The McGraw-Hill Companies

Second Edition

Electronic Devices and Circuits

S Salivahanan
N Suresh Kumar
A Vallavaraj

Published by the Tata McGraw-Hill Publishing Company Limited,
7 West Patel Nagar, New Delhi 110 008.

Second reprint 2008
RYLARRLXRQLRD

This edition can be exported from India only by the publishers,
Tata McGraw-Hill Publishing Company Limited

ISBN-13: 978-0-07-066049-6
ISBN-10: 0-07-066049-2

General Manager: Publishing—SEM & Tech Ed: *Vibha Mahajan*
Asst. Sponsoring Editor: *Shukti Mukherjee*
Jr. Editorial Executive: *Sandhya Chandrasekhar*
Executive—Editorial Services: *Sohini Mukherjee*

General Manager: Marketing—Higher Education & School: *Michael J Cruz*
Product Manager: SEM & Tech Ed: *Biju Ganesan*

Controller—Production: *Rajender P Ghansela*
Asst. General Manager—Production: *B L Dogra*
Senior Production Manager: *P L Pandita*

Typeset at The Composers, 260, C.A. Apt., Paschim Vihar, New Delhi 110 063 and printed at Sai Printo Pack, A-102/4, Okhla Industrial Area, Phase-I,
New Delhi 110 020

Cover: SDR

The McGraw-Hill Companies

Contents

Foreword

It gives me immense pleasure to write a foreword for this book *Electronic Devices and Circuits* authored by Prof S Salivahanan, Mr N Suresh Kumar and Mr A Vallavaraj of Mepco Schlenk Engineering College, Sivakasi, whom I have known very well for many years. Electronics has made rapid advances over the last few decades. Its importance has become more specific with the rapid development of the solid state devices and circuits.

Electronic Devices and Circuits is one of the important core subjects in the field of Electronics, Electrical, Computer Science, Instrumentation and Mechanical Engineering. There has been an increase in the demand for a suitable textbook on this subject. The contents of the book are presented in a simple, precise and systematic manner. Numerous solved examples, self-explanatory sketches and a large number of exercise problems with answers have been presented in each chapter to aid conceptual understanding of the subject. This is an outcome of the several years of teaching experience of the authors.

Professor S Salivahanan has vast experience of teaching various courses in the industry and engineering colleges both in India and abroad. The other two authors too have considerable experience in teaching the subjects over a number of years.

Having been myself a teacher in the field of Electronics and Communication Engineering for more than three decades. I am sure that this book will enrich the students' knowledge in the subject and would be welcomed by the teachers and students of all engineering institutions.

I strongly recommend this book to every Electronics, Electrical, Computer Science, Instrumentation and Mechanical Engineering students and wish to the authors a grand success.

Dr P Balakrishnan
Director of Technical Education
Government of Tamil Nadu
Chennai

Preface to the Second Edition

The past four decades have witnessed several unprecedented and exciting developments in the field of electronics. A large number of solid-state devices and integrated circuits incorporating millions of active devices on a single chip have been invented. A single textbook covering all aspects of Electronic Devices and Circuits satisfying the requirements of engineering college teachers and students is the need of the day. We have made a sincere attempt to bring out such a textbook.

The text is based on a series of lecture courses given to electronics, electrical and computer engineering undergraduates in their first, second and (part of) third years, and hence the coverage of the subject has been organized in a more logically acceptable manner. Since a proper understanding of the subject would involve a serious attempt to solve a variety of problems, a wide variety of problems with their step-by-step solutions are provided for every concept.

The book will serve the purpose of a text to the engineering students of degree, diploma, AMIE and graduate IETE courses and a useful reference for those preparing for competitive examinations. Also, it will meet the pressing need of those interested readers who wish to gain a sound knowledge and understanding of the principles of electronic devices and circuits. Practicing engineers will find the content significantly relevant to their day-to-day functioning.

This book is divided into 25 chapters. The first chapter introduces the physical properties of elements. Chapter 2 covers the passive circuit components, resistors, capacitors and inductors. Chapter 3 deals with electron ballistics. Chapter 4 is devoted to semiconductor diodes and their applications. Chapter 5 explains various types of special diodes including metal-semiconductor junctions. Chapters 6 and 7 deal with Bipolar Junction Transistors and Field Effect Transistors. Chapter 8 describes the different members of the thyristor family such as PNPN diode, SCR, LASER, TRIAC, DIAC and GTO. Chapter 9 deals with midband analysis of small signal amplifiers. Chapter 10 explains the multistage amplifiers. Chapter 11 concentrates on frequency response of amplifiers. Chapter 12 is devoted to large signal amplifiers. Chapter 13 covers the tuned amplifiers. Chapter 14 introduces feedback amplifiers. Chapter 15 deals with oscillators. Chapter 16 describes wave shaping and multivibrator circuits. Chapter 17 gives a brief description on UJT, pulse transformers, blocking oscillators and time base generators. Chapter 18 explains various

types of power supplies. Chapter 19 is devoted to IC fabrication. Chapter 20 deals with operational amplifiers and their applications. Chapter 21 covers different types of transducers. Chapter 22 contains a detailed study of opto-electronic devices including LCD panels and plasma display panels. Chapter 23 discusses a variety of electrical and electronic measuring instruments. Chapter 24 deals with digital circuits and Chapter 25 presents a study of memories, microprocessors and microcontrollers.

All the topics have been profusely illustrated with diagrams for easy understanding. Equal emphasis has been laid on mathematical derivations as well as their physical interpretations. Illustrative examples are discussed to emphasise the concepts and typical applications. Review questions and exercises have been given at the end of each chapter with a view to help the readers increase their understanding of the subject and to encourage further reading.

We sincerely thank the managements of SSN College of Engineering, Chennai, and Velammal College of Engineering and Technology, Madurai, for their constant encouragement and providing necessary facilities for completing this project. We express our deep gratitude to Dr. P. Balakrishnan, Former Director of Directorate of Technical Education, Government of Tamil Nadu for giving a Foreword to the first edition of this book. We thank our colleagues for their useful comments, which have improved the book considerably over the previous edition. We are thankful to Mr. S. Sankar Kumar for efficiently word processing the manuscript of the first edition. Our thanks are also due to Mr. R. Gopalakrishnan for word processing the additional manuscript.

We are grateful to the editorial and production teams including Mrs. Vibha Mahajan, Mr. Ebi John Amos and Ms. Sandhya Chandrasekhar of McGraw-Hill Education (India) Pvt. Ltd., New Delhi, for their initiation and support to bring out this revised edition in a short span of time.

Professor Salivahanan is greatly thankful to his wife, Kalavathy, and sons Santhosh Kanna and Subadesh Kanna; Professor Suresh Kumar expresses his heartfelt thanks to his wife, Andal, and daughters Sree Naga Gowri and Sree Naga Vani; Professor Vallavaraj expresses his warmest thanks to his wife, Radha, and son Nithish Raj; for their spirit of self-denial and enormous patience during the preparation of this book.

We welcome suggestions for the improvement of the book.

S. Salivahanan
N. Suresh Kumar
A. Vallavaraj

Preface to the First Edition

Explosive development and exciting progress has occurred in the field of electronics in the last three decades. A large number of solid-state devices and integrated circuits incorporating millions of active devices on a single chip have been discovered. A single textbook covering all aspects of Electronic Devices and Circuits satisfying the requirements of engineering college teachers and students is the need of the day. We have made a sincere attempt to bring out such a textbook.

The text is based on a series of lecture courses given to electronics, electrical and computer engineering undergraduates in their first, second and (part of) third years and hence the coverage of the subject has been organised in a more logically acceptable manner. The language used in explaining the various concepts is extremely simple. Since a proper understanding of the subject would involve a serious attempt to solve a variety of problems, a wide variety of problems with their step-by-step solutions are provided for every concept.

