space-Time Coding*, John Wiley, IEEE Press, 2002.

92

Ad-Hoc Routing Techniques for Wireless LANs

Ahmed K. Elhakeem
Concordia University

92.1 Introduction

There has been a growing interest in providing mobile users traversing nested wireless LANs with free roaming and minimum call blocking. Existing IEEE 802.11 based wireless LANs [Gier, 1999] have minimal capabilities when it comes to providing wireless connectivity among neighboring wireless LAN segments. These networks have always found applications in the military and disasterous relief fields. However, the future may carry more surprises given the need to deploy the cellular based macrosystems in remote areas, e.g., underground stations of public transportation. The high bandwidth these wireless LANs provide is another attribute to their utilization and interconnection to the generally narrower band macrosystems. This paves the way to wider deployment of the bandwidth-thirsty multimedia applications such as full IP Web access and many other services of the intelligent transportation systems [Elhakeem, 2000]. Many such LAN segments may be interconnected by means of a terrestrial ethernet or other networks utilizing the standard **access point (AP) mode** of operation. However, such connectivity is lacking within the IEEE 802.11 ad-hoc mode. In the latter, users establish peer-to-peer communications without seeking the help of a central station AP. One can envision that the distributed ad-hoc mode will be adopted by the multivendor multidesign based deployment of such nested LANs. This will make the

need for efficient routing of users' packets by wireless means even greater. For mobile users, faster roaming between the microcells of the nested wireless LANs necessitates the devising of efficient routing techniques so as to mitigate the combined effects of logical and physical link failures.

Ad-hoc networks are easier to deploy compared to cellular counterparts which necessitate expensive planning, support ground networks, and expensive base stations. Failure of a few access points (wireless LAN base stations) may not lead to overall network failure. On the other hand, the limited communication range wireless LANs may have is more than offset by the high rate and cheaper operations in the unlicensed 900 MHz, 1900 MHz, 2.4 GHz, and 6 GHz bands. **Ad-hoc routing** techniques for wireless LANs can be categorized as follows [Iwata, 1999]:

- **Global precomputed routing (sometimes called flat or proactive):** All routes to all nodes are computed *a priori* and updated periodically at each node. Distance vector routing DSDV [Perkins, 1994] and links state LS [Jaquet, 2000] routing fit into this category.
- **Hierarchical routing [Kasera, 1997]:** Here, two kinds of nodes exist: end points and switches. End points select the switch (similar to the BS of the cellular system) and form a cell around it. In turn, the switches form clusters among themselves. Cluster heads are appointed and form a higher level cluster and so on.
- **Flooding and limited flooding techniques:** A node hearing a data racket rebroadcasts to all neighbors and so on. To limit the amount of overhead traffic, limited flooding techniques are proposed [Elhakeem, 2000].
- **On-demand routing techniques (sometimes called reactive):** Here, nodes issue route requests and receive route replies as the need arises with no build-up of routing tables nor a forwarding data basis.

92.2 AODV (Ad-Hoc On-Demand Distance Vector Routing) [Perkins, 1997; Assad, 1998]

The Routing Algorithm

This technique is a mix of proactive/reactive ad-hoc routing approaches. It establishes a balance between full proactivity where the wireless node stores all information about every node in the wireless Internet (complete routing tables) and full reactivity where there is no information about any node. In the latter case, excessive routing query-response will have to be carried out leading to long latency (call waiting delay).

The routing algorithm can be summarized as follows:

1. The source node consults its routing table for the intended call destination. If it finds this destination, the call will commence, otherwise a route request (R-request) packet is sent. Figure 92.1 shows the routing related fields in this packet. Other typical fields such as SYNC, length, CRC, etc., are not shown.
2. Every intermediate node hearing one or more R-request packets for the lifetime of the same source-destination will flood only the first heard request (and if and only if it does not know the route to destination). Each intermediate node will also update its routing table to include the source of the R-request in preparation for the reverse-path setup (Fig. 92.2). By flooding, multicast, or broadcast we mean that the packet is physically heard by all neighbors and processed by all neighbors.

Source	Dest	Source Seq_no	Dest Seq_no	# of Hops	Prev Hop

FIGURE 92.1 A typical route request packet of AODV.

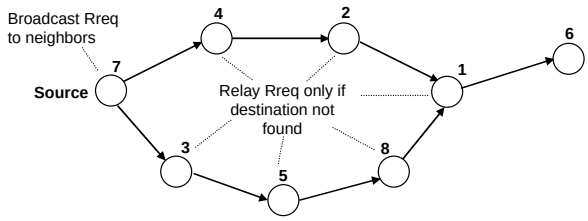


FIGURE 92.2 Route request propagation: reverse path setup.

Reply To	Source	Dest	Dest Seq no	# of Hops	Prev Hop

FIGURE 92.3 A typical route reply message in AODV.

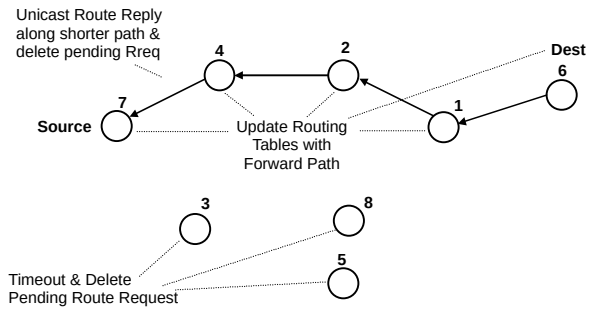


FIGURE 92.4 Route reply propagation: forward path setup.

3. If the node hearing the R-request finds a route to the specified destination in its routing tables or finds itself as the requested destination, it will **unicast** (unicast means that the packet is physically heard by all neighbors but processed only by the intended node) back an R-reply packet to the immediate source through which the R-request was heard; it will not multicast the R-request. Figure 92.3 shows a typical R-reply packet, again not showing the SYNC, CRC, and other fields not related to routing. The sequence number in Fig. 92.2 is the same as that of the corresponding R-request packet.
4. An intermediate node hearing an R-reply packet will repeat it only unicast to the neighboring node through which it has heard the corresponding R-request packet in the first place. It will also insert the destination of the R-reply and the corresponding intermediate node that delivered this R-reply in its routing table. The later information forms the forward path, Fig. 92.4, which is subsequently used for forwarding regular data packets from source to destination.
5. Intermediate nodes hearing an R-request but no corresponding R-reply will time out after a certain time and discard the bending R-request and the corresponding reverse path.
6. Intermediate nodes receiving an R-reply for which no corresponding R-request was received will discard it. However, they will update their routing tables to include the source and destination of the heard R-reply.
7. The source may finally receive many R-reply and, so, many routes to the same destination. The route with smallest number of hops is then selected. However, discovered routes with a more recent sequence number are preferred even if they have more hops as compared to an R-reply with older sequence numbers.
8. Nodes store route information for certain source-destination pairs for a certain lifetime beyond which these routes are eliminated.

Source	Seq_no	List of neighbors

FIGURE 92.5 A typical "hello" message in AODV.

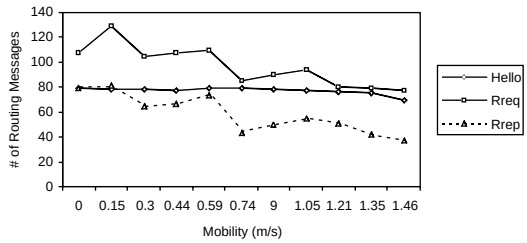


FIGURE 92.6 Routing messages vs. mobility of AODV [Assad, 1998].

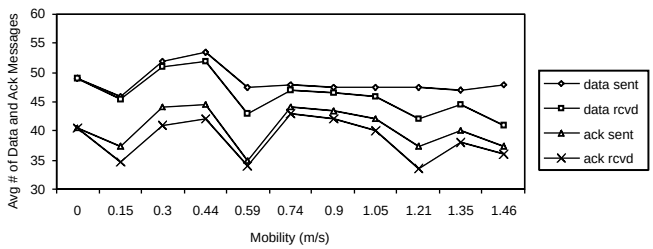


FIGURE 92.7 Data messages vs. mobility of AODV [Assad, 1998].

The Build-up of the Routing Table of AODV

1. Every node periodically sends a hello packet to its neighbors (Fig. 92.5). AODV also has R-request error and other types of packets [Perkins, 1997].
2. All neighbors update their lists of neighbors reacheability set and their routing tables for each hello packet received.
3. Each entry in the reacheability set formed at each node has a lifetime. Upon exceeding this, the entry and all involved routes are erased unless a fresh hello packet is received before expiry of this lifetime. The expiry of these entries is an indication that certain nodes have moved out of reach, in which case link failure may occur and involved source-destination pairs have to be notified.

Performance of AODV

Figures 92.6 and 92.7 show the simulation results in Assad [1998], namely, the control and data performances. The average mobility is the node speed averaged over all nodes and all simulation iterations (time). The number of hello packets does not change much with mobility as compared to numbers of R-request and R-reply packets. Comparing Fig. 92.6 to Fig. 92.7, it seems that the number of control (routing) packets are higher than that of data packets. However, considering the length of control packets finally leads to 1.6% of node time spent on control and 36% spent on data.

Notes

AODV [Perkins, 1997] has other types of packets such as route error (RERR) packets, route reply ACK, and RREP-ACK packets. AODV offers features such as local repair of a route (due to intermediate or end node mobilities). A gratitutous mode, where an intermediate node who knows the route to a certain

destination would unicast this back immediately to the requesting source without further broadcasting the R-request to the destination, also exists in AODV.

92.3 Fisheye Routing (FSR) [Iwata, 1999; Kleinrock, 1971]

The Routing Algorithm

Though the subject node stores routing entries in its table for all possible nodes, it communicates its routing table contents less frequently to more distant nodes. As in Fig. 92.8, the subject node #11 is surrounded by the nodes in the first oval (first scope). Routing information (table of contents) between node #11 and nodes of first scope is exchanged at the highest rate. Routing information is exchanged at a lesser rate between node #11 and member nodes of 2nd scope (2nd oval or second hop) and so on. This way, large latencies (waiting for routes to be exchanged) will occur if the applicable node #11 wishes to communicate with nodes in 2nd or 3rd scopes (hops), but the flooding overhead of link state routing (LS) [Jaquet, 2000] for all nodes (near and far) is eliminated in FSR. This link state LS flooding is caused by the fact that each node has to transmit its routing table contents at the same rate irrespective of distance from the applicable node meaning more routing overhead.

Performance of FSR

The computer simulations in Iwata [1999] are conducted for 5 kinds of routing techniques including FSR. Two scopes (hops) were taken, for a network of 200 nodes. Data packets were generated from a Possion kind of traffic with an inter-arrival period of 2 seconds amounting to a data rate of 4 Kbps per node source. Arriving data packets consist of 10K bits, nodes moved at a speed of 60 km/hour, max radio range is 100 m, and the data rate on channel is 2 Mbps. The routing tables are refreshed at a rate of once every 2 seconds for 1st scope and once every 6 seconds for 2nd scope nodes.

Figures 92.9–92.13 show the routing performance results obtained. Figures 92.9 and 92.10 show that the routing overhead (control) per cluster does not change much with active nodes' numbers or their mobilities (speeds). However, FSR control overhead is relatively higher than the other routing techniques (to be discussed).

Figure 92.11 shows that the average data packet delay does not rise much with nodes' mobilities. FSR performance is the best in this regard compared to other techniques. Similar results are drawn from Fig. 92.12 which shows the average number of hops traveled by a typical data packet.

Figure 92.13 shows that the FSR overhead ratio does not increase as much as other techniques as the network size increases (while keeping the same user density). In comparison, Fig. 92.9 uses the same geographical area while increasing the number of users.

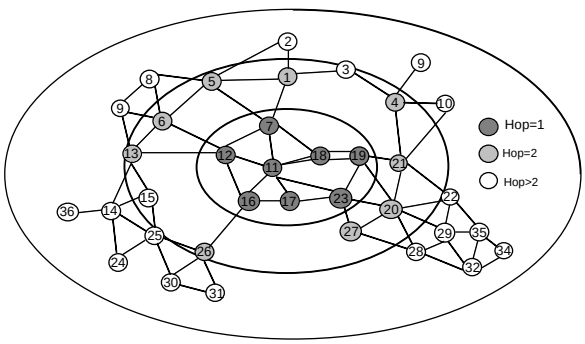


FIGURE 92.8 Scope of fisheye [Iwata, 1999].

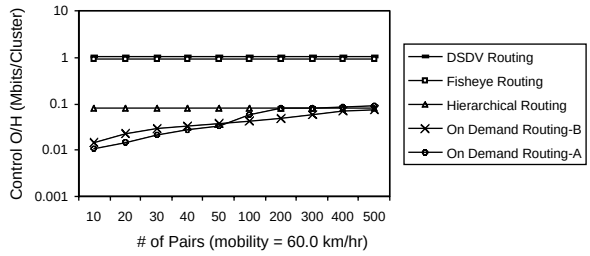


FIGURE 92.9 Control O/H vs. traffic pairs fixed area [Iwata, 1999].

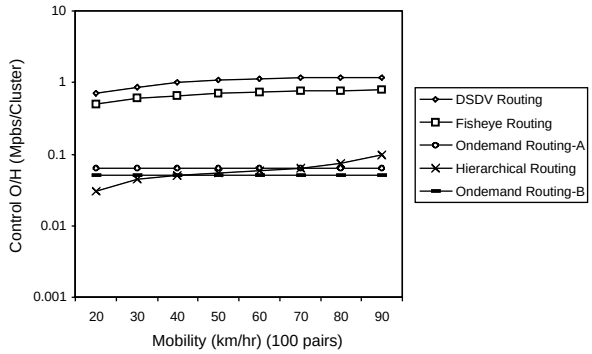


FIGURE 92.10 Control O/H vs. mobility (100 pairs) [Iwata, 1999].

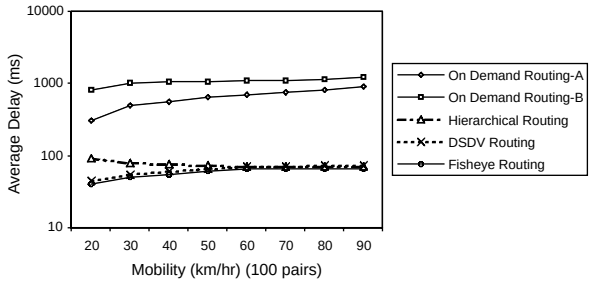


FIGURE 92.11 Average delay vs. mobility (100 pairs) [Iwata, 1999].

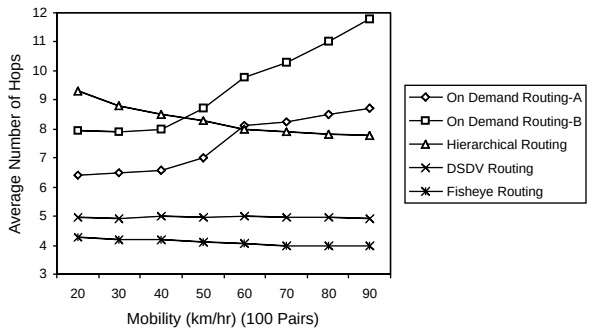


FIGURE 92.12 Average hops vs. mobility (100 pairs) [Iwata, 1999].

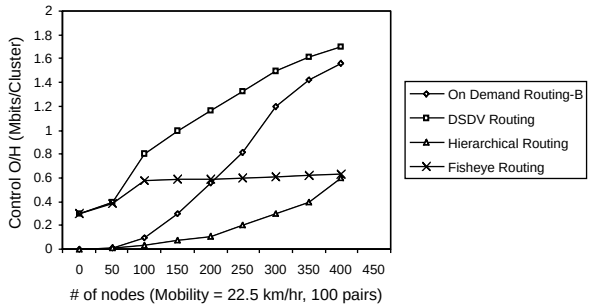


FIGURE 92.13 Control O/H vs. number of nodes (area increases with number of hops) [Iwata, 1999].

Figures 92.9–92.13 also show the performance of simulated on-demand routing techniques such as AODV, DSR, and TORA (to be presented). In types A and B, the on-demand routing tables are updated every 3 and 6 s, respectively.

92.4 Hierarchical State Routing (HSR) [Iwata, 1999]

The Routing Algorithm

HSR combines dynamic, distributed multilevel hierarchical grouping (clustering) and efficient location management. Clustering (dividing nodes into groups and different kinds of nodes) at the MAC and network layers helps the routing of data as the number of nodes grow and, subsequently, the cost of processing (nodes memory and processing required).

HSR keeps a hierarchical topology, where selected cluster heads at the lowest level become members of the next higher level and so on. While this clustering is based on network topology (i.e., physical), further logical partitioning of nodes eases the location management problem in HSR. Figure 92.14 shows four physical level clusters, namely, CO-1, CO-2, CO-3, and CO-4. Cluster heads are selected either manually (by the network administrator) or via the appropriate real time distributed voting mechanism, e.g., [Chiang, 1997 and Gerla, 1995]. In Fig. 92.14, nodes 1, 2, 3, and 4 are assumed to be cluster heads, nodes 5, 9, and 10 are internal nodes, and nodes 6, 7, 8, and 11 are gateway nodes.

Gateway nodes are those belonging to more than one physical cluster. At the physical cluster level, all cluster heads, gateways, and internal nodes use the MAC address, but each also has a hierarchical address as will be discussed.

Cluster heads exchange the link state LS information with other cluster heads via the gateway nodes. For example, in Fig. 92.14, cluster heads 2 and 3 exchange LS information via gateway 8. Cluster heads 4 and 3 via gateway 11 and so on. In the sequel, logical level C1-1 is formed of cluster heads 1, 2, 3, and 4. However, in level C1-1, only 1 and 3 are cluster heads.

Similarly, at level C2-1, only 1 is a cluster head while node 3 is an ordinary node at the C1-2 level. Routing table storage at each node is reduced by the aforementioned hierarchical topology. For example, for node 5 to deliver a packet to node 10, it will forward it to cluster head 1 which has a tunnel (route) to cluster head 3, which finally delivers the packet to node 10. But how would cluster head 1 know that node 10 is a member of a cluster headed by 3? The answer is that nodes in each cluster exchange virtual LS information about the cluster (who is the head, who are the members) and lower cluster (with less details), and the process repeats at lower clusters. Each virtual node floods this LS information down to nodes within its lower level cluster. Consequently, each physical node would have hierarchical topology information (actually, summary of topology including cluster heads and member nodes) rather than the

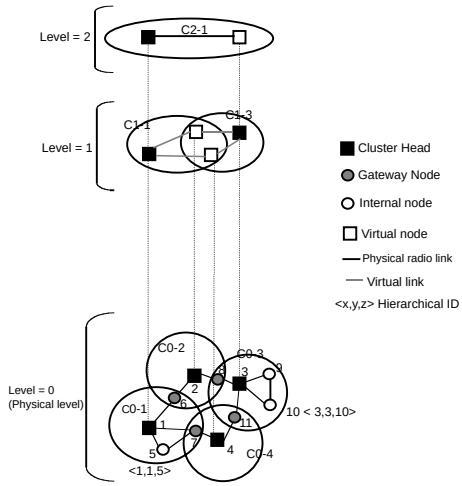


FIGURE 92.14 An example of physical/virtual clustering [Iwata, 1999].

full topology of flat routing where all individual routes to each node in the networks are stored at all nodes and are exchanged at the same rate to all nodes.

The hierarchical address (a sequence of MAC addresses) is sufficient for packet delivery to any node in the network. For example, in Fig. 92.13, node HID (5) = (1, 1, 5) going from the top to the lowest cluster, 1 is the cluster head of clusters C1-1 and CO-1, and node 5 is an interior node of CO-1. Similarly, HID (6) = (3, 2, 6) and HID (10) = (3, 3, 10). Returning back to the example above where node 5 seeks a route to node 10, it will ask 1 (its cluster head). Node 1 has a virtual link or tunnel, i.e., the succession of nodes (1, 6, 2, 8, 3) to node 3 which is the cluster head of the final destination. This tunnel is computed from the LS information flooded down from higher cluster heads, as above. Finally, node 3 delivers the packet to node 10 along the downward hierarchical path which is a mere single hop. Gateways have multiple HID since they belong to more than one physical cluster.

Performance of HSR

Figures 92.9–92.13 show the performance of HSR. The refresh rate of routing tables is 2 seconds. Figure 92.9 reveals that the overheard ratio in HSR is a relatively constant (w.r.t) number of communicating pairs. Also, it is better than that of FSR or DSVD. Figure 92.9 shows that the HSR overhead ratio is better than FSR and DSDV as the nodes increase their speed. Figure 92.11 also shows better delay for HSR, but the average number of hops in Fig. 92.12 does not reveal marked improvement compared to other routing techniques. Finally, Fig. 92.13 shows a superior overhead ratio performance of HSR, as the number of communicating nodes rises (while keeping the user density the same). Figures 92.9–92.13 also show the performance of simulated on-demand routing techniques such as AODV, DSR, and TORA. In types A and B, the on-demand routing tables are updated every 3 and 6 s, respectively.

Drawbacks

Utilization of long hierarchical addresses and the cost of continuously updating the cluster hierarchical and hierarchical addresses as nodes move (memory and processing of nodes) are the shortcomings of HSR. Also, clustering and voting for cluster heads may consume more overhead not to mention creating processing, security, and reliability hazards at cluster heads at different hierarchical levels.

92.5 Destination-Sequenced Distance-Vector Routing (DSDV) [Hedrick, 1988; Perkins, 1994; Iwata, 1999]

Description

DSDV is based on the routing information protocol (RIP) [Hedrick, 1988] used within the Internet. Each node keeps a routing table containing routes (next hop nodes) to all possible destination nodes in the wireless network. Data packets are routed by consulting this routing table. Each entry in the table (Fig. 92.15) contains a destination node address, the address of the next hop node enroute to this destination, the number of hops (metric) between destination and next hop node, and the lifetime of this entry (called the sequence number). Each node must periodically transmit its entire routing table to its neighbors using update packets.

Build-up and update of routing tables is done via the update packets to be described shortly. Figure 92.15 shows how a typical data packet is forwarded from node 1 to final destination node 3. Node 1 consults its routing table and finds node 4 as the immediate node to reach destination node 3. Node 1 then forwards the data packet (Fig. 92.16a) with the destination field equal to 3 and the next hop field set to 4. All nodes in the neighborhood of node 1 hear this data packet (recall the physical broadcast nature of wireless channels), but only node 4 processes this packet. Node 4 consults its routing table (Fig. 92.16b) and finds the next hop node 5 corresponding to destination field set to 3. Node 4 builds a new packet with the destination field set to 3 and the next hop field set to 5 and transmits this packet. This packet is similarly processed upon its reception at node 5 and so on until the packet reaches the final destination.

Routing tables, as such, are updated at a slower rate within the terrestrial based Internet. However, in the case of the mobile wireless LANs and due to nodes mobility, the routing table's update has to occur at faster rate. A great deal of DSDV processing and bandwidth (overhead) is consumed by table management (generation and maintenance). Routing tables have to react faster to topology changes or routing loops may result. The route update packets (Fig. 92.15) contain a destination field which has either the address of the 1st node (called destination) originating the route update or the address of one of its neighbors. The metric field is set to 1 by the 1st node originating the route update. This field denotes the number of wireless hops to the declared destination. This field is incremented by 1 by each node receiving this update and repeating it to its neighbors. The next hop field has the address of the node

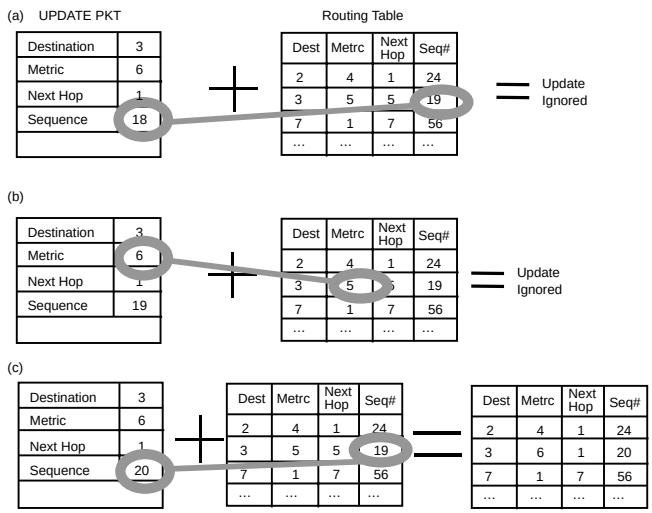


FIGURE 92.15 A node receiving three update packets.

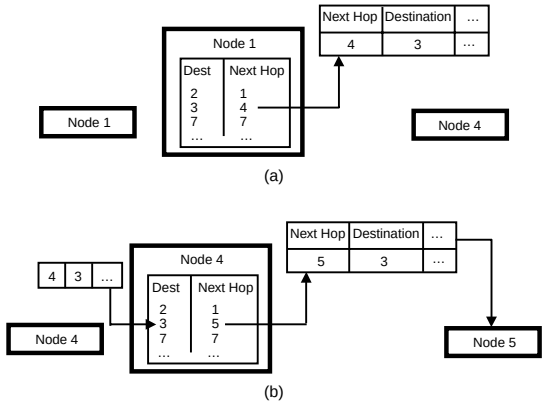


FIGURE 92.16 Routing example in DSDV: (a) node 1 transmits packet to node 4 for forwarding; (b) node 4 looks up the destination in its routing table.

which has just retransmitted the route update packet. The sequence number is set by the originating node. As the node moves, it sends fresh route update packets with higher sequence numbers. Nodes hearing these will update their tables according to the information in the newer packets from the same originator (destination), i.e., from packets having higher sequence numbers (for same destination). Figure 92.15 shows three scenarios corresponding to a certain intermediate node receiving three different update packets from neighboring nodes. In Fig. 92.15a, the received sequence (18) is less than what the subject node has already in routing table (19) to the same destination (3), i.e., it is older and so is discarded.

In Fig. 92.15b, the sequence number is the same in both of the route update packets and the routing table, but the metric is higher in the route update packet. In this case, the route update packet is discarded (meaning that this update packet looped too much and so the smaller metric already existing in the table is preferred).

In Fig. 92.15c, the received route update packet has a sequence number (20) which is higher than the existing value in the routing table (for the same destination). In this case, the received route update packet replaces the existing entry in the routing table even if the metric in the table is smaller than that in the received route update packet. Higher sequence numbers mean fresh information and are preferred even if this new routing information has a shorter path to its destination, i.e., a smaller metric.

Performance of DSDV

This is shown in Figs. 92.9–92.13. Figure 92.9 shows that the DSDV overhead performance is worse than FSR, HSR, etc. Figure 92.10 shows that the DSDV overhead performance vs. speed of nodes is prohibitive. Figures 92.9 and 92.10 also show that DSDV performance is somewhat independent of the nodes traffic conditions; this is due to the fact that all nodes are busy with the transmission of their routing tables. DSDV exhibits good delay and an average number of hops (Figs. 92.11 and 92.12). However, for highly deployed networks where the number of nodes increases as the area enlarges, the overhead ratio is much worse than FSR, HSR, etc. (Fig. 92.13). Figures 92.9–92.13 also show the performance of simulated on-demand routing techniques such as AODV, DSR, and TORA. In types A and B, the on-demand routing tables are updated every 3 and 6 s, respectively.

Comments

DSDV is a flat routing technique that propagates 100% of the routing table contents of all nodes and with the same frequency for all table entries. The frequency of transmitting the routing tables becomes higher as node speeds increase. Perkins [1994] contains procedures for handling different network layer addresses, damping routing oscillations due to topology changes, etc. For example, broken links (a node

does not hear from a neighbor for some time) are repaired by transmitting a new route update packet with a larger sequence number and a metric of ∞.

92.6 Zone Routing Protocol (ZRP) [Pearlman, 1999]

Description

This is a hybrid reactive/proactive routing technique. Routing is flat and not hierarchical (as in HSR) thus leading to reducing the processing and overhead. Nodes keep routing tables of a node's neighborhood and not of the whole network.

Query/reply messaging is needed if the destination is not found in the routing table. The query/reply process is handled only by certain selected nodes (called border nodes of the zone). On the other hand, all interior nodes repeat (physical broadcast) the query/reply packet but do not process it. The border nodes are not gateways nor cluster heads such as the case in HSR. They are plain nodes located at the borders of the routing zone of the applicable node. The routing zone of a certain node of radius $r > 1$ is defined as the collection of these nodes which can be reached via 1, 2, or 4, ..,r hops from the applicable node. Nodes within one hop from the center node (e.g., node S in Fig. 92.17) are those that can directly hear the radio transmission of node S. These are nodes A, B, C, D, and E (one-hop nodes). The higher the power of a typical node S, the larger the number of one hop nodes. This may lead to increasing the control traffic necessary for these nodes to exchange their routing tables. Also, this increases the level of contention of the IEEE 802.11 CSMA based access [Gier, 1999]. The routing zone in Fig. 92.17 corresponds to a routing zone of two hops ($r = 2$). Each node in this routing zone exchanges routing packets only with members in its routing zone (nodes A–K in Fig. 92.17). This takes place according to the Intrazone protocol (IARP) which forms the proactive part of the ZRP routing protocol.

Nodes G–K in Fig. 92.17 are called the peripheral nodes of node S. These again are ordinary nodes but just happened to be at 2 hops, i.e., at the border of the routing zone of node S ($r = 2$, zone radius of routing zone of Fig. 92.17). For a different node, the peripheral nodes will change. The IARP protocol provides nodes with routing tables with a limited number of entries corresponding to nodes in the same routing zone (nodes A–K in Fig. 92.17). If the destination lies outside the routing zone such as node L in Fig. 92.17, the calling node (say node S) will issue a query to search for a route to node L. This query packet will be retransmitted by nodes A–E to the peripheral nodes. These are the only nodes that will process and respond to the route query packet.

This process is called bordercasting and the underlying protocol is called interzone routing protocol (IERP). The identities, distances, and number of hops of the peripheral nodes of the applicable node as well as all other nodes (interior nodes) in its routing zone are assumed to be known to each node in the wireless Internet. Query packets are sent unicast or broadcast to all peripheral nodes if the applicable node does not find the destination in its routing tables. Further, if these peripheral nodes do not find the required destination in their zone, they would forward the query to their own peripheral nodes and so on. If one of the peripheral nodes know the route to destination, it will return a route-reply packet to the node that sent the query. In Fig. 92.18 node S does not find the destination node D in its routing zone.

S bordercasts a query packet to nodes C, G, and H. Each of these retransmits this query packet to its peripheral nodes after finding that D is not in their routing zones. The process repeats until node B finds the destination node D in its routing zone. Node B returns a reply containing the ID sequence S–H–B–D.

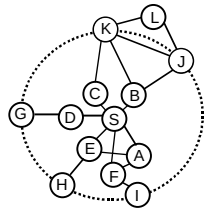


FIGURE 92.17 A routing zone of two hops radius [Pearlman, 1999].

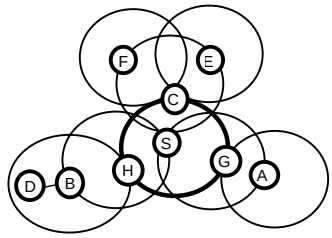


FIGURE 92.18 An example of IERP operation [Pearlman, 1999].

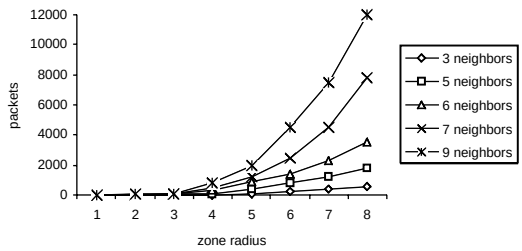


FIGURE 92.19 IARP traffic generated per neighbor [Pearlman, 1999].

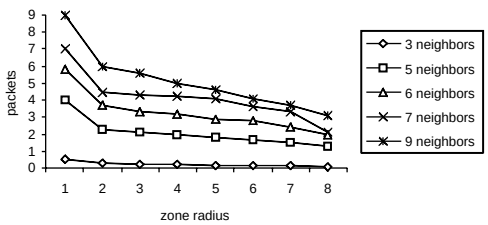


FIGURE 92.20 IERP traffic received by each node per query [Pearlman, 1999].

In this route accumulation routine process, each node adds its ID to the query packets then transmits it to its peripheral nodes. The destination node uses the reversed accumulated sequence of node IDs to send a route reply back to the source node.

The accumulated ID sequence is likened to source routing techniques in terrestrial networks. To alleviate longer route reply packets due to this accumulation, intermediate peripheral nodes may store temporary short routing tables. These contain the IDs of the previous peripheral nodes who have sent the query. The intermediate peripheral nodes will also overwrite the ID of the previous peripheral node from which they have received the query before retransmitting the query packet rather than appending its ID to the query packet. Once this intermediate node receives a route reply, it will send it to the previous peripheral node whose ID is stored in the short temporary table. Needless to say, even with bordercasting, there is still some flooding of the route query (but less than pure flooding) and the source node of the query may select the route with minimum number of hops.

Performance of ZRP

Figures 92.19–92.23 show some of the simulation results in Pearlman [1999]. In all of these, v denotes the speed of the node expressed in neighbor acquisition per second, d is the distance from transmitter to call destination, and Rrout_failure is the rate of route failure per second. Figure 92.19 shows the volume of traffic in packets generated by a neighbor node vs. the zone radius r due to the proactive IARP traffic which increases with the number of neighboring nodes. Figure 92.20 shows a similar result for the IERP

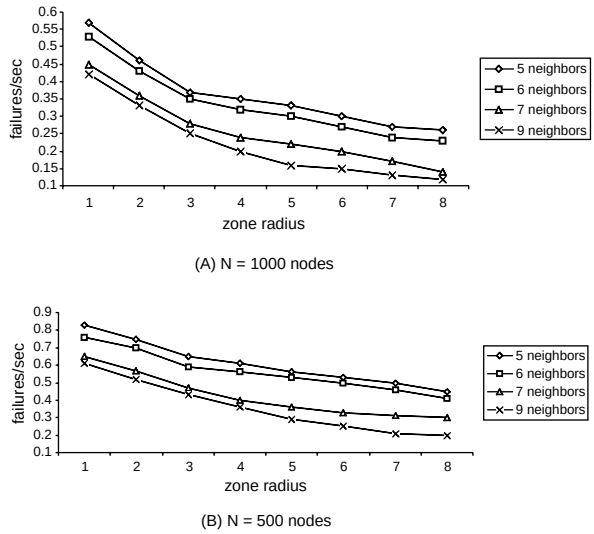


FIGURE 92.21 Route failure rate traffic (normalized to node velocity) [Pearlman, 1999].

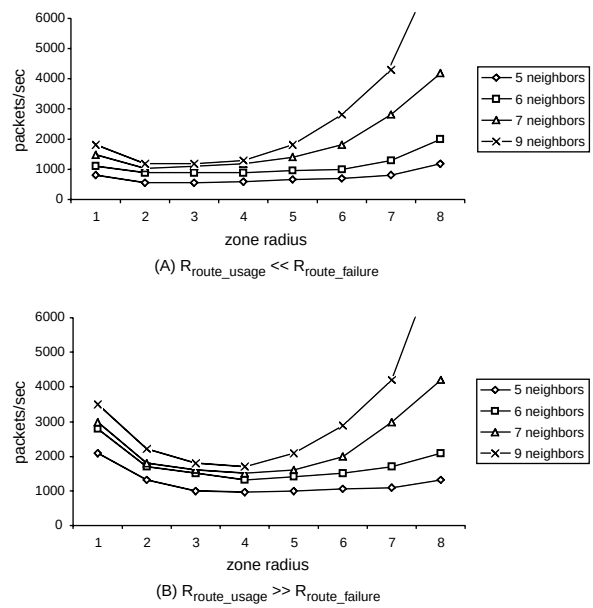


FIGURE 92.22 ZRP traffic per node (N = 1000 nodes, v = 0.5 neighbor/s) [Pearlman, 1999].

traffic volumes (due to route query/reply traffic). However, the IERP traffic decreases with the zone radius as opposed to Fig. 92.19, thus reflecting one of the trade-offs in ZRP and other ad-hoc routing techniques (i.e., proactivity vs. reactivity). Figure 92.21 shows the route failure rate, which decreases with r. The route failure is caused by node mobilities and topology changes. Failure rate increases slightly for higher numbers of nodes. Figure 92.22 shows the ZRP traffic per node vs. r. Optimal values for r are possible, as shown (yielding minimum traffic in packets/s). The dependence of optimal values of zone radius r on nodes speeds is shown in Fig. 92.23.

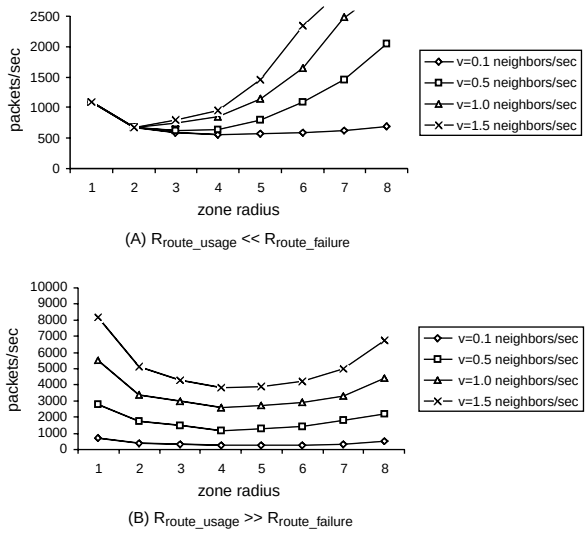


FIGURE 92.23 ZRP traffic per node (N = 1000 nodes, d = 6.0 neighbors) [Pearlman, 1999].

Comments

ZRP is a good example of an ad-hoc routing technique that tries to strike a balance between proactivity and reactivity. The MAC layer neighbor discovery control overhead (which may not be belittled) is called association traffic in IEEE 802.11 [Gier, 1999]. This overhead, being related to the MAC layer, was not accounted for in all of the results above. Bordercasting is seen to decrease the volumes of IARP traffic. Pearlman [1999] also presents excellent ways of adapting/adjusting the zone radius depending on the ratio of reactive IARP traffic to proactive IARP traffic volumes. However, it was noted [Pearlman, 1999] that accuracy of the optimal zone radius is a very complex function of underlying wireless Internet and its too many parameters (mobilities, user density, traffic volumes and types, etc.).

92.7 Dynamic Source Routing (DSR) [Broch, 1998; Maltz, 1999]

Description

In source routing techniques (wireless and wire based), each data packet contains the complete route identification (list of nodes addresses, IDs, or route paths) from source to destination node (Fig. 92.24) DSR does not require periodic route advertisement or link status packets, implying less routing overhead. An intermediate node hearing a data packet, and finding its ID within, preceded by the ID of the node that transmitted the packet will retransmit this data packet. Otherwise this intermediate node discards the data packet. If the intermediate node is listed as the final destination of the data packet, the packet will not be retransmitted. This is seen as an example of physical multicast (broadcast), due to the nature of the wireless channel, but logical unicast since the lonely node meant by this data packet repeats it. Each node has to find and store an end-to-end route to each destination it wishes to communicate with.

Route discovery is the second axiom of DSR, where the source node broadcasts a route request whenever it places a call to this destination. Each node hearing this route request and not finding itself the destination of this request nor finding its address within the route list of the packet (series of intermediate node addresses) will append its address (ID) to the route list then multicasts this packet (physically and logically). The route request hops this way until the destination is reached or one intermediate node has the information regarding the route to destination. In both cases, the route request is

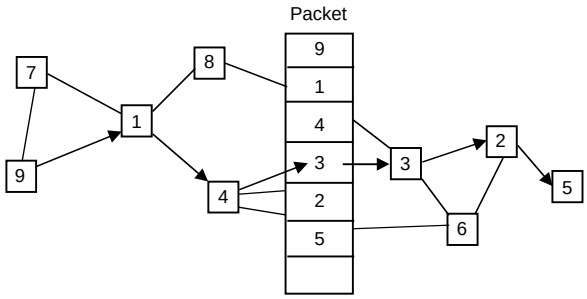


FIGURE 92.24 Path based routing from node 9 to node 5 in DSR.

not retransmitted but a route-reply packet is sent back to the source node along the same (but reversed) path of the route list of the route-request packet which the subject destination node has received. Each intermediate node multicasts a certain route request from a certain source to a certain destination only once and discards similar route requests.

All nodes maintain complete routing information (route path) to the destination they recently communicated with as well as routes for which they have handled the route requests. Similar to wire based networks, the processes above may generate multiple route paths, but the destination node has the option of returning only the first discovered route. Other options include the returning of all possible routes or the best route (which has the minimum number of hops) to the source node.

The third axiom of DSR is route maintenance. Here, a node receiving a data or route request type packet and detecting a route error (e.g., an intermediate node moved and can no longer handle certain routes) will send a route error packet to the source. All intermediate nodes hearing a route error packet will remove from their routing tables the broken link for all involved routes (of other source-destination pairs) which use this broken link. All such routes will then become truncated at this point.

In networks with full duplex connections, route error packets follow the reverse path back to the source. If only half duplex connections are available, Broch [1998] and Maltz [1999] provide other mechanisms for returning route error packets back to the source node.

Route maintenance packets are acknowledged either end-to-end (via the TCP transport layer software, for example) or on a hop-by-hop basis (via the MAC layer).

Performance of DSR

A network with 50 nodes was simulated [Maltz, 1999]. The nodes moved at variable speeds from 0 to 20 m/s. Figure 92.25 shows the packet delivery ratio and routing overhead vs. node mobility (900 in Fig. 92.25 means a stationary node and 0 means a continuously moving node).

Routing overhead include all kinds of route-request, route-reply, etc. Packet delivery ratio is the total number of packets delivered to all destinations during simulations divided by total number of packets actually generated.

Table 92.1 shows the latency of first route-reply (time for route discovery). "Neighbor" refers to all nodes which are at one hop from the subject node. "Cache replies" refers to all route replies returned back by an intermediate rather than the destination node. "Target reply" is the route returned by the destination node. Table 92.2 summarizes the route discovery costs. Non-propagated route requests are defined as those requests forwarded only to neighbors and not transmitted further. Propagated route requests are forwarded to all nodes and are transmitted hop-by-hop until the destination.

"Request Og" is the number of route request originators. "Fw Rep" is the number of replies forwarded and so on. Containment is defined as the percentage of nodes that are not neighbors of the source node. The cost is defined as:

$$(\Sigma \text{Og Req} + \Sigma \text{Fw Req} + \Sigma \text{Og Rep} + \Sigma \text{Fw Rep})/\text{Og Req}$$

TABLE 92.1 Latency of First Route Reply by Type (All Optimal DSR, Rectangular Site, Constant Motion) [Maltz, 1999]

		Latency			
	#Replies	Mean	Min	99 Percentile	Max
Neighbor Replies	3136	7.1 ms	1.3 ms	457 ms	1.78 s
Cache Replies	1524	45.9 ms	1.3 ms	752 ms	2.4 s
Target Replies	12	87.6 ms	23.6 ms	458 ms	458.3 s

TABLE 92.2 Summary of Route Discovery Costs (All Optimal DSR, Rectangular Site, Constant Motion) [Maltz, 1999]

	Nonpropagating	Propagating	Total
Request Og	957	316	1273
Request Fw	0	6115	6115
Reply Og	3912	3215	7127
Reply Fw	0	7002	7002
Containment	77.0%	41.0%	68.0%
Cost	5.06	52.69	16.90

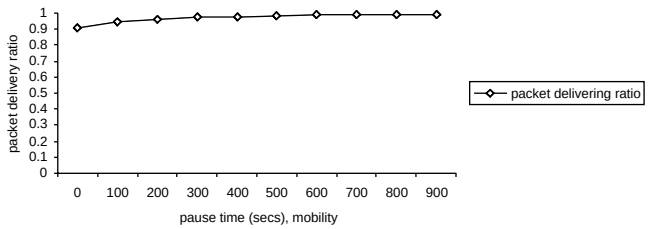


FIGURE 92.25a Baseline DSR packet delivery ratio (all-optimal DSR, rectangular site) [Maltz, 1999].

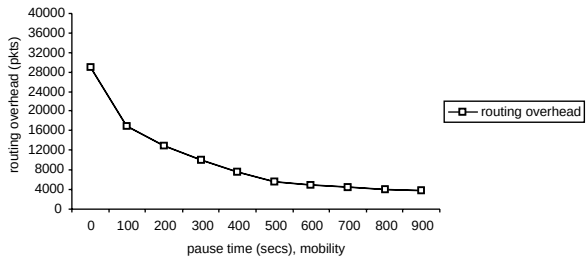


FIGURE 92.25b Baseline DSR routing packet overhead (all-optimal DSR, rectangular site) [Maltz, 1999].

Comments

Intelligent caching techniques [Broch, 1998; Maltz, 1999] can be used to reduce the route discovery and route maintenance overheads at the expense of memory and CPU times of nodes. Yet the greatest overhead loss of DSR is attributed to the long route path in each packet. Compression techniques may be used to cut such overhead, with the penalty of more frequent occurrence of long bursts of bit errors. The security of DSR is also inferior compared to other non-source coding routing techniques.

92.8 Summary

A variety of ad-hoc routing techniques were presented. The ideal or standard ad-hoc routing technique does not exist yet. Extensive research activities are still going on in this important area. Pure proactive or pure reactive routing did not provide the efficiency nor the dynamicity required in interconnected wireless LANs, so most of the recent works are concentrated around hybrid techniques (partly proactive/partly reactive). Establishing the best mix of proactive and reactive techniques is a complex function of the traffic amounts and types, the channel, required QoS, etc.

The difficulty of standardizing the routing performance criteria over the various routing techniques is one of the main obstacles standing in the way of adapting a standard for ad-hoc wireless routing. The processing speeds and memory requirement of the real-time operation of routing candidate remains to be evaluated. It is not uncommon to discard good routing techniques that are too complex for the small battery power, small processors, and memories these handheld devices have. As a matter of fact, for these devices, routing is just one of the many functions the portable node is supposed to handle.

Defining Terms

Access point (AP) mode: Centralized mode of operation of wireless LANs where the AP resembles the functionalities of the cellular system BS.

Ad-hoc routing: Routing techniques amenable to wireless LANs with no predefined borders, no base stations, and with all nodes performing the same distributed routing algorithms based on peer-to-peer communications.

Hierarchical routing: Nodes are sorted into logical groupings; certain nodes assume the functionality of cluster heads and gateways so as to reduce the routing overhead.

Media access control (MAC): Set of techniques and standards for controlling the right to transmit on a certain shared channel such as TDMA, FDMA, CSMA, etc.

On-demand routing: Similar to reactive routing, e.g., AODV [Perkins, 1997; Assad, 1998] and DSR [Broch, 1998].

Proactive (flat) routing: Where modes build and update complete routing tables with entries corresponding to all nodes in the wireless Internet.

Reactive routing: Where nodes issue route requests and receive route replies as the need arises, with no build-up of routing tables nor forwarding data basis.

Unicast: The packet is physically heard by all neighbors but processed by only one intended node and discarded by the other nodes.

References

Assad, S., 1998, http://www.ctr.columbia.edu.

Bellur, B., Ogier, R.G., and Templin, F.L., 2000, Topology broadcast based on reverse-path forwarding (TBRPF), ITEF, Internet draft, draft-itef-manet-tbrpf-00.txt.

Broch, J., Johnson, D.B., and Maltz, D.A., 1998, The dynamic source routing protocol for mobile ad-hoc networks, Internet draft, draft-itef-manet-dsr-01.txt.

Chiang, C.C., Wu, H.K., Liu, W., and Gerla, M., 1997, Routing in clustered multihop, mobile, wireless networks, Proceedings of IEEE Singapore International Conference on Networks, 197.

Elhakeem, A.K., Ali, S.M., Aquil, F., Li, Z., and Zeidi, S.R.A., 2000, New forwarding data basis and ad-hoc routing techniques for nested clusters of wireless LANs, submitted to *Wireless Comm. J.*

Gerla, M. and Tsai, J., 1995, Multiuser, mobile, multimedia radio network, *ACM Baltzer J. Wireless Networks,* 1(3), 255.

Gier, J., 1999, *Wireless LANs, Implementing Interoperable Networks,* Macmillan Technical Publishing, New York.

Hedrick, C., 1988, Routing information protocol, RFC1058, June.

Iwata, A., Chiang, C.C., Pei, G., Gerla, M., and Chan, T.W., 1999, Scalable routing strategies for ad-hoc wireless networks, *IEEE J. Sel. Areas Comm.,* 17(8), 1369.

Jaquet, P., Muhlethaler, M., and Qayyum, A., 2000, Optimized link state routing protocol, ITEF, Internet draft, draft-itef-manet-olsr-01.txt.

Ji, L. and Corson, M.S., 2000, Differential destination multicast (DDM) specification, ITEF, Internet draft, draft-itef-manet-ddm-00.txt.

Kasera, K. K. and Ramanathan, R., 1997, A location management protocol for hierarchically organized multihop mobile wireless networks, Proceedings of IEEE ICUPC'97, 158.

Kleinrock, L. and Stevens, K., 1971, Fisheye: a lenslike computer display transformation, Tech. Report, University of California, Los Angeles, CA.

Lee, S.J., Su, W., and Gerla, M., 2000, On-demand multicast routing protocol (ODMRP) for ad-hoc networks, ITEF, Internet-draft, draft-itef-manet-odmrp-02.txt.

Lin, C.R. and Lui, J.S., 1999, QoS routing in ad-hoc wireless networks, *IEEE J. Sel. Areas Comm.,* 17(8), 1426.

Maltz, D.A, Broch, J., Jetcheva, J., and Johnson, D., 1999, The efforts of on-demand behavior in routing protocols for multihop wireless ad-hoc networks, *IEEE J. Sel. Areas Comm.,* 17(8), 1439.

McDonald, A.B. and Znati, T.F., 1999, A mobilty-based framework for adaptive clustering in wireless ad-hoc networks, *IEEE J. Sel. Areas Comm.,* 17(8), 1466.

Ng, M.J. and Lu, I.T., 1999, A peer-to-peer zone-based two-level link state routing for mobile ad-hoc networks, *IEEE J. Sel. Areas Comm.,* 17(8), 1415.

Park, V.D. and Corson, M.S., 1997, A highly adaptive distributed routing algorithm for mobile wireless networks, Proceedings of INFOCOM '97, 1405.

Pearlman, M.R. and Haas, Z.J., 1999, Determining the optimal configuration for the zone routing protocol, *IEEE J. Sel. Areas Comm.,* 17(8), Aug.

Perkins, C. E., 1997, Ad-hoc demand distance vector routing, ITEF, Internet draft, draft-itef-manet-adov-00.txt.

Perkins, C.E. and Bhagwat, P., 1994, Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers, Proceedings of the SIGCOM'94 Conference on Communication Architectures, Protocols and Applications, Aug., 234.

Sivakumar, R., Sinha, P., and Bharghavan, V., 1999, CEDAR: a core-extraction distributed ad-hoc routing algorithm, *IEEE J. Sel. Areas Comm.,* 17(8), 1454.

Wu, C.W., Tay, Y.C., and Toh, C.K., 2000, Ad-hoc multicast routing protocol utilizing increasing id-numbers (AMRIS) functional specifications, ITEF, Internet draft, draft-itef-manet-amris-spec-00.txt.

Further Information

A core extraction routing algorithm, CEDAR, is presented in Sivakumar [1999]. Both description and QoS performance are presented. GPS facilities are used in Ng [1999] to improve the routing performance of ad-hoc networks. Performance of a new clustering technique that improves the availability of links and leads to route stability is presented in McDonald [1999]. An adaptive routing technique for large and dense wireless networks (TORA) is presented in Park [1999], while Lin [1999] presents a QoS routing for asynchronous transfer mode (ATM) based ad-hoc networks. Miscellaneous Internet drafts such as TBRPF [Belhur, 2000], DDM [Ji, 2000], ODMRP [Lee, 2000], OLSR [Jaquet, 2000], and AMRIS [Wu, 2000] provide a multitude of new algorithms for routing in ad-hoc wireless LANs.

VII

Source Compression

93

Lossless Compression

Khalid Sayood
University of Nebraska–Lincoln

Nasir D. Memon
Polytechnic University

93.1 Introduction

Lossless compression or reversible compression, as the name implies, denotes compression approaches in which the decompressed or reconstructed data exactly match the original. Theoretically, the lossless compression rate in terms of bits per symbol is bounded below by the entropy of the source. For a general source $\mathcal{S}$ with alphabet A = {1, 2,…,m}, which generates a sequence $\{X_1, X_2,\ldots\}$, the entropy is given by

$$H(\mathcal{S}) = \lim_{n\to\infty} \frac{1}{n} G_n \tag{93.1}$$

where

$$G_n = -\sum_{i_1=1}^{i_1=m} \sum_{i_2=1}^{i_2=m} \cdots \sum_{i_n=1}^{i_n=m} P(X_1 = i_1, X_2 = i_2, \ldots, X_n = i_n) \times \log P(X_1 = i_1, X_2 = i_2, \ldots, X_n = i_n) \tag{93.2}$$

and $\{X_1, X_2,\ldots,X_n\}$ is a sequence of length n from the source. The reason for the limit is to capture any structure that may exist in the source output. In the case where the source puts out independent, identically distributed (i.i.d.) symbols, Eq. (93.2) collapses to

$$G_n = n \sum_{i_1=1}^{i_n=m} P(X_1 = i_1) \log P(X_1 = i_1) \tag{93.3}$$

and the entropy of the source is given by

$$H(S) = \sum_{i=1}^{m} P(X = i) \log P(X = i) \tag{93.4}$$

Although taking longer and longer blocks of symbols is certainly a valid method for extracting the structure in the data, it is generally not a feasible approach. Consider an alphabet of size m. If we encoded the output of this source in blocks of n, we would effectively be dealing with an alphabet of size m^n. Besides being impractical, if we assume the source generates strings of finite length, then it has been shown that extending the alphabet by taking longer and longer blocks of symbols turns out to be inherently inefficient.

A more effective strategy is to use a model to capture the inherent structure in the source. This model can be used in a number of different ways. One approach is to use the model to generate a *residual* sequence that is the difference between the actual source output, and the model **predictions**. If the model accurately reflects the structure in the source output, the residual sequence can be considered to be i.i.d. Often, a second stage model is used to further extract any structure that may remain in the residual sequence. The second stage modeling is often referred to as **error modeling**. Once we get (or assume that we have) an i.i.d. sequence, we can use entropy coding to obtain a coding rate close to the entropy as defined by Eq. (93.4).

Another approach is to use the model to provide a **context** for the encoding of the source output and encode sequences by using the statistics provided by the model. The sequence is encoded symbol by symbol. At each step, the model provides a probability distribution for the next symbol to the encoder, based on which the encoding of the next symbol is performed.

These approaches separate the task of lossless compression into a modeling task and a coding task. As we shall see in the next section, encoding schemes for i.i.d. sequences are known that perform optimally and, hence, the critical task in lossless compression is that of modeling. The model that is imposed on the source determines the rate at which we would be able to encode a sequence emitted by the source. Naturally, the model is highly dependent on the type of data being compressed. Later in this chapter we describe some popular modeling schemes for text and image data.

This chapter is organized as follows. We first describe two popular techniques for encoding the residuals which assume knowledge of the probabilities of the symbols, followed by adaptive versions of these techniques. We then describe **universal coding** techniques, which do not require any *a priori* knowledge of the statistics of the source. Finally, we look at two of the three most popular areas for lossless compression, text, and images. The third area, compression of facsimile, is covered in the next chapter.

93.2 Entropy Coders

The idea behind **entropy coding** is very simple: use shorter codes for more frequently occurring symbols (or sets of symbols). This idea has been around for a long time and was used by Samuel Morse in the development of Morse code. As the codes generated are of variable length, it is essential that a sequence of codewords be decoded to a unique sequence of symbols. One way of guaranteeing this is to make sure that no codeword is a prefix of another code. This is called the prefix condition, and codes that satisfy this condition are called prefix codes. The prefix condition, while sufficient, is not necessary for unique decoding. However, it can be shown that given any uniquely decodable code that is not a prefix code, we can always find a prefix code that performs at least as well in terms of compression. Because prefix codes are also easier to decode, most of the work on lossless coding has dealt with prefix codes.

Huffman Codes

The **Huffman coding** algorithm was developed by David A. Huffman as part of a class assignment [Huffman, 1952]. The algorithm is based on two observations about optimum prefix codes:

1. In an optimum code, symbols that occur more frequently will have shorter codewords than symbols that occur less frequently.
2. In an optimum code, the two symbols that occur least frequently will have the same length.

The Huffman procedure is obtained by adding the simple requirement that the codewords corresponding to the two lowest probability symbols differ only in the last bit. That is, if γ and δ are the two least probable symbols in an alphabet, then, if the codeword for γ was $m * 1$, then the codeword for δ would be $m * 0$. Here m is a string of 1s and 0s and $*$ denotes concatenation.

First, the letters of the alphabet are sorted in order of decreasing probability. A new alphabet is generated by combining the two lowest probability letters into a single letter whose probability is the sum of the two probabilities. When this composite letter is decomposed, its constituent letters will have codewords, which are identical except in the final bit. The process of sorting and then combining the two letters with the lowest probabilities to generate a new alphabet is continued until the new alphabet contains only two letters. We assign a codeword of 0 to one and 1 to the other, and we then proceed to decompose the letters. At each step of the decomposition we will get a bit each of two codewords. At the end of the process we will have a prefix code for the entire alphabet. An example of this process is shown in Tables 93.1 and 93.2, where $C(a_k)$ is the codeword for a_k.

The rate for the Huffman code in bits per symbol can be shown to lie in the interval $[H(\mathcal{S}), H(\mathcal{S}) + p_{\max} + 0.086]$, where $p_{\max}$ is the probability of the most probable symbol [Gallagher, 1978]. The lower bound is achieved when the probabilities of the source symbols are all powers of two. If, instead of coding one symbol at a time, we block n symbols together, the bounds on the coding rate are

$$H(\mathcal{S}) \le R_H \le H(\mathcal{S}) + \frac{P_{\max}^n + 0.086}{n} \tag{93.5}$$

Thus, we can get the coding rate arbitrarily close to the **entropy**. Unfortunately, this approach also means an exponential growth in the size of the source alphabet. A technique that avoids this problem of exponential growth with block length is **arithmetic coding**, which is described in the next section.

TABLE 93.1 Composition Process for Huffman Coding

Alphabet (Prob.)	Sorted (Prob.)	Composite (Prob.)	Composite (Prob.)	Sorted (Prob.)	Composite (Prob.)	Sorted (Prob.)
$a_1(0.2)$	$a_2(0.4)$	$a_2(0.4)$	$a_2(0.4)$	$a_2(0.4)$	$a_2(0.4)$	$a_3''(0.6)$
$a_2(0.4)$	$a_1(0.2)$	$a_1(0.2)$	$a_1(0.2)$	$a_3'(0.4)$	$a_3''(0.6)$	$a_2(0.4)$
					$a_3'' \Leftarrow a_3', a_1$	
$a_3(0.2)$	$a_3(0.2)$	$a_3(0.2)$	$a_3'(0.4)$	$a_1(0.2)$		
			$a_3' \Leftarrow a_3, a_4'$			
$a_4(0.1)$	$a_4(0.1)$	$a_4'(0.2)$				
		$a_4' \Leftarrow a_4, a_5$				
$a_5(0.1)$	$a_5(0.1)$					

TABLE 93.2 Code Assignment in Huffman Coding

	$a_3'' \Rightarrow a_3', a_1$	$a_3' \Rightarrow a_3, a_4'$	$a_4' \Rightarrow a_4, a_5$
$C(a_3'') = 0$	$C(a_3') = C(a_3'') * 0 = 00$	$\mathbf{C(a_3)} = C(a_3') * 0 = \mathbf{000}$	$\mathbf{C(a_4)} = C(a_4') * 0 = \mathbf{0010}$
$C(\mathbf{a_2}) = \mathbf{1}$	$\mathbf{C(a_1)} = C(a_3'') * 1 = \mathbf{01}$	$C(a_4') = C(a_3') * 1 = 001$	$\mathbf{C(a_5)} = C(a_4') * 1 = \mathbf{0011}$

Arithmetic Codes

Encoding sequences of symbols is more efficient than encoding individual symbols in terms of coding rate. However, Huffman encoding of sequences becomes impractical because to Huffman code a particular sequence of length m we need codewords for all possible sequences of length m. This latter fact causes an exponential growth in the size of the codebook. What we need is a way of assigning codewords to *particular* sequences without having to generate codes for all sequences of that length.

If we assume that all letters in a given alphabet occur with nonzero probability, the value of the cumulative density function *cdf* $F_X(\boldsymbol{X})$ of a sequence $\boldsymbol{X}$ is distinct from the value of the *cdf* for any other sequence of symbols. If we could impose an ordering on the sequences such that $\boldsymbol{X}_i < \boldsymbol{X}_j$ if $i < j$, then the half-open sets $[F_X(\boldsymbol{X}_{i-1}), F_x(\boldsymbol{X}_i)]$ are disjoint, and any element in this set can be used as a unique tag for the sequence $\boldsymbol{X}_i$. It can be shown that the binary representation of this number truncated to $[\log(1/(P(X))] + 1$ bits is a unique code for this sequence [Cover and Thomas, 1991].

Arithmetic coding is a procedure that generates this unique code in an incremental fashion. That is, the code is developed and transmitted (or stored) as the sequence of symbols develops. The decoding procedure is also incremental in nature. Elements of the sequence can be identified as the code is received, without having to wait for the entire code to be received. The coding and decoding procedure requires the knowledge of the *cdf* $F_X(x)$ of the source (note that we need the *cdf* of the source, not of the sequences).

It can be shown [Cover and Thomas, 1991] that the bounds on the coding rate for arithmetic coding R_A for a sequence of length n are

$$H(S) \le R_A \le H(S) + \frac{2}{n} \tag{93.6}$$

On comparing the upper bound to Eq. (93.5), it seems that Huffman coding will always have an advantage over arithmetic coding. However, recall that Huffman coding is not a realistic alternative for sequences of any reasonable length.

In terms of implementation, Huffman encoding is easier to implement, although the decoding complexities for the two are comparable. On the other hand, arithmetic coding can generally provide a coding rate closer to the entropy than Huffman coding, the exception being when the symbol probabilities are powers of 2, in which case the Huffman code exactly achieves entropy. Other cases where arithmetic coding does not provide much advantage over Huffman coding are when the alphabet size is relatively large. In these cases, $P_{\max}$ is generally small, and Huffman coding will generate a rate close to the entropy. However, when there is a substantial imbalance in probabilities, especially in small alphabets, arithmetic coding can provide significant advantage. This is especially true in the coding of facsimile information where the base alphabet size is two with highly uneven probabilities. Arithmetic coding is also easier to use when multiple codes are to be used for the same source, as the setup cost is simply the generation of multiple *cdf*s. This situation occurs often in text compression where *context modeling* may require the use of a different arithmetic code for different contexts.

Adaptive Huffman Coding

The Huffman code relies on knowledge of source statistics for its efficiency. In many applications, the source statistics are not known *a priori* and the Huffman code is implemented as a two-pass procedure. The statistics of the source are collected in the first pass. These statistics are used to generate the Huffman code, which is then used to encode the source in the second pass. In order to convert this algorithm into a one-pass procedure, the probabilities need to be estimated adaptively and the code altered to reflect these estimates. Theoretically, if we wanted to encode the $(k + 1)$st symbol using the statistics of the first k symbols, we could recompute the code using the Huffman coding procedure each time a symbol is transmitted. However, this would not be a very practical approach due to the large amount of computation involved. The adaptive Huffman coding procedure is a computationally efficient procedure for estimating the probabilities and updating the Huffman code.

In the adaptive Huffman coding procedure, at the start of transmission neither transmitter nor receiver knows anything about the statistics of the source sequence. The tree at both the transmitter and the receiver consists of a single node, which corresponds to all symbols not yet transmitted and has a weight of zero. As transmission progresses, nodes corresponding to symbols transmitted will be added to the tree, and the tree is reconfigured using a computationally efficient update procedure [Gallagher, 1978]. Prior to the beginning of transmission, a fixed code for each symbol is agreed upon between transmitter and receiver and is used upon the first occurrence of the symbol.

The adaptive Huffman coding procedure provides an excellent alternative to Huffman coding in cases where the source statistics are unknown. The drawbacks are increased complexity and substantially increased vulnerability to channel errors. As both transmitter and receiver are building the code as the transmission proceeds, a single error can cause the building of different codes at the transmitter and receiver, effectively stopping any transfer of information. There is some overhead involved in transmitting the fixed codes for the first occurrence of each symbol; however, if the transmission is sufficiently long, the effect of this overhead on the average transmission rate is minimal. Of course, if we try to combat the effect of channel errors with frequent resynchronization, the overhead can become a major factor.