The book will serve the purpose of a text to the engineering students of degree, diploma, AMIE and graduate IETE courses and a useful reference for those preparing for competitive examinations. Also, it will meet the pressing need of interested readers who wish to gain a sound knowledge and understanding of the principles of electronic devices and circuits. Practicing engineers will find the content of significant relevance in their day-to-day functioning.

The book contains 21 chapters. The first chapter introduces the physical properties of elements. Chapter 2 covers the passive circuit components, resistors, capacitors and inductors. Chapter 3 deals with electron ballistics. Chapter 4 is devoted to semiconductor diodes and their applications. Chapter 5 explains various types of special diodes. Chapters 6 and 7 deal with Bipolar Junction Transistor and Field Effect Transistor. Chapter 8 describes the different members of thyristor family such as PNPN diode, SCR, LASER, TRIAC, DIAC and UJT. Chapter 9 deals with small signal models for transistors. Chapter 10 explains the various types of amplifiers. Chapter 11 deals with operational amplifiers and their applications. Chapter 12 introduces feedback amplifiers. Chapters 13 and 14 deal with oscillators and waveform generators respectively. Chapter 15 explains various types of power supplies. Chapter 16 is devoted to IC fabrication. Chapter 17 covers different types of transducers. Chapter 18 contains a detailed study of opto-electronic devices. Chapter 19

of a number of neutral particles called neutrons and a number of positively charged particles called protons. The charge of the nucleus is positive and equal to the number of protons contained in it. The charge of a proton is 1.6×10^{-19} coulomb and that of an electron is -1.6×10^{-19} coulomb. Each atom is ordinarily electrically neutral. Hence, in a neutral atom the number of revolving electrons must equal the number of protons in the nucleus. The mass of an electron is 9.107×10^{-31} kg, while that of a proton or neutron is about 1836 times that of the electron. Thus, almost the entire mass of an atom is concentrated in the nucleus. The radius of an atom is of the order of 1 Angstrom (10^{-10} m).

1.3 HYDROGEN ATOM

The first modern picture of an atom was put forward by Rutherford in 1911. Such an atom has a central nucleus of small dimensions around which move a number of electrons. The electron orbits are of atomic dimensions, i.e. of the order of 10^{-8} cm. The nucleus carries a charge to counteract the combined charge of the electrons. Almost all of the mass of the atom resides in the nucleus. The Rutherford model of the hydrogen atom is shown in Fig. 1.1. The hydrogen atom consists of one proton and one electron revolving around the centre nucleus in a circular orbit of radius r.

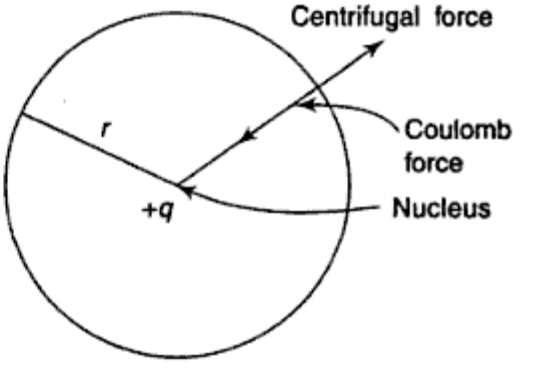


Fig. 1.1 Rutherford model of hydrogen atom

An electron in the circular orbit experiences a centripetal acceleration. According to electromagnetic theory, an accelerated electrical charge must radiate energy in the form of electromagnetic waves. The total energy of an electron in any orbit is the sum of its kinetic and potential energies. The potential energy of an electron is considered to be zero when it is at an infinite distance from the nucleus. The relation between the total energy possessed by an electron and the radius of the circular orbit r, is given below.

$$W = \text{potential energy} + \text{kinetic energy}$$

$$W = \frac{-Zq^2}{4\pi\varepsilon_0 r} + \frac{Zq^2}{8\pi\varepsilon_0 r} = \frac{-Zq^2}{8\pi\varepsilon_0 r} \tag{1.1}$$

where Z is the atomic number ($Z = 1$ for hydrogen); q, the charge of an electron $= 1.6 \times 10^{-19}$ coulomb; and ε_0, the permittivity of free space $= 8.854 \times 10^{-12}$ Farad/m.

If the accelerated electron loses energy by radiation, the total energy of the electron continuously decreases and it must spiral down into the nucleus. Thus, the hydrogen atom cannot be stable. But most of the atoms are stable.

According to classical electromagnetic theory, an accelerating electron must radiate energy at a frequency equal to the mechanical frequency of the orbiting electron and hence proportional to the angular velocity of the electron. Therefore,

as the electron spirals toward the nucleus, the angular velocity tends to infinity and hence, the frequency of the emitted energy will tend to infinity. This will result in a continuous spectrum with all possible wavelengths. But, many atoms like hydrogen emit line spectra of fixed wavelengths only.

The above two points are contradicting to what is normally happening in an atom. These two phenomena could not be explained using the Rutherford atomic model and are considered to be the main drawbacks of Rutherford model of an atom.

1.4 BOHR ATOM MODEL

Bohr atom model could successfully explain many of the atomic phenomena. Bohr took the Rutherford model of the atom and tried to overcome the defects of the model. Bohr proposed that the laws of classical mechanics and electromagnetics broke down within the atom. The basic postulates of Bohr's theory are a combination of the ideas of classical theory and Planck's quantum theory of radiation.

1.4.1 Bohr's Postulates

Bohr postulated three fundamental laws to overcome the inconsistency in the Rutherford model.

Postulate I: Electrons cannot occupy states at all energy levels as given by classical mechanics; electrons can occupy states at only certain discrete energy levels. When an electron occupies a state at one of these discrete energy levels, the electron does not emit radiation and is said to be in stationary or non-radiating state.

Postulate II: When an electron moves one stationary state corresponding to energy W_2 to another stationary state with lower energy W_1, there results emission of radiation at a frequency f given by

$$f = \frac{(W_2 - W_1)}{h} \tag{1.2}$$

where f is the frequency of radiation in hertz; h, the Planck's constant = 6.62×10^{-34} joule s; and W_1 and W_2 are energies in joules.

Postulate III: Any stationary or non-radiating state is determined by the condition that the angular momentum of the electron in this state is quantized and must be an integral multiple of $h/2\pi$. Thus,

$$mvr = \frac{nh}{2\pi} \tag{1.3}$$

where n is an integer. The radii of the various stationary orbits is given by

$$r_n = \frac{\varepsilon_0 h^2 n^2}{\pi m q^2} \tag{1.4}$$

$$= 0.527 \times 10^{-10} n^2 \text{ m} \tag{1.5}$$

The total energy of electron in stationary states in joules is given by

$$W_n = \frac{-mq^4}{8\varepsilon_0^2 h^2 n^2} \tag{1.6}$$

The total energy of electron in stationary states in electron volts is given by

$$W_n = \frac{-mq^3}{8\varepsilon_0^2 h^2 n^2}$$

$$= \frac{-13.6}{n^2} \text{ eV} \tag{1.7}$$

It should be noted that the energy is negative and therefore, the energy of an electron in its orbit increases as n increases. The above expression for the energy of the electron suggests that to remove an electron from the first orbit ($n = 1$) of the hydrogen atom to outside of the atom, that is to ionise the atom, the energy required is 13.6 eV. This is known as the *ionisation energy* or the *ionisation potential* of the atom.

1.4.2 Critical Potentials

The least energy, expressed in electron volts, required to excite a free neutral atom from its ground state to a higher state is called critical potential of the atom. It is usual to distinguish two kinds of critical potentials, namely, excitation potential and ionisation potential.

Excitation potential The energy in electron volts required to raise an atom from its normal state into an excited state is called excitation potential of the state. It is also called as radiation potential or resonance potential.

Ionisation potential It is defined as the energy required to remove an electron from a given orbit to an infinite distance from the nucleus.

For example, the energy associated with an electron in nth orbit of the hydrogen atom is $E_n = -13.6/n^2$ eV. Thus the energies of the first, second, third, ..., ∞ orbits are respectively –13.6, –3.4, –1.51, ..., 0 eV. The energy required to raise the atom from the ground state ($n = 1$) to the first excited state is (13.6 – 3.4) = 10.2 eV. The energy required to raise it to the second excited state is (13.6 – 1.51) = 12.09 eV and so on. It is clear that 10.2 eV, 12.09 eV are excitation potentials, while 13.6 eV is the ionisation potential of the hydrogen atom.

The number of ionisation potential depends on the number of electrons in an atom. For hydrogen atom, there is only one ionisation potential and several excitation potentials. The energy required for the removal of the outermost valence electrons is called the first ionisation potential. The energy required for the removal of the second electron from the influence of the nucleus is called the second ionisation potential and this will be higher than the first ionisation potential.

Example 1.1 Calculate the radii of the first, second and third permitted electron orbits in a Bohr's hydrogen atom.

Radius of the nth orbit for hydrogen, $r_n = \dfrac{\varepsilon_0 h^2 n^2}{\pi m q^2}$

$$= \frac{(8.854 \times 10^{-12})(6.62 \times 10^{-34})^2 n^2}{\pi (9.1 \times 10^{-31})(1.6 \times 10^{-19})^2} = 5.27 \times 10^{-11} n^2 \text{ m.}$$

Therefore, the radius of the first orbit is, $r_1 = 5.27 \times 10^{-11} \times 1^2$
$= 5.27 \times 10^{-11}$ m = 0.527 A.U.

The radius of the second orbit is $r_2 = 5.27 \times 10^{-11} \times 2^2 = 2.108$ A.U.

The radius of the third orbit is $r_3 = 5.27 \times 10^{-11} \times 3^2 = 4.743$ A.U.

1.4.3 Spectral Series of Hydrogen Atom

When an electron jumps from higher orbits to lower orbits, radiation of energy takes place at particular frequencies. When an electron jumps from second, third, ... etc. orbits to the first orbit, the spectral lines are in the ultraviolet region. This is identified as *Lyman* series. When an electron jumps from outer orbits to the second orbit, the series is called *Balmer* series and lies in the visible region of the spectrum. When the transition of an electron is from outer orbits to the third orbits the series is *Paschen* series and lie in the near infrared region. The transition from outer orbits to fourth and fifth orbits are respectively called as *Brackett* series and *Pfund* series. The spectral lines for these two series lie in the very far infra-red region of the hydrogen spectrum.

1.5 ATOMIC ENERGY LEVEL DIAGRAM

In the energy level diagram, the discrete energy states are represented by horizontal lines, and the height of the line represents the total energy E_n as calculated from Eqn. (1.7). Figure 1.2 shows the energy level diagram for hydrogen. The number immediately to the right of a line gives the value of integer n, while the number to the left of each line gives the energy to this level in electron volts. The lowest energy level E_1 is called the normal or the ground state of the atom and the higher energy levels E_2, E_3, E_4, ... are called the excited states. As n increases, the energy levels crowd and tend to form a continuum.

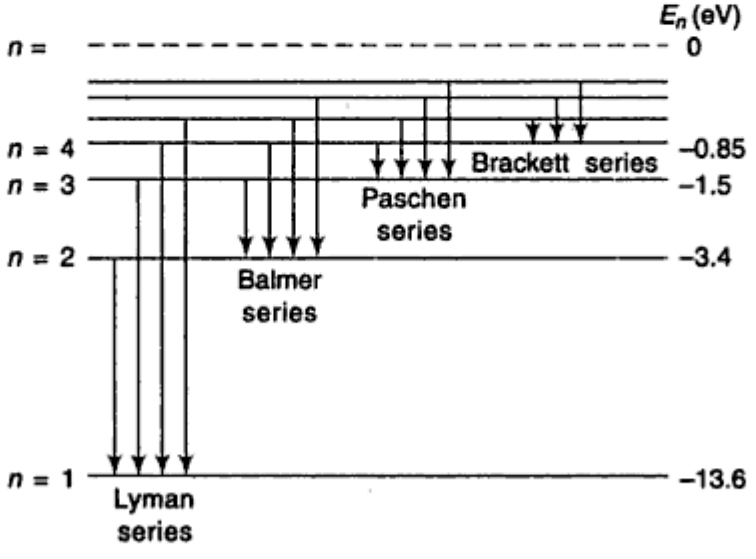


Fig. 1.2 Energy level diagram of hydrogen

Sometimes, it is more convenient to specify the emitted radiation by its wavelength, λ in Angstroms. Equation (1.2) can be rewritten as,

$$\lambda = \frac{12{,}400}{E_2 - E_1} \tag{1.8}$$

where E_2 and E_1 are the energy levels in electron volts.

Example 1.2 Find the wavelength of the photon emitted when a hydrogen atom goes from $n = 10$ state to the ground state.

The wavelength in Angstrom units is given by, $\lambda = \dfrac{12{,}400}{E_2 - E_1}$

Since the hydrogen atom goes from $n = 10$ state to the ground state, $\lambda = \dfrac{12{,}400}{E_{10} - E_1}$

The energy of the 10^{th} state is $E_{10} = \dfrac{-13.6}{10^2} = -0.136$ eV.

The energy in the ground state is $E_1 = -13.6$ eV.

Thus the wavelength of the emitted photon $= \dfrac{12{,}400}{-0.136 - (-13.6)} = 920.97$ Å

Example 1.3 Calculate the wavelength of the Balmer series limit.
When an electron jumps from outer orbits to the second orbit, the series is called Balmer series. Using Eqn. (1.8), the wavelength limit for the Balmer series can be found by calculating the wavelength of the radiation due to the transition of electron from the infinite orbit to the second orbit.

Wavelength of the Balmer series limit $= \dfrac{12{,}400}{E_\infty - E_2}$

Energy of the electron at the infinite orbit, $E_\infty = \dfrac{-13.6}{\infty^2} = 0$

Energy of the electron at the second orbit, $E_2 = -\dfrac{13.6}{2^2} = -3.4$

Therefore, the wavelength limit $= \dfrac{12{,}400}{E_\infty - E_2} = \dfrac{12{,}400}{3.4} = 3647$ A.U.

1.6 ELECTRONIC CONFIGURATION OF ELEMENTS

In order to understand the location and energy of each electron in an atom, four quantum numbers are required. The four quantum numbers are identified as follows:

1. Principal quantum number (n) This number characterises the average distance of an electron from the nucleus and corresponds to the principal energy level in which the electron resides. It gives some idea about the position of the electron around the nucleus. The allowed values of this quantum number n are positive integers starting from 1. Thus n can have values 1, 2, 3, ... The principal energy levels or shells having different values of n are also represented by the letters K, L, M, N and so on. The maximum number of electrons that can reside in a shell corresponding to the principal quantum number n is equal to $2n^2$.

2. Azimuthal quantum number (l) This quantum number is also called as the *orbital angular momentum quantum number.* The quantum number l gives a measure

Any crystal is made up of a space array of atoms of molecules in regular repetition in three dimensions of some fundamental building block. For a gaseous element, the electronic energy levels are the same as for a single free atom because the individual atoms in a gas are well apart and has negligible influence on each other. However, in a crystal the individual atoms are so closely packed that the resulting energy levels are modified due to interaction between the atoms. When atoms form crystals, the energy levels of the inner-shell electrons are not appreciably affected, however, the levels of the outer-shell electrons are considerably altered as these electrons are shared by the adjacent atoms in the crystal. As a result of this interaction between the outer-shell electrons, the energy levels spread up to form a band of energy as shown in Fig. 1.3.