Adaptive Arithmetic Coding

Adapting the arithmetic coder to changing statistics is relatively easy. The simplest way to do this is to use the number of times a symbol is encountered as an estimate of the probability and, hence, the cumulative distribution function. This approach can be refined to provide efficient implementation algorithms. More sophisticated algorithms have been developed for the particular case of binary alphabets. One of the more well-known adaptive arithmetic coders is the *Q-coder*, which is part of the joint photographic expert group (JPEG) and joint bilevel image experts group (JBIG) standards.

93.3 Universal Codes

The coding schemes we have described previously depend on the source statistics being stationary, or at worst, slowly varying. In situations where this assumption does not hold, these coding schemes might provide undesirable results. Coding schemes, which optimally code sources with unknown parameters, are generally known as universal coding schemes [Davisson, 1973]. Although proofs of optimality for these schemes generally rely on asymptotic results, several universal coding schemes have also been shown to perform well in practical situations. We describe some of these in the following.

Lynch–Davisson–Schalkwijk–Cover Codes

In this approach, the input is divided into blocks of n bits where $n = 2^m$ [Cover and Thomas, 1991]. For each block, the Hamming weight w (the number of 1s in the n-bit long sequence) of the block is transmitted using m bits followed by È$\log_2((n/w) - 1)$˘ bits to indicate a specific sequence out of all of the possible n bit sequences with w ones.

Syndrome Source Coding

This approach entails using a block error correcting decoder as a source encoder [Ancheta, 1977]. The source output is represented in binary form and divided into blocks of length n. Each block is then viewed as an error pattern and represented by the syndrome vector of an error correcting code. As different error correcting codes will correct a different number of errors, the syndrome vector is preceded by an index pointing to the code being used.

Golomb Codes

Golomb [1966] codes are unary codes, which are optimal for certain distributions. Given the parameter g for the Golomb code, any integer n is represented by a prefix and a suffix. The prefix is the unary representation of $[n/g]$, and the suffix is the binary representation of $n \bmod m$. Although these codes are optimal for certain exponential distributions, they are not universal in the sense that sources can be found for which the average code length diverges.

Rice Codes

The Rice code can be seen as a universal version of the Golomb codes [Rice, Yeh, and Miller, 1991]. The Rice algorithm first converts the input into a sequence of nonnegative integers. This sequence is then coded in blocks of length J (a suggested choice for J is 16). Each block is coded using a variety of binary and unary codes depending on the properties of the particular block. The Rice code has been shown to be optimal over a range of entropies.

Ziv–Lempel Codes

The **Ziv-Lempel** (LZ) codes are a family of dictionary-based coding schemes that encode strings of symbols by sending information about their location in a dictionary. The basic idea behind these dictionary-based schemes is that for certain sources, certain patterns recur very frequently. These patterns can be made into entries in a dictionary. Then, all future occurrences of these patterns can be encoded via a pointer to the relevant entry in the dictionary. The dictionary can be static or adaptive. Most of the adaptive schemes have been inspired by two papers by Ziv and Lempel in 1977 and 1978 [Bell, Cleary, and Witten, 1990].

The 1977 algorithm and its derivatives use a portion of the already encoded string as the dictionary. For example, consider the encoding of the string *abr acadabra*, where the underlined portion of the string has already been encoded. The string *abra* could then be encoded by simply sending the pair (7,4), where the first number is the location of the previous occurrence of the string relative to the current position and the second number is the length of the match. How far back we search for a match depends on the size of a prespecified window and may include all of the history.

The 1978 algorithm actually builds a dictionary of all strings encountered. Each new entry in the dictionary is a previous entry followed by a letter from the source alphabet. The dictionary is seeded with the letters of the source alphabet. As the coding progresses, the entries in the dictionary will consist of longer and longer strings. The most popular derivative of the 1978 algorithm is the LZW algorithm, a variant of which is used in the UNIX *compress* command, as well as the *GIF* image compression format, and the V.42-bis compression standard.

We now describe the LZW algorithm. The LZW algorithm starts with a dictionary containing all of the letters of the alphabet [Bell, Cleary, and Witten, 1990]. Accumulate the output of the source in a string s as long as the string s is in the dictionary. If the addition of another letter a from the source output creates a string $s * \mathtt{a}$ that is not in the dictionary, send the index in the dictionary for s, add the string $s * \mathtt{a}$ to the dictionary, and start a new string that begins with the letter a. The easiest way to describe the LZW algorithm is through an example. Suppose we have a source that transmits symbols from the alphabet $A = \{a, b, c\}$, and we wish to encode the sequence *aabaabc*…. Assuming we are using 4 bits to represent the entries in the codebook, our initial dictionary may look like this

a	0000
b	0001
c	0010

We will first look at what happens at the encoder. As we begin to encode the given sequence we have the string $s = a$, which is in the dictionary. The next letter results in the string $s * a = aa$, which is not

in the dictionary. Therefore, we send 0000, which is the code for a, and add aa to our dictionary, which is now

a	0000
b	0001
c	0010
aa	0000

and we begin a new string $s = a$. This string is in our dictionary and so we read the next letter and update our string to $s * b = ab$, which is not in our dictionary. Thus, we add ab to our dictionary, send the code for a to the receiver and start a new string $s = b$. Again, reading in the next letter, we have the string $s * a = ba$. This is not in our dictionary and so we send 0001, which is the code for b, add ba to our dictionary, and start a new string $s = a$. This time when we read a new letter, the new string $s * a = aa$ is in our dictionary and so we continue with the next letter to obtain the string $s * b = aab$. The string aab is not in the dictionary and so we send the 0011, which is the code for aa, and add the string aab into the dictionary. The dictionary at this stage looks like this:

a	0000
b	0001
c	0010
aa	0011
ab	0100
ba	0101
aab	0110

Notice that the entries in the dictionary are getting longer. The decoder starts with the same initial dictionary. When it receives the first codeword 0000, it decodes it as a and sets its string equal to a. Upon receipt of the next 4 bits (which would again be 0000), it adds the string aa to the dictionary and initializes the string s to a. The update procedure at the decoder is the same as the update procedure at the encoder and so the decoder dictionary tracks the encoder, albeit with some delay. The only time this delay causes a problem is when the encoder encounters a sequence of the form $a * r * \alpha * r * \alpha * \beta$ is to be encoded, where α and β are symbols, r is a string, and $\alpha * r$ is already in the dictionary. The encoder sends the code for $a * r$, and then adds $\alpha * r\alpha$ to the encoder dictionary. The next code transmitted by the encoder is the code for $a * r\alpha$, which the decoder does not yet have. This problem can easily be handled as an exception.

We can see that after a while we will have exhausted all the 4 bit combinations. At this point we can do one of several things. We can flush the dictionary and restart, treat the dictionary as a static dictionary, or add 1 bit to our code and continue adding entries into the dictionary.

93.4 Text Compression

A text source has several very distinctive characteristics, which can be taken advantage of when designing text compression schemes. The most important characteristic from the point of view of compression is the fact that a text source generally contains a large number of recurring patterns (words). Because of this, directly encoding the output of a text source with an entropy coder would be highly inefficient. Consider the letter u. The probability of this letter occurring in a fragment of English text is 0.02. Therefore, if we designed an entropy coder for this source we would assign a relatively long codeword to this letter. However, if we knew that the preceding letter was q, the probability of the current letter being u is close to one. Therefore, if we designed a number of entropy coders, each being indexed by the previous character encoded, the number of bits used to represent u would be considerably less. Thus, to be efficient, it is important that we treat each symbol to be encoded in the context of the past history, and *context modeling* has become an important approach to text compression.

Another technique for modeling recurring patterns is the use of finite state machines. Although this approach has been less fruitful in terms of generating compression algorithms, it holds promise for the future.

Finally, an obvious approach to text compression is by the use of dictionaries. We have described some adaptive dictionary-based techniques in the section on universal coding. We now briefly describe some of the more popular methods. Descriptions that are more complete can be found in Bell, Cleary, and Witten [1990].

Context Models

In lossless compression, both the encoder and decoder are aware of the same past history. In other words, they are aware of the context in which the current symbol occurs. This information can, therefore, be used to increase the efficiency of the coding process. The context is used to determine the probability distribution for the next symbol. This probability distribution can then be used by an entropy coder to encode the symbol. In practice, this might mean that the context generates an pointer to a code table designed for that particular distribution. Thus, if the letter u follows a space it will be encoded using considerably more bits than if it followed the letter q.

Context-based schemes are generally adaptive as different text fragments can vary considerably in terms of repeating patterns. The probabilities for different symbols in the different contexts are updated as they are encountered. This means that one will often encounter symbols that have not been encountered before for any of the given contexts (this is known as the *zero frequency* problem). In adaptive Huffman coding, this problem was resolved by sending a code to indicate that the following symbol was being encountered for the first time followed by a prearranged code for that symbol. There was a certain amount of overhead associated with it, but for a sufficiently long symbol string the additional rate due to the overhead was negligible. Unfortunately, in context-based encoding, the zero-frequency problem is encountered often enough for overhead to be a problem. This is especially true for longer contexts. Consider a context model of order four (the context is determined by the last four symbols). If we take an alphabet size of 95, the possible number of contexts is 95^4, which is more than 81 million! Of course, most of these contexts will never occur, but still the zero-frequency problem will occur often enough for the effect of the overhead to be substantial. One way to at least reduce the size of the problem would be to reduce the order of the context. However, longer contexts are more likely to generate accurate predictions. This problem can be resolved by a process called *exclusion*, which approximates a blending strategy for combining the probabilities of a symbol with respect to contexts of different orders. The exclusion approach works by first attempting to find if the symbol to be encoded has a nonzero probability with respect to the maximum context length. If this is so, the symbol is encoded and transmitted. If not, an escape symbol is transmitted and the context size is reduced by one and the process is repeated. This procedure is repeated until a context is found with respect to which the symbol has a nonzero probability. To guarantee that this process converges, a null context is always included, with respect to which all symbols have equal probability. The probability of the escape symbol can be computed in a number of different ways leading to different implementations.

Context modeling-based text compression schemes include: *PPMC* [prediction by partial match (PPM), PPMC is a descendant of PPMA and PPMB], *DAFC*, which uses a maximum context order of one, and *WORD*, which uses two alphabets, one which only contains alphanumeric characters, with the other containing the rest. A performance comparison is shown in Table 93.3.

State Models

Context models discard sequential information beyond the size of the context. This information can be preserved and used by finite-state models. However, there is relatively less work available in this area. An adaptive technique for finite-state modeling named dynamic Markov compression (DMC) was given by Cormack and Horspool [Bell, Cleary, and Witten, 1990]. In DMC, one starts with some

TABLE 93.3 Results of Compression Experiments (bits per character)

Text	Size	Diagram	LZB	LZFG	Huffman	DAFC	PPMC	WORD	DMC
Bib (technical bibliography)	111261	6.42	3.17	2.90	5.24	3.84	2.11	2.19	2.28
Book (*Far from the Madding Crowd*)	768771	5.52	3.86	3.62	4.58	3.68	2.48	2.70	2.51
Geo (seismic data 32-b numbers)	102400	7.84	6.17	5.70	5.70	4.64	4.78	5.06	4.77
News (USENET batch file)	377109	6.03	3.55	3.44	5.23	4.35	2.65	3.08	2.89
Obj (executable Macintosh file)	246814	6.41	3.14	2.96	6.80	5.77	2.69	4.34	3.08
Paper (technical, troff format)	82199	5.60	3.43	3.16	4.65	3.85	2.45	2.39	2.68
Progc (source code in C)	39611	6.25	3.08	2.89	5.26	4.43	2.49	2.71	2.98

Source: Bell, T.C., Witten, I.H., and Cleary, J.C. 1989. Modeling for text compression. *ACM Comput. Surveys,* Dec.

fixed initial model and rapidly adapts it to the input data. Adaptation is done by maintaining counts on the transitions from each state and cloning a state when the count exceeds a certain threshold. The size of the model grows rapidly and, hence, this technique has been found to be feasible only for binary input data, in which case each state in the model has only two output transitions. Although at first sight DMC seems to offer the full generality of Markov models, it has been proven that they only generate finite-context models and, hence, are no more powerful then the techniques described earlier. Despite this fact, DMCs provide an attractive and efficient way of implementing finite-context models.

Dictionary-Based Coding

Dictionary-based coding techniques have always been popular for text compression. The dictionary can be static or it can adapt to the source characteristics. Most current adaptive dictionary-based coders are variations of the 1977 and 1978 algorithms due to Ziv and Lempel (LZ77 and LZ78). The LZ77 and LZ78 algorithms are described in the section on universal coding.

The most popular static dictionary-based coding scheme is *diagram coding*. The dictionary consists of the letters of the alphabet followed by the most popular pairs of letters. For example, an 8-bit dictionary for printable ASCII characters could be constructed by following the 95 printable ASCII characters by 161 most frequently occurring pairs of characters. The algorithm would parse the input a pair of characters at a time. If the pair existed in the dictionary, it would put out the code for the pair and advance the pointer by two characters. If not, it would encode the first element of the pair and then advance the pointer by one character. This is a very simple compression scheme; however, compared to the other schemes described here, the compression performance is not very good.

Table 93.3 shows a comparison of the various schemes described here. The LZB algorithm is a particular implementation of the LZ77 approach, whereas the LZFG algorithm is an implementation of the LZ78 approach.

Block Sorting

The block sorting transform (also known as Burrows–Wheeler transform, or BWT) is a relatively recent compression technique. It achieves speeds comparable to fast dictionary-based methods and enables compression competitive with the best context-based like PPM. The BWT transforms a string S containing N characters by forming the N rotations (cyclic shifts) of S, sorting them lexicographically, and extracting the last character of each of the rotations. A string L is formed from these characters, where the ith character of L is the last character of the ith sorted rotation. The algorithm also computes the index I

which is the location of the original string S in the sorted list of rotations. We clarify the algorithm by means of an example. With the input string *abraca*, the rotated strings and the lexicographically sorted list of rotations is as follows:

$$\begin{array}{ll} abraca & aabrac\mathbf{c} \\ bracaa & abrac\mathbf{a} \\ racaab & acaab\mathbf{r} \\ acaabr & \Rightarrow \; braca\mathbf{a} \\ caabra & caabr\mathbf{a} \\ aabrac & racaa\mathbf{b} \end{array}$$

The string of last letters, L, in the sorted list is *caraab*, and, as the original string is the second entry in the sorted list, the index I is 1 (the lowest index value is 0). The original string S can be reconstructed using just L and I. To do this, the inverse transform sorts the string L to form the string F. For this example, the string *caraab* would be sorted to obtain the string *aaabcr*: *caraab* fi *aaabcr*. The algorithm also forms the array T, where T[m] is the index of the first occurrence of L[m] in F added to the number of previous occurrences of L[m] in L. In the above example, T is [4 0 5 1 2 3]. Let J be an index that initially points to the last letter in the reconstructed string R. The inverse BWT procedure copies the value of L[I] to R[J] and sets J to J-1 and I to T[I]. This procedure is repeated until R is filled from back to front. Applying this procedure to the example string, we have J = 5. Therefore, R[5] = L[1] = a. Then the values of I and J are updated; I = T[1] = 0, and J = J - 1 = 4. Then the penultimate reconstructed letter R[4] = L[0] = c. The next value for I is T[0] which is 4, therefore, R[3] = L[4] = a. We have recovered the last three letters of the original string. Continuing in this fashion, we can recover the entire original string.

The structure of the string L that comprises BWT output is similar for typical text files: it has many discrete regions, each with a group of characters with high probability. This is because BWT arranges symbols lexicographically according to their suffixes. For example, strings beginning with *he*° in the list of lexicographically sorted roations of S would be most often *t*. This is because the letter *t* most often precedes *he*. As another example, the suffixes to *WT*... would be most often *B*. For this reason, BWT can be considered a context-based scheme like PPM, with a context of the entire block.

The transform presented thus far only makes its input more compressible. There are many approaches employed for actually compressing BWT output. A simple way of doing this is by using the move-to-front (MTF) coding technique that exploits the locality of occurrence in BWT output. MTF initializes an array with the entire alphabet of L, inputs a character, outputs the position of the character in the array, and moves it to the front of the array. With the string *caraab*, the MTF output is 2 1 3 1 0 3.

93.5 Image Compression

Algorithms for text compression do not work well when applied to images. This is because such algorithms exploit the frequent recurrence of certain exact patterns that are very typical of textual data. In image data, there are no such exact patterns common to all images. Also, in image data, correlation among pixels exists along both the horizontal and vertical dimensions. Text compression algorithms, when applied to images, invariably fail to exploit some of these correlations. Hence, specific algorithms need to be designed for lossless compression of images.

As mentioned before, there are two basic components in a lossless compression technique, modeling and coding. In lossless image compression, the task of modeling is usually split into two stages. In the first step, a *prediction model* is used to predict pixel values and replace them by the error in prediction. If prediction is based on previously transmitted values, then knowing the prediction error and the prediction scheme, the receiver can recover the value of the original pixel.

In the second step, a *prediction error model* is constructed, which is used to drive a variable length coder for encoding prediction errors. As we have already mentioned, coding schemes are known that can

code at rates very close to the model entropy. Hence, it is important to build accurate prediction error models. In the next two sections, we describe various schemes that have been proposed in literature for each of these two steps.

Prediction Models

Linear Predictive Models

Linear predictive techniques (also known as lossless DPCM) usually scan the image in raster order, predicting each pixel value by taking a linear combination of pixel values in a casual neighborhood. For example, we could use $(P[i - 1, j] + P[i, j - 1])/2$ as the prediction value for $P[i, j]$, where $P[i, j]$ is the pixel value in row i, column j. Despite their apparent simplicity, linear predictive techniques are quite effective and give performance surprisingly close to more state-of-the-art techniques. If we assume that the image is being generated by an *autoregressive* (AR) *model*, coefficients that best fit given data in the sense of the L_2 norm can be computed. One problem with such a technique is that the implicit assumption of the data being generated by a stationary source is, in practice, seldom true for images. Hence, such schemes yield very little improvement over the simpler linear predictive techniques such as the example in the previous paragraph [Rabbani and Jones, 1991].

Significant improvements can be obtained for some images by adaptive schemes that compute optimal coefficients on a block-by-block basis or by adapting coefficients to local changes in image statistics. Improvements in performance can also be obtained by adaptively selecting from a set of predictors. An example scheme that adapts in the presence of local edges is given by the median edge detection (MED) predictor used in the JPEG-LS lossless image compression standard. MED detects horizontal or vertical edges by examining the north, west, and northwest neighbors of the current pixel. The north (west) pixel is used as a prediction in the case of a vertical (horizontal) edge. In case of neither, planar interpolation is used to compute the prediction value. The MED predictor has also been called the median adaptive predictor (MAP) and was first proposed by Martucci [1990]. Martucci proposed the MAP predictor as a nonlinear adaptive predictor that selects the median of a set of three predictions in order to predict the current pixel. One way of interpreting such a predictor is that it always chooses either the best or the second best predictor among the three candidate predictors. Martucci reported the best results with the following three predictors: (1) $P[i, j - 1]$, (2) $P[i - 1, j]$, and (3) $P[i, j - 1] + P[i - 1, j] - P[i - 1, j - 1]$. In this case, it is easy to see that MAP turns out to be the MED predictor. In an extensive evaluation [Memon and Wu, 1997], it was observed that the MED predictor gives superior or almost as good a performance as many standard prediction techniques, many of which are significantly more complex. It should also be noted that, strictly speaking, MED is not a linear predictor but it does select adaptively from one of three different linear predictors.

Context Models

In context-based models, data is partitioned into contexts, and statistics are collected for each context. These statistics are then used to perform prediction. One problem with such schemes when compressing gray-scale images is that the number of different contexts grows exponentially with the size of the context. Given the large alphabet size of image data, such schemes quickly become infeasible even for small context sizes. Some clever schemes have been devised for partitioning the large number of contexts to a more manageable size [Wu and Memon, 1997; Todd, Langdon, and Rissanen, 1985]. Such schemes, however, have all been for modeling the prediction errors after using a simple linear predictive scheme. We shall describe them in more detail in the next section. An exception to this rule is binary images, for which context-based prediction schemes are among the most efficient.

Multiresolution Models

Multiresolution models generate representations of an image with varying spatial resolution. This usually results in a pyramid like representation of the image with each layer of the pyramid serving as a prediction model for the layer immediately below. The pyramid can be generated in either a top-down or

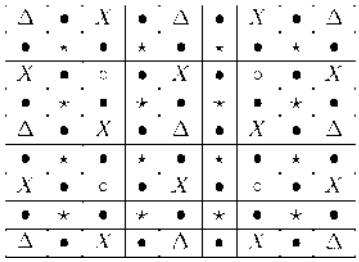


FIGURE 93.1 The HINT scheme for hierarchical prediction.

a bottom-up manner. Transmission is generally done in a top-down manner. One advantage of such a scheme is that the receiver recovers the image level by level. After reconstructing each level, the receiver can get an approximation of the entire image by interpolating unknown pixels. This leads to *progressive transmission*, a technique that is useful in many applications.

One of the more popular such techniques is known as hierarchical interpolation (HINT) and was proposed by Endoh and Yamakazi [Rabbani and Jones, 1991]. The specific steps involved in HINT are as follows. First, the pixels labeled D in Fig. 93.1 are decorrelated using DPCM. In the second step, the intermediate pixels (o) are estimated by linear interpolation and replaced by the error in estimation. Then, the pixels X are estimated from D and o and replaced by the error. Finally, the pixels labeled $\star$ and then $\bullet$ are estimated from known neighbors and replaced. The reconstruction process proceeds in a similar manner.

Lossy Plus Lossless Models

Given the success of lossy compression schemes for efficiently constructing a low-bit-rate approximation, one way of obtaining a prediction model for an image is to use a lossy representation of an image as the prediction model. Such schemes are called *lossy plus lossless schemes* (LPL). Here, a low-rate lossy representation of the image is first transmitted. Subsequently, the difference between the lossy and the original is transmitted to yield lossless reconstruction. Although the lossy image and its residual are usually encoded independently, making use of the lossy image to encode the residual can lead to significant savings in bit rates. Generally, LPL schemes do not give as good compression as other standard methods. However, they are found to be very useful in certain applications. One application is when a user is browsing through a database of images, looking for a specific image of interest.

Modeling Prediction Errors

If the residual image that comprises prediction errors can be treated as an i.i.d. source, then it can be efficiently coded using any of the standard variable length techniques such as Huffman coding or arithmetic coding. Unfortunately, even after applying the most sophisticated prediction techniques, generally the residual image has ample structure that violates the i.i.d. assumption. Hence, in order to encode prediction errors efficiently, we need a model that captures the structure that remains after prediction. This step is often referred to as *error modeling* [Howard and Vitter, 1992]. A very simple technique for modeling and coding the prediction residual in lossless image compression based on the Rice encoder described previously is given in [Rice, Yeh, and Miller, 1991]. However, the error modeling techniques employed by most lossless compression schemes proposed in the literature use a context modeling framework first described in Rissanen and Langdon [1981] and applied in an increasingly refined manner in Todd et al. [1985], Howard and Vitter [1992], Wu and Memon [1997], and Weinberger et al. [1996]. In this approach, the prediction error at each pixel is encoded with respect to a conditioning state or *context*, which is arrived at from the values of previously encoded neighboring pixels. Viewed in this framework, the role of the error model is essentially to provide estimates of the conditional probability of the prediction error given the context in which it occurs. This can be done by estimating the pdf by maintaining counts of symbol occurrences within each context [Todd et al., 1985; Wu and

Memon, 1997] or by estimating the parameters (variance for example) of an assumed pdf (Laplacian, for example) as in Howard and Vitter [1992] and Weinberger et al. [1996]. Each of these uses some clever mechanisms to reduce the number of possible contexts. We describe the technique by Weinberger et al. in more detail as it is used in the lossless compression standard JPEG-LS that is described in more detail in a later section. Contexts in JPEG-LS are formed by first computing the following differences: (1) D1 = NE-N (2) D2 = N-NW and (3) D3 = NW-W. Here N, W, NE, and NW represent the north, west, northeast, and northwest neighbors of the current pixel, respectively. The differences D1, D2, and D3 are then quantized into 9 regions (labeled -4 to +4) symmetric about the origin with one of the quantization regions (region 0) containing only the difference value 0. Further, contexts of the type (q_1, q_2, q_3) and $(-q_1, -q_2, -q_3)$ are merged based on the assumption that $P(e \mid q_1, q_2, q_3) = P(-e \mid -q_1, -q_2, -q_3)$. The total number of contexts turns out to be $9^3 - 1/2 = 364$. These contexts are then mapped to the set of integers [0, 363] in a one-to-one fashion.

As described earlier, contexts are used as conditioning states for encoding prediction errors. Within each state, the pdf of the associated set of events is adaptively estimated by events by keeping occurrence counts for each context. Clearly, to better capture the structure preset in the prediction residuals, one would like to use a large number of contexts or conditioning states. However, the larger the number of contexts, the greater the number of parameters (conditional probabilities in this case) that need to be estimated based on the same data set. This can lead to the "sparse context" or "high model cost" problem. Typically (for example, in the JPEG lossless standard—arithmetic coding version), this problem is addressed by keeping the number of contexts small. One problem with this approach is that keeping the number of conditioning states to a small number fails to effectively capture the structure present in the prediction errors and results in poor performance. Hence, the JPEG-LS baseline algorithm employs a different solution for this problem. First, it uses a relatively large number of contexts to capture the structure present in the prediction errors. However, instead of estimating the pdf of prediction errors, $p(e \mid C)$, within each context C, only the conditional expectation $E\{e \mid C\}$ is estimated using the corresponding sample means within each context. These estimates are then used to further refine the prediction prior to entropy coding by an error feedback mechanism that cancels prediction biases in different contexts. This process is called *bias-cancellation*. Furthermore, for encoding the bias-cancelled prediction errors, instead of estimating the probabilities of each possible prediction error, baseline JPEG-LS essentially estimates a parameter that serves to characterize the specific pdf to be employed from a fixed set of pdfs.

A straightforward implementation of bias-cancellation would require accumulating prediction errors within each context and keeping frequency counts of the number of occurrences for each context. The accumulated prediction error within a context divided by its frequency count would then be used as an estimate of prediction bias within the context. However, this division operation can be avoided by a simple and clever operation that updates the variables in a suitable manner, producing average prediction residuals in the interval [-0.5, 0.5]. For details, the reader is referred to Weinberger et al. [1996] and to the Draft International Standard. In addition, since JPEG-LS uses Rice codes that assign shorter codes to negative residual values than to positive ones, the bias is adjusted such that it produces average prediction residuals in the interval [-1, 0], instead of [-0.5, 0.5]. For a detailed justification of this procedure and other details pertaining to bias estimation and cancellation mechanisms, the reader is referred to Weinberger et al. [1996].

93.6 Lossless Compression Standards

Joint Bilevel Image Experts Group

The JBIG algorithm proposed by the joint bilevel image experts group defines a standard for lossless image compression. The algorithm is mainly designed for bilevel or binary data such as facsimile, but can also be used for image data with up to 256 b/pixel by encoding individual bit planes. The algorithm performs best for data up to 6 b/pixel, and beyond 6 b, other algorithms such as lossless JPEG give better performance. JBIG is described in more detail in the next chapter.

TABLE 93.4 JPEG Predictors for Lossless Coding

Mode	Prediction for $P[I, j]$
0	0 (No Prediction)
1	$P[i - 1, j]$
2	$P[i, j - 1]$
3	$P[i - 1, J - 1]$
4	$P[i, j - 1] + P[i - 1, j] - P[i - 1, j - 1]$
5	$P[i, j - 1] + (P[i - 1, j] - P[i - 1, j - 1)/2$
6	$P[i - 1, j] + (P[i, j - 1] - P[i - 1, j - 1])/2$
7	$(P[i, j - 1] + P[i - 1, j])/2$

Joint Photographic Experts Group

The JPEG committee has developed two standards for lossless image compression. One, which we refer to as the JPEG lossless standard, was developed in conjunction with the well-known DCT based lossy compression standard in 1988. A more recent standard named JPEG-LS was finalized in 1997. Below we briefly describe both.

JPEG Lossless

The well-known JPEG still compression standard [Wallace, 1991] has a lesser-known lossless component that uses linear predictive techniques followed by either Huffman or arithmetic coding of prediction errors. It provides eight different predictive schemes from which the user can select. Table 93.4 lists the eight predictors. The first scheme makes no prediction. The next three are one-dimensional predictors, and the last four are two-dimensional prediction schemes.

In the Huffman coding version, essentially no error model is used. Prediction errors are assumed i.i.d. and are encoded using the Huffman table specified in the bit stream using the specified syntax. The Huffman coding procedure specified by the standard for encoding prediction errors is identical to the one used for encoding DC coefficient differences in the lossy codec.

Unlike the Huffman coding version, which assumes the prediction errors to be i.i.d., the arithmetic coding version uses quantized prediction errors at neighboring pixels as contexts for conditioning the encoding the prediction error. This is a simplified form of error modeling which attempts to capture the remaining structure in the prediction residual. Encoding within each context is done with a binary arithmetic coder by decomposing the prediction error into a sequence of binary decisions. The first binary decision determines if the prediction error is zero. If it is not zero, then the second step determines the sign of the error. The subsequent steps assist in classifying the magnitude of the prediction error into one of a set of ranges, and the final bits that determine the exact prediction error magnitude within the range are sent uncoded. The QM-Coder is used for encoding each binary decision. A detailed description of the coder and the standard can be found in Pennebaker and Mitchell [1993]. Since the arithmetic coded version of the standard is rarely used, we do not delve in the details of the procedures used for arithmetic coding but refer the interested reader to Pennebaker and Mitchell [1993].

JPEG-LS

The JPEG-LS algorithm, like its predecessor, is a predictive technique. However, there are significant differences, as enumerated below:

- Instead of using a simple linear predictor, JPEG-LS uses the MED predictor described earlier, which attempts to detect the presence of edges passing through the current pixel and accordingly adjusts prediction. This results in a significant improvement in performance in the prediction step as compared to the old JPEG lossless standard.
- Like JPEG lossless arithmetic, JPEG-LS uses some simple but very effective context modeling of the prediction errors prior to encoding.

- Baseline JPEG-LS uses Golomb–Rice codes for encoding prediction errors. Golomb–Rice codes are Huffman codes for certain geometric distributions, which serve well in characterizing the distribution of prediction errors. Although Golomb–Rice codes have been known for a long time, JPEG-LS uses some novel and highly effective techniques for adaptively estimating the parameter for the Golomb–Rice code to be used in a given context.
- In order to effectively code low entropy images or regions, JPEG-LS uses a simple alphabet extension mechanism by switching to a run-length mode when a uniform region is encountered. The run-length coding used is again an extension of Golomb codes and provides significant improvement in performance for highly compressible images.
- For applications that require higher compression, JPEG-LS provides a near-lossless mode, which guarantees each reconstructed pixel to be within k counts of the original. Near-lossless compression is achieved by a simple uniform quantization of the prediction error.

The baseline JPEG-LS algorithm described above was subsequently extended in many ways, and the different extensions are collectively known as JPEG-LS Part 2. Part 2 includes many additional features that improve compression but were considered to be too application oriented. Some of the key features that are part of this set of extensions are:

- *Arithmetic Coding:* The biggest difference between JPEG-LS and JPEG-LS Part 2 is in the entropy coding stage. Part 2 uses a binary arithmetic coder for encoding prediction errors, which are binarized by a Golomb code. The Golomb code tree is produced based on the activity class of the context computed from its current average prediction error magnitude. Twelve activity levels are defined. In the arithmetic coding procedure, numerical data is treated in radix 255 representation, with each sample expressed as eight bit data. Probabilities of the more probable symbol (MPS) and the less probable symbol (LPS) are estimated by keeping occurrence counts. Multiplication and division are avoided by approximations that make use of table look-up operations.
- *Prediction:* JPEG-LS baseline is not suitable for images with sparse histograms. In order to deal with such images, Part 2 defines an optional prediction value control mode wherein it is ensured that a predicted value is always a symbol that has actually occurred in the past. This is done by forming the same prediction as JPEG baseline using the MED predictor, then adjusting the predictor to a value that has been seen before.
- *Context Formation:* In order to better model prediction errors, Part 2 uses additional contexts as compared to the baseline.
- *Bias Cancellation:* In baseline JPEG-LS, in the bias cancellation step, prediction errors within each context are centered on -0.5 instead of 0. This, we said, was done as the prediction error mapping technique and Rice–Golomb coding used in the baseline algorithm assigns shorter code words to negative errors as opposed to a positive error of the same magnitude. However, if arithmetic coding is employed, then there is no such imbalance and bias cancellation is used to center the prediction error distribution in each context around zero.
- *Alphabet Extension:* If arithmetic coding is used, then alphabet extension is clearly not required. Hence, in the arithmetic coding mode, the coder does not switch to run-mode on encountering the all zeroes context. However, in addition to this change, Part 2 also specifies some small changes to the run length-coding mode of the original baseline algorithm. For example, when the underlying alphabet is binary, Part 2 does away with the redundant encoding of sample values that terminated a run, as required by the baseline.
- *Near-Lossless Mode:* The near-lossless mode is another area where Part 2 differs significantly from the baseline. Essentially, Part 2 provides mechanisms for a more versatile application of the near-lossless mode. The two main features enabled by Part 2 in the near-lossless mode are:
  - *Visual Quantization:* Near-lossless compression can often lead to annoying artifacts at larger values of k. Furthermore, the baseline does not provide any graceful degradation mechanism between step size of k and $k + 1$. Hence, JPEG-LS Part 2 defines a new visual quantization mode.

In this mode, the quantization step size is allowed to be either k or $k + 1$ depending on the context. Contexts with a larger gradients use a step size of $k + 1$ and contexts with smaller gradients use a step size of k. The user specifies a threshold based on which this decision is made. The standard does not specify how to arrive at the threshold. It only provides syntax for its specification.

- *Rate Control:* By allowing the user to change the quantization step size while encoding an image, Part 2 essentially provides a rate-control mechanism whereby the coder can keep track of the coded bytes, based on which appropriate changes to the quantization step size could be made. The encoder, for example, can compress the image to less than a bounded size with a single sequential pass over the image. Other uses of this feature are possible, including region-of-interest lossless coding, etc.

- *Fixed Length Coding:* There is a possibility that Golomb code causes data expansion and results in a compressed image larger than the source image. To avoid such a case, an extension to the baseline is defined where the encoder could switch to a fixed length coding technique by inserting an appropriate marker in the bit-stream. Another marker is used to signal the end of fixed length coding.

93.7 Compression Packages

The past few years have seen a rapid proliferation of compression packages for use in reducing storage requirements. Most of these are based on the LZ algorithms with a secondary variable length encoder, which is generally a Huffman encoder. These include the following:

- LZ77-based schemes: *arj, lha, Squeeze, UC2, pkzip, zip, zoo*
- LZ78 based schemes: *gif, compress, pak, Stacker, arc, pkarc*
- PPMC based schemes: *ha*
- BWT based schemes: *bzip, bzip2*

Defining Terms

Arithmetic coding: A coding technique that maps source sequences to intervals on the real line.

Compression ratio: Size of original data/size of compressed data.

Context model: A modeling technique often used for text and image compression in which the encoding of the current symbol is conditioned on its context.

Entropy: The average information content (usually measured in bits per symbol) of a data source.

Entropy coding: A fixed to variable length coding of statistically independent source symbols that achieves or approaches optimal average code length.

Error model: A model used to capture the structure in the residual sequence which represents the difference between the source sequence and source model predictions.

Huffman coding: Entropy coding using Huffman's algorithm, which maps source symbols to binary codewords of integral length such that no codeword is a prefix of a longer codeword.

Predictive coding: A form of coding where a prediction is made for the current event based on previous events and the error in prediction is transmitted.

Universal code: A source coding scheme that is designed without knowledge of source statistics but that converges to an optimal code as the source sequence length approaches infinity.

Ziv–Lempel coding: A family of dictionary coding-based schemes that encode strings of symbols by sending information about their location in a dictionary. The LZ77 family uses a portion of the already encoded string as dictionary, and the LZ78 family actually builds a dictionary of strings encountered.

References

Ancheta, T.C., Jr. 1977. Joint source channel coding. Ph.D. Thesis, University of Notre Dame.

Bell, T.C., Cleary, J.C., and Witten, I.H. 1990. *Text Compression*. Advanced reference series. Prentice-Hall, Englewood Cliffs, NJ.

Cover, T.M. and Thomas, J.A. 1991. *Elements of Information Theory,* Wiley series in telecommunications. Wiley, New York.

Davisson, L.D. 1973. Universal noiseless coding. *IEEE Trans. Inf. Th.,* 19:783–795.

Gallagher, R.G. 1978. Variations on a theme by Huffman. *IEEE Trans. Inf. Th.,* IT-24(6):668–674.

Golomb, S.W. 1966. Run-length encodings. *IEEE Trans. Inf. Th.,* IT-12(July):399–401.

Howard, P.G. and Vitter, J.S. 1992. New methods for lossless image compression using arithmetic coding. In *Proceedings of the Data Compression Conference,* Eds. J.H. Reif and J.A. Storer, IEEE Computer Society Press, New York, 257–266.

Huffman, D.A. 1952. A method for the construction of minimum redundancy codes. *Proc. IRE,* 40:1098–1101.

Martucci, S.A. 1990. Reversible compression of HDTV images using median adaptive prediction and arithmetic coding. In *IEEE International Symposium on Circuits and Systems,* IEEE Press, New York, 1310–1313.

Memon, N.D. and Sayood, K. 1995. Lossless image compression—a comparative study. In *Still Image Compression,* SPIE Proceedings, Volume 2418, 8–20.

Memon, N.D. and Wu, X. 1997. Recent developments in lossless image compression. *Computer J.,* 40(2):31–40.

Pennebaker, W.B. and Mitchell, J.L. 1993. *JPEG Still Image Data Compression Standard.* Van Nostrand Reinhold, New York.

Rabbani, M. and Jones, P.W. 1991. *Digital Image Compression Techniques,* Vol. TT7, Tutorial texts series. SPIE Optical Engineering Press.

Rice, R.F., Yeh, P.S., and Miller, W. 1991. Algorithms for a very high speed universal noiseless coding module. Tech. Rept. 91–1, Jet Propulsion Lab., California Institute of Technology, Pasadena.

Rissanen, J. and Langdon, G.G., Jr. 1981. Compression of black-white images with arithmetic coding. *IEEE Trans. Commun.,* COM-29(6):858–867.

Todd, S., Langdon, G.G., and Rissanen, J.J. 1985. Parameter reduction and context selection for compression of gray scale images. *IBM J. Res. Dev.,* 29(March):88–193.

Wallace, G.K. 1991. The JPEG still picture compression standard. *Commun. ACM,* 34(April):31–44.

Weinberger, M.J., Seroussi, G., and Sapiro, G. 1996. LOCO-I: a low complexity context-based lossless image compression algorithm. In *Proceedings of the IEEE Data Compression Conference,* IEEE Press, New York, 140–149.

Wu, X. and Memon, N.D. 1997. Context-based adaptive lossless image coding. *IEEE Trans. on Commun.,* 45(4):437–444.

Further Information

A source for further information on the topics covered in this chapter is the book *Introduction to Data Compression*, by K. Sayood.

A good place to obtain information about compression packages as well as information about recent commercial developments is the frequently asked questions (FAQ) for the group comp.compression in the netnews hierarchy. This FAQ is maintained by Jean-Loup Gailly and is periodically posted in the comp.compression newsgroup. The FAQ can also be obtained by ftp from rtfm.mit.edu in pub/usenet/news.answers/compression-faq/part [1–3]. The information about compression packages was obtained from the FAQ.

94

Facsimile

Nasir D. Memon
Polytechnic University

Khalid Sayood
University of Nebraska—Lincoln

94.1 Introduction

A **facsimile (fax)** image is formed when a document is raster scanned by a light sensitive electronic device, which generates an electrical signal with a strong pulse corresponding to a dark dot on the scan line and a weak pulse for a white dot. In digital fax machines, the electrical signal is subsequently digitized to two levels and processed, before transmission over a telephone line. Modern digital fax machines partition a page into 2376 scan lines, with each scan line comprising 1728 dots. A fax document can, therefore, be viewed as a two-level image of size 2376×1728, which corresponds to 4, 105, 728 bits of data. The time required to transmit this raw data over a telephone channel could be substantial. To reduce the bit rates, some form of compression technique is required. Fortunately, fax images contain sufficient redundancies, and even higher than 15:1 compression can be achieved by state-of-the-art compression techniques.

Facsimile image compression provides one of the finest examples of the importance of the development of efficient compression technology in modern day communication. The field of facsimile image transmission has seen explosive growth in the last decade. One of the key factors behind this proliferation of fax machines has been the development and standardization of effective compression techniques. In the rest of this chapter we describe the different approaches that have been developed for the compression of fax data. For the purpose of discussion, we classify the compression techniques into five different categories and give one or two representative schemes for each. We then describe international standards for facsimile encoding, the development of which has played a key role in the establishment of facsimile transmission as we know it today.

94.2 Facsimile Compression Techniques

Over the last three decades, numerous different techniques have been developed for the compression of facsimile image data. For the purpose of discussion, we classify such compression techniques into five different categories: (1) one-dimensional coding, (2) two-dimensional techniques, (3) multilevel techniques, (4) lossy techniques, and (5) pattern matching techniques. We discuss each approach in a separate subsection and describe one or two representative schemes.

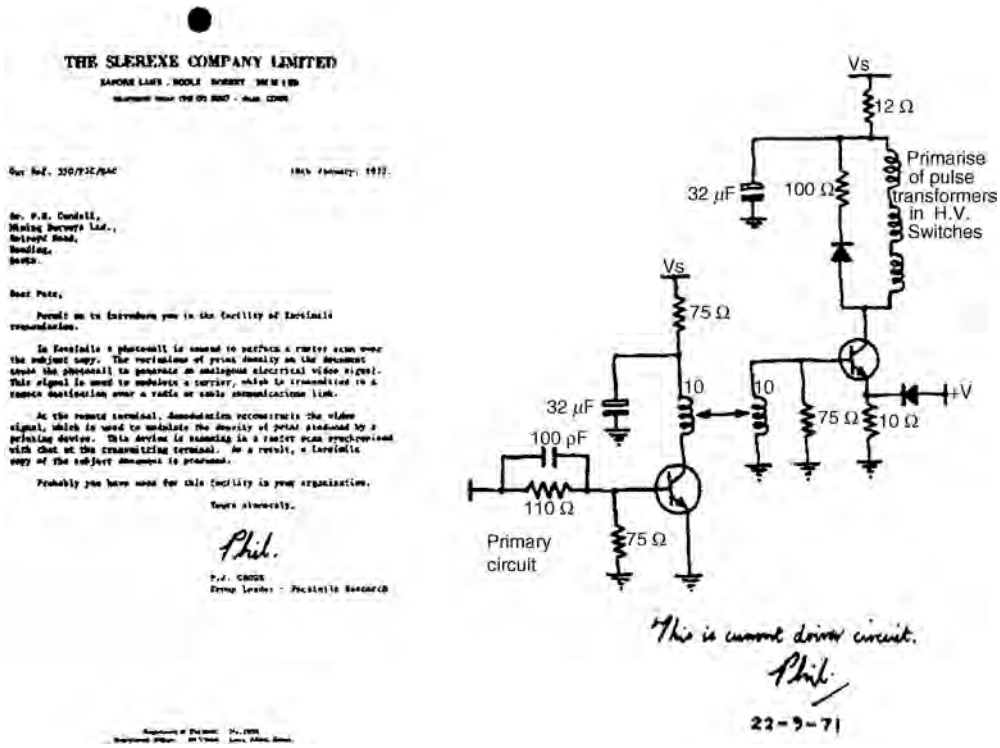


FIGURE 94.1 Example documents from the CCITT group 3 test images.

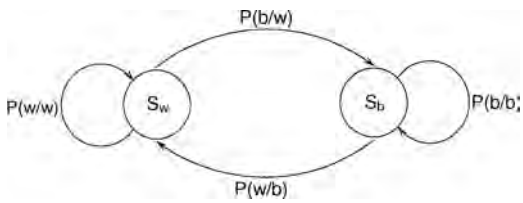


FIGURE 94.2 The Capon model for binary images.

One-Dimensional Coding

Figure 94.1 shows two sample documents that are typically transmitted by a fax machine. One property that clearly stands out is the clustered nature of black (b) and white (w) pixels. The b and w pixels occur in bursts. It is precisely this property that is exploited by most facsimile compression techniques. A natural way to exploit this property is by *run length coding*, a technique used in some form or another by a majority of the earlier schemes for facsimile image coding.

In run length coding, instead of coding individual pixels, the lengths of the runs of pixels of the same color are encoded, following an encoding of the color itself. With a two-level image, just encoding alternating runs is sufficient. To efficiently encode the runlengths we need an appropriate model. A simple way to obtain a model for the black and white runs is by regarding each scan line as being generated by the first-order Markov process, shown in Fig. 94.2, known as the *Capon model* for binary images [Capon, 1959]. The two states S_w and S_b shown in the figure represent the events that the current pixel is a white pixel or a black pixel, respectively. $P(w/b)$ and $P(b/w)$ represent *transition probabilities*. $P(w/b)$ is the probability of the next pixel being a white pixel when the current pixel is black, and $P(b/w)$ is vice versa.

If we denote the probabilities $P(w/b)$ and $P(b/w)$ by t_w and t_b, respectively, then the probability of a run of length r_k in a state s is given by

$$P(r_K|s) = t_s(1-t_s)^{r_k-1} \quad s \in \{S_w, S_b\}$$

which gives us a *geometric distribution* for the run lengths. The expected run length of black and white runs then turns out to be $1/t_b$ and $1/t_w$, respectively. The geometric distribution has been found to be an appropriate model for the run lengths encountered in special classes of facsimile images such as weather maps [Kunt and Johnsen, 1980]. However, for more structured documents such as letters that contain printed text, it turns out to be inadequate. Getting analytical models for run lengths of structured documents is difficult. In practice, models are obtained empirically by analyzing a set of typical images, and optimal variable length codes are then constructed based on the statistics of run lengths in this set. Usually, two distinct sets of codewords are constructed for the black and white runs as the statistics for the two are found to be significantly different. The extra cost involved in maintaining two separate code tables is worth the improvement in compression obtained.

Two-Dimensional Coding Schemes

The amount of compression obtained by one-dimensional coding schemes described in the previous subsection is usually quite limited. This is because such schemes do not take into account vertical correlations, that is, the correlation between adjacent scan lines, typically found in image data. Vertical correlations are especially prominent in high-resolution images that contain twice the number of scan lines per page. There have been many schemes proposed for taking vertical correlations into account. We will discuss a few that are representative.

One way to take vertical correlations into account is by encoding pixels belonging to k successive lines simultaneously. Many different techniques of this nature have been proposed in the literature, including *block coding, cascade division coding, quad-tree encoding*, etc. (for a review, see Kunt and Johnsen [1980] and Yasuda [1980]). However, such techniques invariably fail to utilize correlations that occur across the boundaries of the blocks or bundles of lines that are being encoded simultaneously. A better way to exploit vertical correlations is to process pixels line by line, as in one-dimensional coding, and make use of the information encountered in previous scan lines in order to encode the current pixel or sequence of pixels. Next we list three such techniques that have proven to be very successful.

Relative Element Address Designate (READ) Coding

Since two adjacent scan lines of a fax image are highly correlated, so are their corresponding runs of white and black pixels. Hence, the run lengths of one scan line can be encoded with respect to the run lengths of the previous scan line. A number of schemes based on this approach were developed in the late 1970s. Perhaps the best known among them is the *relative element address designate* (READ) coding technique that was a part of Japan's response to a call for proposals for an international standard [Yasuda, 1980]. In READ coding, prior to encoding a run length, we locate five *reference pixels* on the current and previous scan line. These pixels are denoted by a_0, a_1, a_2, b_1, and b_2, respectively, and are identified as follows:

- a_0: This is the last pixel whose value is known to both encoder and decoder. At the beginning of encoding each line, a_0 refers to an imaginary white pixel to the left of the first actual pixel. Although it often is a transition *pixel*, it does not have to be one.
- a_1: This is the first transition pixel to the right of a_0.
- a_2: This is the second transition pixel to the right of a_0.
- b_1: This is the first transition pixel on the line above the line currently being encoded to the right of a_0 whose color is the opposite of the color of a_0.
- b_2: This is the first transition pixel to the right of b_1 and on the same line as b_1.

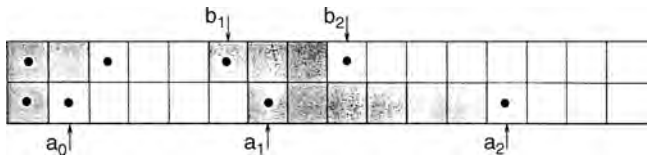


FIGURE 94.3 Two rows of an image; the transition pixels are marked with a dot.

For example, if the second row is the one being currently encoded, and we have encoded the pixels up to the second pixel, then the assignment of the different pixels is shown in Fig. 94.3. Note that while both the transmitter (encoder) and receiver (decoder) know the positions a_0, b_1, and b_2, the positions a_1 and a_2 are known only to the encoder. Coding is done in one of three modes depending on the relative positions of these pixels. If the run lengths on the current and previous lines are similar, then the distance between a_1 and b_1 would typically be much smaller than the distance between a_0 and a_1. Hence, encoding the distance (a1, b1) can specify the current length. This is called *vertical mode* coding. However, when the distance between a_1 and b_1 is large, that is, if there is no similar run on the previous line, then it is better to encode the runs (a_0, a_1) and (a_1, a_2) using one-dimensional run length coding. This type of encoding is known as *horizontal mode* coding. A third type of coding, known as *pass mode*, is performed when the condition $a_0 \le b_1 < b_2 < a_1$ occurs. That is, we go through two runs in the previous line before completing the current run on the current line. In this case, we simply advance the next pixel to be encoded to a_0', which is the pixel on the current line that is exactly under b_2. Before sending any run lengths, a codeword specifying the mode being used is transmitted. Additional details, including the specific codewords to be used, are given in Yasuda [1980].

Two-Dimensional Predictive Coding

In **predictive coding**, the image is scanned in some fixed order and a prediction is made of the current pixel based on the values of previously transmitted pixels. If the neighborhood employed to perform prediction contains pixels from both the previous and current scan lines, then the technique is referred to as two-dimensional prediction. Since prediction is being made on the basis of pixels known to the receiver, only the prediction error needs to be transmitted. With binary images, the prediction error sequence is again binary, with a 0 indicating no error and a 1 indicating that an error in prediction was made. If the prediction scheme is effective, then the prediction error sequence will contain many more zeroes than ones and, hence, can be coded more efficiently. If we fix the neighborhood used for prediction, then, given a specific image, the optimum prediction function that minimizes the probability of prediction error can be computed. However, such an optimum function varies from image to image. This fact limits the practical utility of predictive schemes. Prediction when used as a preprocessing step, however, can often enhance the performance of other facsimile compression techniques such as run length coding [Yasuda, 1980].

Model-Based Coding

If we impose an nth-order Markov model on a binary source, then its entropy is given by

$$\sum_{k=1}^{n} P(S_k)(P(x=0|s_k)\cdot \log_2 P(x=0|s_k) + P(x=1|s_k)\cdot \log_2 P(x=1|s_k))$$

where $s_1, \ldots, s_n$ are the states and x is the current pixel. When coding binary images, the states s_i are simply taken to be the different bit patterns that can occur in a particular neighborhood of n pixels that occur prior to x. Given the conditional probabilities mentioned, the source can be optimally encoded by using *arithmetic coding* [Rissanen and Langdon, 1979]. Note that since *Huffman coding* uses an integral number

of bits to encode each source symbol, it is of little utility for encoding a binary source unless some form of alphabet extension is performed that blocks individual bits to build an extended alphabet set. Hence, model-based coding was not used for binary images until the early 1980s when the development of sophisticated arithmetic coding techniques enabled the encoding of sources at rates arbitrarily close to the entropies. In fact, it has been proven that model-based arithmetic coding is essentially superior to any other scheme that may encode more than one bit at a time [Langdon and Rissanen, 1981]. In practice, however, we do not have the exact conditional probabilities needed by the model. An estimate of these can be adaptively maintained by keeping track of the counts of black and white pixels encountered so far corresponding to every state. The recently finalized **joint bilevel image processing group** (**JBIG**) standard [Hampel et al., 1992] uses model-based arithmetic coding and significantly outperforms the previous standards for facsimile image compression for a wide variety of test images. The compression ratio obtained is especially superior when encoding half-tone images or mixed documents that contain graphics and text [Arps and Truong, 1994].

Multilevel Facsimile Coding

The techniques we have discussed so far can also be applied to facsimile images that have been digitized using more than two amplitude levels. An image containing 2^n gray levels, with $n \geq 2$, can be decomposed into n different bit planes, each of which can then be compressed by any two-level compression technique. Better compression can be obtained if pixel intensities are expressed by using a **Gray code** representation as compared to the standard binary number representation. This is because the Gray code representation guarantees that two numbers that differ in magnitude by one will differ in their representations in only a single bit position.

The bit-plane approach for coding multilevel images can be taken to its extreme by constructing a two-level bit plane for each of the 2^n gray levels in the image. The 2^n resulting *level planes* can then be compressed by some two-level compression technique. Among the 2^n different level planes, a single arbitrary one need not be encoded as it can be completely determined by the remaining $2^n - 1$ level planes. A comparison of level plane and bit plane coding has been made, and it appears that level plane coding performs better than bit plane coding for images that contain a relatively small number of gray levels (typically, 2–4 b per pixel) [Yasuda et al., 1985].

Another approach to coding multilevel facsimile images is to use one of the many techniques that have been developed for encoding gray scale video images. These techniques have been described in the previous chapter. Such techniques typically perform better than bit-plane encoding and level-plane encoding when the number of gray levels present is relatively large (more than 6 b per pixel).

Compression ratios achieved by lossless techniques are usually very modest. Typical state-of-the-art lossless compression techniques can only achieve between 2–1 and 3–1 compression for images that have been acquired by a camera or some similar sensory device. Hence, it is quite common to use *lossy* or noninformation preserving compression techniques for multilevel images. State-of-the-art lossy techniques can easily achieve more than 15–1 compression while preserving excellent visual fidelity. A description of lossy techniques for multilevel images is given in a later section of this chapter.

Lossy Techniques

Besides multilevel facsimile images, lossy techniques can also be used for two-level images. Two types of lossy techniques have been used on two-level images. The first type consists of a large number of pre- and postprocessing techniques that are primarily used for enhancing subsequent lossless compression of two-level images. The scanning and spatial sampling process inherent in digital facsimile systems invariably leads to a high degree of jaggedness in the boundaries between black and white pixels. This jaggedness, besides reducing the visual quality of the reconstructed document, also severely effects the compression ratios that can be obtained by breaking up long runs of uniform color. Hence, preprocessing techniques that filter out noise would not only improve picture quality but also reduce transmission time. Various preprocessing techniques have been developed, a survey of which is given in Yasuda [1980].

A simple preprocessing technique is to remove isolated black points and bridge small gaps of white pixels between a sequence of black pixels. More sophisticated techniques employ morphological operators to modify local patterns such that subsequent compression is increased. Such techniques, however, may introduce significant degradations in the image and, hence, require postprocessing of the reconstructed image at the receiving end. This fact limits their utility in commercial systems, as they require the facsimile equipment at the receiving end to be equipped with circuitry to perform postprocessing. Two approaches that flip pixels when the change will result in higher compression but without a significant loss in quality are described in Martins and Forchhammer [1997].

An alternative approach to reduce jaggedness in a facsimile image is by modifying the **quantizer** that is used to obtain a two-level image from electrical impulses generated while scanning a document. One such quantizer, called the *notchless bilevel quantizer*, has been proposed [Yasuda, 1980] which adaptively adjusts the quantization level on the basis of preceding pixels. It has been shown that images obtained by using the notchless quantizer have considerably lower entropy and better visual quality.

The second class of lossy compression techniques for facsimile image data attempts to approximate the input image by replacing patterns extracted from the image with appropriate patterns from a library. Such schemes form an important special class of facsimile image compression techniques and are discussed in the next subsection.

Pattern Matching Techniques

Since digitized images used in facsimile transmission often contain mostly text, one way of compressing such images is to perform optical character recognition (OCR) and encode characters by their American Standard Code for Information Interchange (ASCII) code along with an encoding of their position. Unfortunately, the large variety of fonts that may be encountered, not to mention handwritten documents, makes character recognition very unreliable. Furthermore, such an approach limits documents that can be transmitted to specific languages, making international communication difficult. However, an adaptive scheme that develops a library of patterns as the document is being scanned circumvents the problems mentioned. Given the potentially high compression that could be obtained with such a technique, many different algorithms based on this approach have been proposed and continue to be investigated [Pratt et al., 1980; Johnsen, Segen, and Cash, 1983; Witten, Moffat, and Bell, 1994].

Techniques based on pattern matching usually contain a *pattern isolater* that extracts patterns from the document while scanning it in raster order. A pattern is defined to be a connected group of black pixels. This pattern is then matched with the library of patterns that has been accumulated thus far. If no close match is formed, then an encoding of the pattern is transmitted and the pattern is added to the library. The library is empty at the beginning of coding and gradually builds up as encoding progresses.

If a close match for the current pattern is found in the library, then the index of the library symbol is transmitted followed by an encoding of an offset with respect to the previous pattern that is needed to spatially locate the current pattern in the document. Since the match need not be exact, the *residue*, which represents the difference between the current pattern and its matching library symbol, also needs to be transmitted if lossless compression is required. However, if the transmission need not be information preserving, then the residue can be discarded. Most practical schemes discard at least part of the residue in order to obtain high compression ratios.

Although the steps just outlined represent the basic approach, there are a number of details that need to be taken care of for any specific implementation. Such details include the algorithm used for isolating and matching patterns, the encoding technique used for the patterns that do not find a close match in the library, algorithms for fast identification of the closest pattern in the library, distortion measures for closeness of match between patterns, heuristics for organizing and limiting the size of the library, etc. The different techniques reported in the literature differ in the way they tackle the issues listed. For a good survey of such techniques, the reader is referred to Witten, Moffat, and Bell [1994].

The international standard ITU T.88, commonly referred to as JBIG2, uses pattern matching to increase the level of compression.

94.3 International Standards

Several standards for facsimile transmission have been developed over the past few decades. These include specific standards for compression. The requirements on how fast the facsimile of an A4 document (210 × 297 mm) is transmitted has changed over the last two decades, and the CCITT, a committee of the International Telecommunications Union (ITU) of the United Nations, has issued a number of recommendations based on the speed requirements at a given time. The CCITT classifies the apparatus for facsimile transmission into four groups. Although several considerations are used in this classification, if we only consider the time to transmit an A4 size document over the phone lines, the four groups are described as follows:

- **Group 1.** This apparatus is capable of transmitting an A4 size document in about 6 min over the phone lines using an analog scheme. The apparatus is standardized in Recommendation T.2.
- **Group 2.** This apparatus is capable of transmitting an A4 document over the phone lines in about 3 min. Group 2 apparatus also use an analog scheme and, therefore, do not use data compression. The apparatus is standardized in Recommendation T.3.
- **Group 3.** This apparatus uses a digitized binary representation of the facsimile. As it is a digital scheme it can, and does, use data compression and is capable of transmitting an A4 size document in about one minute. The apparatus is standardized in Recommendation T.4.
- **Group 4.** The speed requirement is the same as group 3. The apparatus is standardized in Recommendations T.6, T.503, T.521, and T.563.

CCITT Group 3 and 4—Recommendations T.4 and T.6

The recommendations for group 3 facsimile include two coding schemes: a one-dimensional scheme and a two-dimensional scheme. In the one-dimensional coding mode, a run-length coding scheme is used to encode alternating white and black runs on each scan line. The first run is always a white run. If the first pixel is a black pixel, then we assume that we have a white run of length zero. A special end-of-line (EOL) code is transmitted at the end of every line. Separate Huffman codes are used for the black and white runs. Since the number of runlengths is high, instead of generating a Huffman code for each run length r_1, the run length is expressed in the form

$$r_1 = 64 * m + t \quad \text{for } t = 0, 1, \ldots, 63, \quad \text{and} \quad m = 1, 2, \ldots, 27 \tag{94.1}$$

A run length r_l is then represented by the codes for m and t. The codes for t are called the *terminating codes*, and the codes for m are called the *make-up codes*. If $r_l < 63$, then only a terminating code needs to be used. Otherwise, both a make-up code and a terminating code are used. This coding scheme is generally referred to as a **modified Huffman (MH)** scheme. The specific codewords to be used are prescribed by the standard and can be found in a variety of sources including Hunter and Robinson [1980]. One special property of the codewords is that a sequence of six zeroes cannot result no matter how they are concatenated. Hence, the codeword 0000001 is used to indicate end-of-line.

For the range of m and t given, lengths of up to 1728 can be represented, which is the number of pixels per scan line in an A4 size document. However, if the document is wider, the recommendations provide for those with an optional set of 13 codes. The optional codes are the same for both black and white runs.

The two-dimensional encoding scheme specified in the group 3 standard is known as the **modified READ (MR) coding**. It is essentially a simplification of the READ scheme described earlier. In modified READ, the decision to use the horizontal mode or the vertical mode is made based on the distance a_1b_1. If $|a_1b_1| \leq 3$, then the vertical mode is used, otherwise the horizontal mode is used. The codec also specifies a k factor that no more than $k - 1$ successive lines are two-dimensionally encoded; k is 2 for documents scanned at low resolution and 4 for high resolution documents. This prevents vertical propagation of bit errors to no more than k lines.

The group 4 encoding algorithm, as standardized in CCITT recommendation T.6, is identical to the two-dimensional encoding algorithm in recommendation T.4. The main difference between T.6 and T.4 from the compression point of view is that T.6 does not have a one-dimensional coding algorithm, which means that the restriction specified by the *k* factor as described in the previous paragraph is also not present. This slight modification of the modified READ algorithm has earned it the name **modified modified READ (MMR)**. The group 4 encoding algorithm also does away with the end-of-line code, which was intended to be a form of redundancy to avoid image degradation due to bit errors. Another difference in the group 4 algorithm is the ability to encode lines having more than 2623 pixels. Such run lengths are encoded by using a mark-up code(s) of length 2560 and a terminating code of length less than 2560. The terminating code itself may consist of mark-up and terminating codes as specified by the group 3 technique.

Handling Transmission Errors

If facsimile images are transmitted over the existing switched telephone network, techniques for handling transmission errors are needed. This is because an erroneous bit causes the receiver to interpret the remaining bits in a different manner. With the one-dimensional modified Huffman coding scheme, resynchronization can quickly occur. Extensive studies of the resynchronization period for the group 3 one-dimensional coding schemes have been made. It was shown that in most cases the Huffman code specified resynchronizes quickly, with the number of lost pixels being typically less than 50. For a document scanned at high resolution, this corresponds to a length of 6.2 mm on a scan line.

To handle transmission errors, CCITT has defined an optional error limiting mode and an error-correcting mode. In the error-limiting mode, which is used only with MH coding, each line of 1728 pixels is divided into 12 groups of 144 pixels each. A 12-b header is then constructed for the line, indicating an all white group with a 0 and a nonwhite group with a one. The all white groups are not encoded and the nonwhite groups are encoded separately by using MH. This technique limits the effect of bit errors from propagating through an entire scan line.

The error correction mode breaks up the coded data stream into packets and attaches an error detecting code to each packet. Packets received in error are retransmitted as requested by the receiver but only after the entire page has first been transmitted. The number of retransmissions for any packet is restricted to not exceed four.

The Joint Bilevel Image Processing Group (JBIG)

The JBIG is a joint experts group of the International Standards Organization (ISO), International Electrotechnical Commission (IEC), and the CCITT. This experts group was jointly formed in 1988 to establish a standard for the progressive encoding of bilevel images. The work of the committee has led to two standards, the ITU T.82 standard commonly known as the JBIG standard and the ITU T.88 standard commonly referred to as the JBIG2 standard. While both standards address the same problem, the approaches taken are markedly different. In the following subsection we describe the JBIG standard. The JBIG2 standard is discussed in the subsection after that.

The JBIG Standard—T.82

The JBIG standard can be viewed as a combination of two algorithms: a **progressive transmission** algorithm and a lossless compression algorithm. Each of these can be understood independently of the other.