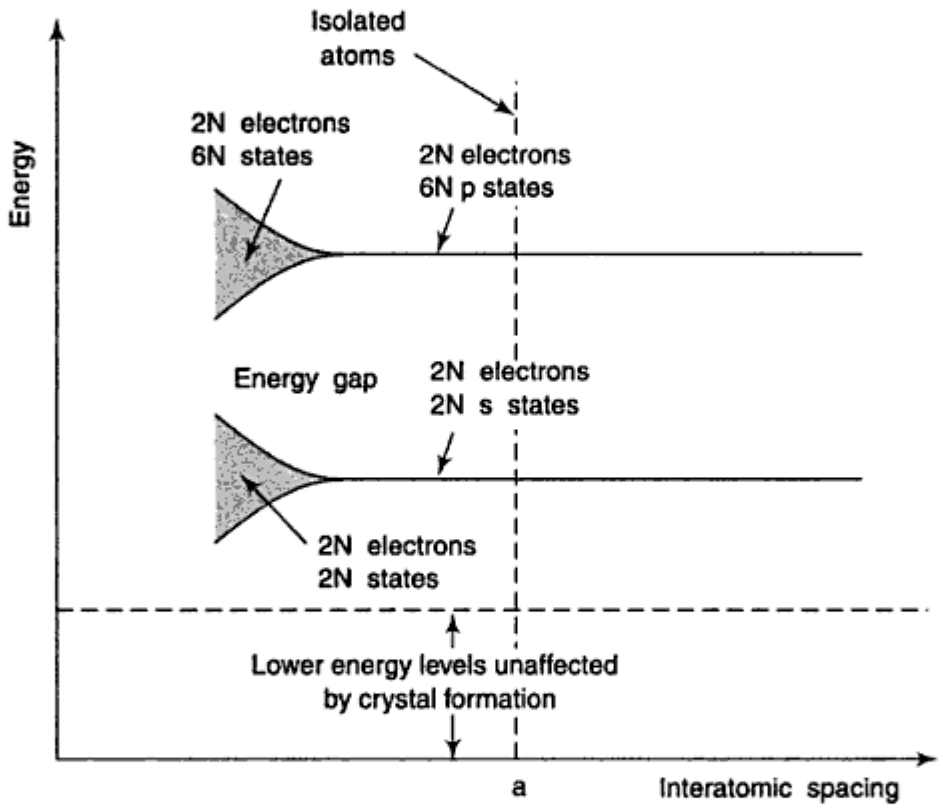


Fig. 1.3 Formation of energy band in a crystal

In order to understand the formation of energy bands in a crystal, let us consider a silicon crystal made up of N atoms. We shall also assume that the interatomic spacing can be varied without affecting the fundamental crystal structure. For very large interatomic spacing, say 'a' in Fig. 1.3, the interaction between adjacent atoms is negligible, and the energy levels are same as that of isolated atoms. In silicon, the outermost subshells, namely, 3s and 3p contain 2 electrons each. Hence, in a silicon crystal consisting of N atoms, the outermost subshells 3s and 3p consist of $2N$ electrons each. Thus, the 3s subshell has $2N$ electrons completely occupying the available $2N$ states, and the 3p subshell has only $2N$ electrons partially occupying the available $6N$ states, all at the same energy level.

If the interatomic spacing is gradually decreased, i.e., moving from right to left in Fig. 1.3, there will be a gradual increase in the interaction between the neighbouring atoms. Due to this interaction, the atomic wave functions overlap, and the crystal

becomes an electronic system which should obey the Pauli's exclusion principle. Hence, the $2N$ s states spread out to form a band of energy. Actually, the separation between levels is small but since N is very large, of the order of 10^{23} cm^{-3}, the total spread between the minimum and maximum energy levels becomes large. This spread will have several electron volts of energy and is referred to as energy band and is indicated by the lower shaded region in Fig. 1.3. The $2N$ states in this band are completely filled with $2N$ electrons.

Similarly, at the same value of interatomic spacing at p-level, $6N$ p states spread up to form a band. This band is shown as the upper shaded region in Fig. 1.3. Though a total of $6N$ states are available in this band, only $2N$ states are occupied and $4N$ states remain unoccupied.

An energy gap exists between the two energy bands. This energy gap is called as *forbidden energy gap*, as no electrons can occupy states in this gap. This forbidden energy gap decreases as the atomic spacing is decreased and becomes zero with further reduction in the interatomic spacing, say at 'b', as shown in Fig. 1.4. That is, the two energy bands will overlap when the interatomic spacing is small enough. Under such circumstances, the $6N$ p states in the upper band merge with the $2N$ s states in the lower band, giving a total of $8N$ states. Half of these $8N$ states are occupied by the $4N$ available electrons. These $4N$ electrons now no longer belong to either p subshell or s subshell but belong to the crystal as a whole. Thus, at this interatomic spacing, each atom in the crystal can contribute 4 electrons to the crystal. The band occupied by these contributed electrons is called the *valence band.*

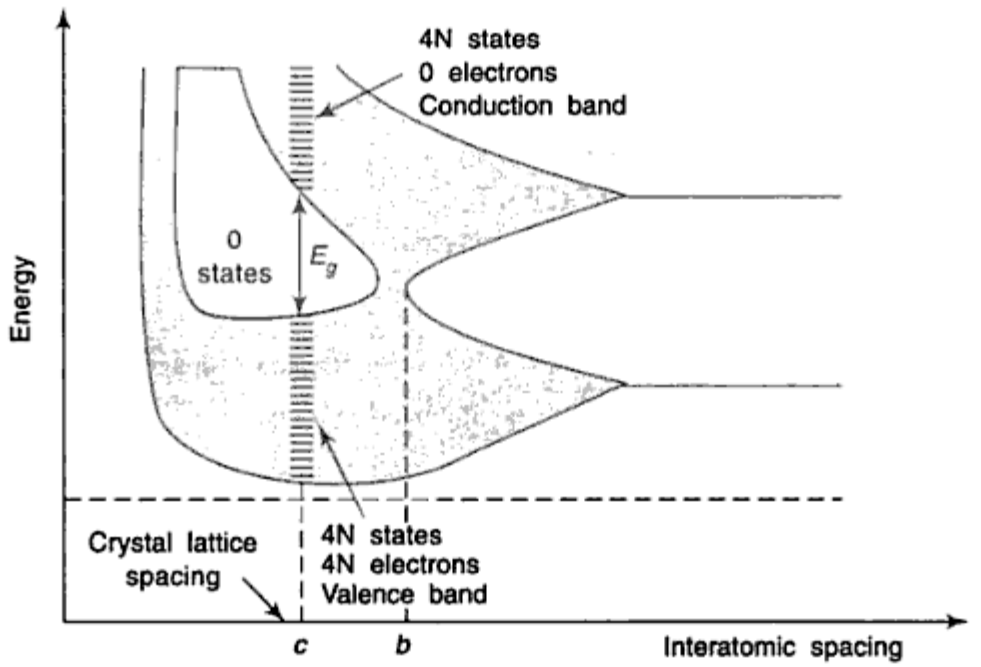


Fig. 1.4 Energy band in crystal as a function of interatomic spacing

If the interatomic spacing is reduced further, say 'c' as shown in Fig. 1.4, the interaction between the atoms becomes extremely large, and the energy band structure assumes the shape shown in Fig. 1.4. The exact energy band structure depends upon (i) the orientation of the atoms relative to one another in space, and (ii) the atomic number of the atom, and may be obtained from solution of Schrodinger's wave equation. For spacing 'c', $4N$ states in the valence band are completely filled

by $4N$ electrons and this valence band is separated from the upper unfilled or empty band by a forbidden energy gap E_g. The forbidden energy gap contains no allowed states. The upper band has $4N$ unfilled states and is referred to as the *conduction band.*

1.8 ENERGY-BAND STRUCTURES AND CONDUCTION IN INSULATORS, SEMICONDUCTORS AND METALS

A very poor conductor of electricity is called an Insulator; an excellent conductor is a metal; and a material whose conductivity lies between these two extremes is a semiconductor. A material may be classified as one of these three depending upon its energy-band structure.

1.8.1 Insulator

An insulator is a material having extremely poor electrical conductivity. The energy-band structure of Fig. 1.4 at the normal lattice spacing is indicated schematically in Fig. 1.5a. The forbidden energy gap is large; for diamond the gap energy is about 6 eV. If additional energy is given to an electron in the upper level of valence band, this electron attempts to cross the forbidden energy gap and enter the conduction band. However in an insulator, the additional energy which may ordinarily be given to an electron is, in general, much smaller than this high value of forbidden energy gap. Hence no electrical conduction is possible. The number of free electrons in an insulator is very small, roughly about 10^7 electrons/m^3.

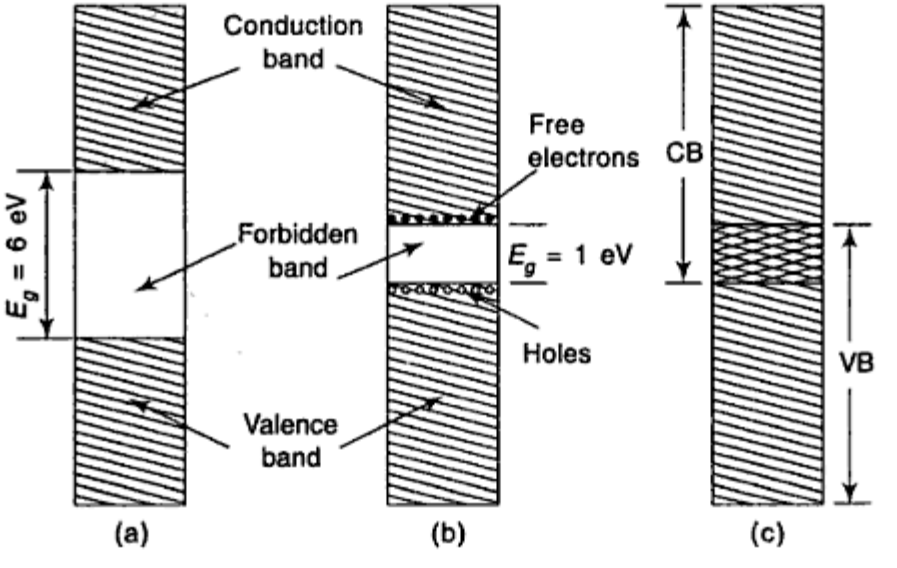


Fig. 1.5 Energy-band gap in (a) Insulators, (b) Semiconductors, and (c) Metals

1.8.2 Metal

The conduction in metals is only due to the electrons. A metal has overlapping valence and conduction bands. The valence band is only partially filled and the conduction band extends beyond the upper end of filled valence band. The outer electrons of an atom are as much associated with one ion as with another, so that the electron attachment to any individual atom is almost zero. The band occupied by the valence electrons may not be completely filled and that there are no forbidden

levels at higher energies. Depending upon the metal, at least one, and sometimes two or three, electrons per atom are free to move throughout the interior of the metal under the action of applied fields. When an electric field is applied, few electrons may acquire enough additional energy and move to higher energy within the conduction band. Thus the electrons become mobile. Since the additional energy required for transfer of electrons from valence band to conduction band is extremely small, the conductivity of metal is excellent.

In *electron-gas theory* description of a metal, the metal is visualized as a region containing a periodic three-dimensional array of heavy, tightly bound ions permeated with a swarm of electrons that may move about quite freely. According to this theory, the electrons in a metal are continuously moving and the direction of flight changes whenever the electron collides with other electrons. The average distance travelled by an electron between successive collisions is called as *mean-free-path* of an electron. In the absence of any applied potential, the average current in a metal is zero because the number of electrons passing through unit area in any direction is almost same as the number of electrons passing through the same unit area in the opposite direction. This can be attributed to the random nature of motion of electrons.

When a constant electric field E (volts per metre) is applied to a metal, the electrons would be accelerated and the velocity would increase indefinitely with time. However, because of collision of electrons, electrons lose energy and a steady-state condition is reached where a finite value of drift velocity v_d is attained. The drift velocity, v_d is in the direction opposite to that of the electric field and its magnitude is proportional to E. Thus

$$v_d = \mu E \tag{1.9}$$

where μ = mobility of the electron, m^2/volt-second. Due to the applied field, a steady-state drift velocity has been superimposed upon the random thermal motion of the electrons. Such a directed flow of electrons constitutes a current. If the concentration of free electrons is n (electrons per cubic meter), the current density J (amperes per square metre) is

$$J = nqv_d = nq\mu E = \sigma E \tag{1.10}$$

where

$$\sigma = nq\mu \tag{1.11}$$

σ is the *conductivity* of the metal in (ohm-metre)$^{-1}$. For a good conductor n is very large, approximately, 10^{28} electrons/m^3. Equation (1.10) can be recognized as Ohm's law which states that the conduction current density is proportional to the applied electric field. The energy acquired by the electrons from the applied field is given to the lattice ions as a result of collisions. Hence, power is dissipated within the metal by the electrons, and the power density (Joule heat) is given by

$$JE = \sigma E^2 \text{ watts/metre}^3 \tag{1.12}$$

1.8.3 Semiconductor

The conductivity of a material is proportional to the concentration of free electrons. The number of free electrons in a semiconductor lies between 10^7 and 10^{28} electrons/m^3. Thus, a semiconductor has conductivity much greater than that of an insulator but much smaller than that of a metal. Typically, semiconductor has

forbidden energy gap of about 1 eV. The most important practical semiconductor materials are germanium and silicon, which have values of E_g of 0.785 and 1.21 eV, respectively, at 0 degree Kelvin. Energies of this magnitude normally cannot be acquired from an applied field. At low temperatures the valence band remains full, the conduction band empty, and these materials are insulators at low temperatures. The conductivity of these materials increases with temperature and hence these materials are called as intrinsic semiconductors. As the temperature is increased, some of the electrons in the valence band acquire thermal energy greater than the gap energy and move into the conduction band. These electrons are now free to move about under the influence of even a small applied field. These free electrons, also called as conduction electrons, constitute for conduction and the material becomes slightly conducting. The current density due to the motion of electrons is given by

$$J_n = n\mu_n qE = \sigma_n E \tag{1.13}$$

where, μ_n is the electron mobility, and the suffix 'n' represents that the respective terms are due to motion of electrons. The absence of an electron in the valence band is represented by a small circle and is called a hole. The hole may serve as a carrier of electricity whose effectiveness is comparable with the free electron. The hole conduction current density is given by

$$J_p = p\mu_p qE = \sigma_p E \tag{1.14}$$

where μ_p is the hole mobility and p is the hole concentration.
Hence, the total current density J in a semiconductor is given by

$$J = (n\mu_n + p\mu_p)qE = \sigma E \tag{1.15}$$

where $\sigma = (n\mu_n + p\mu_p)q$ is the total conductivity of a semiconductor. For a pure semiconductor (intrinsic semiconductor), the number of free electrons is exactly same as the number of holes. Thus, the total current density is

$$J = n_i(\mu_n + \mu_p)\, qE \tag{1.16}$$

where $n_i = n = p$ is the intrinsic concentration of a semiconductor.

The conductivity of an intrinsic semiconductor can be increased by introducing certain impurity atoms into the crystal. This results in allowable energy states which lie in the forbidden energy gap and these impurity levels also contribute to the conduction. Such a semiconductor material is called an extrinsic semiconductor.

Effective mass An electron travelling through a crystal under the influence of an externally applied field hardly notices the electrostatic field of the ions making up the lattice, i.e. it behaves as if the applied field were the only one present. This is the basis of the electron gas approximation and the assumption will now be looked at a little more closely. If the electrons were to experience only the applied field, E, then immediately after a collision it would accelerate in the direction of the field with an acceleration a, proportional to the applied force, i.e.

$$qE = ma \tag{1.17}$$

where the constant of proportionality is the electron mass, m. When quantum theory is applied to the problem of an electron moving through a crystal lattice, it predicts that under the action of an applied field the electron acceleration will indeed be

At high temperatures, the electrons are literally being 'boiled' from the metal surface. This is analogous to the boiling of water.

The commonly used thermionic emitters with low E_w, as given in Table 1.4, are tungsten in high voltage (KV) tubes, thoriated tungsten in high power (KW) tubes and oxide coated metals in low power electron tubes and Cathode Ray Tubes (CRT).