Lossless Compression

The lossless compression algorithm uses a simple *context model* to capture the structure in the data. A particular arithmetic coder is then selected for each pixel based on its context. The context is made up of neighboring pixels. For example, in Fig. 94.4, the pixel to be coded is marked **X**, whereas the pixels to be used as the context are marked **O** or **A**. The **A** and **O** pixels are previously encoded pixels and are

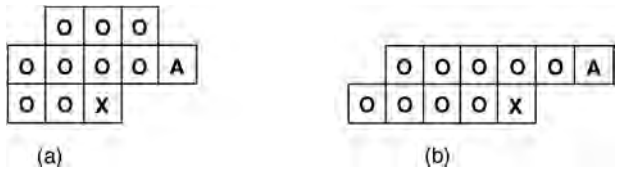


FIGURE 94.4 (a) Three line and (b) two line model template for lowest resolution layer.

available to both encoder and decoder. The **A** pixel can be moved around in order to better capture any structure that might exist in the image. This is especially useful in half-toned images in which the **A** pixels are used to capture the periodic structure. The location and movement of the **A** pixel is transmitted to the decoder as side information.

The arithmetic coder specified in the JBIG standard is a special binary adaptive arithmetic coder known as the QM coder. The QM coder is a modification of an adaptive binary arithmetic coder called the Q coder [Pennebaker and Mitchell, 1993], which in turn is an extension of another binary adaptive arithmetic coder called the *skew* coder [Langdon and Rissanen, 1981]. Instead of dealing directly with the 0s and 1s put out by the source, the QM coder maps them into a more probable symbol (MPS) and less probable symbol (LPS). If 1 represents black pixels and 0 represents white pixels, then, in a mostly black image, 1 will be the MPS, whereas in an image with mostly white regions, 0 will be the MPS. To make the implementation simple, the JBIG committee recommended several deviations from the standard arithmetic coding algorithm. The update equations in arithmetic coding that keep track of the subinterval to be used for representing the current string of symbols involve multiplications, which are expensive in both hardware and software. In the QM coder, expensive multiplications are avoided and rescalings of the interval take the form of repeated doubling, which corresponds to a left shift in the binary representation. The probability q_c of the LPS for context C is updated each time a rescaling takes place and the context C is active. An ordered list of values for q_c is kept in a table. Every time a rescaling occurs, the value of q_c is changed to the next lower or next higher value in the table, depending on whether the rescaling was caused by the occurrence of an LPS or MPS. In a nonstationary situation, it may happen that the symbol assigned to an LPS actually occurs more often than the symbol assigned to an MPS. In this situation, the assignments are reversed; the symbol assigned the LPS label is assigned the MPS label and vice versa. The test is conducted every time a rescaling takes place. The decoder for the QM coder operates in much the same way as the encoder, by mimicking the encoder operation.

Progressive Transmission

In progressive transmission of an image, a low-resolution representation of the image is first sent. This low-resolution representation requires very few bits to encode. The image is then updated, or refined, to the desired fidelity by transmitting more and more information. To encode an image for progressive transmission, we need to create a sequence of progressively lower resolution images from the original higher resolution image. The JBIG specification recommends generating one lower resolution pixel for each two by two block in the higher resolution image. The number of lower resolution images (called layers) is not specified by JBIG. However, there is a suggestion that the lowest resolution image is roughly 10–25 dots per inch (dpi). There are a variety of ways in which the lower resolution image can be obtained from a higher resolution image, including sampling and filtering. The JBIG specification contains a recommendation against the use of sampling. The specification provides a table-based method for resolution reduction. The table is indexed by the neighboring pixels shown in Fig. 94.5 in which the circles represent the lower resolution layer pixels and the squares represent the higher resolution layer pixels.

Each pixel contributes a bit to the index. The table was formed by computing the expression

$$4e + 2(b + d + f + h) + (a + c + g + i) - 3(B + C) - A$$

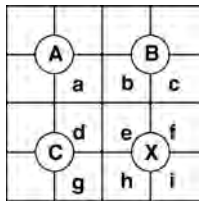


FIGURE 94.5 Pixels used to determine value of lower level pixel.

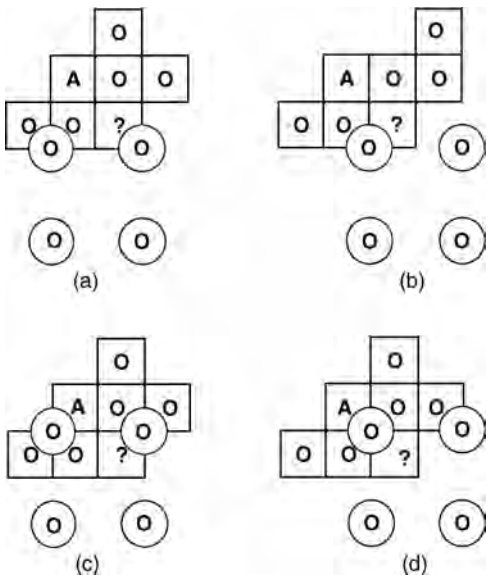


FIGURE 94.6 Contexts used in the coding of higher resolution layers.

If the value of this expression is greater than 4.5, the pixel X is tentatively declared to be 1. The table has certain exceptions to this rule to reduce the amount of edge smearing generally encountered in a filtering operation. There are also exceptions that preserve periodic patterns and dither patterns.

When the progressive mode is used for transmission, information from lower resolution layers can be used to improve compression. This is done by including pixels from lower resolution layers in the context used to encode a pixel in the current layer. The contexts used for coding the lowest resolution layer are those shown in Fig. 94.4. The contexts used in coding the higher resolution layer are shown in Fig. 94.6. In each context, 10 pixels are used. If we include the 2 b required to indicate which context template is being used, 12 b will be used to indicate the context. This means that we can have 4096 different contexts.

The standard does not impose any restrictions on D, the number of resolution layers that are constructed. Indeed, D can be set to zero if progressive coding is of no utility. In this case, coding is said to be *single-progression sequential* or just *sequential*. The algorithm allows some degree of compatibility between the progressive and sequential modes. Images that have been encoded in a progressive manner can be decoded sequentially, that is, as just one layer. Images that have been encoded sequentially, however, cannot be decoded progressively. This compatibility between progressive and sequential modes is achieved by partitioning an image into *stripes*, with each stripe representing a sequence of image rows with user defined height. If the image has multiple bit planes, then stripes from each bit plane can be interleaved. Each stripe is separately encoded, with the user defining the order in which these stripes are concatenated into the output data stream.

The JBIG2 Standard—T.88

Unlike the JBIG standard, the JBIG2 standard, provides for lossy compression of bilevel images, thus providing significantly higher compression. Along with facsimile compression, the target applications of the JBIG2 standard include document storage and archiving, image compression on the Internet, wireless data transmission, and print spooling.

The algorithm first partitions the page into three kinds of regions: a text region, a half-tone region, and an "other" region. The text region contains characters of approximately the same size arranged in rows and columns. In the JBIG2 terminology, these characters are referred to as *symbols*. These symbols can be encoded exactly in a lossless fashion or in a lossy fashion by sending the index in a dictionary of symbols where the match to an entry in the dictionary can be approximate. The halftone region will generally be a region on the page where a grayscale image has been dithered to generate a bilevel representation. The periodic bitmap cells in this region are called *patterns*. The remaining region consists of what is left over. The data contained in this region is referred to as *generic* data. The regions can overlap on a page and the standard provides a syntax for recombining the overlapping regions to reconstitute the page.

The standard requires that all the information be organized into segments where each segment consists of a segment header, a data header, and segment data. The compressed representation can be organized in two ways: sequential or random access. In the sequential organization, the segments are arranged in sequential order with each segment containing all three components. In the random access organization, all the segment headers are collected in the beginning of the file followed by the data and data headers organized in the same order as the segment headers.

The standard assumes that the decoded data will be stored before display or printing. This allows for the existence of internal buffers and considerable processing. As the various regions are decoded, they are stored in the *page buffer*. When the entire page has been decoded, it can be sent to an output file or printing or display device. The standard does not preclude sending out partial pages to accommodate progressive transmission schemes or in response to user input. There are also auxiliary buffers which may hold structures such as dictionaries or other intermediate results.

The standard specifies a control decoding procedure that decodes segment headers. The decoding results in the identification of the segment type, which then results in the invocation of the appropriate decoding procedures.

Generic Region Decoding

There are two main generic decoding procedures. One is for decoding regions encoded using the MMR algorithm described previously. The second decoding algorithm is similar to the algorithm used in JBIG for decoding the bottom layer. Which algorithm is to be used is indicated in the segment header along with the size of the rectangular region of pixels to be decoded.

The standard also provides a generic region refinement procedure that is used to refine an existing region of pixels in the buffer. The refinement procedure uses context based arithmetic coding. The contexts are made up of decoded pixels from the refinement pass as well as pixels from the region being refined.

Text Region Decoding

The text region decoding procedure is a dictionary-based procedure. The decoder first creates a dictionary of symbols which is an indexed set of symbol bitmaps. The symbol bitmaps may have been encoded using the generic procedure or they may be constructed as a refinement of a symbol, or as an aggregation of multiple symbols, already in the dictionary. The text region segment contains indices to the symbol dictionary and the location at which the symbols are to be placed.

Halftone Region Decoding

The halftone decoding functions in a manner similar to the text region decoding procedure. First, a pattern dictionary consisting of fixed size bitmaps is created. These are then placed using location and index information from the halftone region segment.

TABLE 94.1 Comparison of Binary Image Coding Schemes [Arps and Truong, 1994]

Source Description	Original Size, pixels	MH, bytes	MR, bytes	MMR, bytes	JBIG, bytes
Letter	4352 × 3072	20,605	14,290	8531	6682
Sparse text	4352 × 3072	26,155	16,676	9956	7696
Dense text	4352 × 3072	135,705	105,684	92,100	70,703

TABLE 94.2 Comparison of Different Bilevel Coding Standards

Source Description	Group 3, bits	Group 4, bits	JBIG, bits	JBIG2, bits
9200 dpi bilevel images	647,348	400,382	254,223	236,053
9300 dpi bilevel images	1,344,030	749,674	588,418	513,495
4 mask bilevel images	125,814	67,116	54,443	51,453

Comparison of MH, MR, MMR, JBIG, and JBIG2

In the previous subsection we have seen several different facsimile coding algorithms that are part of different international standards. As we might, expect the JBIG2 algorithm performs better than the JBIG algorithm, which performs better than the MMR algorithm, which performs better than the MR algorithm, which in turn performs better than the MH algorithm. The level of complexity also follows the same trend, though one could argue that MMR is actually less complex than MR. A comparison of the first four schemes for some facsimile sources is shown in Table 94.1.

The modified READ algorithm was used with $K = 4$, whereas the JBIG algorithm was used with an adaptive 3 line template and adaptive arithmetic coder to obtain the results in this table. As we go from the one-dimensional MH coder to the two-dimensional MMR coder, we get a factor of two reduction in file size for the sparse text sources. We get even further reduction when we use an adaptive coder and an adaptive model, as is true for the JBIG coder. When we come to the dense text, the advantage of the two-dimensional MMR over the one-dimensional MH is not as significant because the amount of two-dimensional correlation becomes substantially less.

The compression schemes specified in T.4 and T.6 break down when we try to use them to encode half-tone images. This is to be expected as the model that was used to develop these coding schemes is not valid for half tone images. The JBIG algorithm, with its adaptive model and coder, suffers from no such drawbacks and performs well for half-tone images as well [Arps and Truong, 1994].

Another set of comparisons is shown in Table 94.2.

The source images are images scanned using different resolutions and four mask images that were generated from segmenting a page of an electronics catalog. The results clearly show the superiority of the two JBIG algorithms. However, there does not seem to be much difference between the two JBIG algorithms. One should note, however, that these comparisons are only for lossless compression. If lossy compression were allowed, the result could be substantially different.

Defining Terms

Compression ratio: Size of original data/size of compressed data.

Facsimile: The process by which a document is optically scanned and converted to electrical signals.

Facsimile image: The quantized digital image corresponding to the document that has been input to a facsimile machine.

Fax: Abbreviation for facsimile.

Gray code: A binary code for integers in which two integers that differ in magnitude by one differ in only one bit position.

Group 3: Facsimile apparatus capable of transmitting an A4 size document in about one minute. The apparatus is standardized in Recommendation T.4.

Group 4: Facsimile apparatus for sending a document over public data networks with virtually error-free reception. Standardized in Recommendations T.6, T.503, T.521, and T.563.

Joint Bilevel Image Processing Group (JBIG): A group from International Standards Organization (ISO), International Electrotechnical Commission (IEC), and the CCITT. This expert group was jointly formed in 1988 to establish a standard for the progressive encoding of bilevel images. The term JBIG is also used to refer to the coding algorithm proposed by this committee.

Modified Huffman code (MH): One-dimensional coding scheme used by Group 3 equipment.

Modified modified READ code (MMR): Two-dimensional coding scheme used by Group 4 equipment.

Modified READ code (MR): Two-dimensional coding scheme used by Group 3 equipment.

Predictive coding: A form of coding where a prediction is made for the current event based on previous events and the error in prediction is transmitted.

Progressive transmission: A form of transmission in which a low-resolution representation of the image is first sent. The image is then updated, or refined, to the desired fidelity by transmitting more and more information.

Quantizer: The process of converting analog data to digital form.

References

Arps, R. and Troung, T. 1994. Comparison of international standards for lossless still image compression. *Proc. IEEE,* 82(6):889–899.

Capon, J. 1959. A probabilistic model for run-length coding of pictures. *IRE Trans. Inf. Th.* (Dec.):157–163.

Hampel, H. et al. 1992. Technical features of the JBIG standard for progressive bi-level image compression. *Sig. Proc.,* 4(2):103–111.

Hunter, R. and Robinson, A.H. 1980. International digital facsimile standards. *Proc. IEEE,* 68(7):855–865.

Johnsen, O., Segen, J., and Cash, G.L. 1983. Coding of two-level pictures by pattern matching and substitution. *Bell Sys. Tech. J.,* 62(8):2513–2545.

Kunt, M. and Johnsen, O. 1980. Block coding of graphics: a tutorial review. *Proc. IEEE,* 68(7):770–786.

Langdon, G.G., Jr. and Rissanen, J. 1981. Compression of black-white images with arithmetic coding. *IEEE Trans. Commun.,* COM-29(6):858–867.

Martins, B. and Forchhammer, S. 1997. Lossless/lossy compression of bi-level images. In *Proceedings of the IS&T/SPIE Symposium on Electronic Imaging: Science and Technology* 1997, 3018, 38–49, 1997.

Pennebaker, W.B. and Mitchell, J.L. 1993. *JPEG Still Image Compression Standard.* Van Nostrand Reinhold, New York.

Pratt, W., Capitant, P., Chen, W., Hamilton, E., and Wallis, R. 1980. Combined symbol matching facsimile data compression system. *Proc. IEEE,* 68(7):786–796.

Rissanen, J.J. and Langdon, G.G. 1979. Arithmetic coding. *IBM J. Res. Dev.,* 23(2):149–162.

Witten, I., Moffat, A., and Bell, T.C. 1994. *Managing Gigabytes: Compressing and Indexing Documents and Images*. Van Nostrand Reinhold, New York.

Yasuda, Y. 1980. Overview of digital facsimile coding techniques in Japan. *Proc. IEEE,* 68(7):830–845.

Yasuda,Y., Yamakazi, Y., Kamae, T., and Kobayashi, K. 1985. Advances in fax. *Proc. IEEE,* 73(4):707–731.

Further Information

FAX—Facsimile Technology and Applications Handbook, 2nd edition, by K. McConnell, D. Bodson, and R. Schaphorst, published by Artech House, Norwood, MA, is an excellent single source on various aspects of facsimile technology, including compression.

Two comprehensive surveys by Yasuhiko Yasuda et al. and Yasuhiko Yasuda on coding techniques for facsimile have appeared in *Proceedings of the IEEE* in 1980 and 1985, respectively. These surveys summarize most of the research that has been conducted on facsimile coding and contain an extensive list of references. In addition, the two issues that they appear in, July 1980 and April 1985, are both special issues on facsimile coding.

For a description of the CCITT standards, the best sources are the original documents containing the recommendations:

Standardization of Group 3 Facsimile Apparatus for Document Transmission, Recommendation T.4, 1980.

Facsimile Coding Schemes and Coding Control Functions for Group 4 Facsimile Apparatus, Recommendation T.6, 1984.

ITU-T Recommendation T.82 Information Technology—Progressive Bi-Level Image Compression, JBIG Recommendation, 1993.

ITU-T Recommendation T.88 Information Technology—Lossy/Lossless Coding Bi-Level Images, JBIG2 Recommendation, 2000.

These documents can be ordered from the International Telecommunication Union, Place Des Nations 1211, Geneva 20, Switzerland. They are also available from Omnicom, Phillips Business Information, 1201 Seven Locks Road, Suite 300, Potomac, Maryland 20854, fax: 1-800-666-4266.

A more recent survey by Arps and Troung (see reference list) compares the performance of different standards.

95

Speech

Boneung Koo
Kyonggi University

95.1 Introduction

Speech compression refers to the compact representation of speech signals, and speech coding refers to the digital representation of speech signals. Since the primary goal of speech coding is to compress the signal, that is, to reduce the number of bits required to represent it, the two terms, speech compression and speech coding, can be used interchangeably.

Coded speech will be transmitted or stored for a specific application. As the number of bits used in the representation of a signal is reduced, the effective bandwidth of the transmission channel will be increased and the memory space will be reduced. Coded speech must be decoded back to analog form at the receiver. In this coding/decoding process, some amount of distortion results inevitably in the reconstructed speech. Hence, the ultimate goal of speech coding is to compress the signal while maintaining a prescribed level of reconstructed speech quality. Various coding algorithms differ in how to select signals or parameters that represent the speech efficiently. Those selected signals and/or parameters are then quantized and transmitted to the receiver for decoding.

Typical applications of speech coding include the conventional telephone network, wireless personal communication systems, and secure military communications. Speech over the Internet has applications in multimedia, videoconferencing, and Internet phones. Even though prices of processors, memories, and transmission media have become cheaper and processing speed higher than in the past, the importance of speech coding or compression has not diminished because demands for more efficient use of hardware resources are ever increasing.

Since the introduction of 64 kb/s log-pulse-code modulation (PCM) in the AT&T long-distance telephone network in the 1970s, speech coding has been a major research area in telecommunications. In 1984, adaptive differential PCM (ADPCM) was adopted by the Consultative Committee on International Telephony and Telegraphy (CCITT) [now, the International Telecommunications Union—Telecommunications (ITU-T)] as a 32 kb/s international standard. Since the mid-1980s, there have been tremendous activities

in applications and standardization of medium-to-low rate speech coding. An important contribution was made by code-excited linear prediction (CELP), introduced in 1985. Advances in VLSI technology have made it possible to implement complicated coding and signal processing algorithms in real-time. Versions of CELP in the range of 1.6–4.8 kb/s have been adopted in a variety of standards and cellular networks around the world.

Principles and applications of typical speech coding algorithms are reviewed in this chapter. Properties of the speech signal are briefly reviewed, followed by a classification of coder types. Principles of quantization and coding algorithms are then presented. This presentation is not exhaustive; emphasis is given to coders employed in standards. Criteria to be considered for speech coder evaluation and subjective and objective measures for speech quality evaluation are presented. A summary of speech coding standards is given, followed by concluding remarks.

95.2 Properties of Speech Signals

Since speech signal is nonstationary with slowly time-varying characteristics, a speech signal is processed in short time segments, typically of 5–30 ms. Each segment or frame can be classified as voiced, unvoiced, or silence. Voiced sounds are generated by passing quasiperiodic air pulses through the glottis, causing vibrations in the vocal cords. This results in a quasiperiodic nature in the time domain and harmonic structure in the frequency domain. An unvoiced sound is generated by constricting the vocal tract and forcing air through the constriction to create turbulence. This leads to a near-white noise source to excite the vocal tract such that the unvoiced segment looks like random noise in the time domain and broadband in the frequency domain.

In its simplest form, the speech production system can be modeled as a source-system model, that is, a linear system driven by an excitation source. In the conventional two-state approach, the excitation is a train of pulses for voiced segments and white noise for unvoiced segments. In the more recent mixed excitation approach, a segment can be a consequence of both voiced and unvoiced excitation.

In general, energy of the voiced segment is higher than that of the unvoiced one. A speech segment that is not a consequence of speech activity is classified as silence or nonspeech. In telephone speech, approximately 50% of the talk time is known to be silence. This fact is utilized in some wireless cellular systems to increase the effective channel bandwidth via voice activity detection.

The period of the quasiperiodicity in the voiced segment is referred to as the pitch period in the time domain or the pitch or fundamental frequency in the frequency domain. The pitch of the voiced segments is an important parameter in many speech coding algorithms. It can be identified as the periodicity of the peak amplitudes in the time waveform and the fine structure of the spectrum. The pitch frequency of men and women usually lies in the range 50–250 Hz (4–20 ms), and 120–500 Hz (2–8.3 ms), respectively.

The bandwidth of speech rarely extends beyond 8 kHz. In wideband speech coding, bandwidth is limited to 7 kHz, and speech is sampled at 16 kHz. In telephony, bandwidth is limited to below 4 kHz (0.2–3.4 kHz, typically) and speech is sampled, usually, at 8 kHz. It is assumed throughout this chapter that input speech is bandlimited to below 4 kHz and sampled at 8 kHz unless otherwise specified.

The coded speech quality can be roughly classified into the following four categories. *Broadcast* quality refers to wideband speech, *toll* or *network* to narrowband (telephone) analog speech, *communications* to degraded but natural and highly intelligible quality, and, finally, *synthetic* to unnatural but intelligible quality typically represented by the linear predictives coding (LPC) **vocoder**. Speech quality produced by coders to be discussed in this chapter belongs to one of the last three categories.

95.3 Types of Speech Coding Algorithms

Speech coders studied so far can be classified broadly into three categories: **waveform coders**, vocoders, and **hybrid coders**. Waveform coders are intended directly at approximating the original waveform. The reconstructed sound may or may not be close to the original. On the other hand, vocoders are primarily aimed at approximating the sound and, consequently, the reconstructed waveform may or may not be

TABLE 95.1 A Classification of Speech Coders

Type	Coding Algorithm
Waveform Coders	PCM (pulse-code modulation), APCM (adaptive PCM) DPCM (differential PCM), ADPCM (adaptive DPCM) DM (delta modulation), ADM (adaptive DM) CVSD (continuously variable-slope DM) APC (adaptive predictive coding) RELP (residual-excited linear prediction) SBC (subband coding) ATC (adaptive transform coding)
Hybrid coders	MPLP (multipulse-excited linear prediction) RPE (regular pulse-excited linear prediction) VSELP (vector-sum excited linear prediction) CELP (code-excited linear prediction) ACELP (algebraic CELP)
Vocoders	Channel, Formant, Phase, Cepstral, or Homomorphic LPC (linear predictive coding) MELP (mixed-excitation linear prediction) STC (sinusoidal transform coding) MBE (multiband excitation), Improved MBE, Advanced MBE

close to the original. Coders that employ features of both waveform coders and vocoders are called hybrid coders. A classification of coders is shown in Table 95.1.

In waveform coders, all or parts of the original waveform are quantized for transmission. For example, in PCM, input speech sample itself is quantized, and in DPCM and ADPCM, the prediction residual is quantized. Also, adaptive delta modulation (ADM), continuously variable-slope delta modulation (CVSD), adaptive predictive coding (APC), residual-excited linear prediction (RELP), subband coding (SBC), and adaptive transform coding (ATC) are waveform coders. Coders that employ predictors, such as ADPCM, ADM, APC, and RELP, are called predictive coders. SBC and ATC are frequency-domain coders in that coding operations can best be described in the frequency domain. Speech quality produced by waveform coders is generally high, however, at higher bit rates than vocoders.

In vocoders, the speech production model or other acoustic feature parameters that represent perceptually important elements of speech are estimated, quantized, and transmitted to the receiver, where speech is reconstructed based on these parameters. For this reason, vocoders are also called parametric coders and are speech specific in many cases. Vocoder types include channel, formant, phase, cepstral, hormomorphic vocoders, and LPC, sinusoidal transform coding (STC), and multiband excitation (MBE) vocoders. Vocoders can generally achieve higher compression ratios than waveform coders; however, they are known for artificial or unnatural speech quality, except for the recent improvements in STC, MBE, and MELP.

In hybrid coders, the high compression efficiency of vocoders and high-quality speech reproduction capability of waveform coders are combined to produce good quality speech at medium-to-low bit rates. The so-called **analysis-by-synthesis** coders such as MPLP, RPE, VSELP, and CELP are hybrid coders.

95.4 Quantization

The input to the quantizer is generally modeled as a random process with a continuous amplitude. The function of the quantizer is to produce an output with a discrete amplitude such that it can be encoded with a finite number of bits. Depending on the coding algorithm, the quantizer input can be raw speech samples, the prediction residual, or some parameters estimated at the transmitter and required at the receiver to reconstruct the speech. There are two types of quantizers: scalar and vector.

The scalar quantizer operates on a sample-by-sample basis. Parameters of a scalar quantizer are dynamic range, stepsize, input/output step values, and the number of output levels or the number of bits required to represent each output level. The difference between the input and output is called quantization noise.

The quantizer should be designed such that the quantization noise power is minimized for the specified class of input signals. Usually, quantizers are designed according to the probability density function (pdf) of the signal to be quantized. The uniform quantizer works best for a uniform input. Otherwise, a logarithmic or pdf-optimized nonuniform quantizer should be used. Quantizer adaptations to changing input characteristics can take several forms, however, the stepsize adaptation to account for the changing dynamic range or variance is the most common. Detailed information can be found in Jayant and Noll [1984].

According to rate distortion theory, coding vectors is more efficient than coding scalars, and the efficiency is increased as the dimension of the vector is increased at the expense of computational complexity. The input to the vector quantizer (VQ) is a vector formed from consecutive samples of speech or prediction residual or from model parameters. The incoming vector is compared to each codeword in the codebook, and the address of the closest codeword is selected for transmission. The distance is measured with respect to a distortion measure. The simplest and the most common distortion measure is the sum of squared errors. Entries of the codebook are selected by dividing the vector space into nonoverlapping and exhaustive cells. Each cell, called the centroid, is assigned a unique address in the codebook. The quantization process is to find the address of the centroid of the cell that the input vector falls in.

Two important issues associated with vector quantization are the codebook design and the code search procedure. An iterative codebook design procedure of importance is called the Linde, Buzo, and Gray (LBG) algorithm, which involves an initial guess and iterative improvement by using training vectors. The complexity of VQ increases exponentially with the vector dimension and codebook size, however, the code search time can be significantly reduced by using a structured codebook. See Makhoul et al. [1985] for more information.

95.5 Predictive Coders

Speech samples have redundancy due to correlation among samples. In predictive coders, redundancy is removed by using predictors. The difference between the input sample and the predictor output is called the prediction residual. There are two types of predictors widely used in speech coders for removing the redundancy: the short-term predictor (STP) and the long-term predictor (LTP). The STP and LTP model the short-term correlation and the long-term correlation (due to pitch), respectively. The synthesis filter associated with the STP reproduces the spectral coarse or formant structure (spectral envelope). The synthesis filter associated with the LTP is used to reconstruct the spectral fine structure of the voiced segment.

The STP can be used to obtain the prediction residual $e(n)$ shown in Fig. 95.1(a). The difference equation and the transfer function of the loop in Fig. 95.1(a) are $e(n) = s(n) - s_p(n)$ and $1 - A(z)$, respectively. The filter $A(z)$ produces a linear combination of past samples according to the relation

$$s_p(n) = \sum_{i=1}^{p} a_i s(n-i)$$

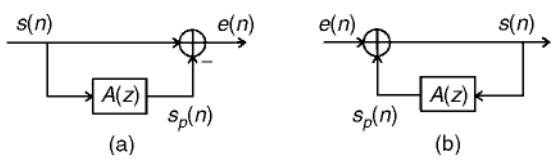


FIGURE 95.1 Linear predictor applications: (a) analysis and (b) synthetics.

and its transfer function is

$$A(z) = \sum_{i=1}^{p} a_i z^{-i} \tag{95.1}$$

where p is the prediction order and the a_i are linear prediction (LP) coefficients. Figure 95.1(b) is the STP synthesis loop, where the original signal is reconstructed by inserting back the redundancy. The transfer function of the synthesis loop in Fig. 95.1(b) is $\{1 - A(z)\}^{-1}$. A variety of algorithms are available for estimating the LP coefficients, and details can be found in Rabiner and Schafer [1978].

Other types of predictors are the all-zero predictor and the pole-zero predictor. The output of an all-zero predictor is a linear combination of past prediction residuals. The pole-zero predictor is a combination of the all-pole predictor and the all-zero predictor. The pole-zero predictor, employed in ADPCM adopted in the CCITT standard, has been known to best model the speech signal. In practice, however, the all-pole predictor has been more widely used because of its adequate performance and computational advantage.

Redundancy due to long-term correlations can be reduced by using the LTP. A general form of the LTP transfer function is given by

$$A_L(z) = \sum_{i=-j}^{J} b_i z^{-i-t} \tag{95.2}$$

where τ is the pitch period in samples and the b_i are parameters called the pitch gain. Here j is usually a small integer of 0 or 1. Several algorithms are available for open-loop estimation of the LTP parameters. In analysis-by-synthesis coding, closed-loop estimation can be used to produce significantly better results than open-loop estimation [Spanias, 1994].

In DPCM coding, the prediction residual instead of the original speech is computed, quantized, and transmitted. The STP is used to compute prediction residual. Because the variance of the prediction residual should be substantially smaller than that of the original speech, a quantizer with fewer bits can be used with little drop in reconstructed speech quality. The transmission rate or the data rate is thus determined by the number of bits required to represent the number of quantizer output levels. The decoder synthesizes the output by adding the quantized prediction residual and the predicted value.

Adaptation of the quantizer stepsize and predictor parameters improves the reconstructed speech quality, which is called ADPCM. The predictor can be forward adaptive or backward adaptive. In forward adaptation, predictor parameters are estimated at the encoder and transmitted to the receiver as side information. In backward adaptation, predictor parameters are estimated from the reconstructed data, which are also available at the receiver, and thus are not required to be transmitted. A variety of adaptive prediction algorithms can be found in Gibson [1980, 1984].

ADPCM was employed in the 32-kb/s CCITT standard, G.721. The coder consists of a 4-b adaptive quantizer and a backward-adaptive predictor with 2 poles and 6 zeros. The coder provides for toll quality comparable to that of 64-kb/s log PCM. The G.721 was replaced by the G.726, which operates at one of the four rates, 40-, 32-, 24-, and 16-kb/s.

DM can be viewed as DPCM with a 1-b quantizer and a first-order predictor. ADM employs an adaptive stepsize to track varying statistics in the input signal. Because of its simple structure in quantization and prediction, ADM generally requires a higher sampling rate than ADPCM. ADM generates very high-quality speech in the range of 32–48 kb/s using a very simple algorithm. A 48-kb/s ADM can produce toll-quality speech with a mean opinion score (MOS) rating of 4.3. A well-known version of ADM is CVSD. More information on ADM and CVSD can be found in Jayant and Noll [1984].

In APC, the LTP or pitch predictor is used in addition to the STP in DPCM, reducing the variance of the prediction residual and allowing fewer bits to be used in the quantizer. Except for the LTP loop, the APC encoder in Fig. 95.2(a) is the same as the ADPCM encoder with an all-pole predictor. Transfer functions of the STP and LTP are shown in Eqs. (95.1) and (95.2), respectively. The prediction residual

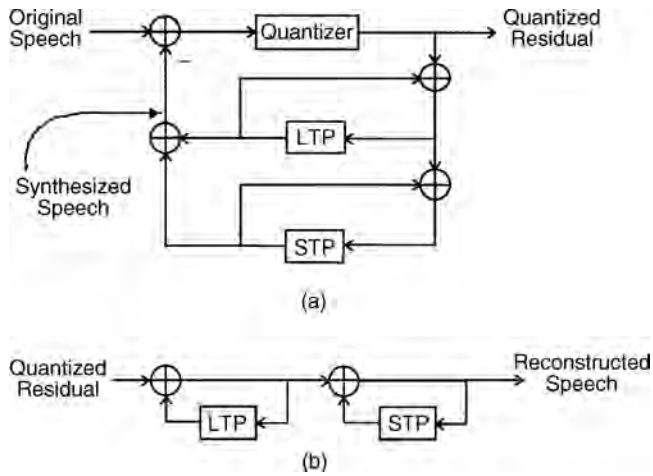


FIGURE 95.2 APC coder: (a) encoder and (b) decoder.

computed by using the predictors is encoded by a scalar quantized on a sample-by-sample basis and transmitted to the receiver. Predictor parameters are estimated, quantized, and also transmitted to the receiver as side information. At the receiver shown in Fig. 95.2(b), speech is reconstructed by passing the quantized prediction residual through the LTP synthesizer and the STP synthesizer. APC provides toll quality speech at 16 kb/s and communications quality at 9.6 kb/s. The International Maritime Satellite Organization (Inmarsat)-B standard employs a 16-kb/s APC coder.

RELP has been studied to reduce the number of bits required for encoding the residual in APC. RELP is basically the same as APC except that only the low-frequency part of the residual is transmitted to the receiver with a decimation factor of 3 or 4. This is based on the observation that perceptually important pitch information is in the low-frequency band. At the receiver, the residual is recovered by nonlinear interpolation and used as an excitation to synthesize speech. The estimated acceptable subjective performance of RELP is limited to 9.6 kb/s or higher.

95.6 Frequency-Domain Coders

SBC and ATC are classified as frequency-domain coders in that coding operations rely on a frequency-domain representation of signals. The basic idea is to divide the speech spectrum into several frequency bands and encode each band separately.

In SBC, the signal is divided into several subbands by filter banks. Each band is then low-pass translated, decimated to the Nyquist sampling rate, encoded, and multiplexed for transmission. The number of bits allocated to each subband is determined by perceptual importance. Usually, more bits are assigned to lower bands to preserve pitch and format information. Any coding scheme such as APCM, ADPCM, or VQ can be used to encode each subband. At the receiver, subband signals are demultiplexed, decoded, interpolated, translated back to their original spectral bands, and then summed up to reproduce the speech. The filter bank is a very important consideration for the implementation and speech quality of the SBC. An important class of filters is quadrature mirror filters (QMF), which deliberately allows aliasing between subbands in the analysis stage and eliminates it in the reconstruction stage, thus resulting in perfect reconstruction except for quantization noise.

An SBC/APCM coder operating at 16 and 24 kb/s was employed in the AT&T voice store-and-forward system, where a five-band nonuniform QMF is used as a filter bank and APCM coders are used as encoders.

More complicated frequency analysis than in SBC is involved in ATC. In ATC, a windowed frame of speech samples is unitary transformed into the frequency domain, and the transform coefficients are

quantized and transmitted. At the receiver, the coefficients are inverse transformed and joined together to reproduce the synthesized speech. The unitary transform produces decorrelated spectral samples, the variances of which are slowly time varying so that redundancy can be removed. The Karhunen–Loeve transform (KLT) is the optimal unitary transform that maximally decorrelates the transform components, however, the discrete cosine transform (DCT) is the most popular in practice because of its near-optimality for first-order autoregressive (AR) processes and computational advantage. The DCT can be computed efficiently by using the fast fourier transform (FFT) algorithm.

The transform block size is typically in the range of 128–256, which is much greater than the number of subbands of the SBC, which is usually 4–16. Bit allocation to each band or transform coefficient can be fixed or adaptive. More bits are allocated to the coefficient that has more variance. Adaptive bit allocation is incorporated in ATC. ATC provides toll quality speech at 16 kb/s and communications quality at 9.6 kb/s.

95.7 Analysis-by-Synthesis Coders

A common feature of coders described so far is that they operate in open-loop fashion. Hence, there is no look back and control over the distortion in the reconstructed speech. Such coders are called **analysis-and-synthesis** coders.

In the **analysis-by-synthesis** class of coders, an excitation sequence is selected in a closed-loop fashion by minimizing the perceptually weighted error energy between the original speech and the reconstructed speech. The conceptual block diagram of an analysis-by-synthesis coder is shown in Fig. 95.3. The coder consists of the excitation generator, pitch synthesis filter, linear prediction synthesis filter, and perceptual weighting filter. The pitch and the LP synthesis filters are, respectively, the same as the LTP and the STP synthesis loops shown in Fig. 95.2(b). The weighting filter is used to perceptually shape the error spectrum by de-emphasizing noise in the formant nulls. The transfer function of the weighting filter often used is given by

$$W(z) = \frac{1 - A(z/\gamma_1)}{1 - A(z/\gamma_2)}, \quad 0 < \gamma_2 < \gamma_1 \leq 1$$

where $A(z)$ is the transfer function of the STP in Eq. (95.1). The value of (γ_1, γ_2) is typically (1, 0.8) in conventional CELP coders and (0.9, 0.4) in the low-delay CELP coder to be described later.

The excitation generator forms or selects the excitation sequence that minimizes the perceptually weighted error criterion, which is usually a weighted sum of squares. The STP and LTP parameters are updated, typically every 10–30 ms and 5–15 ms, respectively. The STP is usually all-pole with a typical order of 10, and parameters are usually converted to another format (line spectrum pairs, for example) and vector quantized for coding efficiency. The pitch lag is usually searched over the range of 20–147 in samples (for 8-kHz sampling). There are several versions of coders depending on the type of excitation model, namely, MPLP, RPE, and CELP coders.

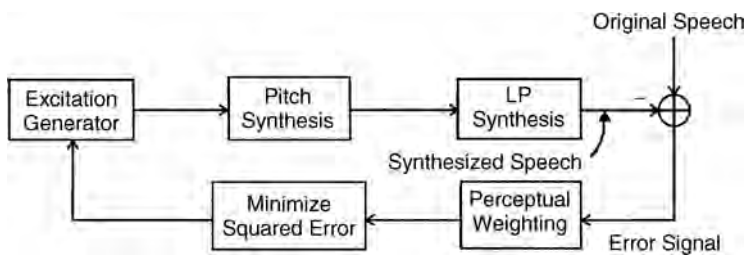


FIGURE 95.3 Analysis-by-synthesis coder.

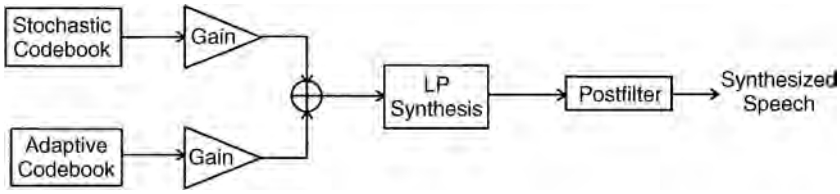


FIGURE 95.4 Speech synthesis in the FS-1016 CELP.

In MPLP coding, an excitation sequence is obtained by forming multiple non-uniformly spaced pulses of different amplitudes. The location and amplitude of each pulse is determined sequentially such that the weighted mean square error is minimized. Typically, 4–6 pulses in each 5-ms frame are used. The pulse amplitudes and locations are transmitted to the receiver. The multipulse technique is used in the 9.6-kb/s Skyphone, an airline communication service, and, in the recent 6.3 kb/s G.723-1, a dual rate (6.3/5.3 kb/s) ITU-T standard for multimedia network applications.

In RPE coding, an excitation sequence is formed by combining multiple pulses of different amplitudes as in MPLP coding; however, pulses are uniformly spaced. Hence, pulse positions are determined by specifying the location of the first pulse. The amplitude of each pulse is determined by solving a set of linear equations such that the weighted mean square error is minimized. Typically, 10–13 pulses in each 5-ms frame are used, and the pulse spacing is 3 or 4. A 13-kb/s RPE-LTP coder, an association of RPE with LTP, is used in the ETSI GSM FR (full-rate) digital cellular standard.

In CELP coding, the optimum excitation sequence is selected from a codebook, perhaps of random Gaussian sequences, which minimizes the perceptually weighted mean squared error. The first CELP coder [Schroeder and Atal, 1985] used 1024 codes (vectors), each of length 40 samples (5 ms in 8-kHz sampling). Early CELP algorithms required heavy computations for the exhaustive codebook search and speech synthesis for each candidate code. A great deal of work followed to reduce computational complexity. Notable developments have been made in structured codebooks, such as sparse, overlapping, and vector-sum excitation, and in the associated fast search algorithms. CELP can provide good quality speech at low rates down to 4.8 kb/s. With the recent developments of high-speed and high-capacity DSP processors, CELP algorithms have become practical and actually adopted in several standards.

A typical example of CELP coding is the 4.8-kb/s U.S. FS-1016 coder. The 10th-order STP parameters are computed every 30 ms over the Hamming windowed frame of incoming speech. Excitation parameters are updated every 7.5 ms. The synthesis part is shown in Fig. 95.4, where the excitation is a combination of codes from a stochastic codebook and an adaptive codebook. The adaptive codebook, which consists of past excitation codes, replaces the LTP synthesis, and the search for the optimum adaptive code is equivalent to search for the optimal pitch lag. The LP synthesis in Fig. 95.4 is same as the STP synthesis loop as shown in Fig. 95.1(b). In this coder and in many modern coders as well, the line spectrum pairs (LSPs), also called the line spectrum frequencies (LSFs), are used as the STP parameters for stability and quantization efficiency. The additive white-type noise contained in the CELP-coded speech can be reduced by using the postfilter of the form $(1 + cz^{-1})W(z)$, which essentially deemphasizes spectral valleys. The range $0.2 \leq c \leq 0.4$ has been found to give reasonable subjective results.

The bit allocation scheme of the 4.8-kb/s DoD-CELP coder is shown in Table 95.2. It can be seen that more than two-thirds of the total bits are allocated to the codebook indices and gains that specify the excitation signal.

VSELP and low-delay CELP (LD-CELP) are versions of CELP coding. In the 8-kb/s VSELP algorithm, adopted in the time-domain multiple access (TDMA) North American digital cellular standard, IS-54, by the Cellular Telecommunications Industry Association (CTIA), the excitation is a combination of vectors (codes) from a pitch adaptive codebook and two highly structured codebooks. The VSELP algorithm was adopted in the 6.4 kb/s Japanese digital cellular standard and the 6.3 kb/s GSM HR (half-rate) coder. The LD-CELP algorithm is employed as a low-delay 16-kb/s international standard, G.728. The LD-CELP has a unique feature in that LTP is not used, and instead, the order of STP is increased to 50 to compensate

TABLE 95.2 Bit Allocations in the FS-1016 CELP

Parameters	Bits	Update, ms	Bit Rate, b/s
STP (10-LSPs)	34	30	1133.33
Stochastic codebook		7.5	1866.67
Index	9		
Gain	5		
Adaptive Codebook		7.5	1600
Index	7		
Gain	5		
Miscellaneous			200
Total			4800

Source: Atal et al. [1991].

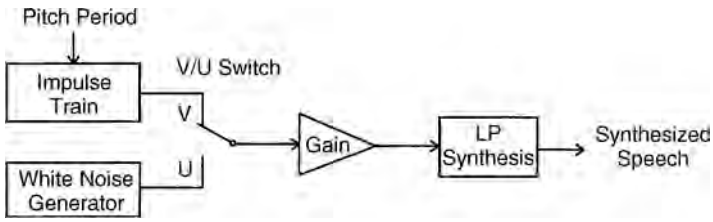


FIGURE 95.5 LPC synthesis model.

for the lack of pitch information. The speech quality of LD-CELP is comparable to that of G.721, 32-kb/s ADPCM, and coding delay is less than 5 ms. More information on DoD-CELP, VSELP, and LD-CELP coders can be found in Atal, Cuperman, and Gersho [1991].

Probably the most prevailing version in recent years is the algebraic CELP (ACELP), which uses an algebraically structured codebook for efficient search. The ACELP is used in the 5.3 kb/s mode of G.723.1. The 8 kb/s G.729 employs ACELP along with the conjugate-structured codebook to reduce complexity. Other examples which adopt ACELP structure include the TIA IS-127 EVRC coder based on the RCELP (relaxed CELP) algorithm and the GSM EFR (enhanced full-rate) coder.

95.8 Vocoders

LPC, probably the most widely-studied of the very low rate vocoding, will be described here due to its historical importance. A classic example is the 2.4-kb/s LPC-10 algorithm adopted as the U.S. FS-1015 standard in 1977, which has been used for secure voice communications. The LPC-10 algorithm uses a 10th-order all-pole predictor and relies on the two-state (voiced and unvoiced) source excitation model. The speech synthesis model of LPC is shown in Fig. 95.5. The linear prediction parameters, gain, pitch period, and voicing/unvoicing (V/U) decision parameters are estimated, quantized, and transmitted to the receiver. At the receiver, speech is synthesized by using the impulse train for the voiced frame and white noise for the unvoiced frame. The impulse train, defined in the standard, is a periodic sequence that has a pitch period similar to a glottal pulse.

Reconstructed speech of an LPC vocoder operating at very low rates such as 2.4 kb/s is of synthetic quality due to the inaccuracy of the source excitation. The quality of the speech produced by LPC-10 is also synthetic in that it sounds artificial, unnatural, and buzzy. The two-state excitation model was improved later by the **mixed excitation** source model, which reduced the buzziness and breathiness in the synthesized speech. A 2.4 kb/s mixed excitation LPC, called the MELP, became the new United States federal standard [McCree et al., 1996]. A 400 b/s LPC vocoder based on LPC-10 was also reported [Atal et al., 1991].

Other types of vocoding are STC and MBE, which belong to the class of sinusoidal coding. In sinusoidal coding, speech is analyzed and synthesized by using sinusoidal components. These coders have received attention recently because of their potential for high-quality speech coding at low bit rates.

The MBE takes a mixed excitation approach in that the short-time spectrum of a speech frame is treated as a collection of multiple bands, each of which is either voiced or unvoiced. The voiced bands are centered around the pitch harmonics and the unvoiced bands contribute to the random nature of the spectrum. The parameters of the MBE speech model consist of fundamental (pitch) frequency, the V/U decision for each band, and spectral magnitudes (envelope). An analysis-by-synthesis technique is used to estimate the parameters that minimize weighted error between the original spectrum and the synthetic spectrum. At the receiver, the voiced bands are synthesized in the time domain using a bank of harmonic sinusoids.

$$v(n) = \sum_{k=1}^{L} A_k \cos(k\omega_o n + \phi_k), \quad n = 0, 1, \ldots, N$$

where A_k, $k\omega_o$, and ϕ_k are magnitudes, frequencies, and phases of L pitch harmonics respectively, ω_o is the pitch frequency, and N is the frame size. The unvoiced bands are synthesized in the frequency domain by scaling the normalized white noise spectrum by the spectral magnitude. Synthesized speech is obtained by adding the voiced bands and inverse Fourier transform of the unvoiced bands spectra. The phase information of the voiced harmonics is determined at the receiver by using partial information transmitted from the encoder and a phase prediction algorithm.

The 4.15 kb/s IBME (improved MBE) coder was adopted as a 6.4-kb/s Inmarsat-M standard, which includes an additional 2.25 kb/s for forward error correction. In this coder, most of the bits are used for scalar quantization of spectral envelope parameters. Bit rate reduction was made possible by representing and vector quantizing the spectral envelope with LSPs that correspond to 10-pole LPC parameters. A new AMBE (advanced MBE) coder reportedly produced near-toll quality at 4-kb/s. Refer to Atal, Cuperman, and Gersho [1991] for details on MBE coders. Information on the STC can be found in Gersho [1994] and Spanias [1994].

95.9 Variable Bit Rate (VBR) Coding

In variable bit rate (VBR) coding, different speech segments are encoded at different bit rates. Variable rate coding can be effective for storage, packet voice, and multiple access channel applications. The two typical approaches for the VBR coding are in the variable update rates and the variable bit allocations for parameters from different speech segments. The VBR coding algorithms are usually based on conventional coders such as ADPCM, SBC, and CELP. A key element of VBR coding is the voice activity detection (VAD) used for voicing decisions and speech/nonspeech classifications.

An example of a VBR coder is a variation of the CELP, called QCELP, which is used in the CDMA North American digital cellular standard, IS-96, adopted by the TIA. QCELP operates at one of four predetermined rates, 8, 4, 2, or 1 kb/s, for each 20-ms analysis frame. Another example is the IS-127 EVRC coder, which encodes speech signals at 8, 4, or 0.8 kb/s depending on the speech activity determined by the VAD algorithm.

95.10 Performance Evaluation

Performance of a coding algorithm is primarily specified by the speech quality at the prescribed coding rate. When a particular speech coder is to be designed or selected for a particular application, other factors should also be considered. Those are coding delay (due to algorithmic and computational factors), complexity [of algorithm and implementation, often measured in million instructions per second (MIPS) of a computing processor], cost (of system and terminal), robustness (to background noise or channel impairments), tandeming (with itself or other coders), transcoding (into another coding format), and voiceband data compatibility (e.g., modem and facsimile). The priority of consideration will depend on the target application. For example, in the public switched telephone network (PSTN), coding delay and

tandeming can be more important than robustness, whereas coding rate and robustness can be more important than coding delay in mobile cellular applications.

Coded speech quality can be measured subjectively and objectively. Subjective measurements are obtained as consequences of listening tests, whereas objective measurements are computed directly from the coded speech parameters.

Typical subjective measures often quoted are MOS, diagnostic acceptibility measure (DAM), and diagnostic rhyme test (DRT). The MOS, probably the most widely used, is obtained by averaging test results rated by a group of listeners who are asked to quantify their perceptual impressions on a five-point scale (bad = 1, poor = 2, fair = 3, good = 4, excellent = 5). The DAM is a more systematic method of assessing speech quality, where the listeners are asked to grade 19 separate scales grouped in three categories on a 100% scale. This test is directed at medium-to-high quality speech and requires trained or experienced listeners. The DRT is directed at testing the intelligibility of the medium-to-low quality speech. In the DRT, the listener is asked to distinguish between word pairs of single syllables, such as meat–beat and sheet–cheat. Such subjective measures, however, can, at best, be a rough indicator of what to expect from a particular coder. More information and references on subjective measures can be found in Jayant and Noll [1984] and Deller et al. [1993].

A good objective measure should be able to consistently predict the subjective quality of the speech and could be used in the design stage in controlling the amount of distortion that can be introduced during the coding process. Widely used objective measures include mean squared error (MSE)–based measures such as signal-to-noise ratio (SNR) and segmental SNR (SEGSNR), and LPC-based measures, such as spectral distance and cepstral distance. Other time- or frequency-domain measures have been introduced and studied; however, a good objective measure that is consistent and provides strong correlation with the subjective quality has yet to be found. More discussion on objective measures can be found in Quackenbush et al. [1988].

95.11 Speech Coding Standards

A summary of typical speech coding standards adopted by industrial, national, or international organizations is shown in Table 95.3. It can be seen that most standards were established in the 1990s, which is a consequence of worldwide demand for bandwidth-efficient mobile voice communications and technological maturity of speech coding algorithms in various applications. Since the listed bit rates of many

TABLE 95.3 Speech Coding Standards

Bit Rate, kb/s	Algorithm	Standard (Year)	MOS
64	log PCM	CCITT, G.711 (1972)	4.3
40/32/24/16	ADPCM	CCITT, G.726 (1990)	4.1 (32)
16	LD-CELP	CCITT, G.728 (1992)	4
16	APC	Inmarsat-B (1985)	—[a]
13	RPE-LTP	ETSI, GSM FR (1991)	3.5
13	ACELP	ETSI, GSM EFR (1996)	4
9.6	MPLP	BTI Skyphone (1990)	3.4
9.6/4.8/1.2	EVRC	TIA, IS-127 (1997)	3.8 (9.6)
9.6/4.8/2.4/1.2	QCELP	TIA, IS-96 (1997)	3.33 (9.6)
8	VSELP	TIA, IS-54 (1993)	3.5
8	CS-ACELP	ITU-T, G.729 (1996)	4
6.7	VSELP	Japanese DMR (1993)	3.4
6.4	IMBE	Inmarsat-M (1993)	3.4
6.3/5.3	MP-MLQ/ACELP	ITU-T, G.723.1 (1996)	3.98/3.7
4.8	CELP	US, FS-1016 (1991)	3.2
2.4	LPC-10	US, FS-1015 (1977)	2.3
2.4	MELP	US, FS (1996)	3.2

[a]MOS data unavailable, however, reportedly near toll quality.
Sources: Spanias [1994], Kondoz [1994], and Spanias [2001].

coders in Table 95.3 include error control bits, actual source coding rates of such coders are smaller than those listed. Coders with MOS ratings of close to or more than 4 can be considered to provide toll quality, whereas the rest provide near-toll or communications quality except for the LPC-10.

95.12 Concluding Remarks

Principles and applications of typical coding algorithms for telephone speech have been reviewed in this chapter. Major research efforts in the last decade have concentrated on the analysis-by-synthesis technique in the bit range of 4–16 kb/s, and some of the results have turned into standards in cellular and mobile voice communications. Important advances can be typified by the standardization of CELP coders in the range of 4.8–16 kb/s and emergence of the sinusoidal vocoding approach (STC, MBE) and MELP in the 2.4–4.8 kb/s range.

Major research efforts in the future are expected to be concentrated on low bit range coding, say 4 kb/s or below, where perceptual factors of the human auditory system become more important [Jayant, 1992]. Continued activities for enhanced coder performance are also expected in the range of 4–8 kb/s for specific applications. In a practical aspect, speech coding algorithms to be included in the next generation of mobile communication systems will be a major issue for the next few years.

Defining Terms

Analysis-and-synthesis: Refers to open-loop operation where the analysis cannot be affected or changed by the synthesis.

Analysis-by-synthesis: Refers to closed-loop operation where the analysis can be affected or changed by the synthesis.

Hybrid coders: Coders that employ features of both waveform coders and vocoders.

Mixed excitation: A speech segment is modeled as a consequence of both voiced and unvoiced excitation.

Vocoders: Coders that are primarily aimed at reproducing the original sound.

Waveform coders: Coders that are primarily aimed at reproducing the original time-domain waveform.

References

Atal, B.S., Cuperman, V., and Gersho, A., Eds. 1991. *Advances in Speech Coding,* Kluwer, Dordrecht, the Netherlands.

Deller, J.R. Jr., Proakis, J.G., and Hansen, J.H.L. 1993. *Discrete-Time Processing of Speech Signals,* Macmillan, New York.

Gersho, A. 1994. Advances in speech and audio compression. *Proc. IEEE,* 82(6):900–918.

Gibson, J.D. 1980. Adaptive prediction in speech differential encoding systems. *Proc. IEEE,* (April):488–525.

Gibson, J.D. 1984. Adaptive prediction for speech encoding. *ASSP Mag.,* 1: 12–26.

Jayant, N.S. 1992. Signal compression: technology targets and research directions. *IEEE Trans. Selec. Areas Commun.,* 10(5):796–818.

Jayant, N.S. and Noll, P. 1984. *Digital Coding of Waveforms: Principles and Applications to Speech and Video,* Prentice–Hall, Englewood Cliffs, NJ.

Kondoz, A.M. 1994. *Digital Speech: Coding for Low Bit Rate Communications Systems,* Wiley, Chichester, England.

Makhoul, J., Roucos, S., and Gish, H. 1985. Vector quantization in speech coding. *Proc. IEEE,* 73(11): 1551–1588.

McCree, A.V., George, E. B., Barnwell, T., and Viswanathan, V. R. 1996. A 2.4 kbit/s MELP coder candidate for the new U.S. federal standard. In *Proc. ICASSP.*

Quackenbush, S.R., Barnwell, T.P., and Clements, M.A. 1988. *Objective Measures of Speech Quality,* Prentice–Hall, Englewood Cliffs, NJ.

Rabiner, L.R. and Schafer, R.W. 1978. *Digital Processing of Speech Signals,* Prentice–Hall, Englewood Cliffs, NJ.

Schroeder, M.R. and Atal, B.S. 1985. Code-excited linear prediction (CELP): high quality speech at very low bit rates. In *Proc. ICASSP,* Tampa, FL, April, 937–940.

Spanias, A.S. 1994. Speech coding: a tutorial review. *Proc. IEEE,* 82(10):1541–1582.

Spanias, A.S. 2001. Speech coding standards. In *Multimedia Communications: Directions and Innovations,* Gibson, J.D., Ed., Academic Press, New York.

Further Information

A comprehensive tutorial can be found in Spanias [1994]. Speech coding standards and references are well summarized in Spanias [2001]. Journals specialized in speech coding are *IEEE Transactions on Speech and Audio Processing* (since 1993), *IEEE Transactions on Signal Processing* (1991–1993), *IEEE Transactions on Acoustics, Speech and Signal Processing* (1974–1990), and *IEEE Transactions on Communications*. The most recent activities can be found in the annual Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP).

96

Video

Eric Dubois
University of Ottawa

96.1 Introduction

Time-varying imagery is a highly effective medium in a variety of applications including entertainment, interpersonal communication, information presentation and retrieval, and many others. We refer to a time-varying image signal in electronic form as a **video signal**. Until recently, video signals were stored and transmitted exclusively in analog form, and this continues to be the most common situation. The principal storage medium is analog video tape, and the main transmission channels are the over-the-air broadcast channel, cable, and satellite. However, due to recent advances in digital hardware and in image compression technology, systems for digital storage and transmission of video signals are becoming more common. It can be expected that digital representations of video will be predominant in the future. Video will be stored in digital form on disks and tape, and will be transmitted digitally over broadband networks as well as wireless networks and broadcast channels.

Video signals are notorious for consuming huge amounts of bandwidth. For example, an analog National Television System Committee (NTSC) video signal uses a 6-MHz bandwidth, whereas an uncompressed digital NTSC video signal requires a bit rate of about 100 Mb/s. To allow economical storage and/or transmission, more efficient representations are required. This process is often viewed as a compression operation, wherein quality is sacrificed; an alternative view is that we seek a representation that will deliver maximum video quality for a given available data rate. At rates of 20–30 Mb/s, this would yield large images of high quality [high-definition television (HDTV)], whereas at low rates such as 64 kb/s and below, this would result in a system providing small pictures of rather low quality and with limited capability to render motion (videophone).

The techniques of source coding seek to exploit both redundancy in the source and the fidelity requirements of the receiver of the information in order to minimize distortion for a given data rate. If a stochastic model of the source and a fidelity criterion for the receiver are known, the optimal performance can be found using rate-distortion theory. However, an implementable coder with acceptable delay giving this optimal performance may not be available. In the case of time-varying imagery, neither general purpose stochastic models of the source nor mathematical fidelity criteria that capture all of the known properties of source and receiver are known at this time. Thus, optimal source coding for time-varying imagery does not currently rest on a firm theoretical basis. Rather, we have available an arsenal of tools that work well and that address, at least qualitatively, the known properties of the source and receiver.

The purpose of this chapter is to identify the properties of source and receiver that are exploited by video source coding algorithms and to describe the tools that are available for source coding. The emphasis is on those aspects that are specific to time-varying imagery, since a previous chapter deals with still imagery. Thus, the use of motion information in the processing and compression of time-varying imagery is a pivotal element for this contribution. An overview of video compression standards is also given, followed by some perspectives for the future. An extensive treatment of video coding can be found in Netravali and Haskell [1995].

96.2 Source Characteristics and Viewer Requirements

A number of factors determine the source characteristics and the level of redundancy that can be exploited. These include the scanning format, the signal-to-noise ratio, the amount of detail, and the characteristics of the motion. Note that the source signal to be coded may not necessarily be the direct output of a video camera; it may be the result of some scan conversion or other preprocessing operations.

The scanning format encompasses the picture size and the sampling structure. We assume that the input to the source coder is fully digital, that is, an ordered sequence of fixed length binary codewords representing the image color at a discrete set of points Ψ This is denoted

$$u_i(x, y, t), \quad (x, y, t) \in \Psi_i, \quad i = 1, 2, 3 \tag{96.1}$$

where u_1, u_2, u_3 represent coordinates in a suitable color space. These could be red–green–blue or luminance-chrominance components. Clearly, black and white images would be represented by a single component. The sampling structure Ψ is a regular array of points such as a lattice, cropped spatially to the image frame size. Picture size may range from small windows on a videophone or workstation, say, of size 64 by 64, to HDTV pictures of size 1920 by 1080 and beyond. Two vertical-temporal scanning structures are widely used for video: **progressive scanning** and **interlaced scanning**. These scanning structures are shown in vertical-temporal profile in Fig. 96.1. In analog systems, where the capture, transmission, and display formats are the same, interlace is a method to achieve higher vertical resolution for a given vertical scan rate and system bandwidth. The vertical scan rate of 50 or 60 Hz used in major worldwide systems is largely determined by the requirement to minimize the visibility of flicker at the display. Although interlace is a good technique in the context of analog systems, it usually leads to spatiotemporal **aliasing** in the presence of vertical motion, which makes subsequent processing, and especially source coding, much more difficult. (See Dubois [1992] for a discussion.) For this reason, there has been increasing focus on progressive scanning in recent system proposals. A detailed discussion of sampling structures and video signal sampling is given in Dubois [1985].

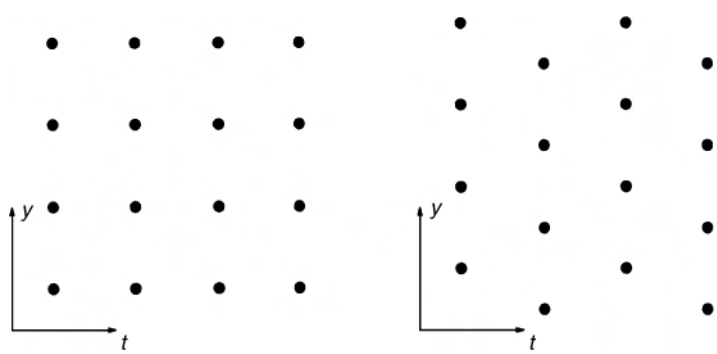


FIGURE 96.1 Profile view of scanning structures: (a) progressive and (b) interlace.

TABLE 96.1 Typical Scanning Formats Used for Digital Video Communications

	Luminance Size		Chrominance Size		Scanning Structure	
Identifier	Hor.	Vert.	Hor.	Vert.	Interlace	Pict. Rate
QCIF	176	144	88	72	1:1	29.97
CIF	352	288	176	144	1:1	29.97
ITU-R 601	720	480	360	480	2:1	59.94
HDTV-P	1280	720	640	360	1:1	60.0
HDTV-I	1920	1080	960	540	2:1	60.0

Although some standards allow considerable flexibility in the choice of picture format, a number of specific formats are in common use or have been proposed. Table 96.1 shows the parameters of some of these formats. The final coded bit rate is very directly related to the source sampling rate, although not linearly.

The signal-to-noise ratio of the source has a significant impact on compression efficiency. In general, noise has little redundancy and cannot be compressed very much; thus, it should be removed prior to coding if possible. There is, of course, a trade-off between coding efficiency and the introduction of artifacts due to noise reduction. Valid assumptions must be made or the processing may remove desired picture content such as snowstorms, fireworks, etc. Motion compensated processing is required for the most effective noise reduction [Dubois, 1992]. However, good image and noise models are necessary. The principle is to filter out those deviations from an assumed image model (e.g., that image intensity is constant along trajectories of motion) that match the characteristics of an assumed noise model (e.g., white noise with a specified variance).

The characteristics of the motion in the scene affect the amount of compression that is possible. Apparent motion in the image can be due to either camera motion (including zooming), object motion, or both. In a videophone application, most apparent motion may be due to movement of the person in front of the camera. In a surveillance application, most of it may be due to the scanning motion of the camera (except during an incident), whereas in a sports scene, both camera motion and scene motion may be significant. Motion may, in general, be low, as in a videophone application, or very active, as in a sports application; it may be highly structured (rigid-body motion) or pseudorandom (rustling leaves). Knowledge of the motion characteristics can affect the coding algorithm; if there is unrestricted motion, a worst-case assumption must be made.

Redundancy in a source is related to the concept of predictability (although they are not the same). Given that a certain portion of the source signal has already been transmitted, the coder may attempt to predict what comes next. If the prediction is at all reliable, it is better to transmit information concerning the deviation from the prediction than the raw information. Predictability implies a certain regularity in the data that can be captured in statistical models. In the spatial dimension, if a certain pattern appears over a certain area, one may assume that the pattern will continue and make the prediction on this basis. Of course, at an object boundary, the prediction will be wrong. However, if it is right most of the time, an overall coding gain can be achieved. In a moving picture, the assumption is that if an object is moving with a certain velocity and acceleration, it will continue to do so. Thus, if the velocity and acceleration are known, the coder can predict what the next frame will look like and only transmit the error between this prediction and reality. This is the essence of motion-compensated coding. Again, at a scene change, the prediction will be completely wrong, and more information must be transmitted. However, on average, there will be a substantial compression. Temporal change is not due solely to motion. The illumination may change, and the luminance of scene points may change as the angles between object surface, illumination source, and camera change. If we can model these effects as well, we can make better predictions and thereby improve coding performance.