Table 1.4 Properties of commonly used thermionic emitters

Thermionic emitter	*Work function in eV*	*Range of operating temperature in °K*	*Emission efficiency in mA/W*
Tungsten	4.52	2500 – 2600	2 – 10
Thoriated tungsten	2.63	1900 – 2000	50 – 100
Oxide-coated	1.10	900 – 1100	100 – 1000

(ii) Field emission When the potential difference between two electrodes is extremely high, electrons are emitted from the negative electrode, even at ordinary temperatures. If the electric field at the metallic surface of the negative electrode is of the order of 10^6 volts per metre, electrons are quite easily pulled out from the surface. This process of electron emission is known as *field emission*. Field emission does not depend on the temperature of the surface. This type of emission is employed in cold cathode devices and mercury and rectifier tube.

(iii) Photo-electric emission When the surface of certain metals are illuminated by a beam of light, electrons are ejected out of the metal by the light photons incident on the metal surface. Such an electron emission is known as *photo-electric emission* and is used in photo-electric cell. The greater the intensity of light beam falling on the metal surface, the greater is photo-electric emission.

(iv) Secondary emission When a beam of high velocity electrons strike a metal surface, the free electrons are ejected out of the metal. This process is known as *secondary emission*.

Review Questions

1. Define the term 'Electronics.'
2. List the major areas of applications of electronics.
3. Give the order of radius of an electron and an atom.
4. Describe Rutherford's model of the atom and what are the drawbacks of this modes?
5. State the postulates of Bohr regarding his atom model. Obtain expressions for the radius and electron-energy of the nth orbit.
6. Calculate the energy of the electron in the nth orbit in hydrogen atom.
   [*Ans.* $E_n = -13.6/n^2$ eV]
7. What is meant by the term ground state of an atom?
8. What are the wavelengths of the first three lines of the Paschen series?
   [Ans. $\lambda_1 = 18750$ Å, $\lambda_2 = 12810$ Å and $\lambda_3 = 10930$ Å]
9. Draw and discuss the energy level diagram of hydrogen atom.
10. What is meant by ionization potential of an atom?

11. Calculate (i) ionisation potential, (ii) first excitation potential of the hydrogen atom. [*Ans.* 13.6 eV, 10.2 eV]
12. What are the four quantum numbers?
13. State and explain Pauli's exclusion principle.
14. What is meant by the terms electronic shell and sub-shell?
15. Give the electronic configuration of Si, Ge, Sn and Pb.
16. Explain why the energy levels in a crystal split up to form energy bands.
17. Describe the energy band theory of a crystal.
18. What is meant by forbidden energy gap?
19. Describe the energy band structures of an insulator, a metal and a semiconductor.
20. Define mean-free-path of an electron.
21. Define drift velocity of an electron.
22. What is meant by effective mass of electron? Explain.
23. What are the three commonly used semiconductors?
24. What are the band gap energies of germanium, silicon and gallium arsenide?
25. Explain how electrons are emitted from metals.
26. Explain the term work function.
27. Discuss briefly the different types of electron emission.
28. Give the Richardson – Dushman equation and explain the effect of various factors in it.

2 Passive Circuit Components

2.1 INTRODUCTION

A passive component is one that contributes no power gain (amplification) to a circuit or system. It has no control action and does not require any input other than a signal to perform its function. The most commonly used passive circuit components in electronic and electrical applications are resistors, capacitors and inductors. The constructional features, characteristics and applications of these components are discussed in this chapter.

2.2 RESISTORS

Physical materials resist the flow of electrical current to some extent. Certain materials such as copper offer very low resistance to current flow, and hence they are called *conductors*. Other materials such as ceramic which offer extremely high resistance to current flow are called *insulators*. In electric and electronic circuits, there is a need for materials with specific values of resistance in the range between that of a conductor and an insulator. These materials are called resistors and their values of resistance are expressed in ohms (Ω).

The resistance R of a given material is proportional to its length L and inversely proportional to its area of cross section A. Thus R varies as L/A and $R = \rho(L/A)$, where ρ is the constant of the material known as its *specific resistance* or *resistivity*.

The various types of resistors are given in Table 2.1.

Table 2.1 Various types of resistors

Fixed resistors	*Variable resistors*
Carbon composition	Potentiometer
Carbon film	Rheostat
Metal film	Trimmer
Wire wound	

2.2.1 Fixed Resistors

Carbon composition This is the most widely used fixed resistor in discrete circuits. The construction of the carbon composition resistor is shown in Fig. 2.1. The

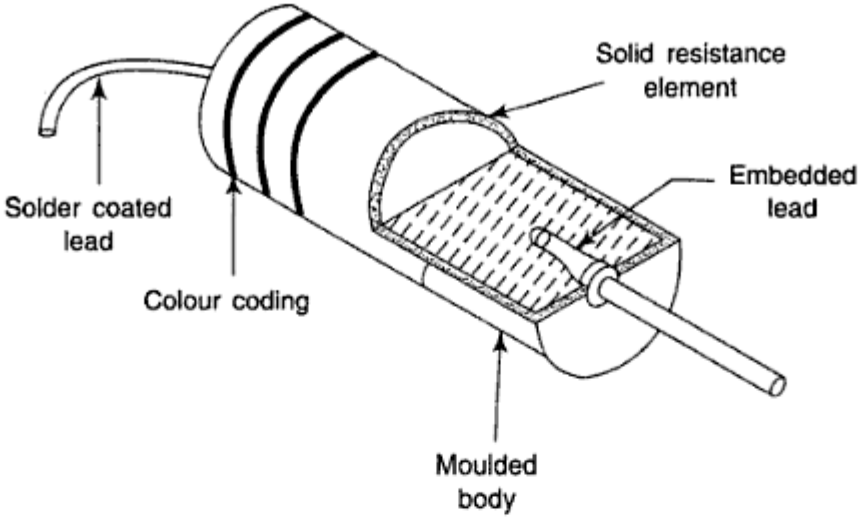


Fig. 2.1 Cutaway view of a carbon composition resistor

carbon resistors are made of finely divided carbon mixed with a powdered insulating material such as resin or clay in the proportions needed for the desired resistance value. These are then placed in a casing (moulded plastic) with lead wires of tinned copper. Resistances of this type are available in the range from few ohms to hundred megaohms and typical power ratings of 1/8 to 2 W.

Carbon film This is yet another type of carbon resistor. Its basic structure is shown in Fig. 2.2. It is manufactured by depositing a carbon film on a ceramic substrate. In this process only approximate values of resistance are obtained by either trimming the layer thickness or by cutting helical grooves of suitable pitch along its length. During this process, the value of the resistance is monitored constantly. Cutting of grooves is stopped as soon as the desired value of resistance is obtained. Contact caps are fitted at both ends, and then the lead wires made of tinned copper are welded to these end caps. This type of resistors are available commercially in the range of 10 Ω to 10 MΩ with a power rating of up to 2 W. Carbon film resistors are less noisy than carbon composition resistors and they are of low cost.

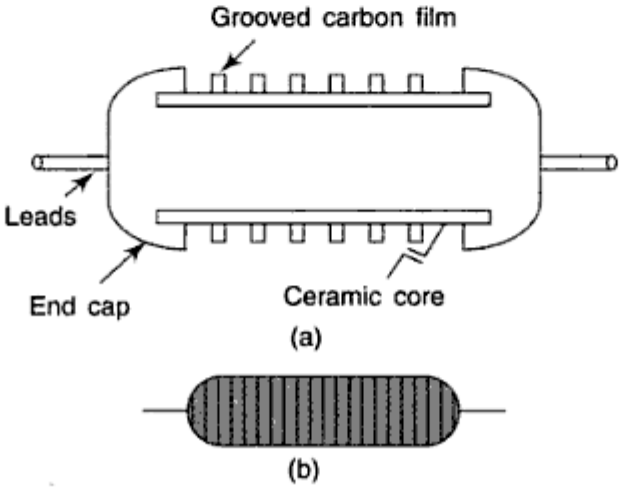


Fig. 2.2 Carbon film resistor (a) Construction, and (b) Carbon film resistor

Metal film resistor Construction of this type of resistor is similar to the carbon film resistors, the only difference being the material used for constructing the film. In metal film resistors, the film usually consists of an alloy of tin and antimony metals. Metal film resistors are available as thin and thick film type components.

(a) Thin film resistor: The resistance element in this type is a film having a thickness of the order of one-millionth of an inch. Typically, the thin film is deposited on a ceramic substrate under a high vacuum and this technique is called vacuum deposition. Metals used for deposition include nickel and chromium. This type of resistors are available in the range of 10 Ω to 1 MΩ with a power rating of up to 5 W.

(b) Thick film resistor: The resistance element in this type is a film having a thickness greater than one-millionth of an inch. Four different types of thick film resistors are available, viz. tin oxide, metal-glaze, cermet and bulk film resistors.

(i) Tin oxide type In this type of resistors, tin oxide in vapour form is usually deposited on a ceramic substrate under high temperature. The vapour reacting with the substrate, which is heated, results in a tightly formed resistance film. This type of resistor is available in the range of a few ohms to 2.5 MΩ with a power rating of up to 2 W.

(ii) Metal-glaze type In this type, a powdered glass and fine metal particle (palladium and silver) mixture is deposited on a ceramic substrate. This combination is then heated to a high temperature, typically 800 °C. This results in a fusion of metal particles to the substrate. This type of resistors are manufactured in the range of a few ohms to 1.5 MΩ with a power rating of up to 5 W.

(iii) Cermet type A cermet film resistor is made by screening a mixture of precious metals and binder materials on a ceramic substrate. The word 'cermet' is derived from ceramic and metal. As in the case of metal glaze resistor, the combination is then heated to a high temperature. It is available in the range of 10 Ω to 10 MΩ with a maximum power rating of 3 W.

(iv) Bulk film type In this type, the metal film is etched on a glass substrate. As metal film and glass have unequal coefficients of expansion, the metal film is compressed slightly by the glass substrate. The compressed film has a negative temperature coefficient which cancels out the inherent positive temperature coefficient of the film. As a result, the bulk film resistor has a temperature coefficient close to zero. This type of resistor is available in the range of 30 Ω to 600 kΩ with a maximum power rating of 1 W.

Wire wound resistor It has a wide range of applications. It is used as an ultra precision resistor in instrumentation and a power resistor in industrial applications. Wire wound resistors are manufactured in two types: (i) Power style wire wound resistor, and (ii) Precision style wire wound resistor.

(i) Power style type: It is made by winding a single layer length of special alloy wire, in the form of a coil around an insulating core. The ends of the winding are attached to metal pieces inserted in the core. Tinned copper wire leads are attached to metal pieces. The unit is then covered with a coating, such as vitreous enamel (an inorganic glass like moisture) or silicone. This coating protects the winding against moisture and breakage. The resistance wire lead must have carefully controlled resistance per unit length of wire and low temperature coefficient, and be able to

formed on aluminium), the resistance wire, usually copper alloy wire for low resistance pots and the nickel chromium for high resistance pots, is wound and then the strip is bent round a cylindrical surface. Contact by means of a slider metal or beryllium copper, which is spring loaded, is made on the inside periphery or the outer edge. Contact from the slider is made through a slip ring or by a coiled spring. The winding is usually of two or three linear resistance sections to approximate ideal taper (such as log). Single turn pots are available in the range from 50 Ω to 5 MΩ and in power ratings of 2 to 3 W. It is commonly used as a gain control element in an amplifier and as brightness and contrast controls in TV receivers.

(ii) Multiturn wire wound potentiometer Multiturn or helical pots are used in applications that require precise setting of a resistance value. An example is the setting of coefficients in an analog computer. Commonly, they have up to 10 turns. As shown in Fig. 2.6(b), the resistance element is wound on a long strip and then formed into a helix and held in a plane using silicone varnish. The contact is of precision metal and has multiple "fingers". Usually, the groove between the helical turns of the element may be used to guide the contact. Because of its construction, the wire wound pot has appreciable stray inductance and capacitance which may be a problem at high frequency operation. They are available in the range from 50 Ω to 250 kΩ with a power rating of up to 5 W.

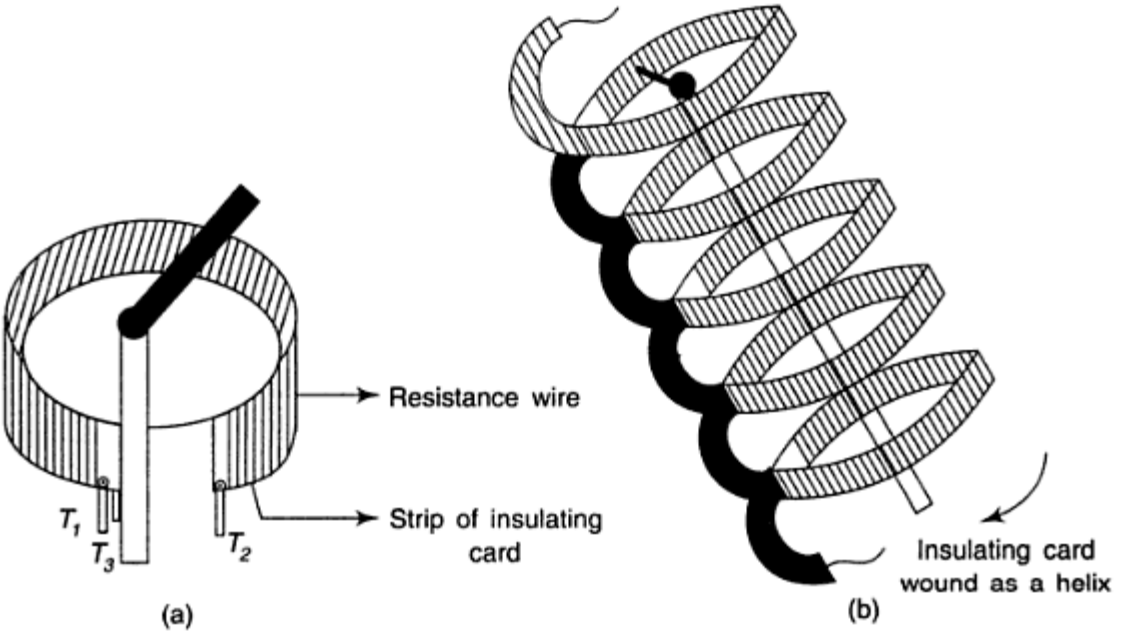


Fig. 2.6 (a) Single turn wire wound potentiometer, and (b) Multiturn (helical) potentiometer

Rheostat A wire wound pot that can dissipate more than 5 W is referred to as a rheostat. As shown in Fig. 2.7, the resistance wire is wound on an open tube of ceramic which is covered with vitreous enamel, except for the track of the movable contact. The rheostat is capable of withstanding temperatures up to 300 °C. It is used to control motor speed, X-ray tube voltages, welding current, ovens and in many other high power applications.