In addition to the structural or statistical redundancy in the source, there is perceptual redundancy (also called irrelevancy). Specifically, information that cannot be perceived by the viewer need not be transmitted. The main phenomena related to this are spatial and temporal **masking**. Spatial masking is

the effect whereby certain types of distortion are less visible in the vicinity of high-contrast image detail than in flat low-contrast areas. Temporal masking is the effect where distortions are less visible in areas of the image that are newly exposed or contain fast motion. An important case here is the scene change. The viewer can tolerate a significant reduction in resolution just after a scene change. This is fortuitous, because this is exactly where the temporal motion-compensated prediction breaks down.

The goal of an efficient source coder is to maximize quality (or minimize distortion) subject to a number of constraints. In most applications of video compression, the quality or distortion are subjective measures determined by the visual properties of the viewer. However, if further signal processing or manipulation of the image sequence is to be performed, a more stringent objective criterion may be more appropriate, since impairments invisible to a viewer could have a negative effect on a signal processing operation such as chroma key. If subjective quality is the main concern, it would be very desirable to have a mathematical function that measures the subjective distortion. This distortion function could then be minimized during the optimization of the coding algorithm. Unfortunately, such a general purpose and widely accepted perceptual distortion measure is not available. Thus, properties of the human visual system are usually ex-ploited based on qualitative or empirical considerations that are obtained under restricted conditions. However, this approach may still be quite effective [Schreiber, 1993; Jayant et al., 1993]. The coding algorithm can then be evaluated using formal subjective testing under controlled conditions [Allnutt, 1983].

Other requirements may affect the type of coding algorithm that can be used. These include delay, random access, playback control, and scalability. Both the coding algorithm and the transmission channel introduce delay. Depending on the application, there may be tight bounds on this delay, especially in interactive applications. The most stringent application is in face-to-face communication, where a maximum end-to-end delay of about 200 ms can be tolerated. If the delay is longer than this, spontaneous interaction becomes difficult. In a database access application, somewhat longer delays of up to a few seconds may be acceptable. Finally, in an application such as video on demand, the delay constraint may be of no practical consequence, that is, any reasonable encoding system would meet the delay constraint.

Random access and playback control are requirements that may exist in services such as database access and video on demand. Some coding schemes are recursive in nature and require that the sequence be decoded continuously from the beginning in order to view a particular frame. This may be acceptable in an application such as video conferencing, but it precludes random access and channel switching. Thus, the encoding algorithm must have *finite memory* in some sense to allow such a random access feature. The playback controls include fast-forward, reverse, and fast-reverse playback as well as pause functions. Again, general coding algorithms may not permit such features, and they must be incorporated as constraints if they are desired.

The coding algorithm must account for the properties of the channel or storage medium. The channel may allow only constant bit rate (CBR) operation, or it may permit variable bit rate (VBR) operation. In the former case, the coder must be designed to produce a fixed number of bits over a specified unit of time. In the latter case, the bit rate may vary over time, although some constraints negotiated at setup time must be respected (e.g., ratio of peak to average bit rate). Another requirement of a coding algorithm is acceptable performance in the presence of typical channel impairments. These channel impairments may vary considerably for different types of channels. For example, a dedicated digital link may have random errors with a very low probability of occurrence, whereas a wireless link may have bursty errors with a high probability of occurrence. Also, depending on the transport mechanism, the errors may occur in isolated bits, or they may involve an entire packet consisting of a large number of bits. In general, when the coding efficiency for redundancy reduction is higher, the coded data is more vulnerable to transmission errors. Thus, overall system performance must be evaluated with the channel impairments taken into account in order to be realistic. If possible, joint source/channel coding should be considered.

In some situations, the same video signal must be delivered to a variety of receivers having different capabilities. Some examples include the broadcast of entertainment television to a population of receivers ranging from small tabletop units to large high-definition displays. Similarly, a database containing video data might be accessed by terminals ranging from relatively inexpensive home computers to

powerful workstations, perhaps with three-dimensional display. The capability to transmit only that portion of the data required by the terminal, or to extract a subset of the data, is known as **scalability**. The most straightforward approach to this is simulcasting or multiple storage of the coded data for each type of receiver. This is complex and wasteful of storage or transmission capacity. Thus, a more efficient approach is the concept of embedded coding, whereby a subset of the encoded data can be used to decode a lower version of the sequence suitable for display on the lower capability display.

Most of the preceding requirements can reduce the coding efficiency achievable, as compared to an unconstrained situation. In each case, the system designer must decide whether the benefits of the feature justify the additional complexity or performance loss.

96.3 Coding Algorithms

The goal of a source coding algorithm is usually to provide the best image quality for a given channel or data rate, subject to constraints on complexity, delay, or other requirements, as discussed. In some circumstances, a certain quality may be prescribed, and a coding algorithm is sought that provides this quality at the lowest possible rate (again, subject to constraints). In time-varying imagery, most incremental information over time is carried by the motion occurring in the scene. Thus, the most highly efficient coding algorithms are based on motion estimation and compensation. This is the main feature distinguishing video coding from still image coding, and it represents the focus of this section. An image sequence is usually encoded by both intraframe coding techniques and interframe coding techniques. The intraframe coding algorithms are very similar to the still picture coding algorithms. They are used for several reasons, including: to permit random access into the sequence, to allow fast forward and reverse playback, to provide robustness to channel errors, and as fallback when temporal coding fails (for example, at a scene change).

Motion Compensation

The three-dimensional motion of objects in a scene induces two-dimensional motion in the image plane. The relationship between them depends on the perspective transformation of the camera and may be complex. The image of a given scene point traces out over time a trajectory in the space–time coordinates of the time-varying image (Fig. 96.2). The basic principle of motion-compensated coding is that the

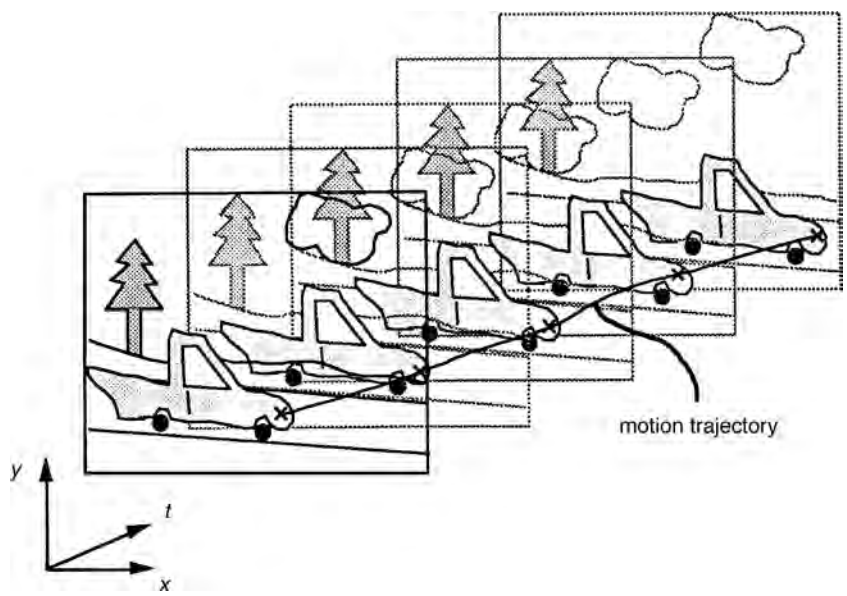


FIGURE 96.2 Trajectory in space time of a scene point.

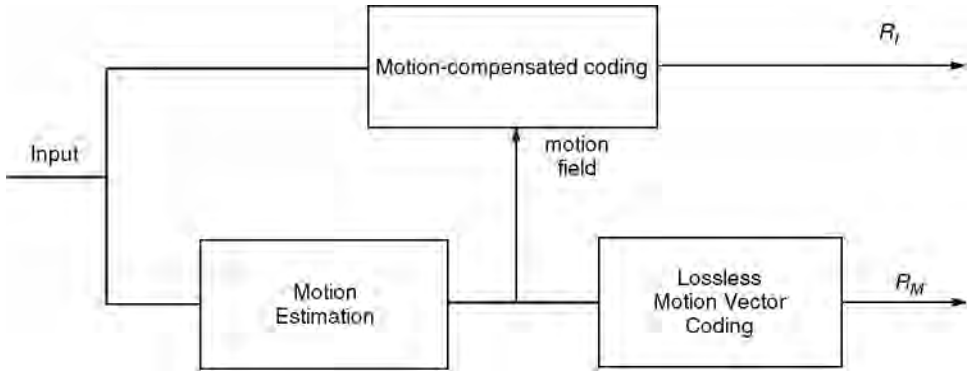


FIGURE 96.3 General structure of a motion-compensated video coder.

redundancy is very high along this trajectory. This is exploited using the coder structure of Fig. 96.3. The input signal is applied to a motion estimation module that determines a **motion field**. This motion field establishes the correspondences between image samples lying on the same trajectory in successive image pictures. Although the motion field specifies estimated displacements for each pixel in the image sequence, it may be specified parametrically over regions, as will be described subsequently. The motion field is then encoded with a lossless code for transmission to the receiver at rate R_M. Any quantization of the motion field is assumed to take place within the motion estimation module. The motion field is also supplied to the motion-compensated coding module, which performs a conditional coding of the image sequence given the estimated and quantized motion field. The output of this module (which includes any lossless coding) is transmitted to the receiver at rate R_I. It is possible for the motion estimator to also make use of previously coded/decoded data. This can be accommodated (conceptually) by the structure of Fig. 96.3 by assuming that the motion estimation module contains a copy of the motion-compensated coder. Of course, the motion-compensated coder would not be duplicated in a practical implementation; common elements would be shared between the motion estimation and motion-compensated coding modules.

The total rate R for coding the video sequence is the sum of R_M, the rate to transmit the motion information, and R_I, the rate to transmit the conditionally coded video information. The best way to partition the overall rate between these two components is an open problem. A greater precision in the motion field may lead to a lower rate R_I for the motion-compensated video but a higher rate R_M for the motion information. The optimal point will depend on the motion coding scheme and the motion-compensated video coding scheme. An extreme situation is when $R_M = 0$, where either no motion information is used, or motion is only computed from previously encoded video data. At the other extreme, we can imagine a situation where the motion field is coded very precisely, so that $R_I \ll R_M$.

There are two basic approaches for estimating the motion field: the pixel-based approach and the parametric approach. In the pixel-based approach, a displacement vector is estimated for each pixel in the picture using some robust algorithm. The resulting motion field is then encoded efficiently to achieve the desired rate. Significant compression is possible since the motion field is, in general, highly redundant (it is very smooth except at object boundaries). The decoded motion field is used in the motion-compensated coder. The parametric approach is more common in video coding, where the motion field is represented parametrically over image regions. The most common technique is to assume a constant displacement over a rectangular block of pixels. Higher order models such as the affine model have also been used. A more general approach is to assume that the motion is represented parametrically over nonrectangular regions. In this case, the specification of region boundaries is part of the motion representation. There are many algorithms to actually estimate the motion, including region matching, gradient-based optimization, and transform domain methods. See Mitiche and Bouthemy [1996] for a survey.

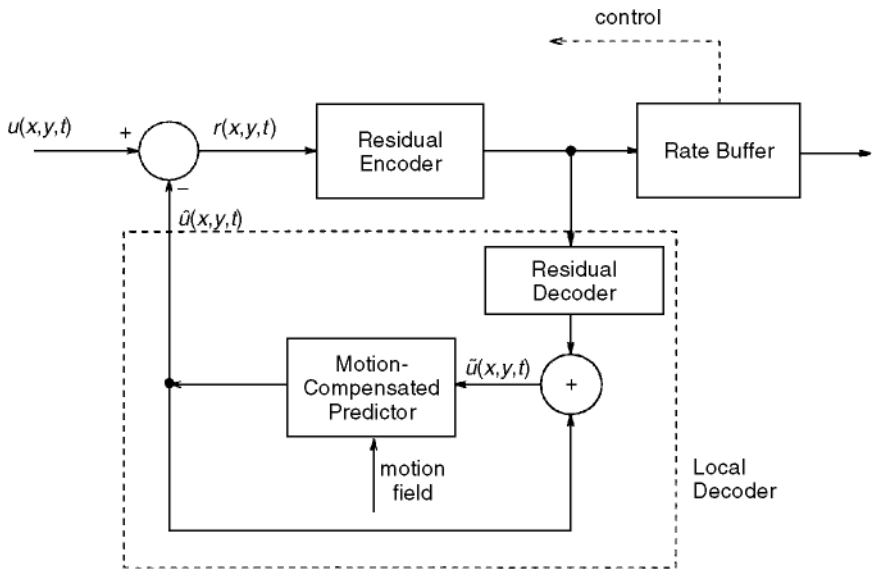


FIGURE 96.4 Motion-compensated DPCM encoder.

The two main techniques to perform conditional coding given motion, **predictive coding** and **interpolative coding**, are discussed in the next two sections.

Predictive Coding

The principle of predictive coding is that, given the motion information and previous decoded pictures, it is possible to form a good estimate of the next picture to be coded. Then, only the deviation from this estimate needs to be encoded and transmitted. A structure well suited to this is differential pulse code modulation (DPCM), shown in Fig. 96.4. In this system, the prediction is formed on the basis of previously reconstructed pictures using a local decoder at the transmitter. The prediction error is then spatially quantized and losslessly encoded at a variable rate for transmission. The types of spatial quantization used are discussed in the section on residual coding.

In the basic mode of operation, the image information at time t is predicted from the suitably displaced information in the previous vertical scan at time $t - T$, where T is the vertical scan time. The prediction of the sampled input image $u(x, y, t)$ is given by

$$\hat{u}(x, y, t) = \tilde{u}(x - d_x(x, y, t), y - d_y(x, y, t), t - T) \tag{96.2}$$

where $(d_x(x, y, t), d_y(x, y, t))$ is the estimated displacement at (x, y, t) and $\tilde{u}$ is the locally reconstructed output of the codec. If $(x - d_x, y - d_y, t - T)$ does not lie on the sampling grid, spatial interpolation must be used to compute u from existing samples of $\tilde{u}$ at time $t - T$. If the video signal is in interlaced format, each vertical scan of the image contains only part of the spatial information. Thus, a more elaborate prediction using at least the two previous pictures is very advantageous, as used in moving pictures experts group-2 (MPEG-2) standards [ISO/IEC, 1994].

A disadvantage of predictive coding is that the quantization error of the residual encoder is fed back to the predictor. This limits the compression efficiency, even if the source is very highly correlated along motion trajectories. One way to counteract this effect is to use interpolative coding for some pictures.

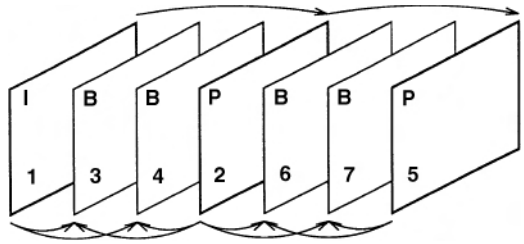


FIGURE 96.5 Example of picture structure for interpolative coding. Every third picture is encoded predictively. Two pictures out of every three are encoded interpolatively. The digit in the lower left corner of each picture specifies the order in which the pictures are coded.

Interpolative Coding

Interpolative coding is sometimes called bidirectional predictive coding. In this method, it is assumed that some pictures spaced M pictures apart, called *reference pictures*, have already been encoded by some other method, say intraframe or predictive coding. Then, the intervening pictures are estimated based on the previous and subsequent reference pictures, using motion compensation. The estimation error is spatially quantized and losslessly coded for transmission. This coding sequence is illustrated with an example in Fig. 96.5. The pixels in a B picture can be estimated from the previous reference picture, the subsequent reference picture, or both, depending on which gives the best results. In this way, occlusion effects can be handled. However, the transmitted motion field must identify which mode is being used and incorporate the necessary motion vectors.

In general, the interpolatively coded pictures have lower error than predictively coded pictures at the same rate. However, the efficiency of DPCM decreases as the spacing between pictures is increased. It has been found that encoding two pictures out of every three by interpolative coding is a good compromise for a wide selection of video sequences. Interpolative coding is an open-loop method; there is no feedback of the quantization error as in DPCM.

Residual Coding

The estimation error in DPCM or interpolative coding must be spatially quantized and losslessly encoded for transmission. This is a key element of the coding algorithm because it is at this point that compression is achieved. The coding scheme should have good rate-distortion performance for the class of residual signals. There should be several versions of the residual coding algorithm available at different points on the rate-distortion curve to allow adaptive coding and buffer control, as discussed in the next section. Although scalar quantization can be used for residual coding, it is more effective to use some form of quantization with memory. The two most popular methods for this purpose are transform coding and subband coding (although the first can be viewed as a special case of the second). The reasons to use quantization with memory are twofold: (1) there is remaining correlation in the residual signal and (2) even for a memoryless source, quantization with memory can achieve greater compression, especially at low rates.

The methods used for residual coding are the same general methods that are applied to spatial coding and are not discussed here. The methods that can be used include block transform coding, and in particular the discrete cosine transform (DCT), subband coding, wavelet coding, and vector quantization. The DCT method has been selected as the residual coding algorithm for all of the main compression standards developed to date. Although the basic algorithm is the same as for still pictures, the exact nature of the quantizers for the transform coefficients may be slightly different. The quantization operation is followed by a reversible code-assignment process. In general, the probability distribution of the quantized residual is highly nonuniform, so that variable-rate codeword assignment can yield a significant reduction in average

bit rate compared with fixed codeword length assignment. The methods that have been found effective include Huffman coding, run length coding, and arithmetic coding. These techniques are discussed in previous chapters. Because of the additional variability in residual statistics due to changing motion, the bit rate process can have significantly more variability in interframe coding than in spatial coding.

Buffer Control

The video coding algorithms described in preceding sections are characterized by a number of parameters, such as the quantizer scale factors, the input sampling structure, prefilter characteristics, etc. These parameters establish the rate and quality at which the coder operates. However, because of the time-varying nature of the video signal characteristics and the use of variable length codes, the short-term rate and quality change with time for a fixed set of parameter values. The rate can also vary as the coding mode changes, for example, the rate in intrapicture mode will be considerably higher than the rate in interpolative mode.

The transmission channel can be either a CBR link or a VBR link. In the former case, the variable rate data emitted by the coder must be smoothed out by a **rate buffer** (as in Fig. 96.4) to deliver data to the channel at a fixed constant rate. Because of constraints on delay and complexity, the buffer size is limited. As a result, there is a significant probability of buffer overflow or underflow. A control mechanism is required to adapt the coder parameters in order avoid buffer overflow while maintaining the highest and most uniform image quality possible. Even for a VBR channel, there will be constraints on the coder output bit rate process that must be satisfied, so that a rate buffer and control algorithm are still required.

The control algorithm periodically monitors the rate buffer occupancy and changes the coder parameters to avoid potential overflow or underflow in such a way as to give the best overall image quality. Large or overly frequent parameter changes causing uneven quality should be avoided. Most techniques for buffer control have been heuristically derived. The occupancy of the rate buffer is used to determine the coding parameters (often only one, a scale factor for the DCT coefficient quantizers). At prespecified time intervals, the buffer fullness is evaluated, and, if it crosses certain thresholds in a given direction, the coding parameters are changed in a way to counteract the change in buffer occupancy. For example, if the buffer level is increasing, the quantization in the coder can be made coarser in order to reduce the number of bits being put into the buffer. The choices of coding modes and buffer switch points are made heuristically. More sophisticated techniques attempt to formulate a cost function that measures performance of the buffer control algorithm and to use the history of the buffer occupancy and the statistics of the bit rate process in different coding modes to choose an optimal parameter adaptation strategy. See Zdepski, Raychaudhuri, and Joseph [1992] and Leduc [1994] for recent work.

Pyramidal Coding

Pyramidal coding is a method that can be used to perform embedded coding for scalable video applications [Girod, 1993]. It is capable of producing coded versions of a given video sequence at different levels of spatial, temporal, and amplitude resolution. The general structure of a pyramidal coder is shown in Fig. 96.6. The input is first coded at the lowest supported rate R_1 by the level 1 coder. The resulting bit stream can be decoded by any decoder supported by the service. This level 1 sequence would be at the lowest spatial, temporal, and amplitude resolution, whereas the input would have full spatial, temporal, and amplitude resolution. Thus, the level 1 coder would need to incorporate appropriate spatial and/or temporal downsampling. To obtain the next highest level in the hierarchy, the level 1 signal is locally decoded and used as *a priori* information in the conditional coding of the input. The incremental information is coded at a rate R_2, and together with the level 1 bit stream specifies the level 2 output. This process is repeated for as many levels as are desired. Conditional coding is usually achieved by interpolating a lower level signal to the same spatial and temporal resolution as the next level. This is then used to form a prediction of this higher level signal; the residual is then further coded for transmission.

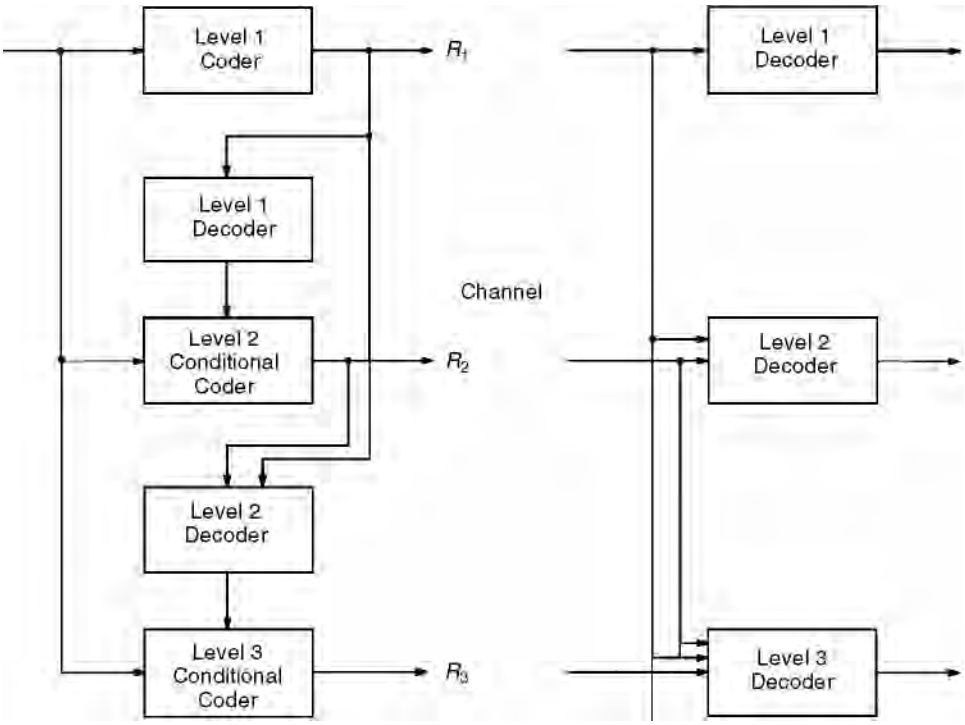


FIGURE 96.6 Pyramidal coder for embedded scalable video coding.

96.4 Standards

Standards play a major role in ensuring the wide availability of image communication products that permit interoperability between equipment provided by different manufacturers. The first important standard for compressed video was the International Telecommunications Union–Telecommunications Standardization Sector (ITU-T) [formerly the Consultative Committee on International Telephony and Telegraphy (CCITT)] H.261 standard for videoconferencing. This was followed by the International Standards Organization (ISO)/IEC MPEG 1 and 2 standards. These standards are summarized in Table 96.2. A brief description of each follows. A more detailed overview of these standards is given in Aravind et al. [1993].

The ITU-T Recommendation H.261 specifies the video codec system as part of a group of standards for audiovisual telecommunication services [ITU-T, 1990]. In 1984, CCITT Study Group XV established a specialist group on coding for visual telephony, and the recommendation was completed and approved in December 1990. This transmission standard was developed for use in videophone and videoconferencing applications. The standard specifies bit rates of the form $p \times 64$ kb/s, where p is an integer between 1 and 30. The standard defines only the functionality of the decoder; the encoder is not explicitly defined, but it must produce a bit stream compatible with the standard decoder. The input signals are in QCIF or CIF format (Table 96.1). The coding algorithm uses motion-compensated temporal DPCM with DCT coding of the residual. The motion compensation is block oriented, with a single motion vector for each 16×16 block of luminance pixels (referred to as a macroblock). The DCT is applied to 8×8 blocks of the residual signal. To meet the strict delay requirements, the prediction is always based on the immediately preceding picture. It is possible to set the temporal prediction to zero, so that the residual is equal to the input signal (interframe coding mode), or to force the motion vector to be zero (interframe coding without motion compensation). The bit stream has a layered structure, with layers for the picture, group of blocks (defined as 3×11 macroblocks), macroblock, and block. The top three layers consist of a

TABLE 96.2 Standards for Compressed Video

Name	Identifier	Application	Typical Bit Rates
$p \times 64$	ITU-T Rec. H.261	Teleconferencing	64 kbit/s–1.92 Mbit/s
MPEG-1	ISO/IEC 11172-2	Digital storage media	1.0–1.5 Mbit/s
MPEG-2	ISO/IEC 13818-2	Broad range	1.5–20 Mbit/s

header containing synchronization information, parameter values, and user defined bits, followed by a number of units of the next lower layer. The lowest layer (block) simply contains encoded transform coefficients and an end-of-block (EOB) code. The coder designer must decide how to regulate the parameters to optimize quality, avoid buffer overflow or underflow, and meet certain constraints imposed by a hypothetical reference decoder (HRD).

MPEG, under the ISO-IEC/JTC1/SC29/WG11, undertook, in 1988, the development of a standard for video coding adapted to rates of about 1–1.5 Mb/s and suitable for digital storage media. The strict delay requirement of H.261 was relaxed, whereas additional features such as random access, fast forward/reverse, etc., were desired. This standard, referred to as MPEG-1, has now been approved as an international standard [ISO/IEC, 1993]. It builds considerably on the H.261 standard and has much commonality. It supports a wider range of input signal formats, although the input sequences are assumed to be progressively scanned. The basic techniques of motion-compensated temporal DPCM and DCT coding of the residual are maintained from H.261. However, the technique of interpolative coding has been added. The concepts of layers used in H.261 has been augmented with the sequence layer and the group of pictures layer. The group of pictures layer is the unit allowing random access into the sequence. The group of blocks layer is replaced with a more flexible slice layer. The macroblock and block layers are similar to H.261. The bit stream is characterized at the sequence layer by a bit rate and buffer size. The MPEG standard establishes constraints on the bit stream to ensure that decoding will occur without buffer overflow for the given buffer size. See Le Gall [1992] for a more detailed overview of the MPEG algorithm.

In the second phase of MPEG, the algorithms have been extended to work with interlaced sequences and to integrate various forms of scalability [ISO/IEC, 1994]. It is considered for application to interlaced video such as the studio standard of ITU-R Recommendation 601 or HDTV. It extends the bit rate range of MPEG-1 up to about 20 Mb/s in currently studied implementations. MPEG-2 has introduced the concept of profiles and levels. A profile is a collection of signal processing tools (i.e., subset of the global syntax), whereas a level is related to the processing complexity and defines constraints on parameter values.

96.5 Perspectives

The video coding techniques described in the previous sections are well understood and will form the basis for many video services in the coming years. However, they do not approach fundamental limits of compression performance. Standards activity is currently addressing high-compression video coding under the phase 4 of MPEG. Techniques such as analysis-by-synthesis, object-oriented coding, and possibly even fractal coding will be brought to bear on the problem.

Defining Terms

Aliasing: An artifact caused by sampling a continuous signal with too low a sampling rate.

Interlaced scanning: A video scanning structure where the vertical position of horizontal scan lines in a given pass through the image is midway between the position of adjacent lines in the previous pass through the image.

Interpolative coding: A compression technique in which the coded signal is given by the quantized difference between the input signal and an estimation based on previously transmitted information occurring before and after the current input in time.

Masking: Reduction of the visibility of distortion by the image content.

Motion field: A vector field that determines the correspondences between image samples lying on the same trajectory in successive image pictures.

Predictive coding: A compression technique in which the coded signal is given by the quantized error between the input signal and a prediction based on previously transmitted information that occurs earlier in time.

Progressive scanning: A video scanning structure consisting of horizontal scan lines that lie in the same vertical position on each pass through the image.

Rate buffer: A first-in–first-out buffer memory that accepts the coded data from the codec at a variable rate and emits it to the channel at a fixed rate (or possibly at a variable rate but with smaller variability).

Scalability: The capability to easily derive multiple versions of the input signal at different spatial, temporal, and amplitude resolutions from the same coded data.

Video signal: A time-varying image signal in electronic form.

References

Allnutt, J. 1983. *Transmitted-Picture Assessment,* Wiley, Chichester, U.K.

Aravind, R., Cash, G.L., Duttweiler, D.L., Hang, H.M., Haskell, B.G., and Puri, A. 1993. Image and video coding standards. *AT&T Tech. J.,* 72(l):67–89.

Dubois, E. 1985. The sampling and reconstruction of time-varying imagery with application in video systems. *Proc. IEEE,* 73(4):502–522.

Dubois, E. 1992. Motion-compensated filtering of time-varying images, *Multidimensional Sys. Signal Proc.,* 3:211–239.

Girod, B. 1993. Scalable video for multimedia workstations, *Comp. Graphics,* 17(3):269–276.

ISO/IEC. 1993. Information technology—coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s. ISO/IEC 11172, International Standards Organization.

ISO/IEC. 1994. Information technology—generic coding of moving pictures and associated audio. JTC1/SC29, ISO/IEC 13818-2 Committee Draft, International Standards Organization.

ITU-T. 1990. Recommendation H.261—video codec for audiovisual services at p ¥ 64 kbit/s. International Telecommunications Union–Telecommunications Standardization Sector, Geneva, Switzerland.

Jayant, N., Johnston, J., and Safranek, R. 1993. Signal compression based on models of human perception. *Proc. IEEE,* 81(10):1385–1422.

Leduc, J.P. 1994. Bit-rate control for digital TV and HDTV codecs. *Sig. Proc. Image Commun.,* 6(l):25–45.

Le Gall, D.J. 1992. The MPEG video compression algorithm, *Sig. Proc. Image Commun.,* 4(2):129–140.

Mitiche, A. and Bouthemy, P. 1996. Computation and analysis of visual motion: A synopsis of current problems and methods. *Int. J. Comp. Vision,* to be published.

Netravali, A.N. and Haskell, B.G. 1995. *Digital Pictures: Representation, Compression, and Standards,* 2nd ed., Plenum, New York.

Schreiber, W.F. 1993. *Fundamentals of Electronic Imaging Systems: Some Aspects of Image Processing,* 3rd ed., Springer–Verlag, Berlin.

Zdepski, J., Raychaudhuri, D., and Joseph, K. 1992. Statistically based buffer control policies for constant rate transmission of compressed digital video. *IEEE Trans. Commun.,* 39(6):947–957.

Further Information

More details on the various aspects of video compression can be found in Netravali and Haskell [1995] and Schreiber [1993]. Current research is reported in a number of journals, including *IEEE Transactions on Image Processing, IEEE Transactions on Circuits and Systems for Video Technology*, and *Signal Processing: Image Communication.* Annual conferences of interest include the SPIE International Conference on Visual Communication and Image Processing, the IEEE International Conference on Image Processing, the European Conference on Signal Processing, and many more.

97

High Quality Audio Coding

Peter Noll
Technische Universität

97.1 Introduction

On a compact disc (CD), today's quasi standard of digital audio, signals with a 10 to 20,000 Hz bandwidth and 44.1 kHz sampling rate are stored with a resolution of 16 bits per sample. The resulting net bit rate is $2 \times 44.1 \times 16 = 1412$ kb/s per stereo channel.[1]

The compact disc has made digital audio popular and serves still as an accepted *audio representation* standard. [In *audio production*, higher resolutions up to and above 24 bits per sample and sampling rates up to 192 kHz are in use.] Lower bit rates than those given by the 16-bit PCM format of the CD are mandatory if audio signals are to be transmitted over channels of limited capacity, in particular over mobile radio channels, or are to be stored on storage media of limited capacity. *Basic requirements* in the design of low bit rate audio coders are: (1) high quality of the decoded signals; (2) robustness to random and bursty channel bit errors and to packet (cell) losses; (3) graceful degradation of quality with increasing bit error rates in mobile radio and broadcast applications; and (4) low complexity and power consumption.

Typical application areas for *digital audio* with bandwidths of not less than 20 kHz are in the fields of audio production, distribution and program exchange, digital audio broadcasting (DAB), digital storage

[1]The CD needs a significant overhead for a runlength-limited line code, for synchronization, and error correction, resulting in a 49-bit representation of each 16-bit audio sample. Hence, the total stereo bit rate is $1.41 \times 49/16 = 4.32$ Mb/s.

(archives, studios, consumer electronics), digital audio on PCs, video-conferencing, multipoint interactive audiovisual communications, and in the field of enhanced quality TV systems.

In the case of *digital storage* of compressed audio, a number of consumer products are on the market. The coded bit streams are stored on magnetic discs or on chip-based memory cards. For example, various **MP3** players make use of the MP3 coding algorithm, a member of a collection of **MPEG** audio coding standards. A second example is Sony's magneto-optical *MiniDisc* (*MD*), which makes use of the adaptive transform acoustic coding (ATRAC) compression algorithm.

Delivery of digital audio signals must be possible over a wide variety of digital channels. In the case of broadcasting, its network chains will include sections with different quality requirements ranging from production quality, where editing, cutting, postprocessing, etc., have to be taken into account, down to emission quality. **ITU-R** has recommended the use of ISO/MPEG-1 Layers II and III (MP3) audio coding algorithms for broadcasting digital audio to the end user.

Of great interest is a high quality transmission of audio bit streams through narrowband **ISDN** networks operating at rates of 64 or 128 kb/s (1 or 2 B channels, respectively). These channels are very useful for a practical distribution of audio signals. Major problems have to be solved to provide high quality for listeners using mobile and portable receivers since multipath interference and selective fading can otherwise cause severe impairments. Recently, the Internet network has become popular for transmitting audio and broadcast programs despite its inherent (and statistical) packet delays and packet losses. The problems of adjusting coders to channel impairments will not be covered in this chapter.

97.2 Auditory Masking

Digital representations of analog waveforms introduce some kind of distortions. A basic problem in the design of source coders is to achieve a given acceptable level of distortion with the smallest possible encoding bit rate. To reach this goal the encoding algorithm must be adapted both to the changing statistics of the source and to auditory perception. Unlike speech, there is no production model for audio signals, therefore audio coding can only exploit auditory perception by transmitting only those details of the signal which are perceptible by ear. Such coders provide high coding quality without providing high signal-to-noise ratios.

Auditory perception is based on critical band analyses in the inner ear where frequency-to-place transformations occur along the basilar membrane. The power spectra are not represented on a linear frequency scale but on frequency bands called *critical bands* with bandwidths in the order of 100 Hz below 500 Hz and with increasing bandwidths (up to 5000 Hz) at high signal frequencies [Zwicker, 1982]. Within critical bands, the intensities of individual tones are summed by the ear. Up to 24,000 Hz bandwidth, twenty-six critical bands have to be taken into account.[2] All audio coders are based on critical-band structured signal processing.

Auditory masking plays a major role in the design of bit-rate efficient coders. It describes the effect that a low level audio signal (the maskee) can become inaudible when a louder signal (the masker) occurs simultaneously. Without a masker, a signal is inaudible if its sound pressure level (SPL)[3] is below the absolute hearing threshold, called *threshold in quiet*. It depends on frequency and covers a dynamic range of more than 60 dB, as shown in the lower curve of Fig. 97.1. In the presence of a masker, a *masking threshold* can be measured, below which any signal will not be audible. This masking threshold, in the context of source coding also known as the *threshold of just noticeable distortion*, depends on the sound pressure level (SPL), the frequency of the masker, and on characteristics of the masker and maskee. If many maskers are present simultaneously in different frequency bands a *global masking threshold* can be derived from the individual masking thresholds.

Consider the example of the masking threshold for the SPL = 60 dB smallband noise masker in Fig. 97.1: around 1 kHz, the SPL of any maskee can be surprisingly high; it will be masked as long as its SPL is within

[2] Telephone bandwidth (up to 3400 Hz) can be described by fifteen or sixteen critical bands.

[3] SPL = 0 dB relates to a sound pressure of 0.02 mN/m^2.

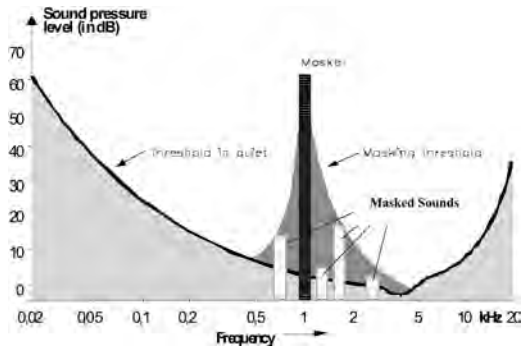


FIGURE 97.1 Threshold in quiet (lower curve) and masking threshold. (Acoustical events in the gray areas will not be audible).

the dark-hatched area. The distance between the level of the masker and the masking threshold is called the *signal-to-mask ratio* (*SMR*). SMR values of around 6 dB are needed to mask a tone by noise; much higher differences of 15 dB and more are needed to have noise being masked by a tone. Coding noise and low level signals will not be audible as long as the SNR is higher than the SMR. The difference between the SMR and SNR is the *noise-to-mask ratio* (*NMR*), which describes the safety margin for non-audible noise and should be of negative sign.

We note in passing that, in addition to simultaneous masking, *temporal masking* plays an important role in human auditory perception. It occurs when two sounds appear within a small interval of time; the stronger one masks the weaker one, regardless of whether the weaker one occurs before or after the stronger. Temporal masking can be used to mask pre-echoes caused by the spreading of a sudden large quantization error over the actual coding block.

97.3 Noise Shaping and Perception-Based Coding

The masking phenomenon depends on the spectral distribution of masker and maskee and on their variation with time. The masked signal may either consist of quantization (coding) noise, of aliasing or channel errors, and/or of low level signal components. A long as these contributions are below the masking threshold, they are irrelevant to the ear and, hence, inaudible.

The dependence of human auditory perception on frequency and the accompanying perceptual tolerance to errors can (and should) directly influence encoder designs; *dynamic noise-shaping* techniques can shift coding noise to frequency bands where it is not of perceptual importance. The noise shifting must be related to the actual short-term input spectrum in accordance with the signal-to-mask ratio and can be done in different ways. In *audio coding*, noise-shaping parameters are dynamically controlled such that simultaneous masking and temporal masking is exploited in an efficient manner. Note that the noise shaping is located in the encoder only, so it does not contribute to the bit rate. Frequency domain coders with *dynamic allocations of bits* (and, hence, of quantization noise contributions) to **subbands** or **transform** coefficients offer an easy and accurate way to control quantization noise [Jayant and Noll, 1984].

Figure 97.2 depicts the structure of a perception-based coder that exploits auditory masking. The encoding process is controlled by the signal-to-mask ratio (SMR) vs. frequency curve from which the needed resolution (and hence the bit rate) in each critical band is derived. The SMR is determined from a high resolution FFT-based spectral analysis of the audio block to be coded. Principally, any coding scheme can be used that can be dynamically controlled by such perceptual information. If the necessary bit rate for a complete masking of distortions is available, the coding scheme will be *perceptually transparent*, i.e., the decoded signal is then *subjectively indistinguishable* from the source signal, e.g., the compact disc reference. In practical designs, we cannot go to the limits of masking or just noticeable

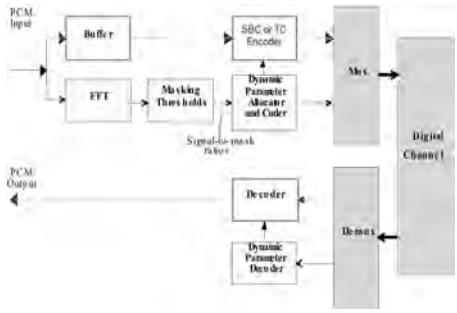


FIGURE 97.2 Block diagram of perception-based coders.

TABLE 97.1 5-Point MOS Impairment Scale

MOS Score	Impairment Scale
5	Imperceptible
4	Perceptible, but not annoying
3	Slightly annoying
2	Annoying
1	Very annoying

distortion, since post-processing of the acoustic signal (e.g., filtering in equalizers) by the end-user and multiple encoding/decoding processes have to be considered. In addition, our current knowledge about auditory masking is very limited. For example, generalizations of masking results, derived for simple and stationary maskers and for limited bandwidths, may be appropriate for most source signals but may fail for others. Therefore, as an additional requirement, we need a sufficient safety margin in practical designs of such perception-based coders.

97.4 Mean Opinion Score

Perceptual coding techniques ask for a distortion measure that specifies *perceptual similarity* between the source signal and its reconstructed version. In audio coding, the **mean opinion score (MOS)** is usually being applied to measure this similarity. Listeners classify the quality of coders on an N-point quality or impairment scale. Table 97.1 shows the *signal impairment scale* and its associated numbering [ITU-R Rec. 562]; see also Jayant and Noll [1984]. The impairment scale is extremely useful if coders with only small impairments have to be graded. The final result of such tests is an averaged judgment called the mean opinion score. In assessing the value of a given MOS rating, the standard deviation of this rating is also important.

Recent results have shown that perception-based objective distortion measures that make use of results from psychoacoustics are quite promising measures of quality. A recent ITU-R standard recommends a "method for objective measurements of subjectively perceived audio quality" [ITU-R, 1999; Thiede, 1999].

97.5 Low Bit Rate Coding

Digital coding at high bit rates is dominantly waveform-preserving, i.e., the amplitude-vs.-time waveform of the decoded signal approximates that of the input signal. At lower bit rates, the error criterion has to be in favor of an output signal that is useful to the human receiver rather than favoring a decoded signal that follows and preserves the input waveform. Basically, any efficient source coding algorithm will

- remove redundant components of the source signal by exploiting correlations between its samples either in time domain by employing adaptive linear prediction (LPC) or in frequency domain by employing high-resolution bandwidth splitting in subband coding (SBC) or transform coding (TC) schemes. Combinations thereof are also used.
- remove components of the source signal, which are irrelevant to the auditory system, i.e., to the ear. Irrelevancy manifests itself as unnecessary amplitude or frequency resolution; all portions of the source signal which are masked are irrelevant. Recently obtained results in audio bit rate compressions result from an efficient identification and subsequent removal of such irrelevant components prior to transmission.

Redundancy reduction can easily be achieved by using frequency-domain coding, such as transform coding, subband coding, and hybrid forms thereof. In addition—and more importantly—frequency-domain coding offers an easy and direct way for perceptual coding, i.e., for noise shaping and suppression of frequency components that need to be transmitted. In the following, we offer brief summaries of a number of audio coding algorithms which are based on such frequency mappings as a first stage of audio coding. The differences are in the number of spectral components, the type of mapping (subband splitting vs. discrete transform), and in the strategies for an efficient quantization of spectral components and masking of the resulting coding errors.

97.6 ITU-T G.722 Wideband Speech Coder

This 64 kb/s wideband speech coder can code signals with a 3 dB-bandwidth of 50–7000 Hz [Mermelstein, 1988]. The ITU-T standard employs a two-band subband technique and a backward-adaptive **ADPCM** coding in each of the two subbands. The subband splitting is based on two identical quadrature mirror filters (QMF) to divide the wideband signal, sampled at 16 kHz, into two 8-kHz sampled components, called low subband and high subband. The filters overlap, and aliasing will occur because of subsampling of each of the components; the synthesis QMF filterbank at the receiver ensures that aliasing products are canceled. However, quantization error components on the two subbands will not be eliminated completely. Therefore 24-tap QMF filters with a stop-band rejection of 60 dB are employed.

The coding of the subband signals is based on ADPCM with backward-adaptive predictors and quantizers, their design is borrowed from [ITU-T Rec.G.726]. The quantizers, however, do not support the slowly adapting mode. High quality coding is provided by a fixed bit allocation, where the low and high subband ADPCM coders use a 6-b/sample and a 2-b/sample quantizer, respectively. In the low subband, the signal resembles the narrowband speech signal in most of its properties. A reduction of the quantizer resolution to 5 or 4 b/sample is possible to support transmission of auxiliary data at rates of 8 and 16 kb/s. Embedded encoding is used in the low subband ADPCM coding, i.e., the adaptations of predictor and quantizer are always based only on the four most significant bits of each ADPCM codeword. Hence, a stripping of one of two least significant bits from the ADPCM codewords does not affect the adaptation processes and cannot lead to mistracking effects caused by different decoding processes in the receiver and transmitter.

Subjective speech coding tests have shown that the 64 kb/s ADPCM-coded version has the same MOS rating (MOS = 4.3) as the uncoded source, implying no measurable difference in subjective quality. In addition, the coder shows a graceful degradation of subjective quality with decreasing bit rate and increasing bit error rate. The coder also provides a good to fair quality for coding of audio signals. Table 97.2

TABLE 97.2 MOS Performance of ITU-T G.722 Coder (7 kHz bandwidth) [Mermelstein, 1988]

Bit Rate in kb/s	MOS Value for Speech	MOS Value for Audio
48	3.8	3.2
64	4.2	3.9

compares MOS values for speech and audio signals at two bit rates. The resulting subjective quality is partly limited by the restriction on the bandwidth. Current activities in wideband speech coding concentrate on coding at 16 kb/s and below, with the 64 kb/s ITU-T standard serving as reference.

MPEG-1 Layer I and II coders are examples of subband coders with a high number of subbands. They are more suitable for perceptual coding; details will be given in a later section.

97.7 Transform Coding

The *conventional adaptive transform coding* (*ATC*), as proposed in the midseventies [Zelinski and Noll, 1975 and 1977],[4] employs a discrete cosine transform (DCT) which maps blocks of N samples into an equivalent number of transform coefficients. The logarithm of the absolute values of these coefficients form the DCT-based short-term log-spectral envelope; its encoded version is used to calculate, for each audio block, the optimum number of bits for each of the N quantizers for the transform coefficients (dynamic bit allocation). The encoder transmits the quantized transform coefficients and the encoded spectral envelope.

The conventional ATC has a number of shortcomings; it produces aliasing effects due to spectral overlap, and it produces pre-echoes because the inverse transform in the decoder spreads large quantization errors—possibly resulting from a sudden attack in the source signal—over the reconstructed block. In addition, inappropriate bit allocations occur if these allocations are based on an unweighted mean-squared coding error criterion. However, it was already stated in the first publication that the ATC performance could be improved if emphasis would be laid on irrelevancy reduction, and that the available bits had to be allocated dynamically such that the *perceived* distortion is minimal [Zelinski and Noll, 1975].

The problem of aliasing can be partly solved by introducing a modified discrete cosine transform (MDCT) with an 50% overlap between adjacent audio blocks. Pre-echoes can be reduced or avoided by using blocks of short lengths. However, shorter blocks cause an insufficient frequency resolution, and a larger percentage of the total bit rate is required for the transmission of side information. A solution to this problem is to switch between block sizes of different lengths of, say, N = 64 to N = 1024 (window switching). The small blocks are only used to control pre-echo artifacts during nonstationary periods of the signal, otherwise the coder switches back to long blocks. It is clear that the block size selection has to be based on an analysis of the characteristics of the actual audio coding block. In addition, window switching has to be carried out such that the aliasing cancellation property of the transform is maintained. MDCT and window switching are used in all advanced audio coders.

97.8 Subband Coding and Hybrid Frequency Mappings

Subband coding and transform coding were first introduced in the context of speech coding. The differentiation between these two categories is mainly due to historical reasons. Both use an analysis filterbank in the encoder to produce subsampled spectral coefficients. In the case of subband coding, this filterbank consists of a typically small number of bandpass filters and each output is critically sampled at a rate that is twice the *nominal* bandwidth of each bandpass. It is easy to show that an N-valued transform can be viewed as a critically decimated bank of N finite impulse response bandpass filters.

Hybrid structures are cascades of a subband and a transform filterbank. MPEG's MP3 and Sony's ATRAC coders employ such hybrid mappings (see later sections). One of the advantages is that different frequency resolutions can be provided in a flexible way and with low complexity.

[4] The conventional ATC used, for the first time, DCT as the appropriate linear transform for speech, it intro-duced the technique of adaptive bit allocation, and it showed that a log-spectral envelope should be computed and parameterized to derive the appropriate bit allocation.

97.9 ISO/MPEG Audio Coding

The moving pictures expert group within the International Organization of Standardization (ISO/MPEG) has provided a generic source coding method (MPEG 1) for storing audio-visual information on digital storage media at bit rates up to about 1.5 Mb/s [ISO/MPEG 1993; Brandenburg and Stoll, 1994; Noll, 1997]. MPEG-1 Audio has been the first international standard in the field of digital audio compression, useful both for audio-visual and audio-only applications.

In all MPEG standards (see next sections), each spectral component is quantized and coded with the goal to keep the quantization noise below the masking threshold. The number of bits for each subband and a scale factor are determined on a block-by-block basis. The number of quantizer bits is obtained from a dynamic bit allocation algorithm that is controlled by a *psychoacoustical model*. The subband codewords, the scale factors, and bit allocation information are multiplexed into one bitstream, together with a header and optional ancillary data. In the decoder, the synthesis filterbank reconstructs a block of 32 audio output samples from the demultiplexed bitstream.

It is worthwhile to note that the normative part of all MPEG standards describes the decoder and the meaning of the encoded bitstream, but that the encoders are not defined, thus leaving room for evolutionary improvements of the encoders. In particular, different psychoacoustic models can be used ranging from very simple (or none at all) to very complex based on quality and implementability requirements (the standard gives two examples of such models). In addition, encoders can be optimized for certain applications. It should be noted that the model is only needed in the encoder which makes the decoder less complex, a desirable feature for audio playback and audio broadcasting applications.

The structure of MPEG coders follows that of perception-based coders. In the first step, the audio signal is converted into spectral components via an analysis filterbank; Layers I and II make use of a subband filterbank, Layer III employs a hybrid transform, and MPEG AAC is purely transform-based.

97.10 MPEG-1 and MPEG-2 Audio

The MPEG 1 standard consists of three layers I, II, and III of increasing complexity and subjective performance. It supports sampling rates of 32, 44.1, and 48 kHz and bit rates per monophonic channel between 32 and 448 kb/s in 32 kb/s steps (Layer I). An additional free format mode allows for any bit rate below 448 kb/s. The coder provides a quality that is subjectively equivalent to compact disc (CD) quality (16 bit PCM) at a stereo rate of 256 kb/s for most audio items.

The MPEG-1 coder uses perception-based frequency-domain coding. In the first stage, the audio signal is converted into 32 spectral subbands components via an analysis filterbank. Each spectral component is quantized, whereby the number of quantizer levels for each component is controlled by a psychoacoustic model that computes signal-to-mask ratios and obtains its information about the short-term spectrum by running an FFT in parallel to the encoder (see Fig. 97.2).

MPEG-1 Layer I contains the basic mapping of the digital audio input into 32 subbands via equally spaced bandpass filters (Fig. 97.3), fixed segmentation to format the data into blocks (8 ms at 48 kHz sampling rate), and quantization with block companding.

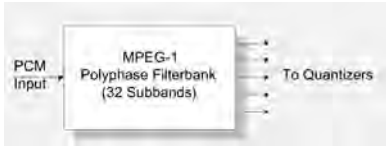


FIGURE 97.3 Structure of MPEG-1 Audio filterbank, Layers I and II.

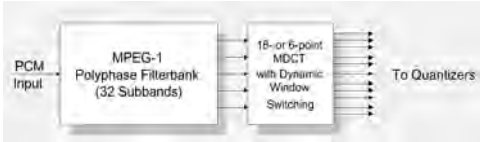


FIGURE 97.4 Hybrid filterbank of MPEG-1 Layer III (MP3) encoder.

At 48 kHz sampling rate, each band has a width of 750 Hz. A polyphase filter structure is used for the frequency mapping; its filters are of order 511, which implies an impulse response of 5.33 ms length (at 48 kHz). In each subband, blocks of 12 decimated samples are formed, and for each block, *signal-to-mask ratios* are calculated via an 512-point FFT. For each subband, the bit allocation selects a quantizer (out of a set of 15) such that both the bit rate requirement and the masking requirement are being met.

The decoding is straightforward: the subband sequences are reconstructed on the basis of 12-sample subband blocks taking into account the decoded scalefactor and bit allocation information. If a subband has no bits allocated to it, the samples in that subband are set to zero. Each time the subband samples of all 32 subbands have been calculated, they are applied to the synthesis filterbank, and 32 consecutive 16-bit PCM format audio samples are calculated.

The MPEG-1 Layer II coder is basically similar to the Layer I coder but achieves a better performance through three modifications. First, the input to the psychoacoustic model is a 1024-point FFT leading to a finer frequency resolution for the calculation of the global masking threshold. Second, the overall scalefactor side information is reduced by exploiting redundancies between the scale factors of three adjacent 12-sample blocks. Third, a finer quantization with a maximum of 16-bit amplitude resolution is provided (which reduces the coding noise).

MPEG-1 Layer III introduces many new features. The Layer III approach has become quite popular, in particular in Internet applications (MP3). MP3 achieves its improvement mainly by an improved frequency mapping, an analysis-by-synthesis approach for the noise allocation, an advanced pre-echo control, and by nonuniform quantization with entropy (Huffman) coding. Finally, a buffer technique, called *bit reservoir*, leads to further savings in bit rate. In order to achieve a higher frequency resolution closer to critical band partitions, the 32 subband signals are subdivided further in frequency content by applying a switched 6-point or 18-point modified DCT (MDCT) with 50% overlap to each of the subbands (see Fig. 97.4). The maximum number of frequency components is $32 \times 18 = 576$ each representing a bandwidth of only $24000/576 = 41.67$ Hz. The term *hybrid filterbank* is used to describe such a cascade of polyphase filterbank and MDCT transform. The decoding follows that of the encoding process.

The MPEG-2 standard includes an extension that adds sampling frequencies of 24, 22.05 and 16 kHz [ISO/MPEG 1995; Stoll, 1996]. These sampling frequencies are useful for the transmission of wideband speech and medium quality audio signals. In addition, a reduction in sampling frequency by a factor of two doubles the frequency solution of the filterbanks and transforms and will therefore improve the performance of the coders, in particular those of low frequency resolution (MPEG-1, Layers I and II). Only small modifications of ISO/MPEG-1 are needed for this extension; the bitstream syntax still complies with the standard.

97.11 MPEG Advanced Audio Coding

More recently, MPEG advanced audio coding (AAC) has been standardized [ISO/MPEG 1997; Bosi et al., 1997]. This standard offers the highest subjective quality and is about to become a universal standard in many application areas with totally different requirements such as consumer electronics, professional audio processing, telecommunications, and broadcasting.

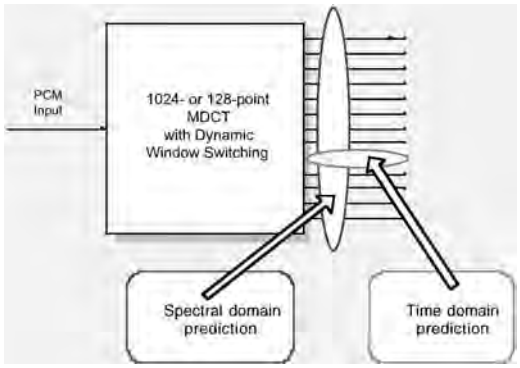


FIGURE 97.5 Spectral and time domain prediction in MPEG-2 advanced audio coding (AAC).

The MPEG-2 AAC standard employs high resolution filterbanks, prediction techniques, and noiseless coding. It is based on recent evaluations and definitions of tools (or modules) each having been selected from a number of proposals. The self-contained tools include an optional preprocessing, a filterbank, a perceptual model, temporal noise shaping, intensity multichannel coding, time domain prediction, M/S stereo coding, quantization, noiseless (Huffman) coding, and a bitstream multiplexer. The filterbank is a 1024-point modified discrete cosine transform (due to an 50% overlap, the transform is taken over 2048 windowed samples). In case of transient signals, the filterbank can switch to a 128-point mode. The perceptual model can be taken from MPEG-1.

The temporal noise shaping (TNS) tool plays an important role in improving the overall performance of the coder (see Fig. 97.5). It performs a prediction of the spectral coefficients of each audio frame. Instead of the coefficients, the prediction residual is transmitted. TNS is very effective in case of transient audio signals since such transients (signal attacks) imply a high predictability in the spectral domain. (Recall that peaky spectra lead to a high predictability in the time domain.) Therefore, the TNS tool controls the time dependence of the quantization noise. Time domain prediction is applied to subsequent subband samples in a given subband in order to further improve coding efficiency, in particular for stationary sounds (see Fig. 97.5). Second-order backward-adaptive predictors are used for this purpose. Finally, for quantization, an iterative method keeps the quantization noise in all critical bands below the global masking threshold.

The MPEG-2 AAC standard offers high quality at lowest possible bit rates. For example, when operating at a 128 kb/s stereo rate, AAC has a mean opinion score (MOS) only 0.2 off the reference value. At that rate, the popular MPEG-1 Layer III coder (MP3) is almost one full score value below the reference value [Meares et al., 1998].

AAC, a joint effort of AT&T, Dolby Laboratories, Fraunhofer Institute, and Sony Corporation provides up to 48 channels (see below) and sampling rates up to 96 kHz. It is supported by a growing number of hardware and software manufacturers.

97.12 Proprietary Audio Coding

Dolby's AC-3 coder (Dolby Digital) is a transform coder with sampling rates of 32, 44.1, and 48 kHz and nineteen mono channel bit rates in the range from 32 to 640 kb/s. The coder supports various multichannel configurations and provides the option of channel down-mixing adjustable by the end user [Fielder et al., 1996]. Unlike other approaches, Dolby's AC-3 coder makes use of hybrid backward/forward adaptive bit allocation to reduce the necessary bit rate for transmitting bit allocation side information. The bit allocation routine, called *parametric bit allocation*, runs both in the encoder and the decoder,

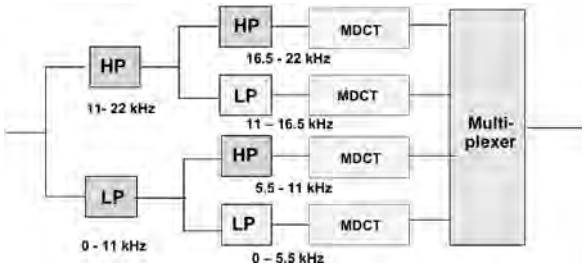


FIGURE 97.6 Block diagram of Sony's ATRAC3 audio encoder.

based on a psychoacoustic model. The forward mode of the algorithm allows the decoder to be re-adjusted by the encoder. If the encoder uses an improved bit allocation routine, it can compare its result with that of the routine used by the decoder and send correction information to fine-tune the decoder's bit allocation. The final bit allocations of the encoder and decoder are, of course, identical. [An earlier version, AC-2, uses a purely backward-adaptive bit allocation based on quantized information. In case of an error-free transmission, this information is also available at the decoder, therefore no side information is required to inform the decoder about the actual bit allocation.] Applications will be covered in the section about multichannel standards.

AC-3 coding serves as an emission or storage coding format optimized for maximum quality at low bit rates, typically for one encode/decode cycle. Within broadcast and postproduction infrastructures, other formats have to be employed. One example is the Dolby E distribution coding system with a sampling rate of 48 kHz and a resolution of 16, 20, and 24 bits per audio sample.

Sony's Adaptive Transform Acoustic Coding (ATRAC) uses a hybrid frequency mapping: the signal is first split into three subbands (0–5.5, 5.5–11.0, and 11.0–22.0 kHz) reflecting auditory perception [Iwadare et al., 1992]. An MDCT with 50% overlap is then applied to all subbands with dynamic window switching: the regular block size is N = 512, to provide optimum frequency resolution, but in the case of possible pre-echoes, N is switched to 64 in the high frequency bands and to N = 128 otherwise. The coder with its 44.1 kHz sampling rate and its 256 kb/s stereo bit rate has been developed for portable digital audio, specifically for the 64 mm optical or magneto-optical disc which has approximately one-fifth of the byte-capacity of a standard CD [Tsutsui et al., 1992].

More recently, ATRAC3 was released. It achieves a higher compression rate mainly by splitting the audio signal into four bands (instead of three) and by employing a 256-point MDCT to all four bands (see Fig. 97.6). In addition, tonal audio components are identified and coded separately from other spectral components. Entropy coding reduces the bit rate further. Finally, a joint stereo coding scheme takes advantage of the correlations between stereo channels.

97.13 Multichannel Standards

In multichannel coding, redundancies and irrelevancies such as intra- and interchannel correlations and masking effects are exploited to reduce the overall bit rate. In addition, stereo-irrelevant components of the multichannel signal may be identified and reproduced in a monophonic format to bring the bit rates further down. A typical multichannel system may employ five channels (L = left, C = center, R = right, LS = left surround, RS = right surround) but up to 48 channels are supported by the AAC standard.

The MPEG standards include two classes of multichannel audio coding, one of which is forward- and backward-compatible with MPEG-1 Audio [ISO/MPEG 1993]. Forward compatibility means that an MPEG-2 multichannel decoder is able to properly decode MPEG-1 mono- or stereophonic signals. Backwards compatibility means that existing MPEG-1 stereo decoders, which can only handle two-channel audio, are able to reproduce a meaningful basic stereo signal from an MPEG-2 multichannel

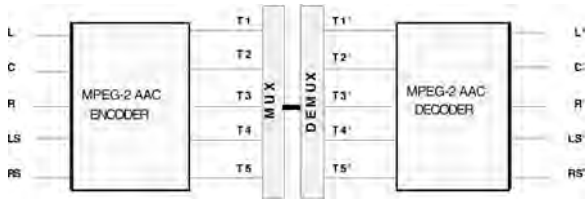


FIGURE 97.7 MPEG-2 advanced audio coding (multichannel configuration; T = transmission channels).

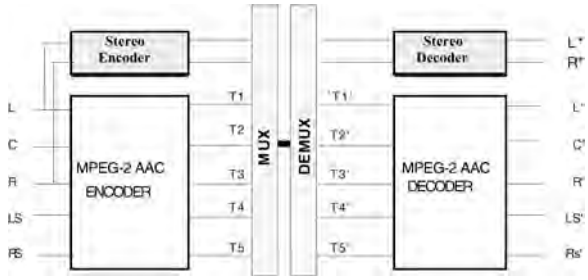


FIGURE 97.8 Backward-compatible multichannel audio coding (simulcast mode).

bitstream so as to serve the needs of users with simple mono or stereo equipment. However, due to its constraints caused by matrixing and dematrixing procedures,[5] the compatible coder will need a higher bit rate for a given audio quality.

Non-backward-compatible multichannel coders will not be able to feed a meaningful bitstream into an MPEG-1 stereo decoder. On the other hand, these codecs have more freedom in producing high quality audio signals. Therefore, we will only cover the standard that supports applications which do not request compatibility with the existing MPEG-1 stereo format. Within the MPEG-2 standard, the multichannel advanced audio coding technique (AAC) offers highest performance at comparably low overall bit rates (see Fig. 97.7).

In extensive formal subjective tests, it turned out that the non-backward-compatible 320 kb/s MPEG-2 AAC coder performed slightly better than the backward-compatible 640 kb/s MPEG-2 Layer II coder. From these subjective tests, it became clear that the concept of backward compatibility implies the need for significantly higher bit rates. The AAC coder satisfies requirements for ITU-R broadcast quality at 320 kbit/s for 5 channels.

As mentioned above, Dolby's AC-3 coding supports up to nineteen mono channel bit rates from 32 to 640 kb/s which allow for a high number of multichannel configurations. A first multichannel application of AC-3 has been for digital optical soundtracks in 35 mm cinema. The soundtrack consists of data between the sprocket hole perforation and presents a 5.1 multichannel audio system operating at a bit rate of only 320 kb/s. AC-3 has been selected as the audio coding standard in the US DTV system with a typical bit rate of 384 kb/s [ATSC, 1995; Fielder et al., 1996]. In DVD Video, it is mandatory to support linear PCM (see section on lossless coding) and AC-3 coded audio (whereas MPEG audio is one of the mandatory audio coding modes for TV systems with 625 lines and 25 frames only).

If bit rates are not of highest concern, a *simulcast transmission* may be employed, where a full stereo bitstream is multiplexed with a full non-backward-compatible multichannel bitstream in order to support backward compatibility without matrixing techniques (Fig. 97.8).

[5] Matrixing describes a down-mix technique of, say, five-channel audio signals into two compatible stereo sig-nals. Three additional transmission channels carry audio information that has to be chosen such that the decoder can recompute the complete 5-channel audio signal (dematrixing).

97.14 MPEG-4 Audio Coding

Activities within the MPEG-4 standardization process have aimed at proposals for a broad field of applications including communication services, interactive services, and broadcast services. The new standard [ISO/MPEG 1999] takes into account that a growing part of information is read, seen, and heard in interactive ways. It supports new forms of communications, in particular for Internet and multimedia applications and in mobile communications. MPEG-1 and MPEG-2 offer only a very limited interaction and control over the presentation and configuration of the system. In addition, an integration of natural and synthetic content is difficult, and access and transmission across heterogeneous networks is not well-supported. MPEG-4 is different: it represents an audiovisual scene as a composition of (potentially meaningful) objects and supports the evolving ways in which audiovisual material is produced, delivered, and consumed. For example, computer-generated content may become part of the production of an audiovisual scene. In addition, interaction of objects with scene is possible, so it will be possible to associate a Web address to a person in a scene.

In the case of audio, MPEG-4 merges the whole range of audio from high fidelity audio coding and speech coding down to synthetic speech and synthetic audio, supporting applications from high-fidelity audio systems down to mobile-access multimedia terminals [Brandenburg et al., 2000; Johnston et al., 2000; Scheirer et al., 2000]. Figure 97.9 indicates the potential of MPEG-4; it describes an audiovisual scene with a number of audio objects: the noise of an incoming train, an announcement, a conversation, and background music. For example, the noise of the train can be described by an eight-channel representation. On the other hand, if the necessary bandwidth is not available, a one-channel representation—or no representation at all—might be used instead. Such a form of scalability will be very useful in future applications whenever audiovisual signals have to be transmitted to and via receivers of differing complexity and channels of differing capacity. In the case of the announcement, one-channel pseudo 3-D and echo effects could be added. The background music may have an AAC format, or it is of synthetic origin [ISO/MPEG 1998].

MPEG-4 supports compression of speech and audio signals at monophonic bit rates ranging from 2 to 64 kb/s. Three core coders are used:

- a parametric coding scheme (vocoder) for low bit rate speech coding (2 to 10 kbit/s)
- a CELP-based analysis-by-synthesis coding scheme for medium bit rates (4 to 16 kb/s)
- a transform-based coding scheme for higher bit rates (up to 64 kbit/s)

The transform-based coder is an AAC coder with a number of modifications. Important examples are (1) a less complex longterm prediction tool, (2) a perceptual noise substitution tool which allows a very efficient coding of noiselike signal components, and (3) bit rate scalability with the potential of adapting the rate dynamically to the time-varying capacity of a channel.

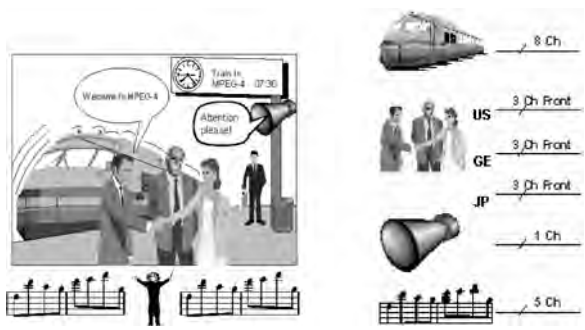


FIGURE 97.9 Audiovisual scene and its audio channels [ISO/MPEG 1998].

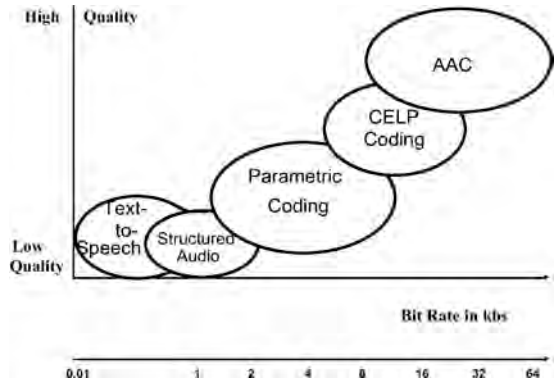


FIGURE 97.10 Range of audio quality and bit rates in MPEG-4.

MPEG-4 not only offers simple means of manipulation of coded data such as time scale control and pitch change, but also a flexible access to coded data and subsets thereof, i.e., scalability of bit rate, bandwidth, complexity, and of error robustness. In addition, MPEG-4 supports not only natural audio coding at rates between 2 and 64 kb/s, but also text-to-speech conversion (TTS) and structured audio [Scheirer et al., 2000]. The standard offers also an interface between TTS and facial animation for synthetic face models to be driven from speech (talking heads).

Figure 97.10 indicates the range of quality and bit rates offered by the new standard. Transform-based audio coders show a very good performance at bit rates down to 16 kb/s, whereas speech coders perform clearly better at rates between 2.4 kb/s and 16 kb/s.

97.15 Lossless Coding

For archiving and processing of audio signals, the coding will have to be *lossless*. Lossless coding does not alter the final decoded audio signal in any way (i.e., in any bit). Digital versatile discs for audio (DVD Audio), with their capacity of 4.7 GB (single layer) or 8.5 GB (double layer) (or up to 17 GB if both sides of the discs are readable), allow storage of audio material without any source coding [Fuchigami et al., 2000]. Therefore, linear PCM was adopted for DVD Audio as one of the mandatory audio coding schemes. Six sampling rates in the range between 44.1 and 192 kHz and amplitude resolutions of 16, 20, and 24 bits per audio sample have been specified.

Lossless coding compresses (packs) PCM-format audio signals in a way that the decoded versions are bit-precise replica of the PCM format. Packed PCM was specified as an additional mandatory coding algorithm for DVD Audio discs [Stuart et al., 2000]. The encoder employs

- a lossless matrix technique to reduce inter-channel correlations
- prediction for each channel
- entropy coding
- buffering to smooth the encoded data rate

In addition, it exploits unused capacity from audio data that is not full scale.

Lossless coding, when applied to 16-bit linear PCM (CD format), results in average bit rates of around 8 bits per sample depending on the audio material. Average values for the complete audio material of the SQAM CD [SQAM, 1988] are 5.0 for the popular linear predictive coding algorithm "Shorten" [Robinson, 1994], and 4.6 for the lossless transform audio compression (LTAC) algorithm of TU Berlin [Liebchen et al., 1999]. Table 97.3 compares coding results for the audio categories of the CD. The results are optimistic since all audio examples start and end with a pause. Please note that data rates in lossless coding are typically variable; the presented values are averages.

TABLE 97.3 Coding Results in Bits per Sample for the Audio Material of the SQAM Disc [Liebchen et al., 1999]

Category	Shorten (Public Domain)	LTAC (TU Berlin)
Alignment signals	6.29	6.06
Artificial signals	2.84	2.60
Single instruments	4.60	4.15
Vocal	5.35	4.83
Speech	5.45	5.36
Solo instruments	5.09	4.52
Vocal & orchestra	6.73	6.14
Orchestra	5.33	5.07
Pop music	6.37	6.03
CD average	4.98	4.56

97.16 Conclusion

Low bit rate digital audio is being applied in many different fields such as consumer electronics, professional audio processing, telecommunications, and broadcasting. Perceptual coding in the frequency domain has paved the way to high compression rates. The ISO/MPEG audio coding algorithms are widely accepted international standards. However, other (propriatory) coding schemes serving different goals and meeting other requirements have evolved and have found many applications. More recently, lossless coding has gained interest for both professional and consumer use where data integrity is mandatory.

Defining Terms

ADPCM: Adaptive differential pulse code modulation, a predictive coding technique with adaptation of quantizer step sizes and predictor coefficients based on the short-term statistics of the source signal.

Auditory masking: Simultaneous masking describes the effect that a low level audio signal (the maskee) can become inaudible when a louder signal (the masker) occurs simultaneously. Temporal masking describes the effect that a low level audio signal can become inaudible before and after a louder signal has been switched on and off.

ISDN: Integrated services digital networks offer end-to-end digital connectivity to support a wide range of communication applications. Customers have physical access to one or two 64 kb/s bearer (B) channels and one data (D) channel (the latter supports signaling at 16 kb/s but can also carry user information). Higher rates than 64 or 128 kb/s are achievable using multiple calls with channel aggregation and may be more generally available in the future as multiple-rate services.

ITU: International Telecommunication Union, a worldwide standardization organization. ITU-R is its radiocommunication section, formerly called CCIR, and ITU-T is its telecommunication section, formerly called CCITT.

Mean opinion score (MOS): A distortion measure that specifies *perceptual similarity* between an uncoded source signal and its coded version.

MPEG: The moving pictures expert group is the working group 11 of the ISO/IEC/JTC1/SC2 standardization body. Its task is to propose algorithms for coding of moving pictures and associated audio.

MP3: Acronym for the MPEG-1, Layer III audio coding standard.

Subband coding: Division of the source signal into bandpass signals (subbands) via a filterbank followed by individual coding of the subband signals.

Transform coding: Deriving transform (spectral) coefficients from blocks of the source signal via a linear transform followed by individual coding of the transform (spectral) coefficients.

References

ACATS (Advisory Committee on Advanced Television Service), 1994, Grand Alliance HDTV System Specification, Version 1.0.

ATSC Doc. A/52, 1995, ATSC Digital Audio Compression Standard (AC-3). [ATSC = Advanced Television Standardization Committee.]

Bosi, M. et al., 1997, ISO/IEC MPEG-2 advanced audio coding, *J. Audio Eng. Soc.,* 45(10), S789–S814.

Brandenburg, K., Kunz, O., and Sugiyama, A., 2000, MPEG-4 natural audio coding, *Signal Proc. Image Commun.,* 15, 423–444.

Brandenburg, K. and Stoll, G., 1994, The ISO/MPEG-audio codec: a generic standard for coding of high quality digital audio, *J. Audio Eng. Soc.,* 42, 780–792.

Fielder, L.D. et al., 1996, AC-2 and AC-3: low complexity transform-based audio coding, in Gilchrist, N. and Grewin, C., Eds., *Collected Papers on Digital Audio Bit-Rate Reduction,* Audio Engineering Society, Inc., New York, 73–82.

Fuchigami, N., Kuroiwa, T., and Suzuki, B.H., 2000, DVD—Audio Specifications, *J. Audio Eng. Soc.,* 48(12).

ISO/MPEG 1993, International Standard ISO/IEC 11172 Part 3 (MPEG-1 Audio).

ISO/MPEG 1995, International Standard ISO/IEC 13818 Part 3 (MPEG-2 Audio).

ISO/MPEG 1997, International Standard ISO/IEC 13818 Part 7 (AAC).

ISO/MPEG 1998, MPEG Document 98/N2431.

ISO/MPEG 1999, International Standard ISO/IEC 14496 Part 3 (MPEG-Audio).

ITU-R BS. 1387, 1999, Method for Objective Measurements of Subjectively Perceived Audio Quality.

Iwadare, M. et al., 1992, A128 kb/s hi-fi audio codec based on adaptive transform coding with adaptive block size, *IEEE J. Sel. Areas Commun.,* 10(1), 138–144.

Jayant, N.S. and Noll, P., 1984, *Digital Coding of Waveforms: Principles and Applications to Speech and Video,* Prentice-Hall, Englewood Cliffs, NJ.

Johnston, J.D. et al., 2000, Review of MPEG-4 general audio coding, in Puri, A. and Chen, T., Eds., *Multimedia Systems, Standards, and Networks,* Marcel Dekker Inc., New York–Basel, 131–156.

Liebchen, T., Purat, M., and Noll, P., 1999, Improved lossless transform coding of audio signals, in *Impulse und Antworten,* Wissenschaft & Technik Verlag, Berlin, 159–170.

Meares, D., Watanabe, K., and Scheirer, E., 1998, Report on the MPEG-2 AAC Stereo Verification Tests, MPEG Document 98/N2006.

Mermelstein, P., 1988, G722, a new CCITT coding standard for digital transmission of wideband audio signals, *IEEE Commun. Mag.,* 8–15.

Noll, P., 1997, MPEG digital audio coding, *IEEE Signal Proc. Mag.,* 14–59–81.

Robinson, T., 1994, Shorten: simple lossless and near-lossless waveform compression, Technical report CUED/F-INFENG/TR.156, Cambridge University Engineering Department.

Scheirer, E.D., Lee, Y., and Yang, J.W., 2000, Synthetic audio and SNHC audio in MPEG-4, in Puri, A. and Chen, T., Eds., *Multimedia Systems, Standards, and Networks,* Marcel Dekker Inc., New York–Basel, 157–178.

SQAM (Sound Quality Assessment Material—Recordings for Subjective Tests), 1988, Technical Centre of the European Broadcasting Union (EBU).

Stoll, G., 1996, ISO-MPEG2 audio: a generic standard for the coding of two-channel and multichannel sound, in Gilchrist, N. and Grewin, C., Eds., *Collected Papers on Digital Audio Bit-Rate Reduction,* Audio Engineering Society, Inc., New York, 43–53.

Stuart, J.R. et al., 2000, MLP lossless compression, AES 9th Regional Convention Tokyo, preprint H-6, 152–155.

Thiede, T., 1999, Perceptual audio quality assessment using a non-linear filterbank, PhD Thesis, Berlin.

Tsutsui, K. et al., 1992, Adaptive transform acoustic coding for MiniDisc, 93rd AES-Convention, San Francisco, preprint 3456.

Zelinski, R. and Noll, P., 1977, Adaptive transform coding of speech signals, *IEEE Trans. Acoustics Speech Signal Proc.,* 299–309.

Zelinski, R. and Noll, P., 1975, Adaptive Blockquantisierung von Sprachsignalen, Heinrich-Hertz-Institute, Technical Report No. 181, Berlin.
Zwicker, E., 1982, *Psychoacoustics,* Springer-Verlag, New York.

Further Information

For audio coding, the state of the art, as of 1996, is well documented in Gilchrist, N. and Grewin, C., Eds., 1996, *Collected Papers on Digital Audio Bit-Rate Reduction*, Audio Engineering Society, Inc., New York. The MPEG-4 standard (including audio) is covered in Puri, A. and Chen, T., Eds., 2000, *Multimedia Systems, Standards, and Networks,* Marcel Dekker Inc., New York–Basel.

There is a significant overlap between speech and audio coding. Therefore we recommend consulting the chapter on speech signal processing in this handbook and overview papers, such as "Advances in Speech and Audio Compression," *Proc. IEEE*, 82(6), 900–918, 1994, by A. Gersho.

Actual contributions to all aspects of digital audio processing and coding can be found at the biannual conventions of the Audio Engineering Society (AES) and the IEEE ASSP Workshops on Applications of Signal Processing to Audio and Acoustics.

98

Cable

Jeff Hamilton
General Instrument Corporation

Mark Kolber
General Instrument Corporation

Charles E. Schell
General Instrument Corporation

Len Taupier
General Instrument Corporation

98.1 Introduction

Cable TV, or community antenna TV (CATV), began shortly after World War II. Television broadcast stations were established in most major cities. The frequencies used for transmission were and are primarily line-of-sight transmission and of a power sufficient to reach consumers in the city and close-in suburbs. As television caught on, this gave rise to a large number of consumers who could not receive good quality pictures because they were in the shadow of a hill or too far from the transmitters.

The answer was to place a large antenna at a good receive site, on top of a mountain, for example, and distribute the signals to an entire community. The earliest such sites were in Astoria, OR and Williamsport, PA. The original cable was RG-11 military surplus from World War II.

Early CATV systems carried only 3 channels in the low-VHF band of 54–88 MHz. Over the years, this evolved to 174, 220, 270, 300, 330, 450, and 550 as broadcast TV added the UHF band channels, pay TV arrived with such services as HBO, and distribution electronics improved to allow high-quality delivery of the ever larger number of channels. The typical new plant is designed for 750 MHz, with a capacity for over one hundred analog video channels.

The advent of digital video compression, based on the moving pictures expert group-2 (MPEG-2) standard, promises to increase this capacity by an order of magnitude. Initially, digital services will be deployed on hybrid plants which carry a mix of analog and digital video signals. Typical plans are to offer 77 analog channels in the 50–550 MHz band which can be received by consumers with only analog equipment in the home. Added to this will be a band of digital audio and video services from 550 to 750 MHz. This 200-MHz digital cable band will provide the same capacity as the popular 32 transponder high-power DBS services.

New hybrid **set-top** converters will be deployed to subscribers of these services that receive both the analog and digital transmissions. Current compression technology provides, typically, 4–10 channels in each 6-MHz band, allowing well over 200 new video services in the initial 200-MHz digital band. As digital receiving equipment rolls out to more and more consumers, cable has the option of evolving to all digital transmission by progressively replacing analog channels with digital carriers.

From the early days of TV, the consumer's choice has been to select from what's playing on a few channels. Cable added dozens more channels and, through its ability to charge on a per channel basis, created the higher value subscription movie channels. Digital compression combined with the delivery of signals over fiber to neighborhoods of a few hundred homes now makes possible an efficient path to provide dedicated bandwidth to each subscriber. This capability amounts to infinite bandwidth in which everything a subscriber might want to view is on, by request. Instead of the consumers choosing among what programmers decide to send them, they will have the ability to tell the system what they want to see.

98.2 Cable System Architecture

The distribution system or cable plant that carries the signals from the **head end** to the subscribers can use various architectures. The common *tree and branch* architecture uses one inch or larger diameter solid aluminum outer conductor coaxial cable to carry the signal from the head end. This trunk cable runs through the system, distributing the broadband signal to each neighborhood. Feeder lines that branch off the trunk typically use smaller diameter cable to carry the signal into the neighborhoods passing each subscriber's home. Drop cables lead from the curb side to each home and are typically 1/4 inch flexible coaxial cable. All CATV cables are 75-W characteristic impedance, which is near optimum for low loss.

Devices called taps are used to extract a portion of the signal from a feeder to supply a drop cable. For example, a 20-dB tap extracts 1/100 of the power from the feeder, causing a 20-dB loss at the tapped output but only a 0.04-dB loss to the main path.

CATV signals are measured in **decibels relative to 1 millivolt** (**dBmV**). For example, 0 dBmV corresponds to 1 mV, and +60 dBmV to 1 V. Signals are normally distributed on the trunk cable at a level from +30 to +50 dBmV per channel. The signal level delivered to the home is usually from 0 to +15 dBmV per channel. In order to maintain the signal level, trunk amplifiers and feeder amplifiers, also known as line extenders, are spaced along the lines about every 1/2 mi. This corresponds to 20–30 dB loss at the highest design frequency. A trunk line may have 40 or more amplifiers in cascade. Since the attenuation of the cables increases with increasing frequency, the amplifiers are equipped with equalizers to provide more gain to the higher frequencies. The ratio of cable loss in decibels between two frequencies is approximately equal to the square root of the ratio of the two frequencies. A cable with 7-dB loss at 55 MHz will have about 20-dB loss at 450 MHz. Some amplifiers include **automatic gain control** (**AGC**) to maintain the desired output level. The amplifier cases can also include diplexers and small upstream amplifiers to separate and amplify the typically 5–30 MHz return path signals. Amplifiers are usually powered by 60 or 90 V AC signals, which are carried along the same coaxial lines.

Many CATV systems now use a *hybrid fiber coax* (HFC) architecture. To minimize the number of trunk amplifiers through which the signal must pass, the signal is carried via fiber optic cables from the head end to multiple remote locations known as a nodes. The head-end optical transmitter normally uses a distributed feedback (DFB) laser diode generating a few milliwatts of optical power at 1310 nm. The entire RF spectrum is used to intensity modulate the optical output of the laser. Since the fiber loss is only about 0.4 dB/km, the node can be located a relatively large distance from the head end, eliminating the need to cascade many amplifiers along a trunk. At the node, a photodetector receiver converts the optical signal back to RF for distribution on the coax. Return path signals are sent from the node back toward the head end through a separate fiber using less expensive Fabry–Perot laser sources.

The HFC architecture improves on the classic tree and branch in several areas. It provides many fewer active devices between the head end and home, which greatly improves system reliability. This reduction in the number of active devices also reduces maintenance costs, problem location, and recovery time.

The fiber cables are not susceptible to noise ingress, which impairs TV pictures, nor do they radiate RF power, which can interfere with nearby radio communications.

The star configuration of HFC fiber plant design provides individual fiber runs to each neighborhood. This means the plant supports dedicated bandwidth from the head end to each node of as few as a few hundred homes. Currently, HFC plants carry the same broadcast TV signals to every node. As interactive and on-demand services are rolled out, the HFC plant will carry the unique signals demanded by the subscribers in each neighborhood on their node alone. This effectively multiplies bandwidth by as many times as there are nodes in a system. Telephony is a primary example of an interactive, on-demand service. A tree and branch cable plant could provide only a limited number of telephone channels before running out of shared bandwidth, particularly upstream. The HFC star architecture can provide dedicated bandwidth to and from each node with ample resources for full service.

Frequency Plans

The 5–30 MHz range, called the subband, is commonly used for return path or upstream (from subscriber to head end) transmissions. This is used for the relaying of remote broadcasts to the head end and for the return of billing and other information from the subscribers' set-top. It can also be used for the upstream portion of telephony and interactive TV applications.

The range of frequencies used by CATV operators for downstream (sending video from the head end to the subscribers home) ranges from 50–200 MHz for older 23-channel systems to 50–550 MHz for 77-channel system. Newer systems are designed to pass up to 750 MHz or even 1000 MHz. Each analog **National Television Systems Committee (NTSC)** TV channel occupies 6 MHz. Using digital compression, several video programs can be transmitted in each 6-MHz channel.

The bands for downstream transmission are:

- Low band: 54–88 MHz (off the air VHF channels 2–6)
- Mid band: 88–174 MHz
- High band: 174–216 MHz (off the air VHF channels 7–13)
- Super band: 216–300 MHz
- Hyper band: above 300 MHz

There are multiple plans for assigning channels within the available frequencies. The Electronics Industry Association (EIA) plan assigns channels to agree with the VHF off the air channel numbers where possible. Additional channels are spaced 6 MHz from these. The harmonically related carriers (HRC) plan places video carriers near multiples of 6 MHz. This helps reduces the visibility of interference from harmonic and intermodulation distortion created in the distribution equipment because the distortion products fall directly on and zero beat with other video carriers. The incrementally related carriers (IRC) plan is similar to the EIA plan except for channels 5 and 6. This plan reduces the visibility of intermodulation distortion products but not harmonic distortion products.

Since many of the frequencies used for CATV channels are assigned to various over-the-air services, it is important for the cable system to maintain shielding integrity, both to keep outside signals from interfering with the CATV system (ingress) and to keep CATV signals from radiating and causing interference to local radio traffic, such as the aircraft communication bands.

Cable TV Frequency Plans

EIA Standard Plan: The standard plan uses xxx.250 MHz for analog video carriers for most channels. Digital center frequencies are 1.75 MHz above the analog video carrier so they will typically fall on xxx.000 MHz. For example, Channel 3 analog video carrier = 61.250 MHz. Channel 3 digital center frequency = 63.000 MHz. This corresponds to the NTSC over the air frequency plan for analog carriers. Analog channels that fall in the aircraft bands are given a +12.5 kHz offset up to xxx.2625 to reduce the probability of interference to aircraft. These aeronautical offsets are not needed for digital carriers because

there is no concentration of energy at the carrier frequency. None of the standard plan carriers are phase locked, therefore all intermodulation and harmonic distortion products can produce visible beats. The carriers are spaced nominally at 6 MHz. Due to the small frequency tolerance errors, the composite second order (CSO) intermodulation products and the composite triple beat (CTB) third order distortion products typically will create visible interference beats.

IRC: Incrementally Related Carrier: Nearly all IRC channels are xxx.2625 and all the carriers are phase locked to a common reference. This causes the CTB products to be nearly invisible because they fall directly on and zero beat the picture carrier. The CTB products are zero beats and invisible in the IRC system, but the CSO products fall 1.25 MHz above and below the analog video carriers and can produce visible interference.

HRC: Harmonically Related Carrier: HRC carriers are all phase locked and are multiples of 6.0003 MHz. This relationship causes both CTB and CSO products to be zero beat and nearly invisible. The 300 Hz or 50 ppm offset provides for protection of aviation services. HRC analog carriers are about 1.25 MHz below the standard analog carriers except channels 5 and 6 which are 0.75 MHz above.

	Second Order CSO Products	Third Order CTB Products
Standard plan	+/-750 kHz or +/-1.25 MHz	+/-15 kHz
IRC Plan	+/-750 kHz or +/-1.25 MHz	zero beat
HRC Plan	zero beat	zero beat

Control Data Channel

In some CATV systems, another signal, known as the out-of-band signal, is also sent from the head end to the subscriber. This signal is used to control addressable set-top **converters**. The converter contains a receiver that is always tuned to this signal regardless of which in-band channel is being viewed or power on/off. This signal is usually located in the FM radio band (88–108 MHz). The dedicated frequency and continuous reception by set-tops provides a reliable path for access control and user features such as program guides and downloaded graphical user interfaces.

98.3 Source Origination and Head End

The signals and programs that are to be delivered to CATV subscribers are all collected at a central location known as the head end. There can be many sources of signals including VHF and UHF off-the-air broadcast signals, analog and digital satellite signals, AM and FM microwave signals, prerecorded video tapes, character generators, digital video servers, and live locally originated programs.

Off-the-Air Broadcast Signals

The reception of distant off-the-air broadcast signals was the original need that created the CATV industry. Today, as populated areas become more spread out, many people subscribe to CATV to improve the quality of their reception of broadcast signals.

At the head end, off-the-air signals are received using high-gain Yagi antennas mounted on a hill or tower to provide the maximum signal. The directional characteristics of the Yagi antenna also help to discriminate against reflections which cause ghosting. Advanced digital techniques are also available to cancel ghosting. A known reference signal is transmitted in the invisible vertical blanking interval section of the TV signal. The receiver compares the received version of the signal to the known reference signal and calculates a ghost profile. This profile is then subtracted from the entire video signal to cancel the ghost effect. Using these techniques, good off-the-air pictures can be received over 60–100 mi away.

Direct Fiber Links

The penetration of cable to the majority of TV viewers in the U.S. means more people watch network TV on cable than off-the-air. This has made direct delivery of broadcast programming to the cable head end an attractive way to maximize the quality of these channels. Fiber is run from either the local broadcast studio or transmitter to the cable head end. There it is directly modulated onto a cable channel for distribution to subscribers.

Analog and Digital Satellite Signals

As the CATV industry grew, the concept of superstations and premium movie channels developed. These are distributed nationwide via geostationary communications satellites located 22,500 mi above the equator. Each programmer transmits their signal up to a satellite via an uplink, usually in the 5.9–6.4 GHz band. The satellite then transmits the signal back to Earth in the 3.7–4.2 GHz range or C band. Most analog CATV programming is concentrated on a few satellites known as cable birds. The transponders on these satellites generate 16–30 W of RE power, which usually requires a 10–30 ft diameter dish antenna at the cable head end. Each satellite contains 24 transponders and can carry 24 analog programs. Although, typically, a separate parabolic dish antenna is needed to receive each satellite, a specially shaped antenna can be used with multiple feeds to simultaneously receive multiple satellites.

Analog satellite signals are received in integrated receiver decoders (IRDs) at the cable head end. These devices tune the desired transponder frequency, demodulate the FM video signal, descramble the video signal, and decrypt and convert the digital audio back to analog form. They are authorized by the program providers **access control** and **security** system. The output analog video and audio go to a cable scrambler, modulator, and channel converter for transmission on the cable plant.

The enormous expense of satellite transponders drove the development of digital video compression for satellite transmission. Digital transmission of compressed video signals provides more efficient use of satellite transponders. Compression systems currently offer 4–10 video programs over a single transponder, greatly reducing the programmer's satellite link costs.

Digital compressed satellite signals are received in a digital IRD, which connects to the head-end systems in much the same way as an analog IRD. The digital IRD performs many more functions than an analog IRD but delivers similar outputs of analog video and audio signals. Differences include higher quality video signals due to digital transmission with no quality degradation of weak signals, high security due to digital encryption, and the availability of many more programs due to the efficiency of transmission.

AM and FM Microwave Signals

In some locations, it is impossible to receive off-the-air broadcast signals with the desired quality. In these cases, a dedicated point-to-point microwave link can be installed. The desired off-the-air signal is received at a location close to the transmitter or, sometimes, the microwave signal can be sent out from the TV station directly. The signal is beamed via very directional antennas directly to the head end. Both AM and FM modulation are used for microwave links.

Commercial Insertion

As special programming intended solely for CATV became available, the concept of local advertisement insertion developed. Local business can purchase advertising air time from the CATV operator. These commercials are inserted during defined commercial breaks in the programming. In the past, advertisements were pre-recorded on video tape and automatically controlled to run at the correct time. Locally originated programs such as locally produced news, school board or local government meetings, and public access programming can also be recorded on video tape for transmission via CATV. New commercial insertion equipment uses compressed digital video sources. This provides greatly enhanced video quality, reliability, insertion tracking, scheduling, and methods for delivery of the program material to

the insertion equipment. Commercial insertion into digital video channels on cable will require a new device called an add–drop multiplexer. This device will replace the compressed video bits of a satellite delivered advertisement with a local ad. This is all accomplished in the digital domain without any need to decompress the content.

Character Generators

Character generators are used to create bulletin board style information channels. Programming, advertising, local government, and weather information can be displayed. Usually, several screens worth of data are stored in the character generator and these are sequentially displayed. Computer graphics can also be included to enhance visual appeal. Expanded forms include characters generated over video stills or live video.

Live Local Programming

Many CATV operators carry local school or government meetings live. Often, a subband frequency (5–30 MHz) is used to carry the signal from the town hall, for example, via the cable plant return path back to the head end. At the head end, the signal is converted to a normal forward path channel and sent to the subscribers via the distribution system.

Digital Satellite to Digital Cable Transcoding

Just as the integrated device called the IRD receives an analog or digital satellite service for use on cable, a device called an integrated receiver transcoder (IRT) receives a compressed digital video multiplex for direct transmission to digital video subscribers without decompression at the head end. The IRT performs almost all of the functions required to pass a compressed multiplex of signals from a digital satellite channel to a digital cable channel.

The IRT tunes a digital satellite signal, demodulates it, performs satellite link error correction, decrypts the multiple video and audio services in the multiplex, replaces the satellite access control information with new cable system control data, re-encrypts the services with cable security, inserts cable error control bits, and modulates the data onto a 64 quadrature amplitude modulation (QAM) carrier. Only a channel converter is required to prepare the signal for CATV transmission. This high density of signal processing means it takes much less rack space in the cable head end to add a channel of digital programming from a satellite than an analog channel. At the same time, the digital channel provides many simultaneous video programs to the subscriber.

Digital Video Servers

An expanded version of the video equipment used for digital commercial insertion can provide full length programs such as **pay-per-view** movies and movies with multiple start times. Video servers can be used to create programming for analog cable plants in place of video tape, but their most efficient application is in digital cable plants. When digital set-tops are deployed, a digital video server can output precompressed services to them for preprogrammed or video on-demand services. The output is encrypted, multiplexed, and modulated in equipment that can also process streams from real-time video encoders.

98.4 Transmission Channel

Channel Characteristics

In order to deliver a high-quality signal, the CATV distribution plant must be designed to add a minimum of noise, distortion, and spurious signals. Various defects in the channel can cause visible impairments to the picture. The in-channel gain and group delay must be constant vs. frequency to avoid ringing or loss of vertical detail. Excessive thermal noise causes snow or graininess in the picture. Spurious signals

TABLE 98.1 Typical Maximum Impairments to Analog Video Signals

Gain flatness	±1dB
Group delay variation	±170 ns
Hum modulation	±1% max
Signal to noise ratio	48 dB
Composite second order (CSO)	-50 dB
Composite triple beat (CTB)	-50 dB

or intermodulation can cause herringbones or diagonal lines that move randomly through the picture. Table 98.1 is a list of some of the relevant parameters and values required for an excellent picture.

Degradation from each part of the system is additive. The consumer's video signal has passed through a great number of active and passive components, each of which degrades the signal. Within the economic constraints on real products, each system component must be designed to minimize its contribution to total impairment.

Digital transmission on cable will, in general, be much more robust, reliable, and easy to monitor than analog transmission. There are a few impairments that will affect digital transmission more than analog. The best example is burst noise. This shows up as a loss of signal for a fraction of one TV scan line in analog TV. This is fleeting and rarely objectionable. With digital compression, all of the redundant information in the video signal has been removed. Any lost data will likely be noticeable. For this reason, the QAM transmission for the cable channels has been designed to be highly robust even for bursts of noise over a full scan line long. Within the design limits, digital video will contain no degradation from passing through the distribution system.

Modulation and Error Correction (Analog and Digital)

Analog Modulation

Unscrambled analog video signals are carried in exactly the same modulation, vestigial sideband–amplitude modulation (AM-VSB) format as broadcast TV. This facilitates the use of TVs as the CATV reception device for simple services. The simplest forms of analog scrambling modify the modulation format by reducing signal power during the synchronizing periods of the video signal. Sync suppression has the added benefit of reducing peak load on the distribution system as the modulated TV signal is designed for peak power during these sync periods. More complex scrambling systems also perform various forms of video inversion. Both of these modified forms of modulation are suitable for CATV applications because of the robust and well-controlled reception conditions. High sync power, for example, is not required because there is no need to receive signals in the microvolt level from distant broadcast stations.

Digital Modulation

Just as analog satellite signals use frequency modulation (FM) digital satellite transmissions use quantenary phase-shift keying (QPSK). Both are wide bandwidth low carrier-to-noise modulation formats. Digital CATV systems use bandwidth-efficient high carrier-to-noise modulation called 64QAM and 256QAM. QAM is similar to QPSK in that two quadrature carriers are used but it adds amplitude modulation on each of the quadrature carriers. For example, 64QAM has eight states for each carrier instead of QPSK's two states. This results in 64 combinations and, therefore, each modulation symbol carries 6 bits of data. For 256QAM, 8 bits are delivered for each symbol. These digital modulation formats deliver about 30 and 43 million bits per second (Mb/s) in each 6-MHz CATV channel.

Digital Forward Error Correction (FEC)

A great advantage of digital transmission is its ability to build on the well-developed field of error correction. Sufficient redundancy is designed into the transmitted signal to allow correction of virtually all errors caused by transmission impairments. The International Telecommunications Union (ITU)

standard J.83 annex B is the standard for digital cable transmission in North America. It defines a concatenated coding system of Reed–Solomon (RS) block error correction with convolutional trellis coded modulation for highly robust performance in the cable channel. Interleaving is provided to correct bursts of errors of up to 95 μs.

After error correction, the 64 and 256QAM modulation modes deliver 27 and 39 Mb/s of near error-free information, respectively, to the receiver.

98.5 Consumer Premises Equipment

The First Converters

The earliest cable systems had trouble carrying VHF channels 7–13 at their broadcast frequencies of 174–216 MHz. Because signal loss in coaxial cable increases at higher frequencies, many more amplifiers were required, causing much higher distortion, than delivering channels 2–6 at 54–88 MHz. To solve this problem, the early systems carried channels 7–13 in the frequency band below channel 2 (54 MHz). This frequency range is called the subband and is numbered T7–T13. The earliest converters block converted these subband frequencies back to the broadcast channels 7–13 near the subscribers' location for reception on standard TVs.

The First Set-Top Converters

Early CATV set-top converters solved two technical problems: cochannel interference and the need for more channels. Without a set-top converter, each channel was received by the TV on its cable channel frequency. Because TVs do not have effective RF shielding, many subscribers received one or more broadcast channels strongly enough at the TVs' input terminals to cause direct pick-up interference. This causes a ghost to be displayed on top of the signal received from the CATV system. Set-top converters allowed the CATV channel to be received in a well-shielded tuner and converted to a TV channel not broadcast in the local area. If the local stations transmitted on channels 2, 4, and 6, the output of the set-top would use channel 3 to the TV set. This eliminated the ghost effect and is still important today.

The second, and possibly most important, use of the set-top was to extend channel capacity. The use of the UHF broadcast channel frequencies was not viable for early CATV systems due to their much higher transmission loss in coaxial cable. To go beyond the 12-VHF channels required the use of non-broadcast channels. This required the use of a new tuner that could tune these cable-only channels. This also provided a rudimentary form of access control. Only those subscribers who had paid for the extended service were given a set-top that could tune the added channels.

Set-Top Designs

Modern CATV converters serve many additional functions. One of these is to provide tuners with good performance in the cable environment. To this day, almost all TV tuners are single conversion with one mixer and one oscillator. TV stations are allocated licenses on nonadjacent channels in each broadcast market. This means, for example, that if channel 3 is active in a certain market, channels 2 and 4 will not be used. In addition, TV tuners are required to receive very low-signal levels in the low-microvolt range to receive distant stations. For this reason, TV tuners must be high-gain, low-noise devices. When presented with strong signals on every channel as occurs in CATV transmission, standard TV tuners overload and generate distortion products, which range from imperceptible to very annoying in the TV picture. The signals in a properly maintained cable system are maintained at above 1 mV.

By around 1978, **descrambling** converters were given unique addresses to enable their control by computers located at the CATV system offices. At the same time, 400-MHz distribution plants were introduced; this boosted channel capacity to 52 channels and added the capability for two-way communications over the cable. At about the same time, wireless remote control of converter channel tuning

was becoming available. By 1980, infrared (IR) control of converters was available to CATV operators. Microprocessors were added to converters that converted the remote control signals to tuning voltages and could interpret the addressable data sent from the cable companies' computers. This was really the start of the modern descrambling converter as we know it today.

Addressable Converters

The addressable cable converter offers many advantages over the older preprogrammed converter. Initially, the most important advantage was the ability to change services at the subscriber location, without requiring a technician to visit the home and replace the equipment. A service representative at the cable system office could enter the subscriber service change into the billing computer and the change was made immediately. It had an added benefit of upgrading the subscribers' billing information at the same time. By about 1983, video demodulators were added to the converter. This allowed processing of baseband video and audio and added volume control as a converter feature. This was a major convenience for the many subscribers who did not have remote control TV sets. Baseband processing added the capability to descramble the higher security baseband scrambled video signals and provide audio privacy capability, which removes the possibly offensive audio signals of channels not wanted by the subscriber. This was also the time many cable plants and converters increased bandwidth to 550 MHz with 77-channel capacity.

Impulse Pay-per-View Converters

The introduction of addressability simplified the process of ordering a one-time **event** over the cable. If a special event, such as a championship boxing match or a live concert, was offered on cable, a subscriber could order the event by calling the cable company, have it transmitted to his home or place of business, and pay a one-time charge for the viewing. This system of pay-per-view was relatively successful but required planning by the subscriber well in advance of the event. It also required a large number of service representatives be made available to staff the telephones.

Between 1983 and 1985, a new technology, **impulse pay-per-view**, was introduced using store-and-forward technology. This allowed a cable company to offer special events or movies 7 days a week, 24 hours a day. Subscribers could **purchase** the event shortly before or after the event started, using only their handheld remote control. The number of subscriber purchases increased dramatically. The converter became a two-way device transmitting the subscriber purchase information back to the cable system via a telephone modem or RF transmitter built into the converter. A record was made of every purchase made by the subscriber. Periodically, the system polled all of its converters and uploaded purchases to the master billing computer at the cable office.

Over the years, many user-friendly features were added to the converter. On-screen displays and downloadable electronic program guides are becoming common. Digital audio reception capability is being added to provide high-quality radio-type services as well as digital transmission of the audio signals for television channels. These added functions continue to make converters easier to use and improve the operators' ability to provide enhanced services.

Digital Cable Converters

Digital cable converters add the ability to receive digital video transmissions on CATV plants. These sophisticated devices start with the most advanced form of analog converter and add QAM reception, MPEG-2 video decompression, Advanced Television System Committee (ATSC) or MPEG audio decompression, and secure digital decryption capabilities. One way to look at them is as a combination of an advanced analog converter and a digital DBS receiver.

Digital cable converters will finally remove the limits on channel capacity and picture quality to the home that have defined cable since its inception. The high security enabled by digital encryption technology will allow new services with very high value, while compression dramatically increases the channel capacity to carry them.

Cable Modems

Cable modems deliver high-bandwidth data service to consumers over the CATV plant. Typical rates for data delivery run to 27 Mb/s for downstream delivery to the consumer. Upstream rates for the various modem products vary widely from symmetrical service, where the upstream transmission matches the downstream rate, to products using telephone return channels limited to 28.8 kb/s. With as much as 1000 times the maximum rate available on dial-up telephone lines, cable modems seek to provide enhanced Internet access and linked digital video and data services to personal computers.

Interdiction

Interdiction is an alternative to **scrambling** channels as a means to selectively deliver CATV services to paying subscribers. It has the advantages of using no set-top converter and delivering all paid services to any cable outlet in the home. This allows the unrestricted and direct connection of multiple VCRs and TVs and the ability to simultaneously watch and record two pay services. Modern interdiction systems are addressable and offer the same services as conventional cable systems. The common interdiction method injects a **jamming** oscillator onto the interdicted channel, making it unwatchable. Interdiction disadvantages include the need for an interdiction module for every subscriber, not just those buying enhanced services as with scrambling, and the piracy threat of delivering clear services to every home up to the point of interdiction.

Each subscriber in an interdiction system will have a complete access control and jamming module in an off-premises enclosure. Enclosures typically service between 1 and 16 subscribers. Each subscriber module has a unique address, which can be addressed by the cable system's master computer. Each module can be authorized for any number of channels or disconnect service entirely.

The heart of the interdiction system is the jamming oscillator. Each oscillator consists of agile **voltage controlled oscillators (VCOs)** which sweep a band of channel frequencies. The number of VCOs is determined by the number of channel bands or the total bandwidth of the cable system to be processed. Typical VCO range is limited to approximately 50 MHz or 8 channels in an NTSC system. Lengthening the VCO range will reduce the on-channel time and will start degrading the effect of jamming on each channel. The VCO will inject a carrier between -250 and +750 kHz of the video carrier of the interdicted channel and skip those channels which the subscriber is authorized to view. Each subscriber module will consist of the jamming oscillator, an oscillator injector, **phase lock loop (PLL)** circuitry to control the oscillator, AGC control of the incoming channel signals, and a microprocessor.

98.6 Access Control and Security

Access control is applied to cable television systems to provide the means to offer services on a individually controlled basis. It allows consumers to choose, and pay for, only those services they desire. Various techniques enable the commonly encountered cable services such as **subscriptions**, events, pay-per-view movies, and **near video on-demand** movies. Access control has developed from simple physical schemes to sophisticated computer-controlled techniques.

Access control techniques themselves do not prevent unauthorized individuals from stealing services. That is the job of the security measures. When piracy becomes widespread, the delivery of enhanced services can become either too expensive for the paying consumer and uneconomical for the operator. The inherent trade-off of a security system is in how much to invest to hold off the pirates who would defeat it. Security measures protect revenues and make it worth deploying new services but increase the investment in set-tops and operating procedures.

Access Control

Access control is a combination of **access requirements** and **authorization rights**, thought of as lock and key, where access requirements are the lock and authorization rights are the key. Access control regulates access to system services, each identified by one or more **service codes**. These services, typically,

are those for which cable customers must pay. Typically, the installation of a set-top in the home enables the system's access control capability.

Service codes are tokens used by the system to identify a particular service. The maximum number of service code values determines the maximum number of services supported by the system at any particular time. For example, if a typical system provides 256 service codes, then it supports up to 256 unique services. Since a single service might use more than one service code, however, the actual number of supported services is often less than 256. Access control systems for digital video offer many thousands of service codes to support the much larger number of digital channels and services.

Access Requirements

In order to access a service, that service must have a service code assigned and be authorized within the set-top box. Fundamentally, these are the only access requirements.

Authorization Rights

The system grants authorization rights to a service by turning on the service code assigned to that service in the set-top. Alternately, the system rescinds authorization rights to a service by turning off that service code.

Physical Access Control Techniques

The physical presence (or absence) of an electronic **trap** located between the cable itself and the converter is an early form of access control used to authorize subscription-type services. Typically, traps are on the utility pole where the residential drop originates. In that case, a set-top, converter may not be required. The cable installer provides access control by inserting or removing the trap device.

Computer-Controlled Access Control Techniques

In modern systems, a microcomputer in the set-top enforces access control by communicating electronically with head-end computers. The set-top only sends services to the television monitor if the service code is authorized. Depending on the type of service and on the capabilities of the set-top, the control system can either directly or indirectly authorize services.

Head-End Authorization. Direct authorizations originate from the system head end. In this case, the head end sends a message to the set-top to authorize or to deauthorize a particular service code. The control computer **tags** the service by attaching this same service code to the signal. When the viewer ultimately selects the tagged service, the set-top reads the attached service code and checks to see if it is authorized. To make a subscription request, for example, the viewer calls an advertised telephone number and speaks to a service representative who makes the authorization. To cancel the subscription (or event), the viewer requests the service representative to deauthorize the service. For subscriptions, the tag value never changes and an authorized viewer always has access. For events, the head end changes the tag value at the end of the event, forcing the set-top boxes to deauthorize and switch to a different screen. After the tag changes, head-end computers transmit messages to deauthorize the service code in all set-top boxes to allow the service code to be used again at a later time.

Self-Authorization. Self-authorizations schemes also exist for viewers to purchase services on impulse without any communication with the head-end. The set-top self-authorizes for a period of time and then either self-deauthorizes or waits for the head-end computer to deauthorize it. This technology requires the set-top to have some means to inform head-end computers of such impulse purchases by the viewer, such as a telephone modem or RF return transmitter. With this capability, the set-top self-authorizes services at any time, knowing that head-end computers will later read them and bill viewers appropriately.

Some form of deauthorization always follows self-authorization so the set-top terminates the service after it is completed, for example, when the movie ends. In a manner similar to authorization there are two forms of deauthorization: deauthorization by the head-end or self-deauthorization via a control timer.

The head-end deauthorization scheme enables services such as impulse pay-per-view. The viewer, on impulse, purchases a movie (or other offering) merely by requesting it via the handheld remote control unit or the set-top's front panel. The set-top tunes the service, reads the service code from the tag, and authorizes it. The box then delivers the service and stores the purchase in memory for later transfer to head-end computers.

At the end of the movie, the head-end changes the tag to a currently deauthorized value, forcing the set-top to stop delivering the service. As in the pay-per-view case previously discussed, head-end computers then deauthorize the service code in all set-tops.

The last case, where the set-top both authorizes and deauthorizes itself, enables a service type known as near video on-demand. In this service, the viewer purchases a movie that is actually running at staggered intervals on multiple cable channels. For example, a 2½ hour movie might be running on five cable channels spaced out at half-hour intervals. At the time of the purchase request, the set-top waits for the nearest movie start time (no more than one-half hour away) and automatically tunes the correct channel.

After tuning the movie's channel, the set-top reads the service code carried in the tag, self-authorizes, and stores purchase information for later transfer to the head end. The authorization time carried in the tag determines how long the authorization is valid. After the authorization time expires, the set-top automatically terminates that service by deauthorizing its service code.

While the timer is running and since all channels carrying the movie have the same tag, the viewer can, in effect, fast-forward or reverse the movie in 1/2 hour intervals by tuning to these different channels until the self-deauthorization occurs.

Security Systems

A security system can be defined as a series of techniques applied to the distribution of cable services that prohibit users from bypassing the system's access control facilities. Optimum security in cable systems is accomplished through two primary venues, legal actions (both litigation and legislation) and secure techniques (both physical and electronic schemes).

Legal Actions

The threat of legal action against the perpetrators is an important part of anti-piracy efforts. The federal law making cable theft a felony is a principle piece of legislation enabling law enforcement agencies to intervene in piracy cases. Many state statutes and some municipal codes also prohibit the sale, and a few even prohibit possession, of pirate devices. Cable equipment suppliers and cable operators support these laws by assisting law enforcement agencies at all levels of government wherever possible.

Secure Techniques

Secure techniques include both physical and electronic means. Physical schemes are used primarily within the head-end and in the distribution paths for set-tops. They are primarily directed at keeping converters away from pirates who may modify them to provide service without payment. Electronic measures are based on coordination between installed set-top boxes and head-end equipment. They are performed in two primary areas:

- Security of the authorizations stored in specific set-top boxes
- Security of the signals transmitted downstream on the cable itself

The cable operator enters the set-top box's serial number into the access control computer. This number relates to a specific physical unit which is now known to the system. During a configuration process, the control computer gives the set-top a unique address along with a system specifier. Using this unique address, the system then authorizes access to services ordered by the viewer.

To scramble an analog television signal, cable operators modify one or more of its key characteristics. Typical scrambled pictures appear faded or torn with the left-hand side of the picture on the right-hand side of the screen and vice versa. Techniques giving rise to these effects suppress the picture synchronization

signals or invert the video information. Less common techniques change the order of the video lines in the picture, individually shift them right or left, or rotate them around a dynamically changing split point. Scrambling techniques may also modify the audio portion of the signal.

Electronic security for digital video signals is implemented as digital encryption of the compressed bit stream. The advanced art of encryption and key delivery can provide very high signal security on all of the digital services when well implemented.

Defining Terms

Access control: A scheme utilized in cable television systems to determine access rights as a function of predefined access requirements.

Access requirements: A set of conditions that must be satisfied prior to granting access to certain services carried on cable television systems. The system first assigns a service code to the service and then grants authorization rights for the service to the set-top box.

Authorization rights: A generic term describing the system's response to various authorization states of a service. If the state is ON, then the system delivers the service. If the state is OFF, then the system blocks the service.

Authorization state: Each service exists within the set-top box in one of two authorization states: on (or authorized) and off (or deauthorized).

Automatic gain control (AGC): An electronic circuit that detects the level of a signal and either increases or decreases its amplitude before passing the signal to the next stage.

Converter: A generic term referring to any device interposed between a cable system and a service-rendering device (for example, a television monitor) which changes the cable signals into a form usable by the rendering device.

Decibel-millivolts (dBmV): A measure of power levels of a signal common in the cable TV field. It can be stated as the level at any point in a system expressed in decibels relative to 1 mV.

Descrambler: An element of some cable converters that is capable of descrambling secure cable signals.

Dither: The repeated sweep in frequency between two extremes.

Event: A type of service available on modern cable systems characterized by fixed start and stop times which typically occurs once. Viewers purchase this service by calling a service agent who authorizes the service for that viewer and who records the purchase for later billing purposes.

Frequency shift keyed (FSK): A type of digital modulation in which a carrier is varied in frequency.

Head end: The point of origination of the cable TV signals before they are transmitted out on the system. Typically, this equipment is located near the business office of the cable company.

Impulse pay-per-view: A type of service available on modern cable systems that enables viewers to purchase the service using their handheld remote units or the set-top front panel. The set-top box typically stores the purchase in memory and transfers it to the control computer at a later time.

Interdiction: The process of denying a particular service to a subscriber.

Jamming: Adding interference to a signal, making it unusable.

National Television Systems Committee (NTSC): The U.S. broadcast standard for TV video transmission also used in Japan and many other countries.

Near video on-demand: A type of service available on modern cable systems that enables viewers to purchase a service at nearly any time and that is always available on the system within a short interval. The set-top box stores the purchase in memory and transfers it to the control computer at a later time.

Pay-per-view: A type of service in which viewers pay to view a specific event, e.g., a movie.

Phase lock loop (PLL): A circuit used to stabilize oscillator circuits to some predetermined frequency by monitoring the actual oscillator frequency and sending a correction voltage to the oscillator tuning circuits.

Purchase: A viewer makes a purchase in one of two ways: either by calling a service agent or on impulse by using the set-top remote unit or front panel controls.

Scrambling: Modifying standard video signals so they can only be received by authorized devices, typically set-top converters.

Security: A generic term referring to one of many techniques used on cable systems to prohibit viewers from bypassing the normal access control feature. Example techniques include legislation, litigation, and various electronic guises.

Service code: A symbol (or token) used by a cable system to identify and control access to a particular service. Service codes are authorized and deauthorized within the set-top box by the control computer.

Set-top: A specific type of converter that follows a cable system's access control constraints. Often, set-tops include a descrambler and/or decryptor for changing scrambled or encrypted cable signals into a form usable by the subscriber.

Subscription: A type of service that, after being authorized, is available to the viewer 24 h per day throughout the billing cycle. Typically, the viewer places a request with a service agent who authorizes the service for the viewer and who records the purchase for later billing purposes. Early systems implemented subscriptions using traps.

Synthesizer: An oscillator capable of producing multiple frequencies from a single clock source.

Tag: Multiple symbols added to a service signal that are detected and used by a set-top box to deliver services. Signals are tagged with, at a minimum, the service code assigned to the service and are used in the set-top box to determine if delivery to the viewer is authorized.

Trap: A device used to control access to subscription services by blocking them at the point of delivery into the home, usually on the telephone pole where the cable drop originates.

Voltage controlled oscillator (VCO): An oscillator whose tank circuit contains a varacter diode, which changes capacity as its bias changes; the most common form of oscillator control in TV receivers and other communication equipment.

99

Video Servers

A. L. Narasimha Reddy
Texas A&M University

Roger Haskin
IBM Almaden Research Center

99.1 Introduction

The recent dramatic advances in communications technology, compression techniques, and computer systems have led to a demand to support multimedia data (digitized video and audio) across data networks. In part, this demand comes from the personal computer marketplace, where the use of video and audio in PC applications and games is now pervasive, and, in part, the demand comes from the entertainment marketplace, where there is a move to use digital video to supplant the traditional analog methods of content creation and broadcast. This intersection of technologies and markets has led to the development of video servers. Several telephone companies and cable companies are planning to install video servers that would serve video streams to customers over telephone lines or television cable. These projects aim to store movies in a compressed digital format and route the compressed video to the customer (through cable, telephone lines, or local area networks) where it can be uncompressed and displayed. Several Web sites already provide low quality video content in compressed digital format that can be viewed by Web browsers.

These projects aim to compete with the local video rental stores with better service: offering the ability to watch any movie at any time, avoiding the situation of all of the copies of the desired movie already being rented out, offering a wider selection of movies, offering other services such as tele-shopping, and delivery of other video content such as interactive video games.

A number of business models have been proposed for such video servers. The business models include complete interactive video games and near-video-on-demand in hotels, among others. The systems' requirements vary considerably depending on the targeted markets. Near-video-on-demand envisions supplying popular/recent movies at regular intervals, say every 5 min, to requesting customers over some delivery medium. This form of service is expected to be popular in hotels, replacing the existing situation of making movies available at only few times during the day. Dozens of movies are expected to be available through such a server. **Video-on-demand** (**VOD**) envisions replacing conventional pay-per-view programming (as available on cable TV), which allows customers to view a small number of movies at widely spaced intervals, with a system that allows customers to choose from a large number of movies and play them within a few minutes or seconds of when the customers request them. Some of these servers plan to provide the ability to supply VCR-like capabilities to pause/resume and fast forward/reverse a movie stream. Such servers are said to provide **interactive TV** (**ITV**) service. Interactive video servers could have

wide applications. Some of the planned applications of interactive video servers include training, education, travel planning, etc. These interactive video servers play a predetermined set of video sequences based on the user's input. Interactive video games require a higher level of interaction with the video server since the playout of the video sequence varies more rapidly based on the user's input. **Video teleconferencing** requires two-way interactive response between the users. **Near video-on-demand**, which delivers video at discrete times, is already avaiable in hotels in large cities. Which of these business models will predominate remains to be seen.

Digital video broadcast (DVB) systems replace analog video channels with compressed digital video. DVB allows more channels to occupy the same frequency spectrum and allows expensive and failure-prone electromechanical devices (e.g., video tape players) to be replaced with more reliable digital storage. In a DVB system, the video storage and transmission are identical to those in a VOD or ITV system. In DVB, however, instead of the video being under the control of a user, a piece of automation equipment controls the sequencing of video playback. Here, sequencing means creating a smooth, continuous video stream out of multiple independent pieces of stored video, for example, a series of commercials followed by a segment of a TV program, a station break, more commercials, etc. A DVB system is actually, in this respect, more comparable to an ITV system than to VOD since the system is required to respond in timely fashion to commands from the automation equipment. Video servers are also being utilized for digital video productions. This includes the management of uncompressed and lightly compressed (HDTV) digital data.

A number of pilot projects have been under way to study the feasibility of these approaches. Several major companies in the computer systems business (IBM, SGI, HPDEC), computer software (Oracle, Microsoft, IBM), telephone carriers (AT&T, US West, Bell Atlantic, Ameritech), cable companies (TCI, Time Warner, Cox Cable), and equipment manufacturers (Scientific Atlanta, General Instrument) have been actively participating in these trials. The technical challenges in making these projects economic/business successes are daunting. In this chapter, we highlight some of the technical problems/approaches in building such video servers, with an emphasis on communication aspects of the problem.

Currently, with **motion pictures expert group-2** (**MPEG-2**) compression at 6 Mb/s, a movie of roughly 90 min duration takes about 4 gigabytes of storage space. Storing a movie at uncompressed HDTV quality takes 260MB/s of storage space. Storing such large data sets in memory is very expensive except for the most frequently requested movies. To limit the costs of storage, movies are typically stored on disks. It is possible to store infrequently accessed movies on tapes or other tertiary media. However, it has been shown that these media are not as cost-effective as magnetic disk drives even for infrequently accessed movies [Chervenak, 1994].

A video server storing about 1000 movies (a typical video rental store carries more) would have to spend about $100,000 just for storing the movies on disk at the current cost of $100/gigabyte for managed storage systems. If higher quality picture is desired and MPEG-2 compression is used at 6.0 Mb/s, then the cost of storage space goes up to $400,000. Other nonrecurring costs include costs for enhancing the delivery medium and providing a **set-top box** for controlling the video delivery. Even as the costs of storage keep falling, requirements for storing larger data sets through better quality standards emerge. This requirement of large amounts of investment implies that the service providers need to centralize the resources and provide service to a large number of customers to amortize costs. However, the ever-decreasing costs of magnetic storage (at the rate of 50% every 18 mo) and computing power (at the rate of 54% every year) may make these cost considerations less of a problem if the technical challenges can be overcome.

99.2 Data Server

A large video server may be organized as shown in Fig. 99.1. A number of nodes act as *storage nodes*. Storage nodes are responsible for storing video data either in memory, disk, tape, or some other medium and delivering the required input/output (I/O) bandwidth to this data. The system also has *network nodes*. These network nodes are responsible for requesting appropriate data blocks from storage nodes and routing them to the customers. Both these functions can reside on the same node, that is, a node

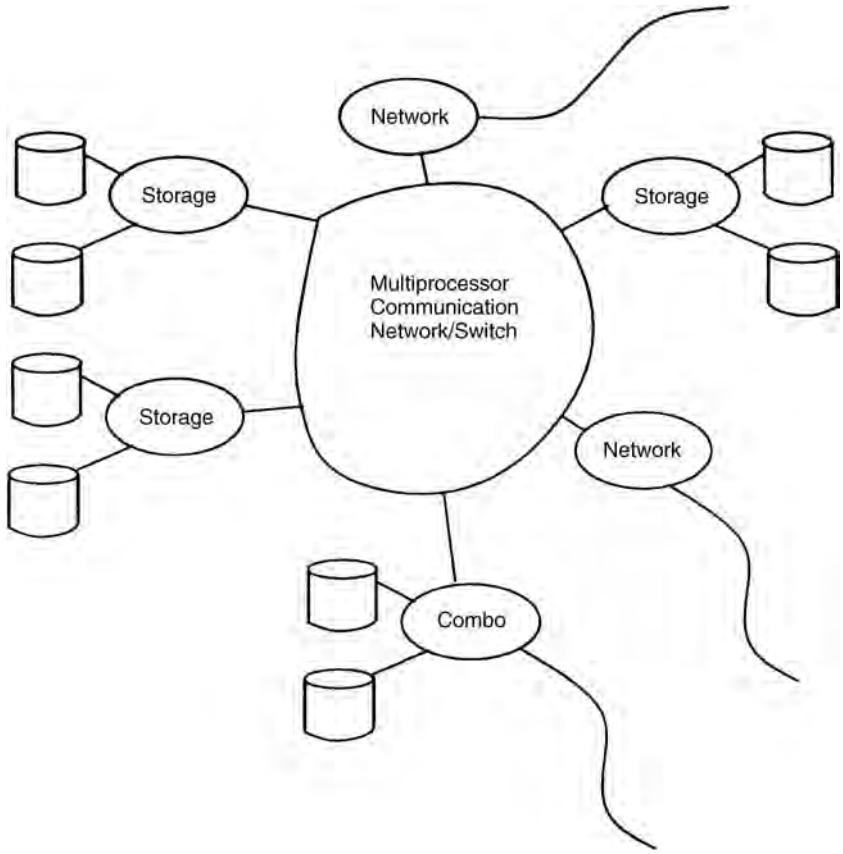


FIGURE 99.1 System model of a multiprocessor video server.

can be a storage node, a network node, or both at the same time. Each request stream would originate at one of the several network nodes in the system, and this network node would be responsible for obtaining the required data for this stream from the various storage nodes in the system. The design of network nodes will change based on the delivery medium. The logical separation (whether or not they are physically separated) of nodes into network nodes and storage nodes makes it easier to adapt the video server organization to different delivery mediums.

The delivered data stream is fed to a set-top box (STB) at the customer's site. The STB decompresses the digitally enoded data into a suitable input form for the TV monitor. The STB may provide additional features such as the ability to pause/resume and fast forward/reverse an incoming video stream. The design of the STB is a hotly discussed topic with proposals ranging from a dumb decoder to intelligent controller with most of the functions of a personal computer. The design of the STB depends on the functions provided, compression standard, bandwidth of the compressed stream, and input standard of the TV monitor, among other things. To enable future adaptation to evolving compression technology, some projects envision doing decompression completely in software by a processor in the set-top box.

To obtain high I/O bandwidth, data has to be **striped** (distributed) across a number of storage nodes. If a movie is completely stored on a single disk, the number of streams requesting that movie will be limited by the disk bandwidth. To enable serving a larger number of streams of a single movie, each movie has to be striped across a number of nodes. As we increase the number of nodes for striping, we increase the bandwidth for a single movie. If all of the movies are striped across all of the nodes, we also improve the load balance across the system since every node in the system has to participate in providing access to each movie. A number of related issues such as fault-tolerance and incremental growth of the

system are closely linked to striping characteristics, and we will discuss these issues later in the chapter. The unit of striping across the storage nodes is called a block. In earlier studies on disk scheduling [Reddy and Wyllie, 1993], 64–256 kilobytes is found to be a suitable disk block size for delivering high real-time bandwidth from the disk subsystem.

As a result of striping, a network node that is responsible for delivering a movie stream to the user may have to communicate with all of the storage nodes in the system during the playback of that movie. This results in a point-to-point communication from all of the storage nodes to the network node (possibly multiple times depending on the striping block size, the number of nodes in the system, and the length of the movie) during the playback of the movie. Each network node will be responsible for a number of movie streams. Hence, the resulting communication pattern is random point-to-point communication among the nodes of the system. It is possible to achieve some locality by striping the movies among a small set of nodes and restricting that the network nodes for a movie be among this smaller set of storage nodes.

The service for a video stream can be broken up into three components: (1) reading the requested block from the disk to a buffer in the storage node, (2) transferring this block from the storage node buffer to a buffer in the network node across the multiprocessor interconnection network, and (3) delivering the requested block to the user over the delivery medium. The critical issues in these three components of service are disk scheduling, interconnection network scheduling, and delivery guarantees over the delivery medium, respectively. We discuss each of these service components briefly here.

Disk Scheduling

Traditionally, disks have used seek optimization techniques such as **SCAN** or shortest seek time first (SSTF) for minimizing the arm movement in serving the requests. These techniques reduce the disk arm utilization by serving requests close to the disk arm. The request queue is ordered by the relative position of the requests on the disk surface to reduce the seek overheads. Even though these techniques utilize the disk arm efficiently, they may not be suitable for real-time environments since they do not have a notion of time or deadlines in making a scheduling decision.

In **real-time** systems, when requests have to be satisfied within deadlines, algorithms such as **earliest deadline first** (**EDF**) and least slack time first are used. The EDF algorithm is shown to be optimal [Liu and Layland, 1973] if the service times of the requests are known in advance. However, the disk service time for a request depends on the relative position of the request from the current position of the read–write head. Moreover, due to the overheads of seek time, strict real-time scheduling of the disk arm may result in excessive seek time cost and poor utilization of the disk.

New disk scheduling algorithms have been proposed recently that combine the real-time scheduling policies with the disk seek-optimizing techniques. Some of these algorithms include SCAN-EDF [Reddy and Wyllie, 1993] and grouped sweeping scheduling [Yu, Chen, and Kandlur, 1993]. These algorithms batch requests into rounds or groups based on request deadlines, serving requests within a round based on their disk track locations to optimize the seek time. These algorithms are shown to significantly improve real-time performance at smaller request blocks and improve nonreal-time performance at larger request blocks [Reddy and Wyllie, 1993]. Figure 99.2 shows the impact of request block size on various scheduling algorithms. Other factors to be considered include multiple data stream rates and variable bit rates.

Multiprocessor Communication Scheduling

The communication scheduling problem deals with the issue of scheduling the network resources of the multiprocessor computer system for minimizing the communication delays. If the nodes in the multiprocessor system are interconnected by a complete crossbar network, there is no communication scheduling problem since any pair of nodes in the system can communicate without a conflict in the network. Regular distribution of video data over the storage networks enables us to guarantee that if source and

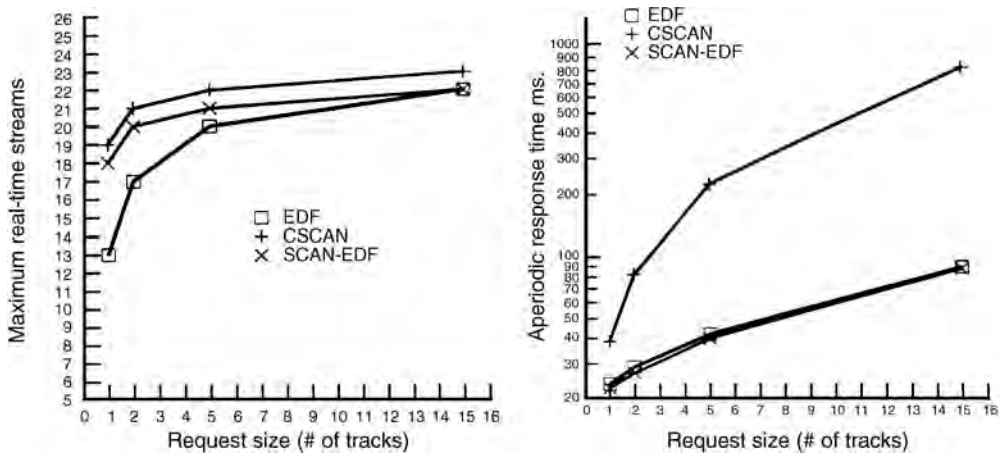


FIGURE 99.2 Performance impact of scheduling policies and request block size.

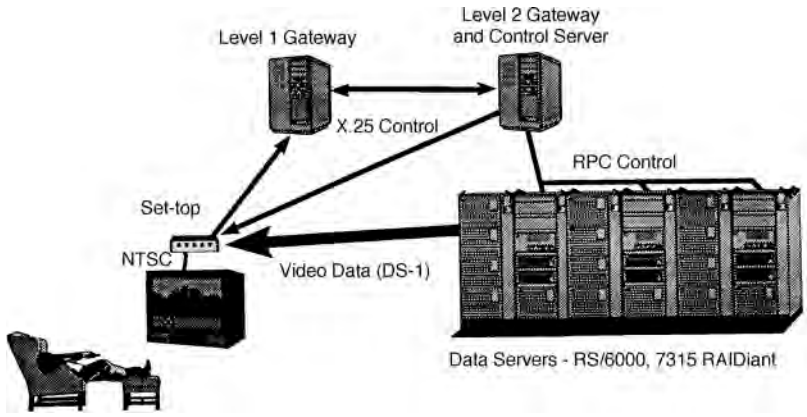


FIGURE 99.3 IBM's Hong Kong video server setup.

destination pairs are scheduled without a conflict for the transmission of the first block of data then the entire playback of that video stream will not see any source and destination conflicts [Reddy, 1995]. Providing similar guarantees within the multiprocessor network depends on the network organization. Some interconnection networks such as Omega networks and hypercubes have been shown to exhibit properties that make it amenable to provide guarantees of no conflicts within the network. A simple solution such as a round-robin distribution of data blocks among the storage nodes suffices to provide such guarantees [Reddy, 1995] in these networks.

The basic approach to providing guarantees of conflict-free transfers relies on time-division multiplexing of the network resources. The network usage is divided into a number of slots. Each slot provides sufficient time to transfer a fixed block across the network under no conflicts. Each stream is assigned a fixed slot. The transfers are repeated after a fixed amount of time (a frame) to transfer the next block of each stream. The regular data distribution and the network properties ensure that if the first block of a stream is scheduled without any conflicts at the source or destination and within the network, the entire stream of blocks of that video stream can be transferred without any conflicts. When the cluster network has a higher bandwidth than the delivery network, the delays across the cluster network may not add significant delays to the end-to-end delay in delivering the data.

Network Delivery

The actual delivery of video data blocks to the customer depends significantly on the network employed. Performance guarantees such as guaranteed bandwidth and delay bounds are necessary to provide high quality video picture to the customer. The current Internet network does not provide any performance guarantees, hence, the resulting low-quality delivery of Web-based video content. It is possible to provide dedicated bandwidth over various networks. These issues will be discussed further in Section 99.4.

Other Issues

In this section, we briefly touch upon a number of issues that have significant impact on the design of a video server. The intent is to give the reader a feel of the problems and to provide starting pointers for further reading.

Buffering

Buffering can make a significant impact on the video server design. Once the video stream is started on the consumer's monitor, the delivery of the video stream has a real-time nature, that is, the data has to be supplied at a constant rate. However, the system can control when the first block of the stream is delivered to the consumer's monitor (latency). Buffering can be used effectively to control the latency of delivery in video-on-demand applications. The more the data is buffered, the greater the stream startup latency and the longer the time to serve a request block at the server. In more interactive applications, long latencies cannot be tolerated. Buffering can be used to increase the request block size or to provide extra time for serving a request of a video stream (deadline extension). Both these options, larger request blocks and deadline extensions, are shown to improve the number of video streams that can be supported [Reddy and Wyllie, 1993].

Interactive Response

Providing a VCR-like capability of pause/resume and fast forward/reverse requires that sufficient resources be available at the server to absorb the variations in bandwidth demands due to such operations [Jayanta et al., 1994]. Pause operation may reduce the bandwidth demand at the server, and the other three operations may increase the bandwidth requirement. To support such operations, the server has to determine that these bandwidth variations do not affect the delivery of other scheduled streams. It is also possible to keep the bandwidth requirements at the server constant by serving alternate copies of video streams that have less quality encoding at the higher frame rates required by the fast-forward and reverse operations [Chen, Kandlur, and Yu, 1994]. Then the server can serve the alternate copies of video data during fast-forward and reverse operations without altering the schedule of the scheduled streams. When interactivity is dominant, EDF-like scheduling may be employed at the cost of lower throughput.

Fault Tolerance

Bandwidth considerations, load balancing, and other issues favor striping data across multiple disks and nodes in a multiprocessor video server. However, this increases the probability of losing service due to a failure of one of the disks or nodes in the system. In *line of business* applications such as ITV, VOD, and DVB, the video server must be highly reliable and tolerate hardware failures with minimal or no system downtime. To illustrate the importance of this, consider a large ITV system supporting 30,000 users (systems of this size were considered in early trials [WSJ, 1995; Comput. Int., 1995; Bus. Wire, 1995]). Assuming the primary application is pay-per-view movies at $5 apiece, the system is generating approximately $75,000/h! This is approximately one year's pay for the service person, who undoubtedly would not collect it for long if outages were frequent or prolonged.

Multilevel encoding has been proposed to tolerate against failures in the delivery medium. In these schemes, the video data is stored in the server in two layers: a base layer and an enhancement layer. When the available bandwidth is affected due to failures, the server can continue to provide a lower quality picture to the consumer by sending only the base layer and thereby putting less load on the network.

A similar idea can be used for reducing the disk overhead by protecting against the loss of the base layer only and by not providing redundancy to protect against the loss of the enhancement layer.

As opposed to a conventional file or database system, loading from tape is not a desirable way to recover from storage device failures. At the 6-Mb/s video data rate common in most ITV systems, a 2-h movie requires 5.2 gigabyte of storage, which takes 1/2 h to restore even from a relatively high-performance 3 megabyte/s tape drive. Some ITV systems in use by customer trials have 1 tetrabyte online [WWW], which would require 97 h (or many tape drives) to restore. This makes redundant storage [mirroring or redundant arrays of inexpensive disks (RAID)] a requirement.

Disk failures can be protected either by making duplicate copies of data (disk mirroring) or by parity protection (RAID) techniques [Patterson, Gibson, and Katz, 1988]. These fault-tolerance techniques increase the availability of data and typically can tolerate the loss of a disk without losing access to data. However, these data protection techniques incur extra load under a failure, and this extra workload needs to be factored into the design for providing guarantees after an failure. To protect against the failure of an SCSI bus (bus that interconnects multiple disks to a system's memory bus), it is possible to connect the disks to two SCSI buses such that the disks may be accessible through a second bus after a failure.

These techniques can guarantee availability of data but cannot guarantee availability of the necessary I/O bandwidth for timely delivery of data after a failure. Scheduling data delivery after a component failure is a difficult problem. Much work needs to be done in this area. Overdesigning the system is one of the possible options so that even after a failure the data can be delivered in a timely fashion. Dynamic resource allocation for tolerating failures is another possible option. In certain cases, only statistical guarantees may be provided as to the capability of providing the required real-time bandwidth after a failure.

Variable Bit Rate Streams

Video can be compressed into constant bit rate (CBR) streams or variable bit rate (VBR) streams. While it is easier to schedule CBR streams, VBR streams tend to produce higher quality at the same bit rates. VBR streams complicate resource scheduling because of the inherent variations in demand. A number of approaches can be taken to account for these variations. The peak bit rate of the video can be considered for scheduling purposes. However, such an approach leads to underutilization of the server and network resources. VBR streams can be smoothed [Salehi, Zhang, Kurose, and Towsley, 1996] to reduce rate variability and to improve the delivered throughput. It is also possible to adjust the starting times of the videos to reduce the overall system load variability over time [Reddy and Wijayaratne, 1999] to improve the total throughput.

Variable Bandwidth Devices

Some of the current disks utilize *zone-bit recording*, which makes the track capacity a variable depending on the location of the track on the disk. This results in a variable data rate depending on the track location. The variable data rate makes it harder to utilize the disk bandwidth efficiently. To guarantee the real-time requirements, the minimum supportable bandwidth can be assumed to be the average deliverable bandwidth. However, this results in inefficient use of the disk bandwidth. It is possible to use the remaining bandwidth for nonreal-time requests while utilizing the minimum disk bandwidth for real-time requests. Request spreading techniques [Birk, 1995] can be used to make sure that the requests are spread uniformly over the surface of the disk to utilize the disk at its average bandwidth.

Integrated Service

Much work has been done recently to enable servers to provide integrated service simultaneously to multiple types of data. While video playback requires periodic guarantees, interactive operations such as fast forward or rewind require quick response from the server. Similarly, interactive video games require the server to provide a quick respone to a user's commands. Providing different guarantees to different types of requests is more challenging than providing bandwidth guarantees required for video playback. A number of projects have recently shown that it is possible to design servers that can provide integrated service.

Fellini [Ozden, Rastogi, and Silberschatz, 1994], Symphony [Shenoy, 1998], and Prism [Wijayaratne, 2001] are examples of experimental file servers that have been designed to provide integrated service to multiple data types.

99.3 Video Networks

The most prominent types of digital networks used for ITV are **asynchronous digital subscriber loop (ADSL), hybrid fiber coax (HFC), asynchronous transfer mode (ATM), and high-speed ethernet**. Although there are major differences between these networks, they all support both a high bandwidth video channel for sending video to the set-top box, and a lower bandwidth control channel (usually bidirectional) for communicating user commands and other control information with the server. ADSL multiplexes the downstream video channel, a bidirectional 16-Kb/s control channel (which usually employs X.25 packet switching), and the existing analog telephone signal over the same copper twisted pair [Bell, 1993a, 1993b]. ADSL is designed to leverage the existing telephone company (telco) wiring infrastructure and, as such, has been used in several VOD trials conducted by telcos [WNT, 1993; Plumb, 1993; Patterson, 1994].

ADSL is used for the connection between the customer premises and the telephone company central switching office. The server can either be located in the central office or at some centralized metropolitan or regional center, in which case the data travels between the server and the customer's central office over the telephone company's digital trunk lines. The video bandwidth the ADSL is capable of supporting depends on the length of the subscriber loop (i.e., the distance between the central office and the customer premises). At present, an ADSL will support 1.544 Mb/s (T1 bandwidth), correspondingly higher bandwidths over shorter distances. MPEG compressed at T1 bandwidth can deliver VCR quality video but is marginal for applications such as sports. The choice of T1 bandwidth is convenient because it is widely used in the telco trunk system and because T1 modems are readily available to connect the video server to the digital network.

As ADSL attempts to leverage the existing telephony infrastructure, hybrid fiber coax [Ciciora, 1995] leverages the existing cable TV plant and, as such, is popular for ITV systems deployed by cable service providers. Standard cable plants deliver some number of analog TV channels (typically 70 channels using a total bandwidth of 450 MHz) which are fanned out to all subscribers. The cable system is structured as a large tree, with repeaters at each of the nodes and subscriber cable decoders at the leaves. However, 70 channels is barely a sufficient number for broadcast and is not nearly sufficient for VOD or ITV. The basic idea behind HFC is to subdivide the tree into smaller subtrees (for example, one subtree per neighborhood) and feed each subtree independently by fiber from the cable company head end. Some number of channels in this subtree are still reserved for broadcast, and the remainder are available for interactive services. A segment of the cable's frequency range (typically 5–30 MHz) is reserved for upstream control signalling, and LAN-like packet switching is used on this upstream channel.

In an HFC system, each user of interactive services is assigned a channel for the period during which he is using the system. The server sends a command over the control channel to tell the STB which channel to select and, subsequently, sends video to the STB over this channel. For security, STBs can prevent interactive channels from being selected manually.

In some early trials, a subset of the analog TV channels were reserved for ITV [UPI, 1993]. MPEG video from the server was converted to analog at the cable head end and multiplexed onto fiber. Since only a small number of channels (those not used for broadcast) on each subtree are available for ITV, the area covered by a subtree is small, and therefore a completely analog system is expensive.

New techniques [vestigial sideband (VSB), quadrature amplitude modulation (QAM)] have been developed that allow sending high-bandwidth digital data over analog TV channels. QAM-64, for example, allows 27 Mb/s of digital data to be carried by one 6-MHz analog TV channel. This payload can carry one high-bandwidth program [e.g., high-definition television (HDTV)], but is normally used to carry an MPEG-2 transport stream with a number of lower-bandwidth programs. For example, QAM-64 allows four 6-Mb/s MPEG-2 programs to be carried on a single 6-MHz analog TV channel, allowing a 70-channel cable TV

plant to carry 280 channels. Higher bandwidth fiber-based systems are being deployed that support a downstream bandwidth of 750 MHz, which supports over 400 6 Mb/s MPEG programs. This is suitable for carrying a full complement of broadcast channels as well as a sufficient number of interactive channels for a neighborhood of 1000 households.

The downstream video path of an HFC video server is somewhat more complicated than its ADSL counterpart. The server sends individual MPEG programs to an MPEG multiplexor, which combines the independent programs into a single MPEG transport stream. The transport stream is then converted (e.g., by a QAM modulator) to a 6-MHz analog signal and then multiplexed onto one of the cable system's analog channels. The STB selects its MPEG program under control of the server from one of the MPEG transport streams.

A number of alternatives exist for transporting video over ATM. ATM defines one standard [ATM adaptive layer-1 (AAL1)] for continuous bit rate traffic and another standard (AAL5) for packet-based data. Perhaps surprisingly, most ITV systems have used the packet-based approach. Even using AAL5, several alternatives exist: the ATM fabric can be used as a fast Internet protocol (IP) network, or a specialized video streaming protocol can be used [ATM, 1995]. Most ATM-based ITV systems implement the ATM Forum standard; examples of IP-based systems are discussed in a later section. Time–Warner's full-service network (FSN) ITV system, deployed in Orlando, Florida [Time, 1993], uses a combination of ATM and HFC. The server sends multiple MPEG program streams via ATM to a QAM modulator. From there, they are transmitted by fiber to the customer STB, as previously discussed. Here, ATM is used primarily as a head-end switching fabric rather than as a distribution mechanism. ATM has been deployed in select campuses and corporations to the desktop but has not found widespread deployment to the curb in residential access.

99.4 Network Multimedia

During the last several years, the cost of computer power, storage devices, and storage media has dropped radically in the personal computer marketplace. This has led to the proliferation of interactive multimedia applications and games for the PC. To date, most multimedia applications run on stand-alone PCs, with digitized video and audio coming from local hard disks and compact disc read-only memories (CD-ROMs). Increasingly, there has been a demand for file servers that support capture, storage, and playback of multimedia data. The reasons for this are identical to the ones that motivate the use of file servers for conventional data: sharing, security, data integrity, and centralized administration.

Servers for Network Multimedia

Although interactive TV is the glamor application for video servers, video playback over a local area network (LAN) is the bread and butter of the industry. A LAN multimedia system includes one or more servers, some number of multimedia client workstations, and a video-capable LAN. Applications for LAN multimedia include training videos, information kiosks, desktop news, and digital libraries. The client workstation software architecture has much to do with the design of the multimedia LAN server. There are two basic client architectures, depending on whether or not the manufacturer-supplied driver software provided with the video codec is used.

Normally, the video codec (e.g., MPEG card and Indeo software codec) comes packaged with driver software that allows video to be played through the operating system multimedia support (e.g., the Windows Media Player interface). The loop to read a buffer of video data and present it on the display is implemented in this driver, which is a black box executable file supplied by the decoder manufacturer. Since this driver reads data through the file system interface, the server must make video available through that interface. With advances in PC speeds, MPEG-2 DVDs are now decoded in software entirely. Figure 99.4 shows a typical multimedia LAN server.

Most client platforms (Windows, Macintosh, UNIX) support the notion of an installable file system (IFS), which the multimedia LAN server can use to provide video data. However, the performance

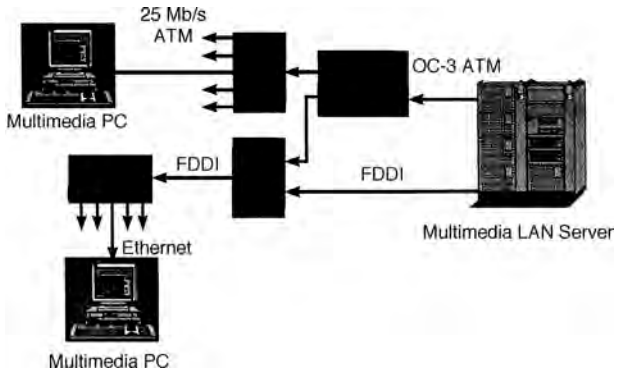


FIGURE 99.4 A multimedia LAN server.

characteristics of the multimedia IFS driver must be equivalent to local disk or CD-ROM. In particular, any read-ahead or extra buffering to smooth out variations in the network transfer rate must be implemented in the IFS driver, not in the video codec device driver (which is unaware that video is coming from the server as opposed to from local disk).

This file-based approach has the advantage of allowing the server to support almost any video codec. Furthermore, applications that use the operating system media player or multimedia programming interfaces (most multimedia applications fall into this class) need not be modified to use the server. For example, such servers allow copying a multimedia CD-ROM to the server, after which it can transparently be used by any number of clients simultaneously.

The server can either implement its own proprietary installable file system or can use one designed to support a standard network operating system (NOS) protocol (e.g., Netware, LAN Manager, or NFS). A proprietary file system allows using more aggressive read-ahead and buffering or a network transport protocol tuned for video. However, a number of commercially available servers use standard NOS IFS drivers (or NOS clients). This has obvious advantages: it allows the server to support a large variety of platforms and makes it unnecessary for the user to install special software to use the server. Several commercial servers have been built using standard NOS clients: examples include IBM OS/2 LAN Server Ultimedia [IBM, 1993], Panasonic Video NFS [MEI, 1995], and IBM Multimedia LAN Server for AIX [IBM, 1996].

Standard NOS network protocols work acceptably for video over a suitable LAN, that is, one with sufficiently high throughput, low latency, and low error rate. The need for throughput is obvious; low latency is required to prevent buffer starvation with the limited amount of read-ahead done by most standard NOS clients. For example, the NFS protocol supports a maximum 8 K read size. Application (e.g., codec driver) read requests larger than 8 K are broken up by the NOS client and sent to the server as multiple 8-K requests. If the NOS client sends these requests synchronously (i.e., one at a time), the time to complete a request must average less than 54 ms to maintain a data rate of 1.2 Mb/s (the data rate for most CD-ROM MPEG). By supporting read-ahead (sending successive requests before previous ones complete), an NOS client can relax this latency requirement somewhat. Most NOS clients support read-ahead. The final requirement for low error rate stems from the fact that in most NOS clients, error recovery is driven by timeouts. These timeouts are typically short enough to allow the occasional dropped packet without a visible interruption in video, but bursts of errors will cause video to degrade.

Modern technology allows configuring LANs to support video in a cost-effective manner. Switched ethernet, fiber distributed data interface (FDDI), and ATM LAN emulation are all suitable for video and with high-performance routers can be scaled up to support a large building or campus. A number of projects have developed technology for ensuring sufficient bandwidth over LAN, for example, Venkataramani and Chiueh [1995].

A number of multimedia LAN server vendors have taken the approach of implementing proprietary streaming protocols to transport video over LANs. Typically, this forces the server vendor to supply special driver software for each video codec it wishes to support. This usually restricts such servers to a limited number of codecs and increases the time required to support new ones. However, a proprietary streaming protocol can, in principle, be designed to work better with a less-than-perfect network. Starlight Networks' Star Works and its MTP protocol is an example of this approach [Starlight, 1996a, 1996b].

Internet Video Servers

The recent explosion of interest in the Internet and the World Wide Web have resulted in a number of attempts to support video over long distances using the Internet. These range from modest attempts to support low-resolution video over dial-up and integrated services digital network (ISDN) connections to more ambitious attempts to transport high-resolution video over the MBONE wide-area network (WAN) and high-speed networks, for example, Numtv [2001].

The earliest attempts to support video and audio over the World Wide Web used store-and-forward. Using this technique, the video file is downloaded to the client workstation in its entirety, after which the operating system media player is launched to play the file. For all but the shortest videos, the delay to download before viewing is unacceptable.

Recently, a number of systems have become available that play video and audio in real-time over the Internet. Most are implemented as a Web browser external viewer or plug-in (the distinction is not important to this discussion), and most work more or less the same way. When the user clicks on a link to video (or audio) on a Web page, the browser loads the contents of the file pointed to by the link and launches the viewer (or plug-in). This file, instead of containing the video data itself, contains a URL pointer to the data, often on a separate, dedicated video server. The viewer starts transferring data from the server, buffering it in random access memory (RAM) until it has a sufficient amount to begin playback. Video is decoded in software and displayed in a window; accompanying audio is played from the data buffer using the low-level audio subsystem programming interfaces (e.g., WaveOutWrite on Windows).

In contrast to the file-based approach, this type of viewer usually does not take advantage of special hardware (e.g., MPEG cards) that might be on the system or that supports a limited number of cards using special drivers. Given the low video data rates possible over the Internet at present, this is usually not a serious consideration. On the other hand, rather than using a conventional network file system protocol (e.g., NFS) or even a conventional streaming protocol [e.g., transport control protocol (TCP)], these viewers usually implement special-purpose network transport protocols (layered on UDP) that are designed specifically for audio or video. Much work has been done in devising suitable protocols for video transmission across networks [for example, Smith, 1996; Dwyer et al., 1998; Floyd et al., 2000]. Some additionally implement proprietary compression algorithms tailored to the Internet's low data rate and high error rate. Commercial products implementing proprietary transport protocols and/or compression algorithms include VDO Live [VDOnet, 1995], Real Audio, and Microsoft Windows Media.

The research community and standards organizations have also been active in transporting multimedia over the Internet. Real-time transport protocol (RTP) defines an encapsulation of any of a number of audio or video data formats into UDP datagrams that allow it to be streamed over the Internet. Rather than retransmitting lost packets as does TCP, RTP timestamps packets to allow missing packets to be detected. RTP contains no flow-control mechanism and leaves the task of recovering from lost packets to the video/audio codec. A number of commercial products are based on RTP (Apple's QuickTime, Precept FlashWare, Netscape LiveMedia, Microsoft's Netmeeting and Real Player). A number of multimedia servers are available from major vendors, for example, IBM's videocharger.

Video Multicast

Since video puts considerable demands on the network bandwidth, there has been much interest in finding efficient ways of distributing video across the network. For transmission of live or stored videos that may be of interest to multiple receivers, multicasting may significantly reduce the bandwidth demands

on the network. Much work has been done recently in devising efficient encoding, transmission techniques for multicasting video. We briefly discuss some of these issues.

The vic video conferencing system supports real-time video multicast using RTP [Schulzrinne et al., 1996]. Although the original version of vic supported only live video, Argonne National Laboratories has extended vic to support playback of stored video at high bandwidth (6 Mb/s) via RTP from IBM's Tiger Shark video server [Haskin, 1998] on an IBM SP-2 parallel computer [IBM, 1995]. Several such experimental video tools, including Xerox PARC's Network Video, INRIA's IVS, and CuseeMe, have been developed for videoconferencing across networks.

When video is multicast across the Internet, receivers with different network bandwidths may wish to participate in the multicast. To accommodate such heterogeneity in network connectivity, video may be transmitted in multiple layers [Brown et al., 1996; McCanne, Jacobson, and Vetterli, 1996]. A receiver would then subscribe to as many layers of video as possible in order to maximize the received video quality within the bandwidth constraints of his/her network connection. Significant amounts of research work have been done to optimize the delivered quality in the presence of receiver heterogeneity and variability in network bandwidth.

In the context of video servers, it is possible to multicast popular videos to several receivers simultaneously. Videos or video objects may be multicast in the form of channels, and a user would subscribe to particular channel(s) to receive the requested video. Depending on the available bandwidth and the popularity of the videos, a video may be multicast at regular intervals to satisfy the user requests. A user would then wait for the beginning of the multicast of the requested video. A number of such video broadcasting schemes have been proposed with various assumptions about network bandwidth, popularity of videos, receivers' network connectivity [Hua and Sheu, 1997; Dan, Sitaram, and Shahabuddin, 1994]. These schemes try to combine as many user requests as possible into one batch to be served by a single multicast channel. In such schemes, a user's latency tolerance for the beginning of a video is a significant factor in determining the multicast efficiency. A number of techniques have been proposed to reduce the user perceived latency in such multicast schedules [Hua, Cai, and Sheu, 1998].

Defining Terms

Asynchronous digital subscriber loop (ADSL): Technology that allows video to be delivered over regular phone lines.

Data server: Computer system that stores and retrieves digital video.

Earliest deadline first (EDF): A real-time scheduling policy.

Hybrid fiber coax (HFC): Technology that allows bidirectional data transfer over cable.

Interactive TV (ITV): Video delivery depends on the interactive commands from the customer.

MPEG: A video compression standard proposed by the motion picture experts group.

Real-time/continuous media: Data that requires delivery within deadlines.

SCAN: A seek optimizing disk scheduling policy.

Set-top box: A decoder used to decompress and convert digital video to an appropriate input form suitable for a customer's monitor.

Striping: Distribution of data over several storage nodes (or disks).

Video on demand (VOD): A service where a video stream is delivered on demand.

References

ATM. 1995. Doc. 95-0012r5. ATM Forum.

Bell, 1993a. ADSL video dial tone network interface requirements. TR-72535, Bell Atlantic, Inc., Dec.

Bell, 1993b. Signalling specification for video dial tone. TR-72540, Bell Atlantic Inc., Dec.

Birk, Y. 1995. Track pairing: a novel data layout for VOD servers with multi-zone recording disks. *Proc. IEEE Conf. Multimedia Comput. Sys.*

Brown, T., Sazzad, S., Schoeder, C., Cantrell, P., and Gibson, J. 1996. Packet video for heterogeneous network using CU-seeme, *Proc. IEEE Int. Conf. Image Proc.,* Sept.

Business Wire, 1995. NEC to supply set-top boxes to Hong Kong telecom IMS. *Business Wire,* Nov. 14.

Chen, M.S., Kandlur, D., and Yu, P.S. 1994. Support for fully interactive playout in a disk-array-based video server. *Proc. ACM Multimedia Conf.,* Oct., 391–398.

Chen, Z., Tan, S.M., Campbell, R.H., and Li, Y. 1995. Real time video and audio in the world wide web. *Proc. 4th Int. World Wide Web Conf.,* Dec. 11.

Chervenak, A. 1994. Tertiary storage: an evaluation of new applications. Ph.D Thesis, University of California, Berkeley, CA.

Ciciora, W.S. 1995. Cable television in the United States—an overview. Cable Television Labs., Inc.

Computergram International. 1995. Hong Kong to get world's first multimedia service. *Computergram Int.,* Nov.

Dan, A., Sitaram, D., and Shahabuddin, P. 1994. Scheduling policies for an on-demand video server with batching, *Proc. ACM Multimedia*.

Dwyer, D., Ha, S., Li, J.R., and Bhargavan, V. 1998. An adaptive transport protocol for multimedia communication, *Proc. IEEE Multimedia Comput. Sys*.

Floyd, S., Handley, M., Padhye, J., and Widmer, J. 2000. Equation-based congestion control for unicast applications, *Proc. ACM SIGCOMM,* August.

Haskin, R. 1998. Tiger Shark: a scalable file system for multimedia. *IBM J. Res. Dev.,* 42(2), 185–198.

Haskin, R. and Stein, F. 1995. A system for the delivery of interactive television programming. *Proc. IEEE 1995 Spring COMPCON,* San Francisco, CA, March, 209–216.

Hua, K., Cai, Y., and Sheu, S. 1998. Patching: a multicast technique for true video-on-demand services, *Proc. ACM Multimedia,* Sept.

Hua, K. and Sheu, S. 1997. A new broadcasting scheme for metropolitan video-on-demand systems, *Proc. ACM SIGCOMM,* Sept.

IBM. 1993. IBM LAN server Ultimedia—multimedia for OS/2, DOS, and Windows clients. IBM Doc. 96F8520, Nov.

IBM. 1995. The medium is the message: IBM, Argonne National Lab. host first bi-coastal "internews" conference. IBM Corp. press release IBM US 490, San Diego, CA, Dec. 6.

IBM. 1996. IBM multimedia server for AIX, version 1.1. IBM Doc. 5967-213, Feb. 20.

Insoft. 1996. Insoft targets internet real time multimedia users and developers with first interactive collaborative environment. Insoft Inc. press release, Mechanicsburg, PA, Jan. 8.

Jayanta, K.D., Salehi, J.D., Kurose, J.F., and Towsley, D. 1994. Providing VCR capabilities in large-scale video server. *Proc. ACM Multimedia Conf.,* Oct., 25–32.

Liu, C.L. and Layland, J.W. 1973. Scheduling algorithms for multiprogramming in a hard real-time environment. *J. ACM,* 46–61.

McCanne, S., Jacobson, V., and Vetterli, M. 1996. Receiver-driven layered multicast, *Proc. ACM SIGCOMM*.

MEI. 1995. Matsushita develops "video server" technology for use on standard computer networks. Matsushita Electric Industrial Co., Ltd. press release, March 22.

Numtv.com. 2001. www.numtv.com.

Ozden, B., Rastogi, R., and Silberschatz, A. 1994. Fellini—a file system for continuous media, Tech. Report 113880-941028-30, AT&T Bell Labs.

Patterson, A. 1994. Hong Kong Telecom, IBM map video effort. *Electron. Eng. Times,* Aug. 1.

Patterson, D.A., Gibson, G., and Katz, R.H. 1988. A case for redundant arrays of inexpensive disks (RAID). *ACM SIGMOD Conf.,* June.

Plumb, L. 1993. Bell Atlantic demonstrates video on demand over existing telephone network. Bell Atlantic press release, June 14.

Reddy, A.L.N. 1995. Scheduling and data distribution in a multiprocessor video server. *Proc. Int. Conf. Multimedia Comput. Sys.,* 256–263.

Reddy, A.L.N. and Wijayaratne, R. 1999. Techniques for improving the throughput of VBR streams. *Proc. Int. Conf. Multimedia Comput. Networking,* Jan.

Reddy, A.L.N. and Wyllie, J. 1993. Disk scheduling in a multimedia I/O system. *Proc. ACM Multimedia Conf.,* Aug.

Salehi, J., Zhang, Z.L., Kurose, J., and Towsley, D. 1996. Supporting stored video: reducing rate variability and end-to-end resource requirements through optimal smoothing. *Proc. ACM SIGMETRICS,* May.

Schulzrinne, H., Casner, S., Frederick, R., and Jacobson, V. 1996. A transport protocol for real-time applications. Internet RFC 1889, Internet Engineering Task Force, Jan.

Shenoy, P. 1998. Symphony: an integrated multimedia file system, Ph.D. Thesis, Dept. of Computer Science, University of Texas.

Smith, B. 1996. Cyclic UDP: a priority-driven best-effort protocol, Tech. Report, Cornell University.

Starlight. 1996a. The challenges of networking video applications. Starlight Networks Tech. Rep., Starlight Networks Inc., Mountain View, CA, April 22.

Starlight. 1996b. Effect of video on LAN data traffic. Starlight Networks Tech. Rep., Starlight Networks Inc., Mountain View, CA, April 22.

Time. 1993. Time Warner cable completes first phase of full service network. Time Warner press release, Dec. 16.

UPI. 1993. Cox to launch interactive tests in Omaha. United Press International, Dec. 1.

VDOnet. 1995. VDOlive will enable motion video on the internet. VDOnet Corp. press release (http://www.vdolive.com), Oct. 27.

Venkataramani, C. and Chiueh, T.C. 1995. Design, implementation and evaluation of a software driven real-time ethernet protocol, *Proc. ACM SIGCOMM*.

Wijayaratne, R. 2001. Prism: a file server architecture for providing integrated services, Ph.D. Thesis, Dept. of Computer Science, Texas A & M University.

WNT. 1993. Bell Atlantic, IBM announce agreement for video on demand server. *World News Today,* Jan. 8.

WSJ. 1995. *Wall Street Journal,* Nov. 10.

WWW. Basic plan for the multimedia experiments at the Tokyo metropolitan waterfront subcenter multimedia experiments. http://www.tokyo-teleport.co.jp/english/atms/atmsj20l.htm.

Yu, P.S., Chen, M.S., and Kandlur, D.D. 1993. Grouped sweeping scheduling for DASD-based multimedia storage management. *Multimedia Sys.,* 1:99–109.

Further Information

Readers are encouraged to refer to *Readings in Multimedia Computing and Networking*, edited by K. Jeffay and H. Zhang, published by Morgan Kaufmann publishers, August 2001.

100

Videoconferencing

Madhukar Budagavi
Texas Instruments

100.1 Introduction

Videoconferencing enables groups of people and individuals in different locations to communicate with each other using the following media: (1) video, with which participants can see motion pictures of each other, (2) audio, with which participants can hear each other, and (3) **data**, with which participants can exchange still images/slides and share files, documents, and applications. Early videoconferencing systems were expensive and required large amounts of bandwidth (mainly because of the bandwidth required for transmitting video). Also, high speed networks necessary to transfer videoconferencing data were not widely available. All these factors hindered widespread deployment of videoconferencing systems. But with advances in microelectronics, video and audio compression techniques, and network technologies, videoconferencing is becoming more and more economical and popular. Current videoconferencing systems can be broadly divided into room-based systems and desktop-based systems. In room-based videoconferencing, participants at a site have to go to conference rooms specially fitted with videoconferencing equipment (typically with a large television monitor) in order to videoconference with others in similarly equipped rooms at remote sites. In contrast, desktop videoconferencing takes videoconferencing to the familiar desktop personal computers (PC) of the users. Desktop videoconferencing enables users to sit at their own desk and videoconference with others using their personal computers. It leads to a cheaper, more convenient, and flexible way of implementing a system for videoconferencing. With the proliferation of advanced wireless networks, another way of videoconferencing—wireless videoconferencing using mobile handheld terminals—is slowly emerging now.

The next section gives an overview of the various components that make up a videoconferencing system. Videoconferencing over the integrated services digital network (ISDN), over the common analog

TABLE 100.1 Videoconferencing Standards Discussed in This Chapter

Standard	Name	Networks Supported
H.320	Narrowband visual telephone systems and terminal equipment	ISDN
H.324	Terminal for low bit rate multimedia communication	GSTN, Circuit-switched wireless network, ISDN
H.323	Packet-based multimedia communications systems	IP Networks
SIP	Session Initiation Protocol	IP Networks, Packet-switched wireless networks

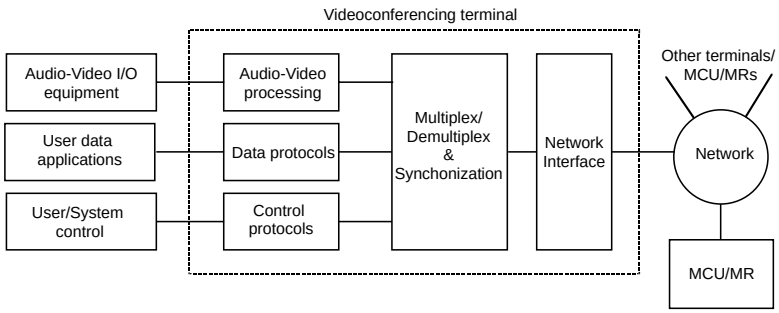


FIGURE 100.1 A general videoconferencing system configuration.

general switched telephone network (GSTN), and over packet-based networks (specifically over networks based on the Internet protocol [IP]) are described next. The emphasis of this chapter will be on discussing relevant standards for videoconferencing, since videoconferencing standards have played a very significant role in accelerating the widespread acceptance of videoconferencing. Table 100.1 lists the videoconferencing standards discussed in this chapter. Both room-based and desktop-based videoconferencing systems use the same set of videoconferencing standards. The chapter concludes with a list of recent developments and extensions to the videoconferencing standards—one of the extensions that standardizes videoconferencing over wireless networks is discussed in more detail.

100.2 Overview

A videoconferencing system consists of three important components: the videoconferencing terminals, the network, and the **multipoint** control unit/**multicast** router (MCU/MR). Figure 100.1 shows a typical videoconferencing system configuration. The videoconferencing terminal is the device that provides the necessary videoconferencing facilities to the users. It interfaces with input/output devices producing/consuming videoconferencing data for/from other users. In the case of desktop-based systems, the videoconferencing terminal is usually the user's desktop computer or workstation. In the case of room-based videoconferencing, the videoconferencing terminal is housed in cabinets or on carts that can be wheeled into conference rooms. The network is used for transporting videoconferencing data between participants, whereas the MCU/MR is responsible for coordinating and directing the flow of information between three or more videoconferencing users who wish to simultaneously communicate to each other.

Videoconferencing Terminal

The videoconferencing terminal consists of four processing blocks: (1) audio and video, (2) data, (3) control, and (4) multiplex–demultiplex-synchronization blocks, as shown in Fig. 100.1.

- **Audio and video processing block.** This block interfaces with audio and video (AV) input/output (I/O) devices. In the case of desktop-based systems, video input is usually through a desktop video

camera connected to the PC through special purpose video processing cards inside the PC. In the case of room-based systems, the video camera is more sophisticated and, typically, its direction of focus can be controlled by electronic means. The users at the remote end (and also at the local end) can change the direction the camera is pointing to in order to focus on different people. Electronic controllability also enables implementation of advanced features such as voice activated camera control, where the camera automatically trains on the person speaking. The video output is usually on the display monitor of the PC for desktop-based systems and on a large television monitor for room-based systems. The audio I/O equipment usually consists of microphones and headphones for desktop-based systems. For room-based systems, the audio output is usually through speakerphones. The AV processing block performs another important function: compression/decompression of the AV signals. Compression is necessitated by the fact that current networks used to transport videoconferencing data have limited bandwidth. Compression/decompression is typically done using AV codec chips or special purpose digital signal processing chips. In the case of desktop-based systems, compression/decompression can be done in software too, using the local processor of the PC.

- **Data processing block.** The term data here is used to specify information arising out of applications such as electronic whiteboards, still images/slides, shared files and applications, and documents. The data processing block enables users to exchange such information and to work jointly on applications. Since applications usually reside on the PC of the user, data sharing is more convenient in desktop videoconferencing when compared to the room-based videoconferencing systems.
- **Control block.** As visible to the user, control is in the form of some graphical user interface (GUI) window on the PC screen or on the television monitor, with which a user can initiate conferences and set up preferred data rates and compression methods for audio and video. The user can typically select the options on the GUI using a mouse for desktop-based systems or through a wireless keypad for room-based systems. At the system level, the control unit is responsible for **end-to-network signaling** for accessing the network and **end-to-end signaling** for proper operation of the desktop terminal.
- **Multiplex–demultiplex-synchronization block.** This block multiplexes the resulting audio, video, data, and control signals into a single stream before transmission on the network. Similarly, the received bit stream is demultiplexed to obtain the audio, video, data, and control signals which are then passed on to their respective processing blocks. This block is also responsible for synchronization of the different media (audio/video/data) being transmitted and for receiver feedback for **quality of service** (QoS) monitoring. The multiplex–demultiplex block accesses the network through a suitable network interface. Depending on the network interface, multiplexing and demultiplexing may not be required sometimes.

Network

The network is the entity used for transporting videoconferencing data between participants. For videoconferencing applications it is desirable that the network provide a low and predictable delay connection; this is essential for transmission of audio and video data, which are real-time signals with strict delay requirements. There are two types of networks, circuit-switched networks and packet-switched networks, for which videoconferencing standards and applications currently exist. Circuit-switched networks, which allocate a dedicated amount of bandwidth, provide a predictable delay connection, whereas packet-switched networks, which packetize data and transmit them over shared nondedicated bandwidth networks, cannot guarantee predictable timing of data delivery. Commonly used networks are the ISDN and GSTN (which are circuit-switched networks) and local area networks (LANs) and the Internet (which are packet-switched networks). The network interfaces provided by ISDN and GSTN networks mandate the use of a multiplex/demultiplex block in videoconferencing terminals. However, for packet-switched networks, the network interface is such that the multiplex/demultiplex block is not essential—multiplexing and demultiplexing occurs at networking layers below the layer exposed at the network interface.

Multipoint Control Unit/Multicast Router

Videoconferencing can be classified as point-to-point or multipoint, based on the number of participants in a given session. A point-to-point videoconference has only two users who communicate with each other via audio, video, and shared data applications, whereas multipoint conferencing involves more than two participants and multimedia information from each participant has to be multicast to all others. Special network servers are required to coordinate the distribution of the videoconferencing data amongst the multiple participants in a multipoint videoconference. These servers are called multipoint control units in the terminology of circuit-switched networks. The equivalent function in packet-switched networks can be carried out by multicast routers. Note that some packet-switched networks such as LANs inherently have multicasting capabilities.

100.3 Videoconferencing over Integrated Services Digital Network (ISDN)

The integrated services digital network (ISDN), a digital communication channel, is currently the most widely used channel in videoconferencing for business related purposes. Most of the videoconferencing systems over ISDN incorporate the H.320 standard [1], proposed by the International Telecommunications Union—Telecommunications Standardization Sector (ITU-T). ITU-T H.320 is a suite of standards and operates at rates from 64 kb/s to about 1920 kb/s. H.320 makes use of a number of other ITU-T standards, which can be integrated, as shown in Fig. 100.2, to form the complete videoconferencing system. The standards for audio, video, control, and multiplex are summarized in Table 100.2.

TABLE 100.2 H.320 Series of Recommendations

H.320	Narrowband visual telephone systems and terminal equipment
H.261	Video codec for audiovisual services at $p \times 64$ kb/s
G.711	Pulse coded modulation of voice frequencies
G.722	7-kHz audio coding within a 64-kb/s channel
G.728	Coding of speech at 16 kb/s using low-delay code excited linear prediction
H.221	Frame structure for a 64–1920 kb/s channel in audiovisual teleservices
H.242	System for establishing communication between audiovisual terminals using digital channels up to 2 Mb/s
H.230	Frame synchronous control and indication signals for audiovisual systems

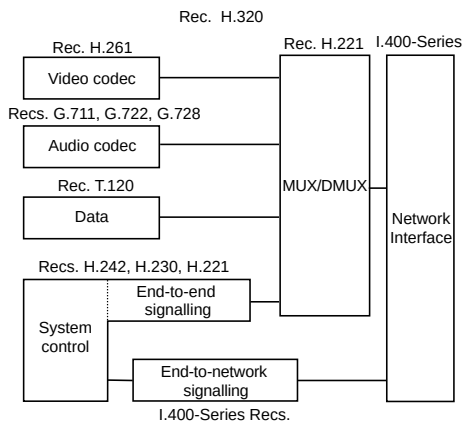


FIGURE 100.2 H.320 protocol stack.

TABLE 100.3 Picture Formats for H.261 and H.263

Picture Format	Recommendations	Luminance (pixels per line ¥ lines)	Chrominance (pixels per line ¥ lines)
Sub-QCIF	H.263	128 × 96	64 × 48
QCIF	H.261 and H.263	176 × 144	88 × 72
CIF	H.261 and H.263	352 × 288	176 × 144
4 CIF	H.263	704 × 576	352 × 288
16 CIF	H.263	1408 × 1152	704 × 576

Audio

H.320 supports both telephone quality (G.711 and G.728) as well as **wideband** quality (G.722) speech [2]. G.711 does the traditional A-law/μ-law quantization of speech whereas G.728 is based on the **analysis-by-synthesis** approach and makes use of the low-delay **code excited linear prediction** (LD-CELP) algorithm to compress speech to 16 kb/s. The wideband standard, G.722, splits the wideband speech into low and high bands and does adaptive differential pulse code modulation (ADPCM) encoding of the two subband sequences. The type of audio coding used in a videoconferencing session is signaled by using procedures defined in the control standards of H.320 described later. For further details on G.722 and G.728 and other speech coding standards, the reader is referred to [2].

Video

H.261 specifies the video coding standard for use in H.320. H.261 operates on video signals, which have to be in either of the two standard picture formats: the common intermediate format (CIF) or the quarter-CIF (QCIF). The **luminance** and **chrominance** resolutions of the CIF and QCIF formats are summarized in Table 100.3. Pictures in H.261 are coded either in intraframe or in interframe modes. In intraframe coding, the image is encoded without any relation to the previous image, whereas in interframe coding, the current image is predicted from the previous image using **block motion compensation** and the difference between the current image, and the predicted image is encoded. A single **motion vector** is used for each *macroblock* (data corresponding to 16 × 16 blocks in an image). Depending on the mode of coding used, the macroblocks of either the image or the residual image are split into blocks of size 8 × 8, which are then transformed using a two-dimensional discrete cosine transform (DCT). The resulting DCT coefficients are quantized, run-length encoded, and finally Huffman encoded to give a bit rate of $p \times 64$ kb/s, where p is an integer between 1 to 30. An overview of the H.261 standard is given in [3].

Multiplex

The H.221 standard determines the way in which audio, video, data, and control is multiplexed in H.320 before transmission on the ISDN line. Data on each 64 kb/s ISDN channel is transmitted in 80 octet (bytes) frames, as specified in H.221. Each 80 octet frame is split into 8 subchannels, as shown in Fig. 100.3. Multiplexing is achieved by allocating video, audio, data, and control information to different subchannels. For example, if G.728 (16 kbps) is used for speech coding, the resulting compressed speech would be transmitted in subchannels 1 and 2, and video, data, and control would be sent in the remaining subchannels. A fixed part of subchannel 8 is always reserved for signaling the service channel (SC) in all H.221 frames. The SC in a frame consists of 8 bits each of bit-rate allocation signal (BAS), frame alignment signal (FAS), and encryption control signal (ECS), as shown in Fig. 100.3; the remaining bits (bits 25–80) of the SC can be used for transmitting videoconferencing data. The first 8 bits of the SC form the FAS, which helps the receiver synchronize to the H.221 frame boundaries. The BAS channel consists of bits 9–16 of SC and is used to signal commands and **capabilities** to the other end. It is also used to transmit information describing the bit allocation between video, audio, and data channels in H.221 frames.

Bit number								
1	2	3	4	5	6	7	8 (SC)	
								1 Octet number
S	S	S	S	S	S	S	FAS	:
u	u	u	u	u	u	u		8
b	b	b	b	b	b	b		9
-	-	-	-	-	-	-	BAS	:
c	c	c	c	c	c	c		16
h	h	h	h	h	h	h		17
a	a	a	a	a	a	a	ECS	:
n	n	n	n	n	n	n		24
n	n	n	n	n	n	n		25
e	e	e	e	e	e	e		.
l	l	l	l	l	l	l		.
#	#	#	#	#	#	#	#	.
1	2	3	4	5	6	7	8	80

FAS Frame alignment signal
BAS Bit-rate allocation signal
ECS Encryption control signal

FIGURE 100.3 H.221 frame.

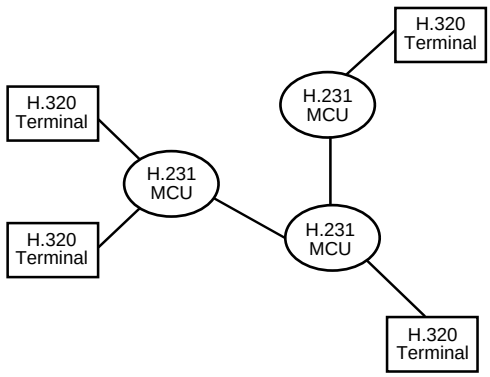


FIGURE 100.4 Multipoint videoconferencing using H.320.

Control

ITU-T H.242 defines the procedure and protocols to be used for negotiating capabilities of the terminals in order to determine the mode of operation (i.e., the type of audio/video codec, etc.) to use. It uses BAS codes defined in H.221 to negotiate the mode.

Other Standards Related to H.320

ITU-T H.231 and H.243 cover procedures for use by MCUs for multipoint conferences. Figure 100.4 shows a topology of H.320 terminals in a multipoint conference. ITU-T H.233 and H.234 specify encryption and key exchange for H.320. Encryption is used if secure videoconferencing, conferences that cannot be monitored as they pass through the network, is desired.

100.4 Videoconferencing over General Switched Telephone Networks (GSTN)

General switched telephone networks (GSTN) are the common analog telephone line networks. ITU-T H.324 specifies the standard for videoconferencing over the GSTN by using V.34 data modems [4,5]. The bit rate targeted is very low at around 33.6 kb/s. The H.324 protocol stack is shown in Fig. 100.5. Like H.320, H.324 also invokes other ITU-T standards to provide complete videoconferencing functionality. Table 100.4 gives a list of these standards. The audio coding standard G.723.1 [6] and the video coding

TABLE 100.4 H.324 Series of Recommendations

H.324	Terminal for low bit rate multimedia communication
H.263	Video coding for low bit rate communication
G.723.1	Dual rate speech coder for multimedia communication transmitting at 5.3 and 6.3 kb/s
H.223	Multiplexing protocol for low bit rate multimedia communication
H.245	Control protocol for multimedia communication

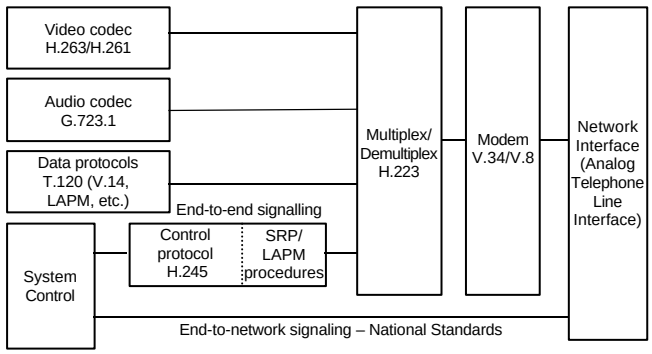


FIGURE 100.5 H.324 protocol stack.

standard H.263 [7] were specifically developed for the low bit rate operation on GSTN targeted by H.324. All H.324 terminals should support at least the V.34 modem, H.223 multiplex, and H.245 system control protocol and, depending on the application, could support audio and video codecs and data protocols.

Audio

G.723.1 is the recommendation for speech coding in the H.324 family of standards [6]. G.723.1 specifies a dual rate speech coder with bit rates of 5.3 and 6.3 kb/s. The higher bit rate coder provides better quality than the lower bit rate coder. Both rates are a mandatory part of the encoder and decoder. The variable bit rate operation allows for dynamic adjustment of the speech bit rate depending on the requirements of other bit streams (such as video) which have to be simultaneously transmitted. G.723.1 is based on the principles of **linear prediction** (LP) and analysis-by-synthesis coding. The coder operates on speech frames of 30 ms duration (which, in turn, is split into four subframes of 7.5 ms each). The LP coefficients are vector quantized and transmitted once per frame. The excitation to the LP filter is modeled as a combination of two components: one from a fixed codebook and the other from an adaptive codebook. The adaptive codebook contribution corresponds to long-term prediction and reduces redundancy present in the speech signal due to the pitch periodicities. A fifth-order pitch predictor is used, and information concerning the pitch predictor (pitch lags and gains) is transmitted once every subframe. The main difference in coding for the two rates comes from the type of excitation used in the fixed codebook. The high rate coder uses multipulse maximum likelihood quantization (MP-MLQ) and the low rate coder uses algebraic codebook excited linear prediction (ACELP). Both schemes essentially model the excitation as a sequence of pulses, and information regarding the fixed codebook excitation is transmitted in the form of pulse position and pulse signs once every subframe.

Video

H.263 is the video coding standard for low bit rate communication in H.324 [7]. H.263 is based on the H.261 recommendation and improves upon it to provide an equivalent subjective quality at less than half the bit rate. H.263 consists of a base level coder and four optional modes that are defined in four

annexes to the standard. H.263 specifies five standardized picture formats, as listed in Table 100.3. The H.263 recommendation mandates that all decoders be able to operate using at least the sub-QCIF and QCIF formats and that the encoder be able to operate on at least either of the two formats, sub-QCIF or QCIF. Use of other picture formats is optional. Most of the improvements at the base level come from using motion vectors with half-pixel resolution and from using improved variable length coding for coding DCT coefficients. Also, motion vectors and other overhead information are coded more efficiently than in H.261. The four optional modes of H.263 further improve compression efficiency. The reader is referred to [7] for more details on the optional modes.

Multiplex

Video, audio, data, and control information is transmitted in H.324 on distinct logical channels. ITU-T H.223 determines the way in which the logical channels are mixed into a single bit stream before transmission over the GSTN. The H.223 multiplex consists of two layers, the adaptation layer that sits on top of the multiplex layer. The adaptation layer adapts the information stream provided by the applications above it to the multiplex layer below it by adding, where appropriate, additional octets for the purposes of error control and sequence numbering. The type of error control used depends on the type of information (audio/video/data/control) being conveyed in the stream. The adaptation layer provides error control support in the form of cyclic redundancy check (CRC) bits and retransmission protocols. The multiplex layer is responsible for multiplexing the various logical channels. It transmits the multiplexed stream in the form of packets. Each packet transmitted by the multiplex layer contains a one octet header and a payload consisting of a variable number of information field octets carrying videoconferencing information. The packets are delimited by an 8 bit HDLC (high-level data link control) flag (01111110). The header octet includes a multiplex code, which specifies, by indexing to a multiplex table, the logical channels to which each octet in the information field belongs. To prevent emulation of the HDLC flag in the payload, bitstuffing is adopted.

Control

H.324 systems contain exactly one control channel in each direction. The messages and procedures used on these channels are specified in Recommendation H.245. The control channel is always carried on logical channel 0. The information carried on the control channel includes capabilities data, multiplex table entries, and mode preferences and flow control messages. The control channel is also used for opening and closing logical channels. More details on H.245 are provided in the next section.

100.5 Videoconferencing over Internet Protocol (IP) Networks

Unlike videoconferencing over ISDN and GSTN discussed earlier, videoconferencing over the Internet is through packet-switched networks. Connectivity on the Internet is provided through the Internet protocol (IP). Two transport layer protocols were developed for use with IP: the transmission control protocol (TCP) and the user datagram protocol (UDP). TCP was designed to provide a reliable point-to-point service for delivery of packet information in proper sequence, whereas the UDP was designed to simply provide a service for delivering packets to the destination without bothering about congestion control. TCP uses retransmission of lost packets to guarantee delivery. TCP is found to be inappropriate for real-time transport of audio and video information because retransmissions may result in indeterminate delays leading to discernible distortions and gaps in the real-time playout of the audio and video streams. In contrast, UDP does not have the problem of indeterminate delays because it does not use retransmission. However, the problem in using UDP for transmitting media is that UDP does not provide sequence numbers to transmitted packets. Hence, UDP packets can get delivered out-of-order and packet loss might go undetected.

Another important factor in determining the transport to use for media transmission purposes is their support for multicasting. The problem that was initially faced in using IP directly for large-scale

videoconferencing was that IP was a unicast (point-to-point) protocol and did not support efficient multicasting. Clearly, unicasting separate copies of videoconferencing data to each of the recipients is highly inefficient and leads to flooding of the network with redundant data. To overcome this problem, the IP multicast model was proposed by Deering and Cheriton [8]. In IP multicast, the network does the replication as paths to different receivers diverge. Support for IP multicast is required both in network routers and switches and in the end terminals. TCP does not support multicast, whereas UDP supports it.

In practice, UDP is used in videoconferencing for transmitting audio and video because it does not have the problem of indeterminate delays (as in TCP) and because it supports multicasting. However, as was noted earlier, UDP does not provide information about sequencing of packets. It also does not provide support for intermedia and intramedia synchronization. These problems motivated the development of the real-time transport protocol (RTP) [9] for transporting real-time audio and video data over the Internet. RTP was proposed by the Audio Video Transport Working Group (AVTWG) of the Internet Engineering Task Force (IETF). RTP is typically layered over UDP and relies on the UDP to provide it services such as framing and multiplexing. RTP consists of two parts: a data part and a control part. The data part handles the transportation of the actual videoconferencing media (audio and video) and takes care of timing and sequencing of media packets. The control part, real-time transport control protocol (RTCP), is responsible for sending control information related to the RTP streams. RTCP packets include information that helps in synchronizing different media (e.g., lip-syncing of audio and video). They also include receiver feedback of packet loss. The receiver feedback is used to monitor QoS of the network. The QoS information may be used to adaptively change the source coder parameters so as to either decrease or increase the data rate as required.

Based on these developments, videoconferencing standards have since been standardized both in ITU-T, which standardized the H.323 protocol [10–12], and in IETF, which standardized the session initiation protocol (SIP) [13,14]. Both these standards are described in this section. Both these standards make use of RTP for transporting the videoconferencing media.

H.323

Unlike H.320 and H.324, H.323 describes not only the videoconferencing terminals but also other components on the network. Hence, before going into the details of H.323, it is useful to first look at the various components in a H.323 enabled network. Figure 100.6 shows a H.323 enabled network. The four main components are: (1) H.323 terminal: the videoconferencing terminal where the user initiates a videoconference or receives a videoconferencing call, (2) MCU: the multipoint control unit responsible for multipoint videoconferencing, (3) Gatekeeper: controls access of H.323 terminals to the network and regulates the bandwidth that the terminals can use, and (4) Gateway: the endpoint in the network that

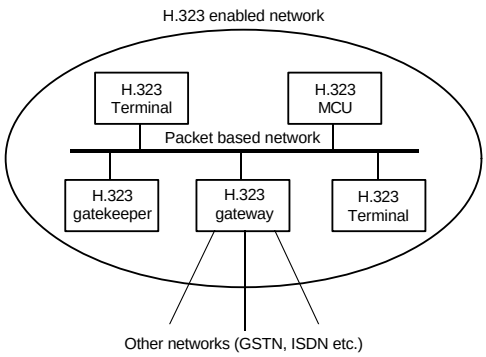


FIGURE 100.6 H.323 enabled network.

TABLE 100.5 H.323 Series of Recommendations

H.323	Packet-based multimedia communications systems
H.261, H.263	Video coding standards. See Tables 100.2 and 100.4
G.711, G.722, G.723.1, G.728	Speech coding standards. See Tables 100.2 and 100.4
G.729	Coding of speech at 8 kbit/s using conjugate structure algebraic code-excited linear-prediction (CS-ACELP)
H.225.0	Call signaling protocols and media stream packetization for packet based multimedia communication systems
H.245	Control protocol for multimedia communication

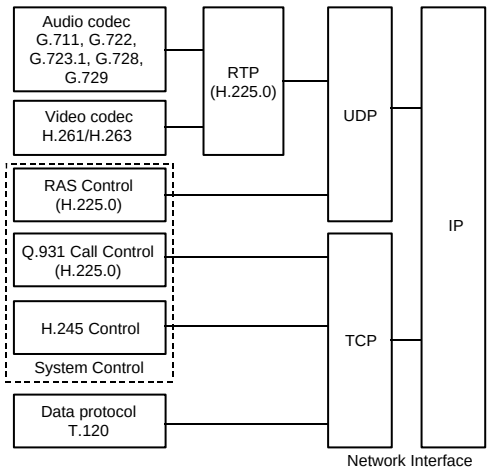


FIGURE 100.7 H.323 protocol stack.

provides services for H.323 terminals to interoperate with other terminals on other networks (such as ISDN, GSTN etc.). The H.323 standard basically specifies the signaling between all these entities and it also specifies the media codecs that can be used. Table 100.5 lists the main standards that fall under the umbrella of H.323. Figure 100.7 shows the H.323 protocol stack.

Audio and Video

The video and audio coding standards described in the previous sections are supported in H.323 too. ITU-T G.729 is an additional audio codec that is supported in H.323. G.729 is a 8 kbps speech coder based on a linear prediction analysis-by-synthesis technique that uses ACELP. More details on G.729 can be found in [15].

Media Transmission

Standard H.225.0 specifies that RTP should be used for media transmission. There is no multiplexing of audio and video data within a RTP packet. Audio and video data are sent separately in different RTP packets in separate streams. Multiplexing of audio and video packets occurs at the network level.

Control

There are basically three control and signaling protocols:

- **Registration admission and status (RAS):** This protocol is defined in H.225.0 and takes care of terminal to gatekeeper signaling. The messages exchanged deal with bandwidth allocation, call admission, etc.

- **Call control signaling:** This protocol is defined in H.225.0 and is used for call-setup and teardown between H.323 terminals. This protocol makes use of the Q.931 signaling protocol that is commonly used for signaling in ISDN and GSTN.
- **End-to-end control using H.245:** H.245 is used in H.323, too, for exchanging terminal capability, opening and closing of logical channels, and other control purposes. The H.245 standard defines a message syntax and a set of protocols to exchange multimedia messages. The message syntax is defined using ASN.1 [16]. ASN.1 encodes messages in a binary format. It is instructive to look at a H.245 message. Figure 100.8 shows an example H.245 message where both the binary encoding and the corresponding decoded message are shown (both were obtained using the H323 stack from www.openh323.org). The decoded message has been printed in a form that aids user readability. The H.245 message shown is an open logical channel message that requests the opening of a μ-law quantized G.711 audio channel. The media codec information is sent in the data type field. This open logical channel message also includes other information such as the IP address to use for the RTCP connection (this appears in the media control channel field). Note that the open logical channel message in a H.324 system will be different from the one shown in Fig. 100.8.

The interplay of the various protocols can be understood by tracing the steps of a typical H.323 call. Figure 100.9 shows a typical call sequence where two H.323 endpoints place a call through the gatekeeper. There are seven phases. In the first phase, the endpoints make use of RAS to request permission from the gatekeeper to make/receive a call. In the second phase, Q.931 is used to place a call from the calling endpoint to the called endpoint. H.245 is then used to negotiate capabilities between the two endpoints. H.245 is also used to open the logical channels for transmitting audio and video using RTP. The call is said to have been established at this stage. In the next phase (the fourth phase), the endpoints use RTP to send audio and video data to each other to achieve two-way communications. Once the videoconferencing conversation is over, the call can be shutdown by either endpoints. The logical channels are first closed using H.245. This forms the fifth phase of the call. In the sixth phase, the call is torn down using Q.931. In the final phase of the call, the endpoints signal the gatekeeper using RAS to indicate that the

```
H.245 - actual sequence of bytes transmitted (Hex notation):
03 00 00 64 0c 60 1d 80 0b 0d 00 01 00 80 f7 63 23 13 89 80

H.245 Decoded message:
request openLogicalChannel {
  forwardLogicalChannelNumber = 101
  forwardLogicalChannelParameters = {
    dataType = audioData g711Ulaw64k 30
    multiplexParameters = h2250LogicalChannelParameters {
      sessionID = 1
      mediaGuaranteedDelivery = FALSE
      mediaControlChannel = unicastAddress iPAddress {
        network = 4 octets {
          80 f7 63 23
        }
        tsapIdentifier = 5001
      }
      silenceSuppression = TRUE
    }
  }
}
```

FIGURE 100.8 An H.245 message example.

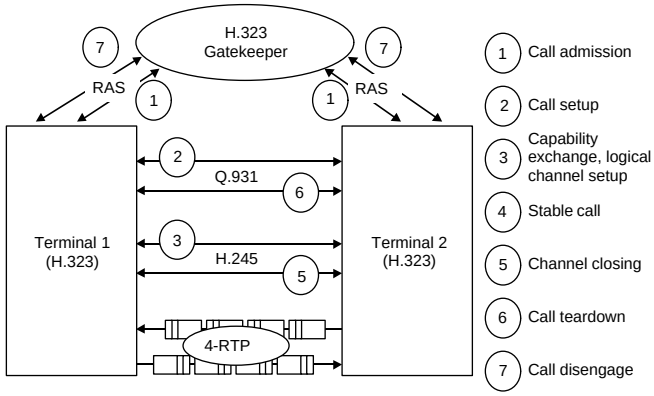


FIGURE 100.9 A gatekeeper enabled H.323 call.

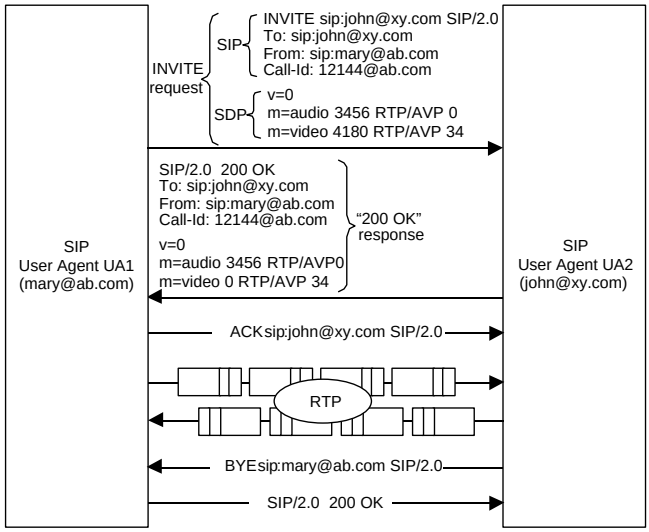


FIGURE 100.10 A simple SIP call.

resources (e.g., bandwidth) used for the call are no longer in use. The gatekeeper then captures the resources and waits for the next call.

Session Initiation Protocol (SIP)

SIP was standardized by IETF [13,14]. The videoconferencing functionality supported by SIP is more or less the same as that supported by H.323. SIP is becoming popular because it is simpler than H.323 and can be more easily modified to support newer features. The SIP protocol is basically a call signaling protocol and, as such, does not mandate the procedures to use in deciding on the media streams to use in the videoconference nor does it specify the audio and video codecs to use. Current implementations of SIP use the session description protocol (SDP) [17] to describe multimedia sessions and use RTP to transmit audio and video. SDP provides the functionality of H.245 in SIP. Note, however, that H.245 is much more sophisticated, and also much more complex, than SDP. A videoconferencing terminal in SIP is called a user agent. SIP uses textual encoding of its messages with a syntax very similar to HTTP [18]. The working of SIP is explained using the message trace of a simplified call between two user agents—UA1 and UA2, as shown in Fig. 100.10 (Note that the messages shown in Fig. 100.10 have been truncated

to simplify the narrative and to highlight the important parts). UA1 places a call to UA2 by sending an INVITE message. The "From" field indicates the originator of the call, the "To" field indicates the recipient of the call, and the "Call-ID" field provides a unique session ID for the call. The INVITE message also includes the SDP message in its body, as shown in Fig. 100.10. The "m=" lines in the SDP message indicate the type of media that UA1 is able to receive and transmit. In this case "m=audio 3456 RTP/AVP 0" indicates capability to receive G.711 audio using RTP and "m=video 4180 RTP/AVP 34" indicates capability to receive and transmit H.263 video using RTP. On receipt of the INVITE message, UA2, if it decides to accept the call, returns a "200 OK" message indicating willingness to participate in the call. The "200 OK" response also contains an SDP message indicating the media that UA2 can receive and transmit. UA1 acknowledges the receipt of "200 OK" message from UA2 by sending an "ACK" message. At this stage, the videoconferencing call is considered to be established. Media transmission using RTP then takes place between the two user agents. At the time of termination of the call, the user agent that wants to terminate the call sends a "BYE" message. More details on SIP can be found in [14].

SIP is also being standardized for use on third generation packet-switched wireless networks [19].

100.6 Recent Developments and Extensions to Videoconferencing Standards

Many of the videoconferencing standards described in the previous section have been extended from the time they were first finalized. These extensions include addition of new features, support for other network types, and modifications for improvement in calling speed. They also include support for newer media codecs such as MPEG-4 for video [20], H.263 Version 2 (which includes more optional annexes) for video [21], and some GSM speech coders [22]. This section highlights some of the extensions.

Wireless Extensions to H.324

The wireless channel is a noisy fading channel characterized by long bursts of errors on the channel. Hence, the data that is transmitted on the wireless channel needs to be error protected. H.324 was initially designed for use on the benign GSTN. It was later extended to cover mobile wireless networks by adding error protection to H.223 [20,23]. Three annexes (referred to as Levels 1–3, respectively) were defined to provide improved levels of error protection. The initial specification of H.223 (See Section 100.4) is referred to as Level 0. Together, Levels 0–3 provide for a trade-off of error robustness against the overhead required, with Level 0 being the least robust and using the least amount of overhead and Level 3 being the most robust and also using the most amount of overhead.

- **H.223 Level 1 (Annex A)**—Communication over low error-prone channels—A longer 16-bit pseudo-noise synchronization flag was introduced in Level 1 to allow for more reliable detection of packet boundaries. The input bitstream is correlated with the synchronization flag and the output of the correlator is compared with a correlation threshold. Whenever the correlator output is equal to or greater than the threshold, a flag is detected. The other improvement incorporated in Level 1 is that bitstuffing is removed. The use of bitstuffing leads to poor performance in the presence of errors. Since bitstuffing is not performed, it is possible to have the synchronization flag emulated in the payload. However, the probability of such an emulation is low and is outweighed by the improvement gained by not using bitstuffing over error-prone channels.
- **H.223 Level 2 (Annex B)**—Communication over moderate error-prone channels—When compared to the Level 1 operation, Level 2 increases the protection on the packet header. A multiplex payload length (MPL) field, which gives the length of the payload in bytes, is introduced into the header to provide additional redundancy for detecting the length of the video packet. A (24,12,8) extended Golay code is used to protect the multiplex code and the MPL fields. Use of error protection in the header enables robust delineation of packet boundaries. Note that the payload data is not protected in Level 2.

- **H.223 Level 3 (Annex C)**—Communication over highly error-prone channels—Level 3 goes one step above Level 2 and provides for protection of the payload data. Rate compatible punctured convolutional (RCPC) codes, various CRC polynomials, and ARQ techniques are used for protection of the payload data. Level 3 allows for the payload error protection overhead to vary depending on the channel conditions. RCPC codes are used for achieving this adaptive level of error protection.

The wireless extensions of H.324 have been adopted for use on third-generation circuit-switched wireless networks [24].

Extensions to H.323

H.323 has been updated three times so far, resulting in specifications commonly called H.323 Version 2, H.323 Version 3, and H.323 Version 4. The original H.323 is called Version 1. Some of the modifications and improvements have been:

- **Fast Connect:** Reduces call setup time by combining the Q.931 and H.245 phases of the call.
- **Supplementary services:** A new set of standards—the H.450 series of standards—was defined to support supplementary services such as call transfer, call forwarding, call waiting, name identification, etc., in H.323 enabled networks. These services are similar to those available on GSTN currently.
- **Use of UDP for call signaling transport (Annex E/H.323):** Signaling in H.323 (H.245 and Q.931) is normally transferred on top of TCP. Because of the handshake mechanism inherent in establishing a TCP connection, TCP takes more time to connect to a remote endpoint when compared to UDP. Therefore, use of UDP reduces call setup time.
- **Simple endpoint type (SET) (Annex F/H.323):** The SET was introduced to reduce the cost of application-specific hardware devices that do not require the full functionality and generality of H.323. A SET device utilizes only a small portion of the full H.323 protocol stack but is still able to establish audio calls with other H.323 endpoints.
- **Multiplexed stream transmission:** Currently, both audio and video are sent in separate RTP packets in two separate RTP streams. Synchronization of audio and video becomes difficult because of this. Annex F of H.225.0 allows both audio and video to be multiplexed into a single RTP stream. This enables endpoints to achieve better synchronization of audio and video. The H.223 multiplex is basically used for this multiplexing. The multiplexed packets that come out of H.223 are encapsulated in RTP packets and transmitted.

Defining Terms

Analysis-by-synthesis technique: In this technique, the excitation to the linear prediction filter is chosen from a (predefined) set of possible excitation vectors such that the perceptually weighted error between the input speech signal and its reconstructed version (which is the output of the linear prediction filter) is minimized.

Block motion compensation: Motion compensated prediction is done on a block basis, that is, blocks of pixels are assumed to be displaced spatially in an uniform manner from one frame to another.

Capability: Describes the terminal's ability to process a particular kind of information stream.

Code excited linear prediction (CELP): Uses a "codebook"—a set of vectors available at both the encoder and the decoder. An index into this codebook is used to signal the excitation vector to be used for reconstructing the speech.

Data: This term is also used to specify information arising in applications such as electronic whiteboards, application sharing, and still images/slides transmission.

End-to-network and end-to-end signaling: End-to-network signaling is the signaling between the terminal and the network access point to gain access to the network. End-to-network signaling is used to place calls to the remote end. End-to-end signaling consists of system messages that are

exchanged between terminal endpoints, e.g., capability messages exchanged to determine a common mode of operation.

Linear prediction: Models the speech as parameters of a linear filter wherein the current sample of speech is predicted from the past (typically) 10–12 samples of speech.

Luminance and chrominance: Luminance is the brightness information in a video image, whereas chrominance is the corresponding color information. For H.261 and H.263, color is encoded as two color differences commonly called Cb and Cr components.

Motion vector: Specifies the spatial displacement of a block of pixels from one frame to another.

Multicast: Transmission of the same information to multiple recipients. Terminology more commonly used in packet-switched networks such as the Internet.

Multipoint communication: Simultaneous communication between more than two participants. Terminology usually used in circuit-switched network applications.

Quality of service (QoS): Provided by Internet, is measured in terms of parameters such as packet loss and packet delay variation.

Wideband speech: Traditional telephone quality speech is bandlimited to about 3 kHz. Wideband speech is bandlimited to 7 kHz and gives more natural sounding speech.

References

1. International Telecommunications Union-Telecommunications Standardization Sector, Narrowband visual telephone systems and terminal equipment, Recommendation H.320, 1993.
2. Spanias, A.S., Speech coding: a tutorial review, *Proc. IEEE,* 82, 1541–1582, Oct. 1994.
3. Schafer, R. and Sikora, T., Digital video coding standards and their role in video communications, *Proc. IEEE,* 83, 907–924, June 1995.
4. International Telecommunications Union-Telecommunications Standardization Sector, Terminal for low bit rate multimedia communications, Recommendation H.324, 1998.
5. Lindbergh, D., The H.324 multimedia communication standard, *IEEE Commun. Mag.,* 34, 46–51, Dec. 1996.
6. Cox, R.V. and Kroon, P., Low bit-rate speech coders for multimedia communication, *IEEE Commun. Mag.,* 34, 34–51, Dec. 1996.
7. Rijkse, K., H.263: video coding for low-bit-rate communication, *IEEE Commun. Mag.,* 34, 42–45, Dec. 1996.
8. Deering, S.E. and Cheriton, D.R., Multicast routing in datagram internetworks and extended LANs. *ACM Trans. Comp. Sys.,* 8, 85–110, May 1990.
9. Schulzrinne, H., Casner, S., Frederick, R., and Jacobson, V., RTP: a transport protocol for real-time applications, Request for Comments 1889, Internet Engineering Task Force, Jan. 1996.
10. International Telecommunications Union–Telecommunications Standardization Sector, Packet-based multimedia communications systems, Recommendation H.323, 1999.
11. Thom, G.A., H323: the multimedia communications standard for local area networks, *IEEE Commun. Mag.,* 34, 52–56, Dec. 1996.
12. Liu, H. and Mouchtaris, P., Voice over IP signaling: H.323 and beyond, *IEEE Commun. Mag.,* 38, 142–148, Oct. 2000.
13. Handley, M., Schulzrinne, H., Schooler, E., and Rosenberg, J., SIP: Session initiation protocol, Request for Comments 2543, Internet Engineering Task Force, Mar. 1999.
14. Schulzrinne, H. and Rosenberg, J., The session initiation protocol: Internet-centric signaling, *IEEE Commun. Mag.,* 38, 134–141, Oct. 2000.
15. Salami, R. et al., Design and description of CS-ACELP: a toll quality 8 kb/s speech coder, *IEEE Trans. Speech Audio Proc.,* 6, 116–130, Mar. 1998.
16. International Telecommunications Union–Telecommunications Standardization Sector, Information technology—Abstract Syntax Notation One (ASN.1)—Specification of basic notation, Recommendation X.680, 1994.

17. Handley, M. and Jacobson, V., SDP: session description protocol, Request for Comments 2327, Internet Engineering Task Force, Apr. 1998.
18. Fielding, R. et al., Hypertext Transfer Protocol—HTTP/1.1, Request for Comments 2616, Internet Engineering Task Force, June 1999.
19. 3rd Generation Partnership Project, Packet switched conversational multimedia applications: default codecs, Technical Specification 26.235 V5.0.0, 2001.
20. Budagavi, M. and Talluri, R., Wireless video communications, in *The Mobile Communications Handbook,* Gibson, J.D., Ed., CRC Press, Boca Raton, FL, 1996, Chap. 31.
21. Cote, G. et al., H.263+: video coding at low bit rates, *IEEE Trans. Circuits Syst. Video Technol.,* 8, 849–866, Nov. 1998.
22. Budagavi, M. and Gibson, J.D., Speech coding in mobile radio communications, *Proc. IEEE,* 86, 1402–1412, July 1998.
23. Farber, N., Girod, B., and Villasenor, J., Extensions of ITU-T recommendation H.324 for error-resilient video transmission, *IEEE Commun. Mag.,* 36, 120–128, Jun. 1998.
24. 3rd Generation Partnership Project, Codec for circuit switched multimedia telephony service: Modification to H.324, Technical Specification 26.111 V3.4.0, 2000.
25. Netravali, A.N. and Haskell, B.G., *Digital Pictures: Representation and Compression,* Plenum, New York, 1988.
26. Deller, J.R., Jr., Proakis, J.G., and Hansen, J.H.L., *Discrete-Time Processing of Speech Signals,* Macmillian, New York, 1993.
27. Schaphorst, R., *Videoconferencing and Videotelephony: Technology and Standards,* 2nd ed., Artech House, Norwood, MA, 1999.
28. Gibson, J.D. et al., *Digital Compression for Multimedia: Principles and Standards,* Morgan Kaufmann, San Francisco, CA, 1998.
29. Sun, M.T. and Reibman, A.R., Eds., *Compressed Video over Networks,* Marcel Dekker, New York, 2001.

Further Information

The ITU-T standards covered in this chapter can be procured from http://www.itu.int/ and the IETF standards referred to in this chapter can be found at http://www.ietf.org/. Third-generation partnership project (3GPP) standards for wireless videoconferencing [19,24] can be found at http://www.3gpp.org/. Fundamental level books on video and speech processing include the books by Netravali and Haskell [25] and by Deller, Proakis, and Hansen [26], respectively. Books containing more information on videoconferencing include those by Schaphorst [27], Gibson et al. [28], and the book edited by Sun and Reibman [29].

Current research relevant to desktop videoconferencing is reported in a number of journals including *IEEE Transactions on Speech and Audio Processing, IEEE Transactions on Image Processing, IEEE Transactions on Circuit and Systems for Video Technology*, and *Signal Processing*. Conferences of interest include the IEEE International Conference on Image Processing (ICIP), the IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), the IEEE Infocomm, and ACM Multimedia.

VIII

Data Recording

101

Magnetic Storage

Jaekyun Moon
University of Minnesota

101.1 Introduction

Magnetic data storage is a critical component in modern computers, taking up a significant fraction of the overall computer system cost. From the early 1990s to around 1997, the storage density of commercial hard disk drives grew at an annual compound rate of 60% a year. In more recent years, the growth rate has been over 100% a year. The demand for increased storage capacity will continue in the Internet era, given the need to store massive amounts of data, including graphics and images, on central servers as well as personal computers. The growth of the E-commerce industry will only stimulate the appetite for more storage space. Magnetic storage is also competing with flash memory in portable electronics devices. Microdrives are available that weigh just 20 grams and store 1 gigabyte of data. The size of a quarter, these miniature drives are poised for deployment in digital cameras and various handheld devices. Application to future mobile phones is an intriguing possibility. Personal TV is another visible and growing application area for magnetic disk storage.

In the year 2000, commercial disk drives could store up to 18 gigabits of data per square inch of disk surface. Internal media transfer rates are up to 700 Mbits/sec. The reliability requirement is stringent: after rereads and full error correction, the unrecoverable bit error rate is kept below 10^{-14} while the probability of miscorrection is typically less than 10^{-21}.

For a successful design of high-performance, high-capacity magnetic storage devices, advanced signal processing techniques are playing an increasingly important role. There are several important factors one must consider in developing signal processing schemes suitable for magnetic storage systems. Among them are cost constraint, power consumption, and the data rate requirement.

Starting with the communication channel modeling of the read/write process, this chapter reviews some of the important coding and detection techniques employed in today's magnetic storage devices. Techniques that have a potential for improving future storage systems are also discussed.

101.2 Communication Channel Model of the Read/Write Process

Figure 101.1 shows a block diagram of general data storage systems. Error correction coding (ECC) is first applied to information bits. In commercial magnetic storage devices, the Reed–Solomon (RS) codes are used almost universally for outer error correction. The information bits and the RS parity bits are then subject to modulation coding. The modulation code imposes runlength constraints on the bit sequence to facilitate gain control and timing recovery during the retrieval operation. Constrained codes can also be designed to improve distance properties at the detector output. The modulation encoded bits or the symbol sequence are converted into a rectangular current waveform and then written in the medium in the form of a magnetization waveform. This "write" process is accomplished by a small magnetic element called the write head, moving at a constant velocity above the magnetic storage medium. In modern digital recording, the written magnetization takes the form of a rectangular waveform as the medium is saturated either along or opposite to the track direction. At the time of data retrieval, the stored bit sequence is reproduced as the read head senses the magnetic flux (either the flux itself or its gradient depending on the type of read head) emitted from the written magnetization pattern. The signal at the output of the read head appears as a continuous time voltage waveform. The amplified version of the read head output waveform is then passed through what is generally called the read channel in the data storage community. The read channel consists of some type of bandlimiting filter, a sampler driven by a timing circuit, an equalizer, and a symbol detector (in a broader definition, the read channel also includes the modulation encoder/decoder and, possibly, an inner error correction or error detection encoder/decoder). The detected symbol or written bit sequence is then applied to the modulation decoder and finally to the error correction decoder.

A mathematical model of the read/write (R/W) process is useful in design, analysis, and simulation of signal processing schemes. To begin, note that there are two different formats for generating the current waveform given a binary data pattern. With the non-return-to-zero-interleaved (NRZI) format, a binary 1 in the data pattern produces a transition and a binary 0 results in the absence of transition. This is shown in Fig. 101.2a. With the non-return-to-zero (NRZ) method, the amplitude level in the waveform directly reflects the given binary information, as shown in Fig. 101.2b. In the past, data recovery was typically carried out by a simple peak detector which determined whether a transition was present within the given symbol interval. The NRZI recording format is a preferred method with this peak detection, as an error made in a given interval has no effect in the future decisions (i.e., no error propagation). In contrast, serious error propagation will result when the NRZ waveform is reconstructed based on the peak detector outputs. With more sophisticated partial response maximum likelihood (PRML) detectors, however, the NRZI method offers no distinct advantage over the NRZ recording format.

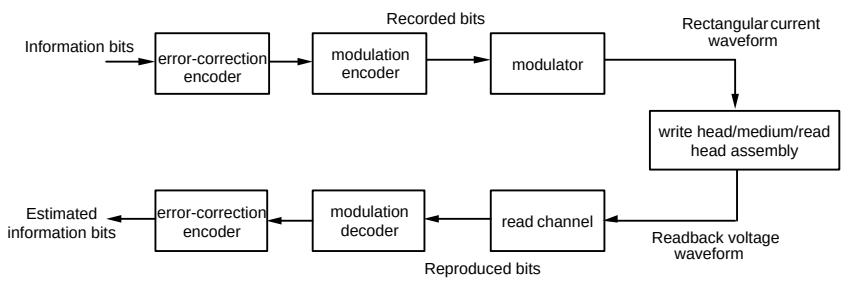


FIGURE 101.1 Block diagram of a data storage system.

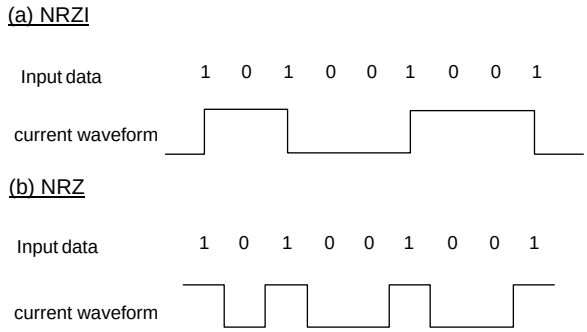


FIGURE 101.2 NRZI and NRZ recording formats.

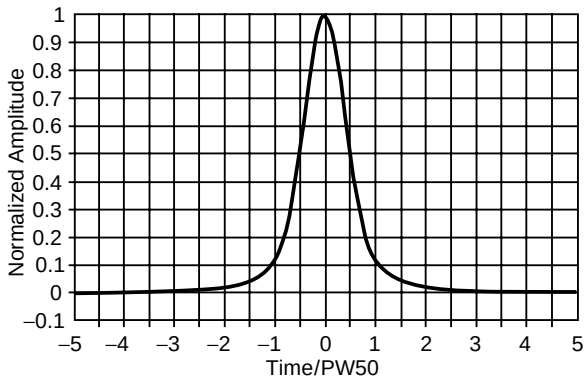


FIGURE 101.3 A typical step response in disk drive systems.

In digital communication systems based on pulse amplitude modulation (PAM), the received waveform is commonly expressed as

$$z(t) = \sum_k b_k s(t - kT) + n(t) \tag{101.1}$$

where b_k is the digital data sequence transmitted, $s(t)$ is the channel impulse response, $n(t)$ represents the additive noise observed at the receiver, and T is the symbol period. An analogous mathematical description exists for the R/W process of magnetic recording. The rectangular current waveform generated according to the given binary data pattern (whether NRZI or NRZ) is viewed as the input to the write-head/medium/read-head assembly. The voltage waveform generated at the output of the read head during the read process is taken as the output of the same write-head/medium/read-head assembly and is essentially a differentiated and low-pass-filtered version of the applied input waveform. See Fig. 101.3. The reproduced waveform is also corrupted by noise. The overall write/read process can be modeled mathematically as:

$$z(t) = \frac{dw(t)}{dt} * h(t) + n(t) \tag{101.2}$$

where $z(t)$ is the reproduced or readback voltage waveform; $w(t)$ denotes the rectangular current waveform that swings between 1/2 to –1/2; $h(t)$ represents the low-pass filter type of response arising from the frequency dependent signal loss terms due to the head-medium spacing in the read process, the read

head gap effect, the medium thickness effect and the imperfect writing of magnetic transitions; $n(t)$ is the additive white Gaussian noise due to the read head and electronics; and * denotes the convolution. The response $h(t)$ can be obtained experimentally by measuring the read voltage after writing a single magnetic step or transition. In this sense, $h(t)$ is called the transition response or step response. The differentiation occurs under the assumption of an inductive read element, but the modeling results will essentially be the same with the shielded magneto-resistive (MR) read head. Recording systems using thin-film media also exhibit a highly localized, non-additive type of noise. We shall discuss this type of noise shortly.

Let us for the time being assume that the NRZ recording format is used. Then, the input current waveform, $w(t)$, is generated by multiplying each bipolar input b_k taking the value +1 or −1 by a rectangular pulse of duration T and height 1/2:

$$w(t) = \sum_k b_k p(t - kT) \tag{101.3}$$

Combining Eqs. (101.2) and (101.3) and realizing that

$$\frac{dw}{dt} = \sum_k b_k \cdot \frac{\delta(t - kT) - \delta(t - T - kT)}{2} \tag{101.4}$$

we can express the readback waveform as

$$z(t) = \sum_k b_k \cdot \frac{h(t - kT) - h(t - T - kT)}{2} + n(t) \tag{101.5}$$

Comparing Eqs. (101.1) and (101.5), we see that the effective impulse response of the magnetic recording channel is $[h(t) - h(t - T)]/2$. This is usually called the pulse response since it is the head/medium response to a rectangular current pulse. Note that the response $h(t) - h(t - T)$ is traditionally called the dibit response in magnetic recording community since it is the response to two consecutive magnetic transitions or two nonzero NRZI bits. Equation (101.5) can be alternatively written as

$$z(t) = \sum_k a_k h(t - kT) + n(t) \tag{101.6}$$

where

$$a_k = (b_k - b_{k-1})/2 \tag{101.7}$$

The symbol a_k takes values 1, 0, and −1, which represent a positive transition, no transition, and a negative transition, respectively. Note that the polarity must alternate between any two successive nonzero a_k symbols.

The step response is often analytically modeled as the Lorentzian function:

$$h(t) = \sqrt{\frac{4E_t}{\pi PW50}} \cdot \frac{1}{1 + (2t/PW50)^2} \tag{101.8}$$

where E_t is the energy of a transition, i.e., $\int |h(t)|^2 dt = E_t$, and PW50 is the width of the pulse at half height. Figure 101.3 shows a more realistic step response arising from micromagnetic modeling of a thin-film disk and a magneto-resistive (MR) head.

Equation (101.5) can be used to describe the input/output relationship of a recording system based on the NRZI format after replacing b_k with the precoded sequence b'_k. The sequence b'_k is obtained by passing b_k through a precoder with transfer function $1/(1 \oplus D)$, where D represents a delay and $\oplus$ is the modulo-2 addition. Therefore, when the NRZI format is assumed (as is done in many modulation code designs), the detector can still be designed and evaluated based on the PAM-like channel description of Eq. (101.5). In this case, however, the decision is made on the precoded sequence b'_k, and the original bit sequence b_k is recovered through an inverse precoding operation: $b_k = b'_k \oplus b'_{k-1}$. The modulo-2 addition is applied, of course, to (1,0) binary representations of b_k and b'_k.

Media Noise

Today's hard disk drive recording systems invariably employ thin-metallic media. In such systems, the step response may change from one transition to the next due to random variations in the geometry of magnetic transitions. This random deviation of the step pulse gives rise to what is known as transition noise. This type of noise depends on the written data pattern and, strictly speaking, cannot be modeled as additive noise. A simple model for transition noise is obtained by introducing random variations to both width and position parameters of the step response. Let $h(t,w)$ denote the readback response to a noise-free transition located at $t = 0$, where w is a parameter that represents the physical width of the written transition. While it is clear that w is a certain fraction of the parameter PW50, the exact relationship between the two parameters depends on physical characteristics of the particular head/disk combination. Now, let $h_k(t,w)$ denote the transition response for the kth symbol interval. The subscript k emphasizes the underlying assumption that the step response is in general different from transition to transition. Then, the read response to a noisy transition in the kth symbol interval can be written as

$$h_k(t, w) = h(t - kT + j_k, w + w_k) \tag{101.9}$$

where j_k and w_k are random parameters representing deviations in the position and width, respectively, from the nominal values. Taking an nth order Taylor series expansion, the above expression can be approximated as a linear sum of the noise-free response and residual responses due to deviations around the nominal position and width of the pulse. Figure 101.4 shows the second order model in both continuous time and discrete time, assuming only the position jitter component. The width variation components can be included in a similar fashion. The model in Fig. 101.4 also includes additive noise. As for the accuracy of series expansion relative to Eq. (101.9), simulation results show that the first order model is sufficient in evaluating PRML techniques. However, to analyze detection techniques optimized for medium noise, the second order components must be included [Moon and Park, 2001]. Note that the matched filter, although not optimum in the presence of medium noise, does provide virtually sufficient statistics since at reasonably high densities both the signal and medium noise are effectively bandlimited. The use of matched filter also yields nice closed-form expressions for discrete-time channel representation based on the widely used Lorentzian step response.

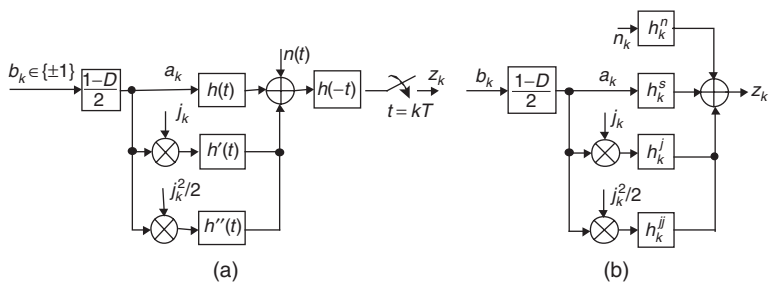


FIGURE 101.4 Channel model with second order position jitter. (a) Continuous time and (b) discrete-time.

Nonlinearity

The above channel model assumes that the signal portion of the received waveform can be constructed from linear superposition of isolated step responses. However, as density increases, nearby transitions start to interact, resulting in significant nonlinear distortion. An important source of nonlinearity is the shift in transition positions, which occurs as the demagnetizing field of the previous transition influences the head field writing the current transition. Another type of nonlinearity is the broadening of the current transition as the head field gradient is reduced by the demagnetizing field from a previous transition. Thus, a transition is shifted earlier in time and tends to broaden due to the effect of the transition written in the previous symbol interval. Yet another form of nonlinearity, which becomes significant as transitions approach each other even more closely, corresponds to partial erasure of adjacent transitions. In the read voltage, this appears as a sudden reduction of the signal amplitude. Unlike transition noise, nonlinearity is a repeatable phenomenon. The nonlinear distortion can be characterized using the Volterra functional series. Other nonlinearity characterization methods exist which rely more heavily on physical understanding of the write process. While different signal processing techniques have been developed which are geared specifically toward nonlinearity, they are mostly exploratory at present. In today's recording systems, nonlinear precompensation is aimed only at mitigating pattern-dependent transition position shifts.

101.3 SNR Definition and Code Rate Penalty

Care must be taken in defining the signal to noise ratio (SNR) that can provide a basis for comparing various coding and modulation strategies that operate at different symbol rates or densities. The E_b/N_o ratio, which is universally used in power-constraint communication channels, cannot be employed here. Part of the reason is that saturation recording is not a power-constrained (in the sense of neither average nor peak) channel. Furthermore, finding the right SNR definition is particularly confusing due to the fact that the statistical properties of medium noise depend on the written pattern as well as the symbol density.

SNR Definition in AWGN

First consider only additive noise. In communication channels with additive white Gaussian noise (AWGN) and an average transmitter power constraint, the E_b/N_o ratio, the information bit energy to noise spectral height ratio, can be used to compare coding/modulation schemes based on different symbol rates. Assuming the transmitted symbol is binary, the symbol error probability is mainly determined by the ratio

$$\beta_{com} = \frac{E_s}{N_o} = \frac{rE_b}{N_o} = \frac{P_s}{N_o/T} \tag{101.10}$$

where E_s is the symbol energy, r is the code rate, and P_s denotes the average transmitter power. This ratio is sometimes referred to as the detection SNR. For a given E_b/N_o ratio, the detection SNR is proportional to the rate of the code r.

The magnetic recording channel should be viewed as the $h(t)$-constrained channel, where $h(t)$ is the transition response discussed above; there is no average or peak power constraint here, but for a given head/medium interface, $h(t)$ is fixed and coding and signal processing engineers must live with it. At low to medium recording densities (where single errors are more likely than multiple errors), the detection SNR of an optimal sequence detector is given by

$$\beta_{mag} = \frac{\int_{-\infty}^{\infty}[h(t) - h(t-T)]^2 dt}{4N_o} \tag{101.11}$$

For the Lorentzian channel, this reduces to

$$\beta_{mag} = \frac{1}{D_s^2 + 1} \cdot \frac{E_t}{N_o} = \frac{r^2}{D_u^2 + r^2} \cdot \frac{E_t}{N_o} \tag{101.12}$$

where D_s denotes the symbol density defined as $PW50/T$ and D_u is the user density given by $r \cdot D_s$. Equation (101.12) suggests that E_t/N_o, the isolated transition pulse energy to noise spectral height ratio, play the E_b/N_o-like role for magnetic recording. It can be seen that for a given E_b/N_o, β_{mag} is roughly proportional to r^2 (assuming $D_u^2 \gg r^2$), indicating that the code rate penalty is more severe with magnetic recording than with the average-power-constrained communication channel. At high densities where multiple errors dominate, β_{mag} drops even faster than r^2.

SNR Definition in the Presence of Transition Noise

The E_t/N_0 ratio, however, does not provide a meaningful SNR definition when noise includes a signal-dependent component. The following SNR definition can handle mixed noise channels:

$$SNR = \frac{E_t}{N_0 + M_0} = \frac{E_t}{N_\alpha} \tag{101.13}$$

where M_0 is twice the average energy of the medium noise associated with each transition, $N_\alpha = N_0 + M_0$, and α denotes the medium noise power expressed as a percentage of the total in-band noise power, i.e.,

$$\alpha = \frac{M_0}{N_0 + M_0} \times 100 \tag{101.14}$$

We can further write $N_0 = [(100 - \alpha)/100]N_\alpha$ and $M_0 = (\alpha/100)N_\alpha$ with $0 \le \alpha \le 100$. With $\alpha = 0$, the SNR reduces to E_t/N_0. Letting $n_m(t)$ denote the medium noise voltage waveform associated with a given transition, the average medium noise energy per transition can be defined as

$$M_0/2 = \int_{-\infty}^{\infty} \overline{n_m^2(t)}\,dt \tag{101.15}$$

where the statistical average is taken over transitions written at different positions in the disk. It is safe to assume that the right hand side of Eq. (101.15) is finite. If the medium noise is statistically independent among transitions (as in the case of disks exhibiting a linear increase of noise power vs. linear density), then the total medium noise power arising from the all-transitions (or 1T) pattern is easily shown to be $\int_{-\infty}^{\infty} \overline{n_m^2(t)}\,dt/T$. This means that

$$\frac{M_0}{2T} = (\text{total medium noise power in 1T pattern}) \tag{101.16}$$

and M_0 can be viewed as the equivalent (single-sided) band-limited white noise spectrum of the medium noise for this pattern.

For the first-order position jitter and width variation (1PW) noise model, M_0 can be written as

$$M_0 = 2\sigma_j^2 I_t + 2\sigma_w^2 I_w \tag{101.17}$$

where σ_j^2 and σ_w^2 are the variances of position jitter and width variation, respectively, and

$$I_t = \int_{-\infty}^{\infty} [\partial h/\partial t]^2 dt \tag{101.18}$$

$$I_w = \int_{-\infty}^{\infty} [\partial h/\partial w]^2 dt \tag{101.19}$$

With the above definitions, specifying E_t/N_α with a particular value of α determines both N_o and M_0 [assuming $h(t, w)$ is given as well]. Once M_0 is determined, the noise parameters necessary for BER simulation or analysis are specified. For example, with the 1PW model, both σ_j^2 and σ_w^2 are specified once M_0 is known, provided the ratio σ_j^2/σ_w^2 is also known. Once N_o, σ_j^2, and σ_w^2 as well as the symbol period are specified, one can proceed with performance analysis or BER simulations to compare codes/detectors operating at different symbol or user densities. Note that the key assumption here is that M_0 is independent of the symbol period T so that the noise composition parameter α is also independent of T. This means that the E_t/N_α definition is free of symbol density, as desired. This assumption will hold for any recording systems where the total integrated medium noise power increases linearly with $1/T$. For the Lorentzian channel and assuming $w = \text{PW50}/2$, it can be shown that

$$I_t = I_w = \frac{2E_t}{PW50^2} \tag{101.20}$$

As a point of reference, to meet a 10^{-5} target BER for the 1PW Lorentzian channel at user density 2.5, EPR4ML with a rate 16/17 RLL code requires an E_t/N_o slightly above 24 dB and an E_t/N_{90} around 23 dB with medium noise comprised only of independent position jitter.

101.4 Coding and Detection

The magnetic recording channel can be characterized as a severe ISI channel. It has been suggested that for ISI channels, the capacity can be approached through coding with the relatively simple decision feedback equalizer (DFE) replacing complex sequence detectors and utilizing transmitter precoding to eliminate error. However, transmitter precoding is applicable only to multi-level input channels and cannot be used for modern magnetic storage channels, which are binary-input channels. At present, it seems that the only way to utilize a considerable fraction of the channel capacity available from the given heads and media technologies is to rely on some type of sequence detection.

A well-known example of the application of a sequence detector to magnetic storage is the PRML technique, which combines Viterbi detection with partial response linear equalization [Fisher, 1996]. The basic idea of the PRML method is illustrated in Fig. 101.5. The sampled read signal is applied to a linear equalizer, which forces the transfer function of the combined channel filter and equalizer to a $(1 - \text{D}) \times (1 + \text{D})^n$ polynomial, where D represents a symbol delay and n is a positive integer [Thapar and Patel, 1987]. The equalizer output signal is then fed to the Viterbi detector tuned to the ISI coefficients corresponding to the $(1 - \text{D})(1 + \text{D})^n$ target function. While offering a reasonably good match to the natural channel response, this particular family of target functions yields a smaller number of distinct signal levels at the equalizer output than more general target functions with the same ISI length. This, combined with the feature that the signal levels are integer-valued, significantly reduces the number of

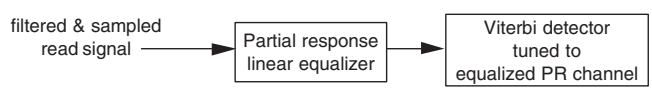


FIGURE 101.5 Partial response maximum likelihood (PRML) detection.

multiplies required in the implementation of the Viterbi detector. In the absence of noise enhancement and noise correlation, the Viterbi detector will perform maximum likelihood detection. But, there is an obvious loss of optimality associated with the PR equalizer. A close match between the target function and the natural channel will make this loss small. For this reason, more general PR targets that allow fractional coefficients are also being considered. Other types of PR targets with integer-coefficients are possible, such as $(1 - D)(1 + D)(2 + 2D + D^2)$ considered in Conway [1998]. Equalization is typically done using a finite impulse response (FIR) digital filter, but it can also be done in continuous time before sampling occurs. In disk drives, equalizers are typically trained at the factory floor and the trained taps are stored in shift registers that reside within the read channel IC. In disk recording, as the disk spins at a constant angular velocity, the linear velocity of the head relative to the data track is reduced as the head travels from the inner tracks to the outer ones. Despite this, storage density is kept more or less constant across the entire surface of the platter by forcing the data rate to increase at outer tracks. This technology is called zone bit recording (ZBR) and is crucial in achieving the maximum possible storage density. For practical reasons, however, the data rate is not allowed to change in every track; rather, entire tracks are divided into a number of bands and the data rate changes only from one band of tracks to another. Thus, in practice, the equalizer taps are trained and stored for each band separately. At the time of data retrieval, as the head travels to the target track, the right set of taps is loaded from the shift register.

The State-of-the Art Detector: EPR4ML Combined with Parity Code

Although variations exist, it is fair to say that the current industry standard in the read channel architecture is based on extended class IV (EPR4) equalization followed by the Viterbi algorithm (dubbed EPR4ML), incorporating a single bit parity code on top of the usual $(0,k)$ runlength-limited (RLL) code. The $(0,k)$ constraint limits the runs of consecutive NRZI 0s to k to ensure frequent non-zero signal levels at the read point to aid gain control and timing recovery [Siegel and Wolf, 1991]. EPR4 equalization is based on the $(1 - D)(1 + D)^2$ target polynomial. The basic RLL code is a block code that converts one or two bytes of user data (as well as the RS parity bits) to 9 or 17 bit-long codewords, respectively. In some cases, uncoded user blocks are interspersed among the RLL codewords to effectively raise the code rate at the expense of an increased k parameter. In addition to allowing a lower complexity encoder/decoder for a given rate, this strategy mitigates error propagation associated with long codewords. A single error event typically contaminates the whole code block after modulation decoding, but it may also corrupt two consecutive blocks if it occurs on the codeword boundary. As the block gets large, error propagation will undermine the error correction capability of the outer RS code. Once the RLL code is determined, a single parity check (SPC) bit is appended to each RLL codeword. For example, a rate 32/34 RLL/SPC code can be constructed by first attaching two uncoded user bytes to a 16/17 $(0,k)$ RLL code to effectively form a rate 32/33 $(0,k + 16)$ RLL code and then appending a single parity bit to the resulting RLL codeword.

The single bit parity is able to detect isolated occurrences of all error events consisting of an odd number of NRZ bit errors, such as $\pm\{2\}, \pm\{2, -2, 2\}, \pm\{2, -2, 2, -2, 2\}$, and so on. These types of error events indeed represent the errors frequently observed in high density recording. Two different methods may be used to handle detected errors. One way is to add a post processor that attempts to correct errors whenever the error flag is raised [Conway, 1998]. The attempt for error correction is based on keeping a list of likely error patterns and their associated metrics for all positions within the parity code block. When the parity is violated at the end of a given block, the post processor identifies the position and the actual pattern of the most likely error event from the stored list and attempt to correct the error accordingly. The metric or the likelihood of a candidate error pattern can easily be obtained by correlating the corresponding noiseless signal pattern with the noisy received signal [Conway, 1998]. While the required hardware complexity is relatively small with this approach, associated latency may be an issue.

Another approach to error correction for the single parity code is to directly incorporate the parity constraint in the Viterbi trellis. In the conventional EPR4 trellis, the state at time k is given by three most recent input bits $\{b_{k-1}, b_{k-2}, b_{k-3}\}$. With binary inputs, this gives rise to an 8-state trellis. With the single parity constraint, however, the state is completely specified by $\{b_{k-1}, b_{k-2}, b_{k-3}; S\}$, where S is the modulo-2

running sum of input bits since the beginning of the block. At the end of each block, S is forced to take on either 1 or 0, depending on the type of parity imposed. The result is a time-varying trellis that starts with 8-states at the beginning of each parity code block, grows to 16 states, and then reduces back to 8 states at the end of the block. While the difference may be small, this approach yields a better performance than the post processor technique since the probability of miscorrection is smaller. Although the actual gain varies depending on the channel properties including operating linear density and noise characteristics, the single parity code typically provides a dB or so SNR gain.

Attempts have been made to use more than one parity bit, but the additional SNR gain observed at the detector output tends to disappear at the output of the RS decoder. The reason might be the increased error propagation associated with multiple parity bits [Feng et al., 2000].

Post ECC Performance Metric

It is important to realize that the performance measure based on detector symbol error rates is not the ideal metric for the read channel reliability. A more meaningful performance metric is the sector retry rate (SRR), the probability that the RS code fails to correct all the errors in a sector before attempting any reread. When the byte errors observed at the detector output are independent, the translation from a BER to an SRR is straightforward using the binomial probability distribution for the byte errors. If the byte error statistics are the same for all code/detector strategies, the BER can be used to rank different inner coding/detection strategies since the SNR gain of one scheme over another at a given BER remains more or less intact at the corresponding SRR level. However, in reality certain read channel strategies tend to produce more burst errors than others. Since the RS code performance degrades more rapidly with burst errors, a BER level gain does not necessarily translate to an SRR level gain. This is one of the reasons why feedback-based detectors such as the DFE [Tyner and Proakis, 1993] and fixed delay tree search with decision feedback (FDTS/DF) [Moon and Carley, 1990] have not found much success in commercial applications so far. These techniques are prone to error propagation, although they offer significant SNR advantages over PRML techniques at a given BER. While feedback-based detectors are not well suited to the ECC parameters of today's hard disk systems, they can provide meaningful performance gains over PRML techniques with more aggressive ECC parameters.

Distance-Enhancing Codes

It is easy to see that the single bit parity code itself does not have any error correction capability. One way of ensuring that given types of error events never occur is to apply a constrained code. One example is the maximum transition run (MTR) code [Moon and Brickner, 1996]. The basic idea of the MTR code is to eliminate critical bit patterns that cause most errors in sequence detectors. More specifically, the MTR code eliminates input patterns that contain j or more consecutive transitions in the corresponding current waveform. As a result, for $j = 2$, the NRZ error events of the form $\pm\{2, -2, 2\}$ (and any error event that contains this pattern) are eliminated. This substantially improves the performance of a broad class of near-optimal sequence detectors at high linear densities. The code can be viewed as having a distance-enhancing property for high linear density recording. In addition to enhancing distances, the MTR codes are useful in controlling nonlinearities and other write-side anomalies associated with crowded transitions. As a simple example of an MTR code with $j = 2$, 16 codewords required to implement the rate 4/5 block code are given in Table 101.1. The NRZI format is assumed in mapping the codewords to the magnetization waveform. It can be seen that consecutive runs of 1s are limited to 2 both within

TABLE 101.1 Codewords for a Rate 4/5 MTR Block Code with $k = 8$

00001	00110	01100	10010
00010	01000	01101	10100
00100	01001	10000	10101
00101	01010	10001	10110

the codewords and at the codeword boundaries. The runs of consecutive 0s are also limited (to 8 in this example) for timing recovery purposes. Bit error rate (BER) simulation results indicate that the MTR code can yield significant performance improvement over the uncoded case when used with high-order PRML and FDTS detectors. The rate 6/7 MTR $j = 2$ code has been constructed using a state-dependent encoding strategy [Brickner and Moon, 1997]. A higher-rate location-dependent MTR code has also been designed which allows three or more consecutive transitions starting at every other bit position [Fitzpatrick, 1997].

There also exist other types of code constraints that posses similar distance-enhancing properties for high-order PRML and other ISI channels. For the interested reader, Karabed et al. [1999] provides a comprehensive treatment on general distance-enhancing codes for severe ISI channels.

101.5 Glimpse of the Future

In order for the data storage industry to sustain the growth rate trend in density, it is important that the read channel provides substantial SNR gains in future storage systems. Two techniques that will potentially yield large SNR gains are low density parity check (LDPC) codes and pattern-dependent noise prediction.

Low Density Parity Check Codes

As in a wide variety of communication channels, there is considerable interest in the magnetic recording community in employing turbo [Berrou et al., 1993; Ryan et al., 1998] or LDPC codes [Gallager, 1962; Fan et al., 1999] to data storage devices. This interest is driven by the promise of very large SNR gain. Unlike turbo codes, LDPC codes do not seem to exhibit error floors and, thus, may offer a more attractive solution to storage systems where the error rate requirements are very low. Also, the overall latency in the decoding operation is considerably smaller in LDPC codes.

Figure 101.6a shows the BER simulation results of uncoded EPR4ML vs. LDPC coded EPR4 at a user density of 2.5, assuming the noise is 100% AWGN. The Lorentzian step response is assumed. The rate of the code is 8/9, and the frame length is about 5000 bits. The encoder is an irregular LDPC code with column weight four and an overlap constraint that removes the length-4 cycles. The soft channel information is generated using the forward-backward trellis search algorithm. The decoder is based on the belief-propagation/message-passing algorithm. There is no iteration between the channel detector and the LDPC decoder. Figure 101.6b shows similar curves except that the medium noise is 90% of the total noise [in the sense of Eq. (101.14)]. It can be seen that the relative gain of the LDPC code is larger with medium noise than with additive noise. Results on turbo code suggest similar gains.

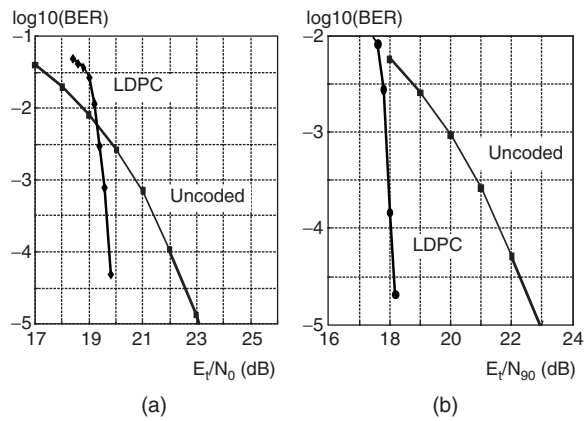


FIGURE 101.6 BER vs. SNR for (a) 100% AWGN and (b) 10% AWGN/90% medium noise (all position jitter).

Both fundamental and practical issues must be addressed to overcome major obstacles in employing these types of codes in real storage systems. With turbo or LDPC coding, a large coding gain is achieved only when the block size becomes comparable to the present sector size of 512 K bytes. In disk drives, the sector failure rate is more important than the BER. This means that the block error rate of turbo or LDPC codes must be low enough to meet the demanding reliability requirement of storage systems without further error correction. Therefore, the error flooring effects that are common in turbo coded systems are simply not acceptable here. As computer simulation is not feasible at the low target error rates of typical disk drives, the performance must be reasonably ensured using error bound approaches. In light of recent developments in algebraic LDPC codes with known minimum distance properties [Kou et al., 2000], it might be possible to ensure a desired level of error probability through analytical error bounds, although more work is needed in this direction. Another issue is the implementation complexity. With the stringent power and speed constraint, the complexity is a major obstacle to the deployment of these codes in commercial storage devices. Efforts must be directed to both decoder and encoder implementation. The encoder complexity, which is a major issue in the implementation of Gallager's original LDPC code, seems to be no longer a problem with, for example, the cyclic encoder of Kou et al. [2000]. As for decoding, while the implementation speed can be made extremely fast using the inherently parallel message-passing algorithm, the complexity is proportional to the frame length and can be excessively high. Variations of the turbo/LDPC code that allow both practical encoding and decoding may also be possible. One example is a turbo product code based on the single parity constituent codes. This code can be viewed as an LDPC code whose generator matrix is also sparse, allowing flexible decoder architecture options ranging from a serial forward-backward operation to a massively parallel message-passing algorithm [Oenning and Moon, 2001].

Another remaining question is robustness. Not all recording channel impairments are well characterized. Thermal asperity and offtrack interference are just two examples of channel disturbances that can lead to totally unexpected results on read channel performance testing. It is important that the code does not fail in the presence of some unforeseen disturbances. Performance verification and sensitivity studies are needed in the presence of thermal asperity, offtrack, and nonlinearity.

Pattern-Dependent Noise Prediction

Unless the code is extremely powerful as in turbo and LDPC coded systems, it is difficult to overcome the severe rate penalty of the magnetic channel. In light of this observation, one effective strategy to achieve an SNR gain is to improve the detector capability. Especially when medium noise dominates, there is a significant room for improvement in detection performance. A viable strategy along this direction is to incorporate pattern-dependent noise prediction into the Viterbi algorithm [Caroselli et al., 1997; Kavcic and Moura, 2000; Moon and Park, 2001]. This approach is based on the observation that the noise power changes from one branch (state-transition) to the next in the Viterbi trellis because of the pattern-dependent nature of medium noise. It can be shown that if the noise can be predicted on each branch separately based on predetermined local noise statistics, the overall immunity against medium noise can be improved significantly. This method is particularly promising in the design of future high density storage systems where medium noise is likely to be a major limiting factor.

101.6 Other Considerations

Although not discussed here, servo signal processing is important to achieve accurate positioning of the head over the given track. As track density increases, data written in adjacent tracks may interfere with the processing of the main track signal. This type of interference, called offtrack interference, may become a major source of errors in future storage systems. Gain control and timing recovery loops are also critical components in the overall read channel design. In storage systems that employ flexible media such as tape, signal fades due to medium asperity and mechanical fluctuation of head-tape spacing can be a

major cause of errors. For these channels, requirements on equalization and ECC can be considerably different from the hard disk channel.

101.7 Summary

A brief description is given on the characteristics of the magnetic recording channel including signal-dependent transition noise and nolinearity. Some of the coding and detection techniques of today's magnetic storage systems are reviewed. Future possibilities for the read channel strategy also have been discussed.

References

Berrou, C., Glavieux, A., and Thitimajshima, P. 1993. Near Shannon limit error-correcting coding and decoding: turbo-codes. *Proc. IEEE Intl. Conf. Commun.,* Geneva, May, 1064.

Brickner, B. and Moon, J. 1997. Design of a rate 6/7 maximum transition run code. *IEEE Trans. Magnet.,* 33(5).

Caroselli, J., Altekar, S. A., McEwen, P., and Wolf, J. K. 1997. Improved detection for magnetic recording systems with media noise. *IEEE Trans. Magnet.,* 33(5), 2779.

Conway, T. 1998. A new target response with parity coding for high density magnetic recording channels. *IEEE Trans. Magnet.,* 34(4), 2382.

Fan, J. L., Kurtas, E., Friedman, A., and McLaughlin, S. 1999. Low-density parity-check codes for magnetic storage. *Allerton Conference,* Sept.

Feng, W., Vityaev, A., Burd, G., and Nazari, N. 2000. On the performance of parity codes in magnetic recording systems. Proc. IEEE Global. Tel. Comm. Conf.

Fisher, K. D., Abbott, W. L., Sonntag, J. L., and Nelson, R. 1996. PRML detection boosts hard-disk drive capacity. *IEEE Spectrum,* 70.

Fitzpatrick, K. K. 1997. Time-varying MTR codes for high density magnetic recording. Proc. IEEE Global. Tel. Comm. Conf.

Gallager, R. G. 1962. Low-density parity check codes. *IRE Trans. Inform. Theory,* IT-8, 21.

Karabed, R., Siegel, P. H., and Soljanin, E. 1999. Constrained coding for binary channels with high intersymbol interference. *IEEE Trans. Inform. Theory,* 45(6), 1777.

Kavcic, A. and Moura, J. M. F. 2000. The Viterbi algorithm and Markov noise memory. *IEEE Trans. Inform. Theory,* 46, 291.

Kou, Y., Lin, S., and Fossorier, M. 2000. Low-density parity-check codesbased on finite geometries: a rediscovery. Proc. IEEE Int. Symp. on Inform. Theory.

Moon, J. and Brickner, B. 1996. Maximum transition run codes for data storage systems. *IEEE Trans. Magnet.,* 32(5).

Moon, J. and Carley, L. R. 1990. Performance comparison of detection methods in magnetic recording. *IEEE Trans. Magnet.,* 26(6), 3155.

Moon, J. and Park, J. 2001. Near-maximum likelihood sequence detection in pattern dependent noise. *IEEE J. Sel. Areas Commun.,* to appear.

Oenning, T. and Moon, J. 2001. A low density generator matrix interpretation of parallel concatenated single bit parity codes. *IEEE Trans. Magnet.*

Ryan, W., McPheters, L., and McLaughlin, S. 1998. Combined turbo coding and turbo equalization for PR4-equalized Lorentzian channels. Proc. Conf. on Information Systems and Sciences.

Siegel, P. H. and Wolf, J. K. 1991. Modulation and coding for information storage. *IEEE Commun. Mag.,* 29(12), 68.

Thapar, H. and Patel, A. M. 1987. A class of partial response systems for increasing storage density in magnetic recording. *IEEE Trans. Magnet.,* 23(5).

Tyner, D. J. and Proakis, J. G. 1993. Partial response equalizer performance in digital magnetic recording channels. *IEEE Trans. Magnet.,* 29(6), 4194.

Further Information

For a comprehensive treatment on constrained codes used for data storage channels, see *Codes for Mass Data Storage Systems*, K. A. S. Immink, Shannon Foundation Publishers, 1999. For an overview of magnetic recording technology including computer tape and disk recording as well as consumer video and audio recording, consult *Magnetic Recording Technology* and *Magnetic Storage Handbook*, Mee and Daniel, Second Edition, McGraw-Hill, 1996. For more detailed information on some coding and signal processing ideas for data storage, see the January 1992 issue of the *IEEE Journal on Selected Areas in Communications*. A new *JSAC* issue devoted to the same topic will appear sometime early in the year 2001. Also, each year, the Global Telecommunications Conference and the International Conference on Communications run sessions dedicated to data storage. In addition, either the September or the November issue of the IEEE *Transactions on Magnetics* includes papers presented at the annual International Conference on Magnetics and contains a significant number of papers on communications and signal processing issues of data storage technology. Finally, every third year, the Magnetic Recording Conference is devoted to systems and signal processing issues for storage. The last TMRC dedicated to signal processing was in the year 2000, and some of the papers presented there appeared in the January 2001 issue of the *IEEE Transactions on Magnetics*.

102

Magneto-Optical Disk Data Storage

Masud Mansuripur
University of Arizona

102.1 Introduction

Since the early 1940s, magnetic recording has been the mainstay of electronic information storage worldwide. Audio tapes provided the first major application for the storage of information on magnetic media. Magnetic tape has been used extensively in consumer products such as audio tapes and video cassette recorders (VCRs); it has also found application in backup/archival storage of computer files, satellite images, medical records, etc. Large volumetric capacity and low cost are the hallmarks of tape data storage, although sequential access to the recorded information is perhaps the main drawback of this technology. Magnetic hard disk drives have been used as mass storage devices in the computer industry ever since their inception in 1957. With an areal density that has doubled roughly every other year, hard disks have been and remain the medium of choice for secondary storage in computers.[1] Another magnetic data storage device, the floppy disk, has been successful in areas where compactness, removability, and fairly rapid access to the recorded information have been of prime concern. In addition to providing backup and safe storage, inexpensive floppies with their moderate capacities (2 megabyte on a 3.5-in-diam platter is typical nowadays) and reasonable transfer rates have provided the crucial function of file/data transfer between isolated machines. All in all, it has been a great half-century of progress and market

[1] At the time of this writing, achievable densities on hard disks are in the range of 10^7 b/cm^2. Random access to arbitrary blocks of data in these devices can take on the order of 10 ms, and individual read/write heads can transfer data at the rate of several megabits per second.

dominance for magnetic recording devices, which are only now beginning to face a potentially serious challenge from the technology of optical recording.

Like magnetic recording, a major application area for optical data storage systems is the secondary storage of information for computers and computerized systems. Like the high-end magnetic media, optical disks can provide recording densities in the range of 10^7 b/cm^2 and beyond. The added advantage of optical recording is that, like floppies, these disks can be removed from the drive and stored on the shelf. Thus, the functions of the hard disk (i.e., high capacity, high data transfer rate, rapid access) may be combined with those of the floppy (i.e., backup storage, removable media) in a single optical disk drive. Applications of optical recording are not confined to computer data storage. The enormously successful audio **compact disc (CD)**, which was introduced in 1983 and has since become the de facto standard of the music industry, is but one example of the tremendous potential of the optical technology.

A strength of optical recording is that, unlike its magnetic counterpart, it can support read-only, write-once, and erasable/rewritable modes of data storage. Consider, for example, the technology of optical audio/video disks. Here the information is recorded on a master disk, which is then used as a stamper to transfer the embossed patterns to a plastic substrate for rapid, accurate, and inexpensive reproduction. The same process is employed in the mass production of read-only files [CD read only memory (CD-ROM), and optical ROM (O-ROM)], which are now being used to distribute software, catalog, and other large data bases. Or consider the write-once read-many (WORM) technology, where one can permanently store massive amounts of information on a given medium and have rapid, random access to them afterwards. The optical drive can be designed to handle read only, WORM, and erasable media all in one unit, thus combining their useful features without sacrificing performance and ease of use or occupying too much space. What is more, the media can contain regions with prerecorded information as well as regions for read/write/erase operations, both on the same platter. These possibilities open new vistas and offer opportunities for applications that have heretofore been unthinkable; the interactive video-disk is perhaps a good example of such applications.

In this chapter, we lay out the conceptual basis for optical data storage systems; the emphasis will be on disk technology in general and magneto-optical disk in particular. The next section is devoted to a discussion of some elementary aspects of disk data storage including the concept of track and definition of the access time. Section 102.3 describes the basic elements of the optical path and its functions; included are the properties of the semiconductor laser diode, characteristics of the beam shaping optics, and certain features of the focusing **objective lens**. Because of the limited depth of focus of the objective and the eccentricity of tracks, optical disk systems must have a closed-loop feedback mechanism for maintaining the focused spot on the right track. These mechanisms are described in Sections 102.4 and 102.5 for automatic focusing and automatic track-following, respectively. The physical process of thermomagnetic recording in magneto-optic (MO) media is described next, followed by a discussion of the MO readout process in Section 102.7. The final section describes the properties of the MO media.

102.2 Preliminaries and Basic Definitions

A disk, whether magnetic or optical, consists of a number of **tracks** along which the information is recorded. These tracks may be concentric rings of a certain width, W_t, as shown in Fig. 102.1. Neighboring tracks may be separated from each other by a guard band whose width we shall denote by W_g. In the least sophisticated recording scheme imaginable, marks of length Δ_0 are recorded along these tracks. Now, if each mark can be in either one of two states, present or absent, it may be associated with a binary digit, 0 or 1. When the entire disk surface of radius R is covered with such marks, its capacity C_0 will be

$$C_0 = \frac{\pi R^2}{(W_t + W_g)\Delta_0} \text{ bits per surface} \tag{102.1}$$

Consider the parameter values typical of current optical disk technology: $R = 67$ mm corresponding to 5.25-in.-diam platters, $\Delta_0 = 0.5\ \mu$m, which is roughly determined by the wavelength of the read/write laser

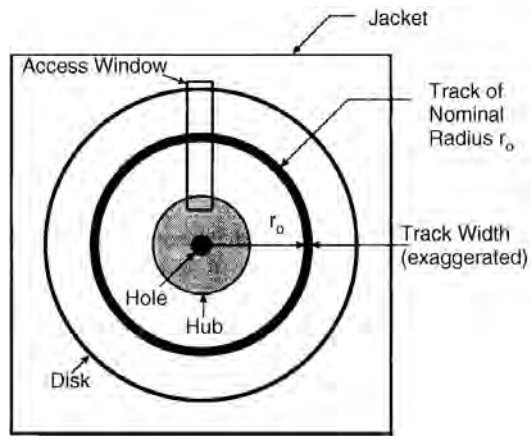


FIGURE 102.1 Physical appearance and general features of an optical disk. The read/write head gains access to the disk through a window in the jacket; the jacket itself is for protection purposes only. The hub is the mechanical interface with the drive for mounting and centering the disk on the spindle. The track shown at radius r_0 is of the concentric-ring type.

diodes, and $W_t + W_g = 1\ \mu$m for the track-pitch. The disk capacity will then be around 28×10^9 b, or 3.5 gigabytes. This is a reasonable estimate and one that is fairly close to reality, despite the many simplifying assumptions made in its derivation. In the following paragraphs, we examine some of these assumptions in more detail.

The disk was assumed to be fully covered with information-carrying marks. This is generally not the case in practice. Consider a disk rotating at N revolutions per second (r/s). For reasons to be clarified later, this rotational speed should remain constant during the disk operation. Let the electronic circuitry have a fixed clock duration T_c. Then only pulses of length T_c (or an integer multiple thereof) may be used for writing. Now, a mark written along a track of radius r, with a pulse-width equal to T_c, will have length ℓ, where

$$\ell = 2\pi \mathbf{N}_r T_c \tag{102.2}$$

Thus, for a given rotational speed N and a fixed clock cycle T_c, the minimum mark length ℓ is a linear function of track radius r, and ℓ decreases toward zero as r approaches zero. One must, therefore, pick a minimum usable track radius, $r_{\min}$, where the spatial extent of the recorded marks is always greater than the minimum allowed mark length, Δ_0. Equation (102.2) yields

$$r_{\min} = \frac{\Delta_0}{2\pi \mathbf{N} T_c} \tag{102.3}$$

One may also define a maximum usable track radius $r_{\max}$, although for present purposes $r_{\max} = R$ is a perfectly good choice. The region of the disk used for data storage is thus confined to the area between $r_{\min}$ and $r_{\max}$. The total number N of tracks in this region is given by

$$N = \frac{r_{\max} - r_{\min}}{W_t + W_g} \tag{102.4}$$

The number of marks on any given track in this scheme is independent of the track radius; in fact, the number is the same for all tracks, since the period of revolution of the disk and the clock cycle uniquely determine the total number of marks on any individual track. Multiplying the number of usable tracks

N with the capacity per track, we obtain for the usable disk capacity

$$C = \frac{N}{\mathbf{N}T_c} \tag{102.5}$$

Replacing for N from Eq. (102.4) and for NT_c from Eq. (102.3), we find

$$C = \frac{2\pi r_{\min}(r_{\max} - r_{\min})}{(W_t - W_g)\Delta_0} \tag{102.6}$$

If the capacity C in Eq. (102.6) is considered a function of $r_{\min}$ with the remaining parameters held constant, it is not difficult to show that maximum capacity is achieved when

$$r_{\min} = \frac{1}{2} r_{\max} \tag{102.7}$$

With this optimum $r_{\min}$, the value of C in Eq. (102.6) is only half that of C_0 in Eq. (102.1). In other words, the estimate of 3.5 gigabyte per side for 5.25-in disks seems to have been optimistic by a factor of two.

One scheme often proposed to enhance the capacity entails the use of multiple zones, where either the rotation speed N or the clock period T_c are allowed to vary from one zone to the next. In general, zoning schemes can reduce the minimum usable track radius below that given by Eq. (102.7). More importantly, however, they allow tracks with larger radii to store more data than tracks with smaller radii. The capacity of the zoned disk is somewhere between C of Eq. (102.6) and C_0 of Eq. (102.1), the exact value depending on the number of zones implemented.

A fraction of the disk surface area is usually reserved for **preformat** information and cannot be used for data storage. Also, prior to recording, additional bits are generally added to the data for **error-correction coding** and other housekeeping chores. These constitute a certain amount of overhead on the user data and must be allowed for in determining the capacity. A good rule of thumb is that overhead consumes approximately 20% of the raw capacity of an optical disk, although the exact number may vary among the systems in use. Substrate defects and film contaminants during the deposition process can create bad **sectors** on the disk. These are typically identified during the certification process and are marked for elimination from the sector directory. Needless to say, bad sectors must be discounted when evaluating the capacity.

Modulation codes may be used to enhance the capacity beyond what has been described so far. Modulation coding does not modify the minimum mark length of Δ_0 but frees the longer marks from the constraint of being integer multiples of Δ_0. The use of this type of code results in more efficient data storage and an effective number of bits per Δ_0 that are greater than unity. For example, the popular (2, 7) modulation code has an effective bit density of 1.5 b per Δ_0. This or any other modulation code can increase the disk capacity beyond the estimate of Eq. (102.6).

The Concept of Track

The information on magnetic and optical disks is recorded along tracks. Typically, a track is a narrow annulus at some distance r from the disk center. The width of the annulus is denoted by W_t, whereas the width of the guard band, if any, between adjacent tracks is denoted by W_g. The track pitch is the center-to-center distance between neighboring tracks and is therefore equal to $W_t + W_g$. A major difference between the magnetic floppy disk, the magnetic hard disk, and the optical disk is that their respective track-pitches are presently of the order of 100, 10, and 1 μm. Tracks may be fictitious entities, in the sense that no independent existence outside the pattern of recorded marks may be ascribed to them. This is the case, for example, with the compact audio disk format where prerecorded marks simply define their own

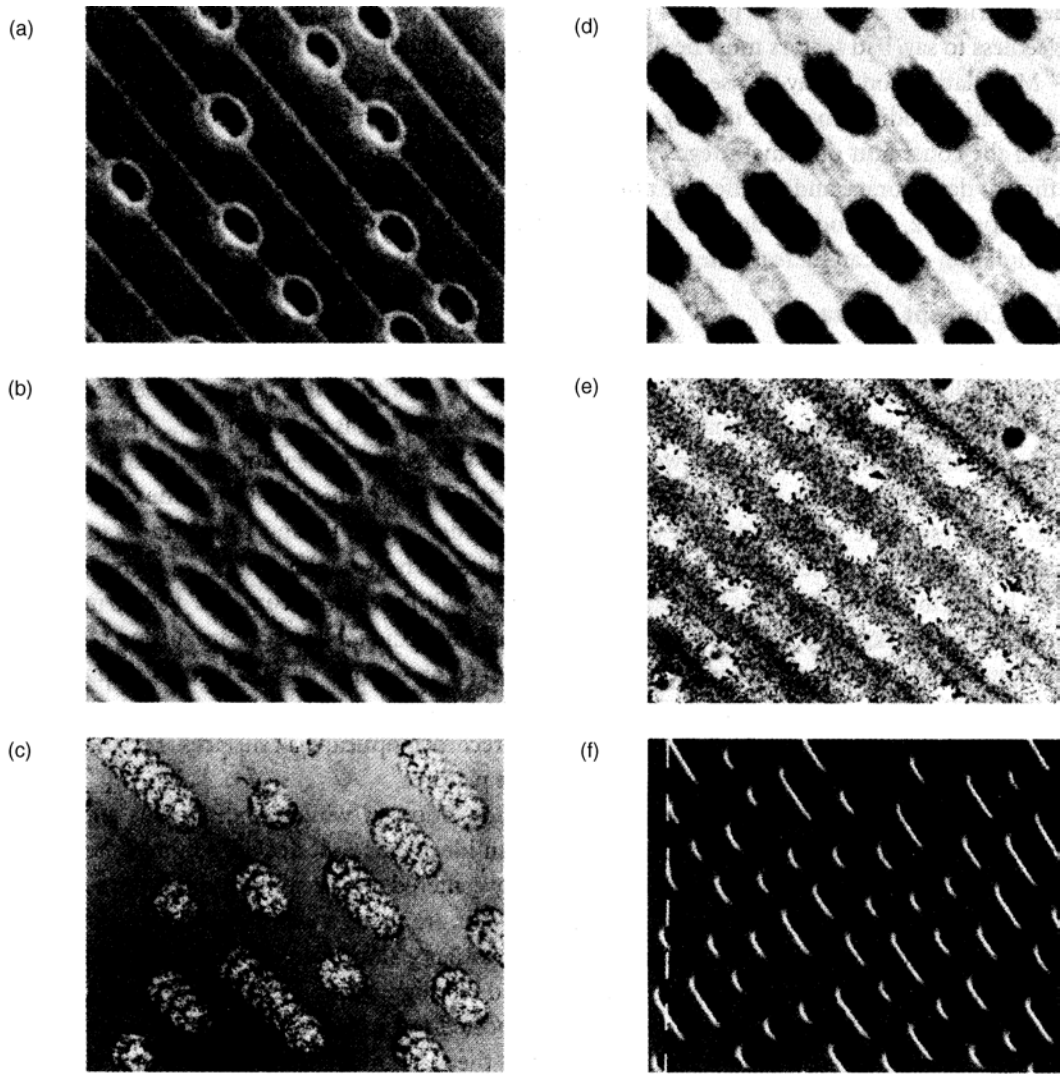


FIGURE 102.2 Micrographs of several types of optical storage media. The tracks are straight and narrow (track pitch = 1.6 μm), with an orientation angle of $\simeq -45°$. (a) ablative, write-once tellurium alloy; (b) ablative, write-once organic dye; (c) amorphous-to-crystalline, write-once phase-change alloy GaSb; (d) erasable, amorphous magneto-optic alloy GdTbFe; (e) erasable, crystalline-to-amorphous phase-change tellurium alloy; and (f) read-only CD-audio, injection-molded from polycarbonate with a nickel stamper. (*Source: Ullmann's Encylopedia of Industrial Chemistry*, 5th ed., Vol A14, VCH, Weinheim, 1989, 196. With permission.)

tracks and help guide the laser beam during readout. In the other extreme are tracks that are physically engraved on the disk surface before any data is ever recorded. Examples of this type of track are provided by pregrooved WORM and magneto-optical disks. Figure 102.2 shows micrographs from several recorded optical disk surfaces. The tracks along which the data are written are clearly visible in these pictures.

It is generally desired to keep the read/write head stationary while the disk spins and a given track is being read from or written onto. Thus, in an ideal situation, not only should the track be perfectly circular, but the disk must also be precisely centered on the spindle axis. In practical systems, however, tracks are neither precisely circular, nor are they concentric with the spindle axis. These eccentricity problems are solved in low-performance floppy drives by making tracks wide enough to provide tolerance for misregistrations and misalignments. Thus, the head moves blindly to a radius where the track center is nominally expected to be and stays put until the reading or writing is over. By making the head narrower than the track pitch, the track center is allowed to wobble around its nominal position without significantly

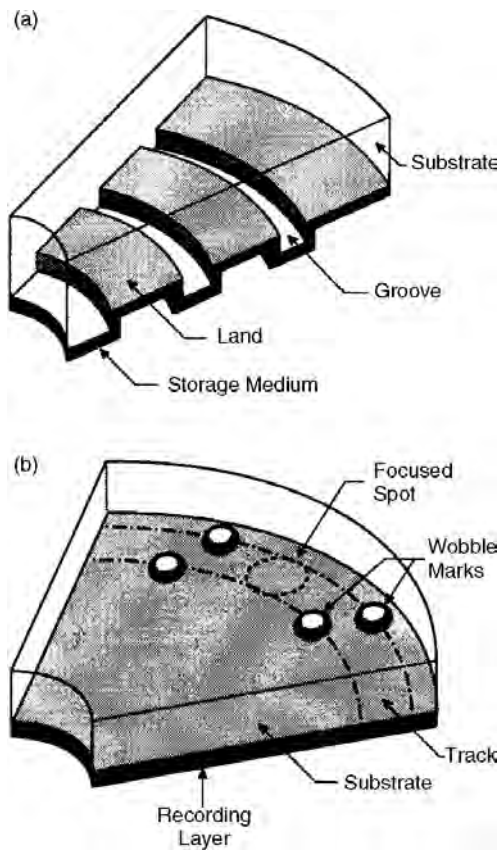


FIGURE 102.3 (a) Lands and grooves in an optical disk. The substrate is transparent, and the laser beam must pass through it before reaching the storage medium. (b) Sampled-servo marks in an optical disk. These marks, which are offset from the track-center, provide information regarding the position of focused spot.

degrading the performance during the read/write operation. This kind of wobble, however, is unacceptable in optical disk systems, which have a very narrow track, about the same size as the focused beam spot. In a typical situation arising in practice, the eccentricity of a given track may be as much as +50 μm while the track pitch is only about 1 μm, thus requiring active track-following procedures.

One method of defining tracks on an optical disk is by means of pregrooves that are either etched, stamped, or molded onto the substrate. In **grooved media of optical storage**, the space between neighboring grooves is the so-called land [see Fig. 102.3(a)]. Data may be written in the grooves with the land acting as a guard band. Alternatively, the land regions may be used for recording while the grooves separate adjacent tracks. The groove-depth is optimized for generating an optical signal sensitive to the radial position of the read/write laser beam. For the push–pull method of track error detection, the groove depth is in the neighborhood of $\lambda/8$, where λ is the wavelength of the laser beam.

In digital data storage applications, each track is divided into small segments or sectors intended for the storage of a single block of data (typically either 512 or 1024 bytes). The physical length of a sector is thus a few millimeters. Each sector is preceded by header information such as the identity of the sector, identity of the corresponding track, synchronization marks, etc. The header information may be preformatted onto the substrate, or it may be written on the storage layer prior to shipping the disk. Pregrooved tracks may be carved on the optical disk either as concentric rings or as a single continuous spiral. There are certain advantages to each format. A spiral track can contain a succession of sectors without interruption, whereas concentric rings may each end up with some empty space that is smaller than the required length for a sector. Also, large files may be written onto (and read from) spiral tracks without jumping to the next track, which occurs when concentric tracks are used. On the other hand, multiple

path operations, such as write-and-verify or erase-and-write, which require two paths each for a given sector or still-frame video are more conveniently handled on concentric-ring tracks.

Another track format used in practice is based on the sampled-servo concept. Here the tracks are identified by occasional marks placed permanently on the substrate at regular intervals, as shown in Fig. 102.3(b). Details of track-following by the sampled-servo scheme will follow shortly; suffice it to say at this point that servo marks help the system identify the position of the focused spot relative to the track-center. Once the position is determined it is fairly simple to steer the beam and adjust its position.

Disk Rotation Speed

When a disk rotates at a constant angular velocity ω, a track of radius r moves with the constant linear velocity $V = r\omega$. Ideally, one would like to have the same linear velocity for all of the tracks, but this is impractical except in a limited number of situations. For instance, when the desired mode of access to the various tracks is sequential, such as in audio and video disk applications, it is possible to place the head in the beginning at the inner radius and move outward from the center thereafter while continuously decreasing the angular velocity. By keeping the product of r and ω constant, one can thus achieve constant linear velocity for all of the tracks.* Sequential access mode, however, is the exception rather than the norm in data storage systems. In most applications, the tracks are accessed randomly with such rapidity that it becomes impossible to adjust the rotation speed for constant linear velocity. Under these circumstances, the angular velocity is best kept constant during the normal operation of the disk. Typical rotation speeds are 1200 and 1800 r/m for slower drives and 3600 r/m for the high-data-rate systems. Higher rotation rates (5000 r/m and beyond) are certainly feasible and will likely appear in future storage devices.

Access Time

The direct-access storage-device (DASD) used in computer systems for the mass storage of digital information is a disk drive capable of storing large quantities of data and accessing blocks of this data rapidly and in arbitrary order. In read/write operations, it is often necessary to move the head to new locations in search of sectors containing specific data-items. Such relocations are usually time consuming and can become the factor that limits performance in certain applications. The access time τ_a is defined as the average time spent in going from one randomly selected spot on the disk to another. Also, τ_a can be considered the sum of a seek time τ_s, which is the average time needed to acquire the target track, and a latency τ_l, which is the average time spent on the target track while waiting for the desired sector. Thus,

$$\tau_a = \tau_s + \tau_l \tag{102.8}$$

The latency is half the revolution period of the disk, since a randomly selected sector is, on the average, halfway along the track from the point where the head initially lands. Thus for a disk rotating at 1200 r/m, $\tau_l = 25$ ms, whereas at 3600 r/m, $\tau_l \simeq 8.3$ ms. The seek time, on the other hand, is independent of the rotation speed of the disk but is determined by the traveling distance of the head during an average seek, as well as by the mechanism of head actuation. It can be shown that the average length of travel in a random seek is one-third of the full stroke. (In our notation, the full stroke is $r_{max} - r_{min}$.) In magnetic disk drives where the head/actuator assembly is relatively light-weight (a typical Winchester head weighs about 5 g), the acceleration and deceleration periods are short, and seek times are typically around 10 ms in small format drives (i.e., 5.25 and 3.5 in). In optical disk systems, on the other hand, the head, being an assembly of discrete elements, is fairly large and heavy (typical weight $\simeq$ 100 g), resulting in values of τ_s that are several times greater than those obtained in magnetic recording systems. The seek times reported

*In compact audio disk players, the linear velocity is kept constant at 1.2 m/s. The starting position of the head is at the inner radius $r_{min} = 25$ mm, where the disk spins at 460 r/m. The spiral track ends at the outer radius $r_{max} = 58$ mm, where the disk's angular velocity is 200 r/m.

for commercially available optical drives presently range from 20 ms in high performance 3.5-in. drives to about 80 ms in larger (and slower) drives. We emphasize, however, that the optical disk technology is still in its infancy; with the passage of time, the integration and miniaturization of the elements within the optical head will surely produce lightweight devices capable of achieving seek times of the order of only a few milliseconds.

102.3 The Optical Path

The **optical path** begins at the light source which, in practically all laser disk systems in use today, is a semiconductor GaAs diode laser. Several unique features have made the laser diode indispensable in optical recording technology, not only for the readout of stored information but also for writing and erasure. The small size of this laser has made possible the construction of compact head assemblies, its coherence properties have enabled diffraction-limited focusing to extremely small spots, and its direct modulation capability has eliminated the need for external modulators. The laser beam is modulated by controlling the injection current; one applies pulses of variable duration to turn the laser on and off during the recording process. The pulse duration can be as short as a few nanoseconds, with rise and fall times typically less than 1 ns. Although readout can be accomplished at constant power level, that is, in continuous wave (CW) mode, it is customary for noise reduction purposes to modulate the laser at a high frequency (e.g., several hundred megahertz during readout).

Collimation and Beam-Shaping

Since the cross-sectional area of the active region in a laser diode is only about 1 μm, diffraction effects cause the emerging beam to diverge rapidly. This phenomenon is depicted schematically in Fig. 102.4(a). In practical applications of the laser diode, the expansion of the emerging beam is arrested by a collimating lens, such as that shown in Fig. 102.4(b). If the beam happens to have aberrations (astigmatism is particularly severe in diode lasers), then the collimating lens must be designed to correct this defect as well.

In optical recording, it is most desirable to have a beam with circular cross-section. The need for shaping the beam arises from the special geometry of the laser cavity with its rectangular cross-section. Since the emerging beam has different dimensions in the directions parallel and perpendicular to the junction, its cross-section at the collimator becomes elliptical, with the initially narrow dimension expanding more rapidly to become the major axis of the ellipse. The collimating lens thus produces a beam with elliptical cross-section. Circularization may be achieved by bending various rays of the beam at a prism, as shown in Fig. 102.4(c). The bending changes the beam's diameter in the plane of incidence but leaves the diameter in the perpendicular direction intact.

Focusing by the Objective Lens

The collimated and circularized beam of the diode laser is focused on the surface of the disk using an **objective lens**. The objective is designed to be aberration free so that its focused spot size is limited only by the effects of diffraction. Figure 102.5(a) shows the design of a typical objective made from spherical optics. According to the classical theory of diffraction, the diameter of the beam, d, at the objective's focal plane is given by

$$d \simeq \frac{\lambda}{NA} \tag{102.9}$$

where λ is the wavelength of light, and NA is the numerical aperture of the objective.*

*Numerical aperture is defined as $NA = n \sin\theta$, where n is the refractive index of the image space and θ is the half-angle subtended by the exit pupil at the focal point. In optical recording systems, the image space is air whose index is very nearly unity; thus, for all practical purposes $NA = \sin\theta$.

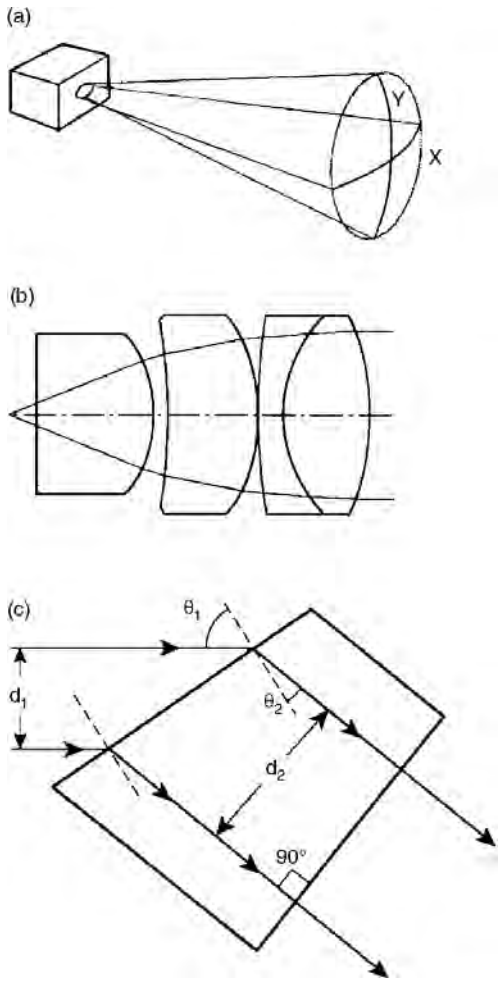


FIGURE 102.4 (a) Away from the facet, the output beam of a diode laser diverges rapidly. In general, the beam diameter along X is different from that along Y, which makes the cross-section of the beam elliptical. Also, the radii of curvature R_x and R_y are not the same, thus creating a certain amount of astigmatism in the beam. (b) Multielement collimator lens for laser diode applications. Aside from collimating, this lens also corrects astigmatic aberrations of the beam. (c) Beam-shaping by deflection at a prism surface. θ_1 and θ_2 are related by Snell's law, and the ratio d_2/d_1 is the same as $\cos\theta_2/\cos\theta_1$. Passage through the prism circularizes the elliptical cross-section of the beam.

In optical recording it is desired to achieve the smallest possible spot, since the size of the spot is directly related to the size of marks recorded on the medium. Also, in readout, the spot size determines the resolution of the system. According to Eq. (102.9) there are two ways to achieve a small spot: first, by reducing the wavelength and second, by increasing the numerical aperture of the objective. The wavelengths currently available from GaAs lasers are in the range of 670–840 nm. It is possible to use a nonlinear optical device to double the frequency of these diode lasers, thus achieving blue light. Good efficiencies have been demonstrated by frequency doubling. Also recent developments in II–VI materials have improved the prospects for obtaining green and blue light directly from semiconductor lasers. Consequently, there is hope that in the near future, optical storage systems will operate in the wavelength range of 400–500 nm. As for the numerical aperture, current practice is to use a lens with $NA \simeq 0.5$–0.6. Although this value might increase slightly in the coming years, much higher numerical apertures are unlikely, since they put strict constraints on the other characteristics of the system and limit the tolerances. For instance, the working distance at high numerical aperture is relatively short, making access to the recording layer through the substrate more difficult. The smaller depth of focus of a high numerical aperture lens will make attaining/maintaining proper focus more of a problem, whereas the limited

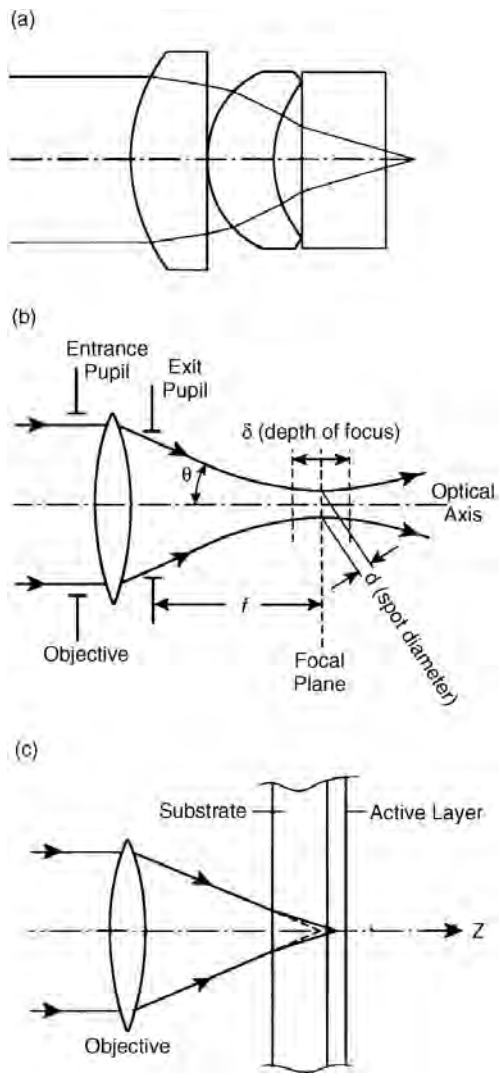


FIGURE 102.5 (a) Multielement lens design for a high numerical aperture video disc objective. (*Source:* Kuntz, D. 1984. Specifying laser diode optics, *Laser Focus,* March. With permission.) (b) Various parameters of the objective lens. The numerical aperture is $NA = \sin\theta$. The spot diameter d and the depth of focus δ are given by Eqs. (102.9) and (102.10), respectively. (c) Focusing through the substrate can cause spherical aberration at the active layer. The problem can be corrected if the substrate is taken into account in the design of the objective.

field of view might restrict automatic track-following procedures. A small field of view also places constraints on the possibility of read/write/erase operations involving multiple beams.

The depth of focus of a lens, δ, is the distance away from the focal plane over which tight focus can be maintained [see Fig. 102.5(b)]. According to the classical diffraction theory,

$$\delta \simeq \frac{\lambda}{NA^2} \tag{102.10}$$

Thus for a wavelength of $\lambda = 700$ nm and $NA = 0.6$, the depth of focus is about ± 1 μm. As the disk spins under the optical head at the rate of several thousand revolutions per minute, the objective lens must stay within a distance of $f \pm \delta$ from the active layer if proper focus is to be maintained. Given the conditions under which drives usually operate, it is impossible to make rigid enough mechanical systems to yield the required positioning tolerances. On the other hand, it is fairly simple to mount the

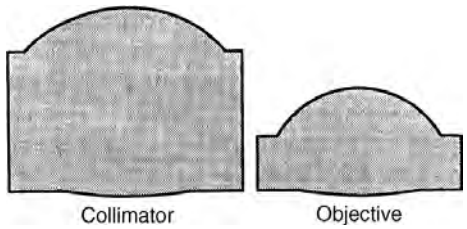


FIGURE 102.6 Molded glass aspheric lens pair for optical disk applications. These singlets can replace the multi-element spherical lenses shown in Figs. 102.4(b) and 102.5(a).

objective lens in an actuator capable of adjusting its position with the aid of closed-loop feedback control. We shall discuss the technique of automatic focusing in the next section. For now, let us emphasize that by going to shorter wavelengths and/or larger numerical apertures (as is required for attaining higher data densities), one will have to face a much stricter regime as far as automatic focusing is concerned. Increasing the numerical aperture is particularly worrisome, since δ drops with the square of NA.

A source of spherical aberrations in optical disk systems is the substrate through which the light must travel to reach the active layer of the disk. Figure 102.5(c) shows the bending of the rays at the disk surface that causes the aberration. This problem can be solved by taking into account the effects of the substrate in the design of the objective, so that the lens is corrected for all aberrations including those arising at the substrate. Recent developments in molding of aspheric glass lenses have gone a long way in simplifying the lens design problem. Figure 102.6 shows a pair of molded glass aspherics designed for optical disk system applications; both the collimator and the objective are single-element lenses and are corrected for aberrations.

102.4 Automatic Focusing

We mentioned in the preceding section that since the objective has a large numerical aperture ($NA \geq 0.5$), its depth of focus δ is rather shallow ($\delta \simeq \pm 1$ μm at $\lambda = 780$ nm). During all read/write/erase operations, therefore, the disk must remain within a fraction of a micrometer from the focal plane of the objective. In practice, however, the disks are not flat and they are not always mounted rigidly parallel to the focal plane, so movements away from focus occur a few times during each revolution. The peak-to-peak movement in and out of focus may be as much as 100 μm. Without **automatic focusing** of the objective along the optical axis, this runout (or disk flutter) will be detrimental to the operation of the system. In practice, the objective is mounted on a small motor (usually a voice coil) and allowed to move back and forth in order to keep its distance within an acceptable range from the disk. The spindle turns at a few thousand revolutions per minute, which is a hundred or so revolutions per second. If the disk moves in and out of focus a few times during each revolution, then the voice coil must be fast enough to follow these movements in real time; in other words, its frequency response must extend to several kilohertz.

The signal that controls the voice coil is obtained from the light reflected from the disk. There are several techniques for deriving the focus error signal, one of which is depicted in Fig. 102.7(a). In this so-called obscuration method, a secondary lens is placed in the path of the reflected light, one-half of its aperture is covered and a split-detector is placed at its focal plane. When the disk is in focus, the returning beam is collimated and the secondary lens will focus the beam at the center of the split detector, giving a difference signal ΔS equal to zero. If the disk now moves away from the objective, the returning beam will become converging, as in Fig. 102.7(b), sending all the light to detector 1. In this case, ΔS will be positive and the voice coil will push the lens toward the disk. On the other hand, when the disk moves close to the objective, the returning beam becomes diverging and detector 2 receives the light [see Fig. 102.7(c)]. This results in a negative ΔS that forces the voice coil to pull back in order to return ΔS to zero. A given focus error detection scheme is generally characterized by the shape of its focus error

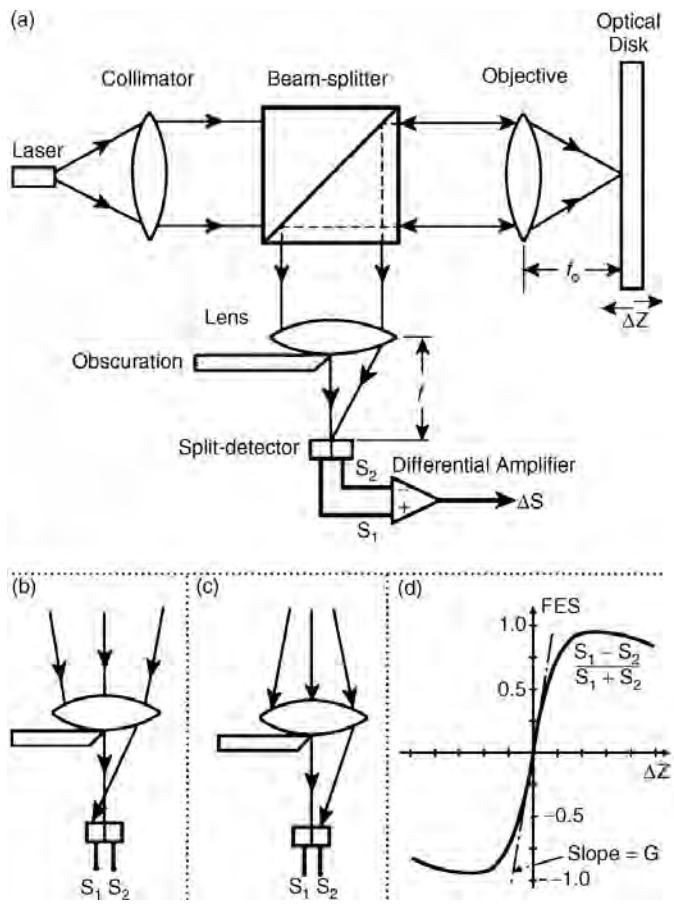


FIGURE 102.7 Focus error detection by the obscuration method. In (a) the disk is in focus and the two halves of the split detector receive equal amounts of light. When the disk is too far from the objective (b), or too close to it (c), the balance of detector signals shifts to one side or the other. A plot of the focus error signal (FES) vs. defocus is shown in (d), and its slope near the origin is identified as the FES gain G.

signal ΔS vs. the amount of defocus Δz; one such curve is shown in Fig. 102.7(d). The slope of the focus error signal (FES) curve near the origin is of particular importance since it determines the overall performance and stability of the servo loop.

102.5 Automatic Tracking

Consider a track at a certain radial location, say r_0, and imagine viewing this track through the access window shown in Fig. 102.1. It is through this window that the head gains access to arbitrarily selected tracks. To a viewer looking through the window, a perfectly circular track centered on the spindle axis will look stationary, irrespective of the rotation rate. However, any eccentricity will cause an apparent radial motion of the track. The peak-to-peak distance travelled by a track (as seen through the window) depends on a number of factors including centering accuracy of the hub, deformability of the substrate, mechanical vibrations, manufacturing tolerances, etc. For a typical 3.5-in. disk, for example, this peak-to-peak motion can be as much as 100 μm during one revolution. Assuming a revolution rate of 3600 r/m, the apparent velocity of the track in the radial direction will be several millimeters per second. Now, if the focused spot remains stationary while trying to read from or write to this track, it is clear that the beam will miss the track for a good fraction of every revolution cycle.

Practical solutions to the described problem are provided by **automatic tracking** techniques. Here the objective is placed in a fine actuator, typically a voice-coil, which is capable of moving the necessary

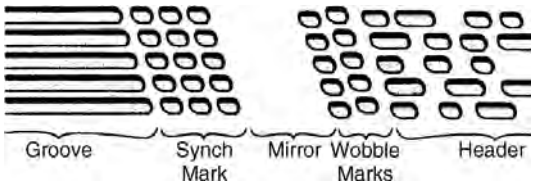


FIGURE 102.8 Servo fields in continuous/composite format contain a mirror area and offset marks for tracking. (*Source:* Marchant, A.B. 1990. *Optical Recording,* Addison–Wesley, Reading, MA, 264. With permission.)

radial distances and maintaining a lock on the desired track. The signal that contols the movement of this actuator is derived from the reflected light itself, which carries information about the position of the focused spot. There exist several mechanisms for extracting the track-error signal (TES); all of these methods require some sort of structure on the disk surface in order to identify the track. In the case of read-only disks (CD, CD-ROM, and video disk), the embossed pattern of data provides ample information of tracking purposes. In the case of write-once and erasable disks, tracking guides are carved on the substrate in the manufacturing process. As mentioned earlier, the two major formats for these tracking guides are pregrooves (for continuous tracking) and sampled-servo marks (for discrete tracking). A combination of the two schemes, known as continuous/composite format, is often used in practice. This scheme is depicted in Fig. 102.8, which shows a small section containing five tracks, each consisting of the tail-end of a groove, synchronization marks, a mirror area used for adjusting focus/track offsets, a pair of wobble marks for sampled tracking, and header information for section identification.

Tracking on Grooved Regions

As shown in Fig. 102.3(a), grooves are continuous depressions that are either embossed, etched, or molded onto the substrate prior to deposition of the storage medium. If the data is recorded on the grooves, then the lands are not used except for providing a guard band between neighboring grooves. Conversely, the land regions may be used to record the information, in which case grooves provide the guard band. Typical track-widths are about one wavelength of the light. The guard bands are somewhat narrower than the tracks; their exact shape and dimensions depending on the beam size, the required track-servo accuracy, and the acceptable levels of cross talk between adjacent tracks. The groove depth is usually around one-eighth of one wavelength ($\lambda/8$), since this depth can be shown to give the largest TES in the push–pull method. Cross-sections of the grooves may be rectangular, trapezoidal, or triangular.

When the focused spot is centered on the track, it is diffracted symmetrically from the two edges of the track, resulting in a balanced far-field pattern. As soon as the spot moves away from the center, the symmetry breaks down and the light distribution in the far field tends to shift to one side or the other. A split photodetector placed in the path of the reflected light can therefore sense the relative position of the spot and provide the appropriate feedback signal. This strategy is depicted schematically in Fig. 102.9(a); also shown in the figure are intensity plots at the detector plane for light reflected from various regions of the disk. Note how the intensity shifts to one side or the other depending on the direction of motion of the spot.

Sampled Tracking

Since dynamic track runout is usually a slow and gradual process, there is actually no need for continuous tracking as done on grooved media. A pair of embedded marks, offset from the track center as in Fig. 102.3(b), can provide the necessary information for correcting the relative position of the focused spot. The reflected intensity will indicate the positions of the two servo marks as two successive short pulses. If the beam happens to be on track, the two pulses will have equal magnitudes and there shall be no need for correction. If, on the other hand, the beam is off track, one of the pulses will be stronger than the other. Depending on which pulse is the strong one, the system will recognize the direction in which it has to move and will correct the error accordingly. The servo marks must appear frequently enough along the track to ensure proper track

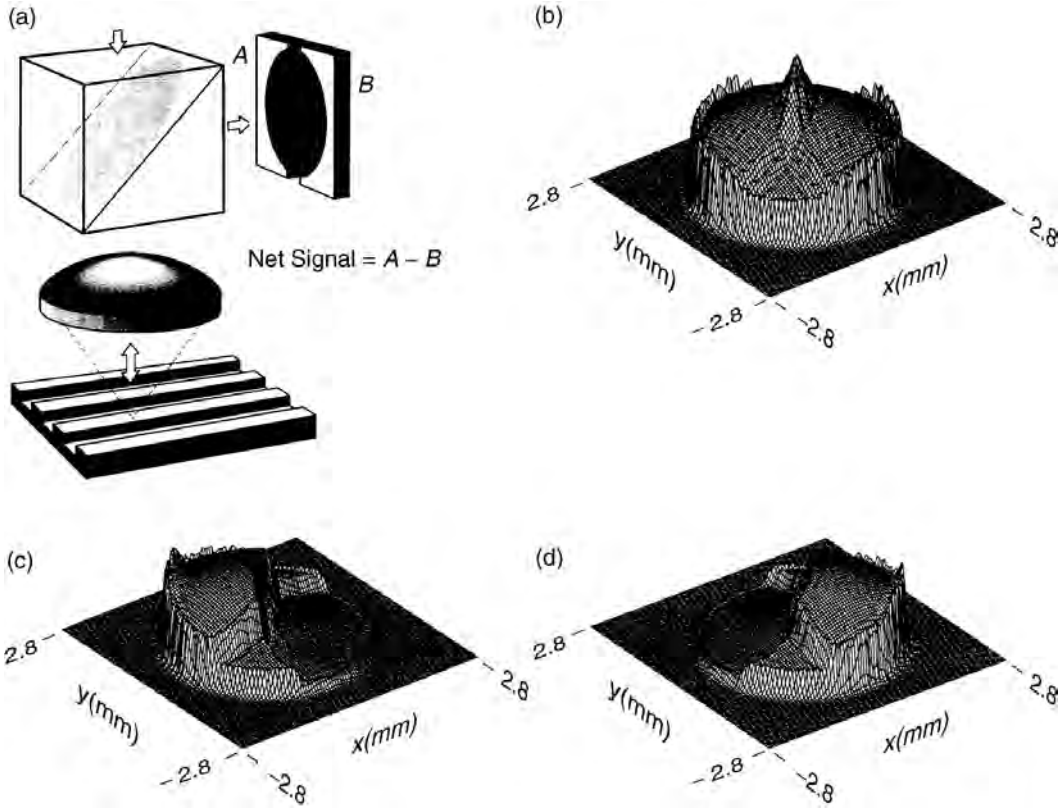


FIGURE 102.9 (a) Push–pull sensor for tracking on grooves. (*Source:* Marchant, A.B. 1990. *Optical Recording,* Addison–Wesley, Reading, MA, 1990, 175. With permission.) (b) Calculated distribution of light intensity at the detector plane when the disk is in focus and the beam is centered on track. (c) Calculated intensity distribution at the detector plane with disk in focus but the beam centered on the groove edge. (d) Same as (c) except for the spot being focused on the opposite edge of the groove.

following. In a typical application, the track might be divided into groups of 18 bytes, 2 bytes dedicated as servo offset areas and 16 bytes filled with other format information or left blank for user data.

102.6 Thermomagnetic Recording Process

Recording and erasure of information on a magneto-optical disk are both achieved by the **thermomagnetic process**. The essence of thermomagnetic recording is shown in Fig. 102.10. At the ambient temperature, the film has a high magnetic coercivity* and, therefore, does not respond to the externally applied field. When a focused beam raises the local temperature of the film, the hot spot becomes magnetically soft (i.e., its coercivity drops). As the temperature rises, coercivity drops continuously until such time as the field of the electromagnet finally overcomes the material's resistance to reversal and switches its magnetization. Turning the laser off brings the temperatures back to normal, but the reverse-magnetized domain remains frozen in the film. In a typical situation in practice, the film thickness may be around 300 Å, laser power at the disk 10 $\simeq$ mW, diameter of the focused spot $\simeq$1 μm, laser pulse duration $\simeq$50 ns, linear velocity of the track $\simeq$10 m/s, and the magnetic field strength $\simeq$200 G. The temperature may reach a peak of 500 K at the center of the spot, which is sufficient for magnetization

*Coercivity of a magnetic medium is a measure of its resistance to magnetization reversal. For example, consider a thin film with perpendicular magnetic moment saturated in the $+Z$ direction. A magnetic field applied along $-Z$ will succeed in reversing the direction of magnetization only if the field is stronger than the coercivity of the film.

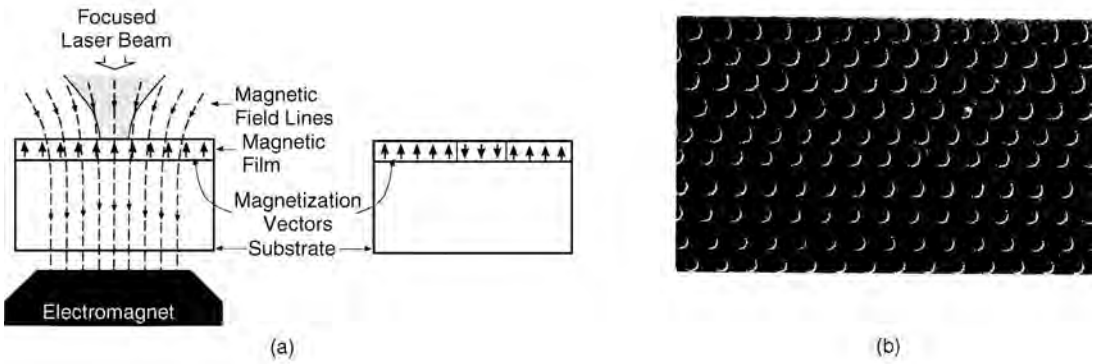


FIGURE 102.10 (a) Thermomagnetic recoding process: the field of the electromagnet helps reverse the direction of magnetization in the area heated by the focused laser beam; (b) Lorentz micrograph of domains written thermomagnetically. The various tracks shown here were written at different laser powers, with power level decreasing from top to bottom. (*Source:* Greidanus, F. et al., 1989. Paper 26B-5, International Symposium on Optical Memory, Kobe, Japan, Sept. With permission.)

reversal but is not nearly high enough to melt or crystalize or in any other way modify the material's atomic structure.

The materials of magneto-optical recording have strong perpendicular magnetic anisotropy. This type of anisotropy favors the up and down directions of magnetization over all other orientations. The disk is initialized in one of these two directions, say, up, and the recording takes place when small regions are selectively reverse magnetized by the thermomagnetic process. The resulting magnetization distribution then represents the pattern of recorded information. For instance, binary sequences may be represented by a mapping of 0s to up-magnetized regains and 1s to down magnetized regions [nonreturn to zero (NRZ)]. Alternatively, the NRZI scheme might be used, whereby transitions (up-to-down and down-to-up) are used to represent the 1s in the bit sequence.

Recording by Laser Power Modulation (LPM)

In this traditional approach to thermomagnetic recording, the electromagnet produces a constant field, whereas the information signal is used to modulate the power of the laser beam. As the disk rotates under the focused spot, the on/off laser pulses create a sequence of up/down domains along the track. The Lorentz electron micrograph in Fig. 102.10(b) shows a number of domains recorded by laser power modulation (LPM). The domains are highly stable and may be read over and over again without significant degradation. If, however, the user decides to discard a recorded block and to use the space for new data, the LPM scheme does not allow direct overwrite; the system must erase the old data during one disk revolution cycle and record the new data in a subsequent revolution cycle.

During erasure, the direction of the external field is reversed so that up-magnetized domains in Fig. 102.10(a) now become the favored ones. Whereas writing is achieved with a modulated laser beam, in erasure the laser stays on for a relatively long period of time, erasing an entire sector. Selective erasure of individual domains is not practical, nor is it desired, since mass data storage systems generally deal with data at the level of blocks, which are recorded onto and read from individual sectors. Note that at least one revolution period elapses between the erasure of an old block and its replacement by a new block. The electromagnet therefore need not be capable of rapid switchings. (When the disk rotates at 3600 r/m, for example, there is a period of 16 ms or so between successive switchings.) This kind of slow reversal allows the magnet to be large enough to cover all of the tracks simultaneously, thereby eliminating the need for a moving magnet and an actuator. It also affords a relatively large gap between the disk and the magnet, which enables the use of double-sided disks and relaxes the mechanical tolerances of the system without overburdening the magnet's driver.

The obvious disadvantage of LPM is its lack of direct overwrite capability. A more subtle concern is that it is perhaps unsuitable for the pulse width modulation (PWM) scheme of representing binary waveforms. Because of fluctuations in the laser power, spatial variations of material properties, and lack of perfect focusing and track following, etc., the length of a recorded domain along the track may fluctuate in small but unpredictable ways. If the information is to be encoded in the distance between adjacent domain walls (i.e., PWM), then the LPM scheme of thermomagnetic writing may suffer from excessive domain-wall jitter. Laser power modulation works well, however, when the information is encoded in the position of domain centers [i.e., pulse position modulation (PPM)]. In general, PWM is superior to PPM in terms of the recording density, and, therefore, recording techniques that allow PWM are preferred.

Recording by Magnetic Field Modulation

Another method of thermomagnetic recording is based on magnetic field modulation (MFM) and is depicted schematically in Fig. 102.11(a). Here the laser power may be kept constant while the information signal is used to modulate the magnetic field. Photomicrographs of typical domain patterns recorded in the MFM scheme are shown in Fig. 102.11(b). Crescent-shaped domains are the hallmark of the field modulation technique. If one assumes (using a much simplified model) that the magnetization aligns itself with the applied field within a region whose temperature has passed a certain critical value, T_{crit}, then one can explain the crescent shape of these domains in the following way: with the laser operating

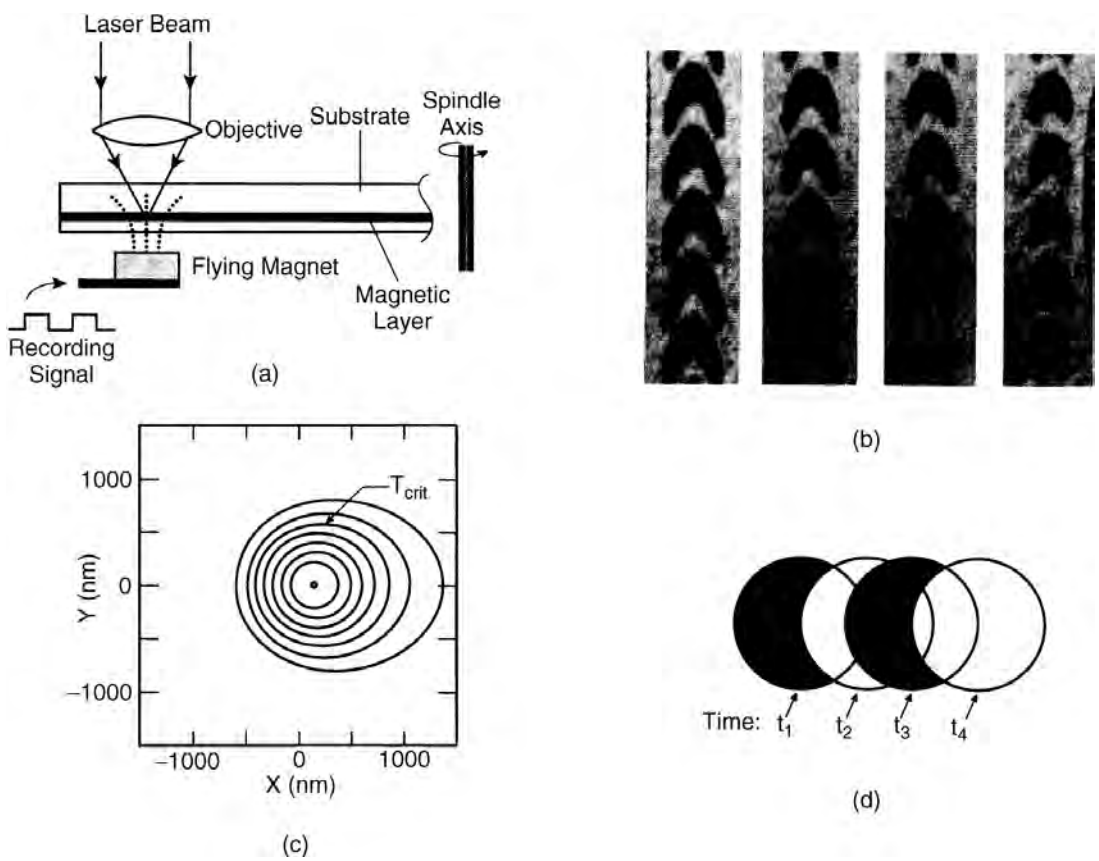


FIGURE 102.11 (a) Thermomagnetic recording by magnetic field modulation. The power of the beam is kept constant, while the magnetic field direction is switched by the data signal. (b) Polarized-light microphotograph of recorded domains. (c) Computed isotherms produced by a CW laser beam, focused on the magnetic layer of a disk. The disk moves with constant velocity under the beam. The region inside the isotherm marked as T_{crit} is above the critical temperature for writing, that is, its magnetization aligns with the direction of the applied field. (d) Magnetization within the heated region (above T_{crit}) follows the direction of the applied field, whose switchings occur at times t_n. The resulting domains are crescent shaped.

in the CW mode and the disk moving at constant velocity, temperature distribution in the magnetic medium assumes a steady-state profile, such as that shown in Fig. 102.11(c). Of course, relative to the laser beam, the temperature profile is stationary, but in the frame of reference of the disk, the profile moves along the track with the linear track velocity. The isotherm corresponding to T_{crit} is identified as such in the figure; within this isotherm, the magnetization aligns itself with the applied field. Figure 102.11(d) shows a succession of critical isotherms along the track, each obtained at the particular instant of time when the magnetic field switches direction. From this picture it is easy to infer how the crescent-shaped domains form and also understand the relation between the waveform that controls the magnet and the resulting domain pattern.

The advantages of magnetic field modulation recording are that (1) direct overwriting is possible, and (2) domain wall positions along the track, being rather insensitive to defocus and laser power fluctuations, are fairly accurately controlled by the timing of the magnetic field switchings. On the negative side, the magnet must now be small and fly close to the disk surface if it is to produce rapidly switched fields with a magnitude of a hundred gauss or so. Systems that utilize magnetic field modulation often fly a small electromagnet on the opposite side of the disk from the optical stylus. Since mechanical tolerances are tight, this might compromise the removability of the disk. Moreover, the requirement of close proximity between the magnet and the storage medium dictates the use of single-sided disks in practice.

102.7 Magneto-Optical Readout

The information recorded on a perpendicularly magnetized medium may be read with the aid of the polar **magneto-optical Kerr effect**. When linearly-polarized light is normally incident on a perpendicular magnetic medium, its plane of polarization undergoes a slight rotation upon reflection. This rotation of the plane of polarization, whose sense depends on the direction of magnetization in the medium, is known as the polar Kerr effect. The schematic representation of this phenomenon in Fig. 102.12 shows that if the polarization vector suffers a counterclockwise rotation upon reflection from an up-magnetized region, then the same vector will rotate clockwise when the magnetization is down. A magneto-optical (MO) medium is characterized in terms of its reflectivity R and its Kerr rotation angle θ_k. R is a real number (between 0 and 1) that indicates the fraction of the incident power reflected back from the medium at normal incidence. Here θ_k is generally quoted as a positive number, but is understood to be positive or negative depending on the direction of magnetization; in MO readout, it is the sign of θ_k that carries the information about the state of magnetization, that is, the recorded bit pattern.

The laser used for readout is usually the same as that used for recording, but its output power level is substantially reduced in order to avoid erasing (or otherwise obliterating) the previously recorded

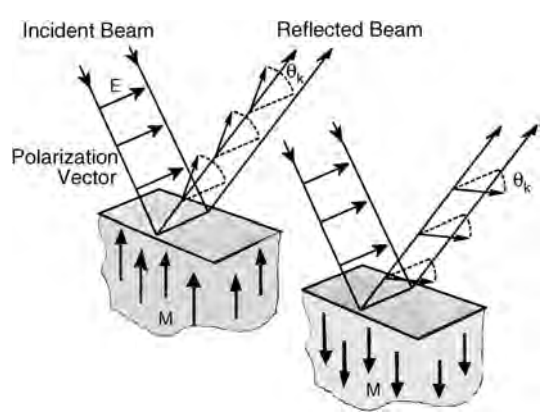


FIGURE 102.12 Schematic diagram describing the polar magneto-optical Kerr effect. Upon reflection from the surface of a perpendicularly magnetized medium, the polarization vector undergoes a rotation. The sense of rotation depends on the direction of magnetization $\boldsymbol{M}$ and switches sign when $\boldsymbol{M}$ is reversed.

information. For instance, if the power of the write/erase beam is 20 mW, then for the read operation the beam is attenuated to about 3 or 4 mW. The same objective lens that focuses the write beam is now used to focus the read beam, creating a diffraction-limited spot for resolving the recorded marks. Whereas in writing the laser was pulsed to selectively reverse magnetize small regions along the track, in readout it operates with constant power, that is, in CW mode. Both up- and down-magnetized regions are read as the track passes under the focused spot. The reflected beam, which is now polarization modulated, goes back through the objective and becomes collimated once again; its information content is subsequently decoded by polarization-sensitive optics, and the scanned pattern of magnetization is reproduced as an electronic signal.

Differential Detection

Figure 102.13 shows the differential detection system that is the basis of magneto-optical readout in practically all erasable optical storage systems in use today. The beam-splitter (BS) diverts half of the reflected beam away from the laser and into the detection module.* The polarizing beam-splitter (PBS) splits the beam into two parts, each carrying the projection of the incident polarization along one axis of the PBS, as shown in Fig. 102.13(b). The component of polarization along one of the axes goes straight through, whereas the component along the other axis splits off and branches to the side. The PBS is oriented such that, in the absence of the Kerr effect, its two branches will receive equal amounts of light. In other words, if the polarization, upon reflection from the disk, did not undergo any rotations whatsoever, then the beam entering the PBS would be polarized at 45° to the PBS axes, in which case it would split equally between the two branches. Under this condition, the two detectors generate identical signals, and the differential signal ΔS will be zero. Now, if the beam returns from the disk with its polarization rotated clockwise (rotation angle = θ_k), then detector 1 will receive more light than detector 2, and the differential signal will be positive. Similarly, a counterclockwise rotation will generate a negative ΔS. Thus, as the disk rotates under the focused spot, the electronic signal ΔS reproduces the pattern of magnetization along the scanned track.

102.8 Materials of Magneto-Optical Data Storage

Amorphous rare earth transition metal alloys are presently the media of choice for erasable optical data storage applications. The general formula for the composition of the alloy may be written $(Tb_yGd_{1-y})_x(Fe_zCo_{1-z})_{1-x}$, where terbium and gadolinium are the rare earth (RE) elements, whereas iron and cobalt are the transition metals (TM). In practice, the transition metals constitute roughly 80 atomic percent of the alloy (i.e., $x \simeq 20$). In the transition metal subnetwork, the fraction of cobalt is usually small, typically around 10%, and iron is the dominant element ($z \simeq 90$). Similarly, in the rare earth subnetwork, Tb is the main element ($y \simeq 90$), whereas the gadolinium content is small or it may even be absent in some cases. Since the rare earth elements are highly reactive to oxygen, RE-TM films tend to have poor corrosion resistance and, therefore, require protective coatings. In multilayer disk structures, the dielectric layers that enable optimization of the medium for the best optical/thermal behavior also perform the crucial function of protecting the MO layer from the environment.

The amorphous nature of the material allows its composition to be continuously varied until a number of desirable properties are achieved. In other words, the fractions x, y, z of the various elements are not constrained by the rules of stoichiometry. Disks with very large areas can be coated uniformly with thin films of these media, and, in contrast to polycrystalline films whose grains and grain boundaries scatter the beam and cause noise, amorphous films are continuous, smooth, and substantially free from noise. The films are deposited either by sputtering from an alloy target or by cosputtering from multiple elemental targets. In the latter case, the substrate moves under the various targets and the fraction of a given element in the alloy is determined by the time spent under each target as well as the power applied to that target. During film deposition, the substrate is kept at a low temperature (usually by chilled water)

*The use of an ordinary beam-splitter is an inefficient way of separating the incoming and outgoing beams since half the light is lost in each pass through the splitter. One can do much better by using a so-called leaky beam-splitter.

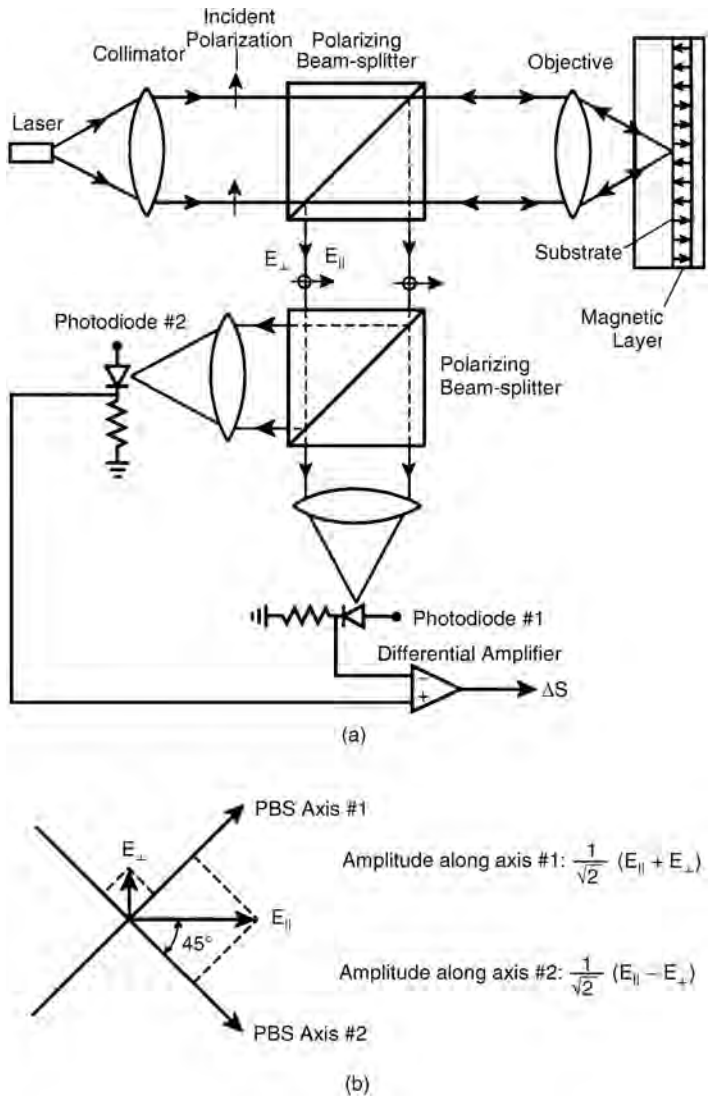


FIGURE 102.13 Differential detection scheme (a) utilizes a polarizing beam splitter and two photodetectors in order to convert the rotation of polarization to an electronic signal. $E_{\parallel}$ and $E_{\perp}$ are the reflected components of polarization; they are, respectively, parallel and perpendicular to the direction of incident polarization. The diagram in (b) shows the orientation of the PBS axes relative to the polarization vectors.

in order to reduce the mobility of deposited atoms and to inhibit crystal growth. The type of the sputtering gas (argon, krypton, xenon, etc.) and its pressure during sputtering, the bias voltage applied to the substrate, deposition rate, nature of the substrate and its pretreatment, and temperature of the substrate all can have dramatic effects on the composition and short-range order of the deposited film. A comprehensive discussion of the factors that influence film properties will take us beyond the intended scope here; the interested reader may consult the vast literature in this field for further information.

Defining Terms

Automatic focusing: The process in which the distance of the disk from the objective's focal plane is continuously monitored and fed back to the system in order to keep the disk in focus at all times.

Automatic tracking: The process in which the distance of the focused spot from the track center is continuously monitored and the information fed back to the system in order to maintain the read/write beam on track at all times.

Compact disc (CD): A plastic substrate embossed with a pattern of pits that encode audio signals in digital format. The disk is coated with a metallic layer (to enhance its reflectivity) and read in a drive (CD player) that employs a focused laser beam and monitors fluctuations of the reflected intensity in order to detect the pits.

Error correction coding (ECC): Systematic addition of redundant bits to a block of binary data as insurance against possible read/write errors. A given error correcting code can recover the original data from a contaminated block, provided that the number of erroneous bits is less than the maximum number allowed by that particular code.

Grooved media of optical storage: A disk embossed with grooves of either the concentric-ring type or the spiral type. If grooves are used as tracks, then the lands (i.e., regions between adjacent grooves) are the guard bands. Alternatively, lands may be used as tracks, in which case the grooves act as guard bands. In a typical grooved optical disk in use today, the track width is 1.1 μm, the width of the guard band is 0.5 μm, and the groove depth is 70 nm.

Magneto-optical Kerr effect: The rotation of the plane of polarization of a linearly polarized beam of light upon reflection from the surface of a perpendicularly magnetized medium.

Objective lens: A well-corrected lens of high numerical aperture, similar to a microscope objective, used to focus the beam of light onto the surface of the storage medium. The objective also collects and recollimates the light reflected from the medium.

Optical path: Optical elements in the path of the laser beam in an optical drive. The path begins at the laser itself and contains a collimating lens, beam shaping optics, beam splitters, polarization-sensitive elements, photodetectors, and an objective lens.

Preformat: Information such as sector address, synchronization marks, servo marks, etc., embossed permanently on the optical disk substrate.

Sector: A small section of track with the capacity to store one block of user data (typical blocks are either 512 or 1024 bytes). The surface of the disk is covered with tracks, and tracks are divided into contiguous sectors.

Thermomagnetic process: The process of recording and erasure in magneto-optical media, involving local heating of the medium by a focused laser beam, followed by the formation or annihilation of a reverse-magnetized domain. The successful completion of the process usually requires an external magnetic field to assist the reversal of the magnetization.

Track: A narrow annulus or ringlike region on a disk surface, scanned by the read/write head during one revolution of the spindle; the data bits of magnetic and optical disks are stored sequentially along these tracks. The disk is covered either with concentric rings of densely packed circular tracks or with one continuous, fine-pitched spiral track.

References

Arimoto, A., Ojima, M., Chinone, N., Oishi, A., Gotoh, T., and Ohnuki, N. 1986. Optimum conditions for the high frequency noise reduction method in optical video disk players. *Appl. Opt.,* 25:1398.

Beck, J.W. 1970. Noise considerations of optical beam recording. *Appl. Opt.,* 9:2559.

Bouwhuis, G., Braat, J., Huijser, A., Pasman, J., Van Rosmalen, G., and Immink, K.S. 1985. *Principles of Optical Disk Systems,* Adam Hilger, Bristol, Chaps. 2 and 3.

Carslaw, H.S. and Jaeger, J.C. 1954. *Conduction of Heat in Solids,* Oxford University Press, London.

Chikazumi, S. and Charap, S.H. 1964. *Physics of Magnetism,* Wiley, New York.

Egashira, K. and Yamada, R. 1974. Kerr effect enhancement and improvement of readout characteristics in MnBi film memory. *J. Appl. Phys.,* 45:3643–3648.

Hansen, P. and Heitman, H. 1989. Media for erasable magneto-optic recording. *IEEE Trans. Mag.,* 25:4390–4404.

Heemskerk, J. 1978. Noise in a video disk system: experiments with an (AlGa) As laser. *Appl. Opt.,* 17:2007.

Huth, B.G. 1974. Calculation of stable domain radii produced by thermomagnetic writing *IBM J. Res. Dev.,* pp. 100–109.

Immink, K.A.S. 1986. Coding methods for high-density optical recording. *Philips J. Res.,* 41:410–430.
Kivits, P., deBont, R., and Zalm, P. 1981. Superheating of thin films for optical recording. *Appl. Phys.,* 24:273–278.
Kryder, M.H. 1987. Data-storage technolgoies for advanced computing. *Sci. Am.,* 257:116–125.
Maissel, L.I. and Glang, R., Eds. 1970. *Handbook of Thin Film Technology,* McGraw–Hill, New York.
Malozemoff, A.P. and Slonczewski, J.C. 1979. *Magnetic Domain Walls in Bubble Materials,* Academic Press, New York.
Mansuripur, M. 1989. Certain computational aspects of vector diffraction problem. *J. Opt. Soc. Am. A,* 6:786–806.
Mansuripur, M., Connell, G.A.N., and Goodman, J.W. 1982a. Laser-induced local heating of multilayers. *Appl. Opt.,* 21:1106.
Mansuripur, M., Connell, G.A.N., and Goodman, J.W. 1982b. Signal and noise in magneto-optical readout. *J. Appl. Phys.,* 53:4485.
Marchant, A.B. 1990. *Optical Recording,* Addison–Wesley, Reading, MA.
Patel, A.M. 1988. Signal and error-control coding. In *Magnetic Recording,* Vol. II, Eds. C.D. Mee and E.D. Daniel, McGraw–Hill, New York.
Pershan, P.S. 1967. Magneto-optic effects. *J. Appl. Phys.,* 38:1482–1490.
Smith, D.O. 1965. Magneto-optical scattering multilayer magnetic and dielectric films. *Opt. Acta,* 12:13.
Special Issue. 1978. *Appl. Opt.,* July 1 (video disks).
Weissler, G.L. and Carlson, R.W., Eds. 1979. *Vaccum Physics and Technology, Vol. 14, Methods of Experimental Physics,* Academic Press, New York.
Wolf, E. 1959. Electromagnetic diffraction in optical systems. I. An integral representation of the image field. *Proc. R. Soc. Ser. A,* 253:349–357.

Further Information

Proceedings of the Optical Data Storage Conference are published annually by SPIE, the International Society for Optical Engineering. These proceedings document the latest developments in the field of optical recording each year. Two other conferences in this field are the International Symposium on Optical Memory (ISOM), whose proceedings are published as a special issue of the *Japanese Journal of Applied Physics*, and the Magneto-Optical Recording International Symposium (MORIS), whose proceedings appear in a special issue of the *Journal of the Magnetics Society of Japan*.