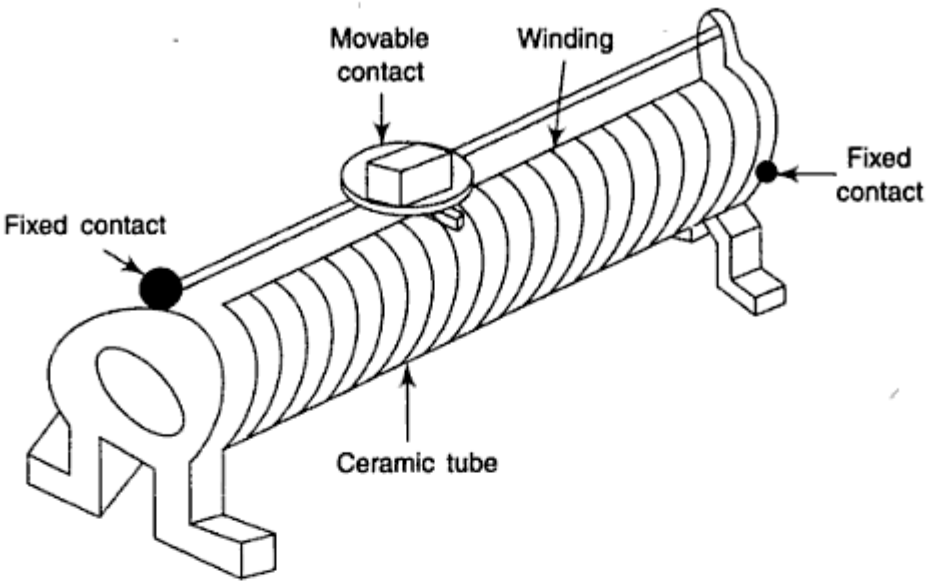


Fig. 2.7 Rheostat

Trimmer A trimmer or trimming potentiometer is used where the resistance must be adjustable but not continuously variable. It finds a very important place in calibration and balancing of electronic equipments. These are screw activated and the resistance can be adjusted by a screw driver. Figure 2.8 shows the commonly used rotary trimmer. Here, the resistance track is made of carbon and cermet. The carbon track along with the movable slide is housed in a steel casing. Typically, their resistance range is from a few ohms to 5 MΩ and the power rating is 1 W.

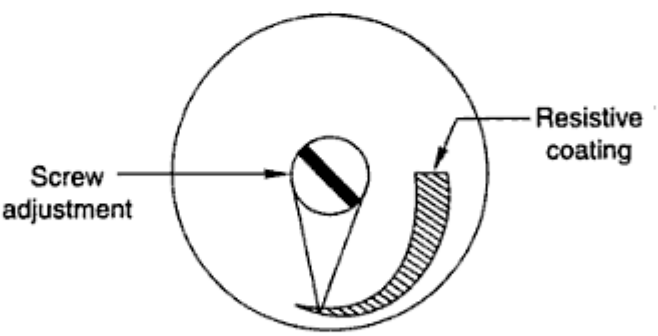


Fig. 2.8 Rotary trimmer

2.2.3 Tolerance

In the manufacture of resistors where thousands of resistors are made in a day, it is not possible to adjust every ordinary resistor to an exact value. The term tolerance denotes the acceptable deviation in the resistance value of a resistor. The usual specified tolerances are 5%, 10% and 20% for ordinary resistors, while precision resistors have a tolerance close to 0.1%. For example, the resistance marked as 1 kΩ with a 10% tolerance can have any value between 900 Ω 1100 Ω.

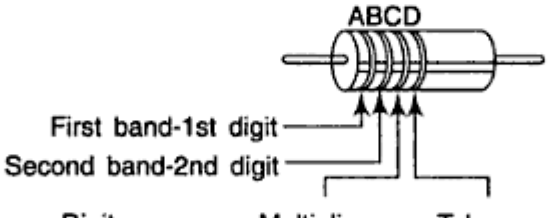
Table 2.2 gives the tolerance values for the various types of resistors.

Table 2.2 Tolerances for various resistors

Type of resistor	*Tolerance*
Fixed	
1. Carbon composition	5% or above
2. Carbon film	5% or above
3. Metal film	
(a) Thin film	< 0.5%
(b) Thick film	
(i) Tin oxide	< 1%
(ii) Metal-glaze	< 1%
(iii) Cermet type	≥ 1%
(iv) Bulk film	≥ 0.005%
4. Wire wound	
(a) Power style	5% to 20%
(b) Precision style	< 0.5%
Variable	
1. Potentimeter	
(a) Carbon type	10% to 20%
(b) Wire wound	
(i) Single turn	10% to 20%
(ii) Multi turn	3%
2. Trimmer	10%

2.2.4 Colour Coding of Resistors

Resistors are coded to indicate the resistance value and tolerance. As shown in Fig. 2.9, there are four colour bands (A, B, C and D), one by the side of the another starting from the left end. The first two bands (A and B) denote the first and second



Color	Digit	Multiplier	Tolerance
Black	0	1	—
Brown	1	10	±1%
Red	2	100	±2%
Orange	3	1000	—
Yellow	4	10000	—
Green	5	100000	—
Blue	6	1000000	—
Violet	7	10000000	—
Gray	8	—	—
White	9	—	—
Gold	—	0.1	±5%
Silver	—	—	±10%
No color	—	—	±20%

Fig. 2.9 Standard color code for composition and some axial type resistors

digits of the resistance value and the third band (C) indicates how many zeros follow the first two digits. Tolerance is given by the fourth band (D).

Figure 2.9 shows the colour code and tolerance values for the various colours. For example, a resistor with the following colour bands sequence

Yellow-Violet-Orange-Gold

denotes a 47000 Ω with 5% tolerance resistor.

2.2.5 Connecting Resistors

Resistors in series Figure 2.10 shows n fixed resistors connected in series. The current I flowing through each resistor is the same in a series circuit. The total equivalent resistance of the series circuit R_{TS} is equal to the sum of the individual resistances, i.e.

$$R_{TS} = R_1 + R_2 + ... + R_n$$

When all resistors have equal value, then $R_{TS} = n\,R$

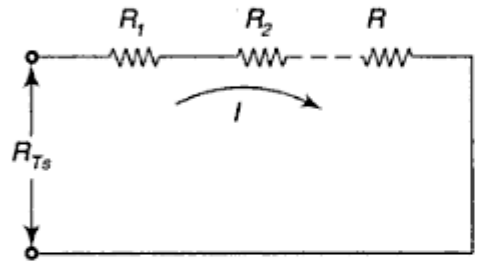


Fig. 2.10 n fixed resistors connected in series

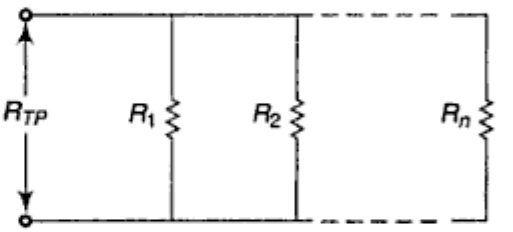


Fig. 2.11 n fixed resistors connected in parallel

Resistors in parallel Figure 2.11 shows n fixed resistors connected in parallel. Here, the voltage V across each resistor is the same. The total equivalent resistance of the parallel circuit R_{TP} can be calculated by

$$\frac{1}{R_{TP}} = \frac{1}{R_1} + \frac{1}{R_2} + ... + \frac{1}{R_2}$$

Considering two resistors R_1 and R_2 connected in parallel, the total equivalent resistance of the parallel circuit R_{TP} is equal to their product divided by their sum, i.e.

$$R_{TP} = \frac{R_1 R_2}{R_1 + R_2}$$

When all resistances are equal, then $R_{TP} = \frac{R}{n}$.

2.3 CAPACITORS

Next to resistors, capacitors are the most widely used passive elements in circuits. Capacitors are the devices which can store electric charge. Capacitors may be used to build up sufficient electric charge to fire, for example, flash tube or a laser. They are used in tuned circuits, timing circuits, filters, amplifier circuits, oscillator circuits and relay circuits. They are also used for power factor correction and for starting single phase motors.

The reactance of a capacitor of capacitance C at frequency f is given by

$$X_C = \frac{1}{2\pi f C}\ \Omega$$

The capacitive reactance (X_C) varies inversely with the frequency of the applied a.c. voltage. Therefore, the capacitor allows higher frequency currents more easily than lower frequency currents. For d.c. voltages, i.e. $f = 0$, $X_C = \infty$. Hence, a capacitor blocks (cannot conduct) the d.c. voltage or current.

A capacitor essentially consists of two conducting plates separated by a dielectric material as shown in Fig. 2.12. The capacitance of a parallel-plate capacitor is given by

$$C = \varepsilon_o \varepsilon_r \frac{A}{d}$$

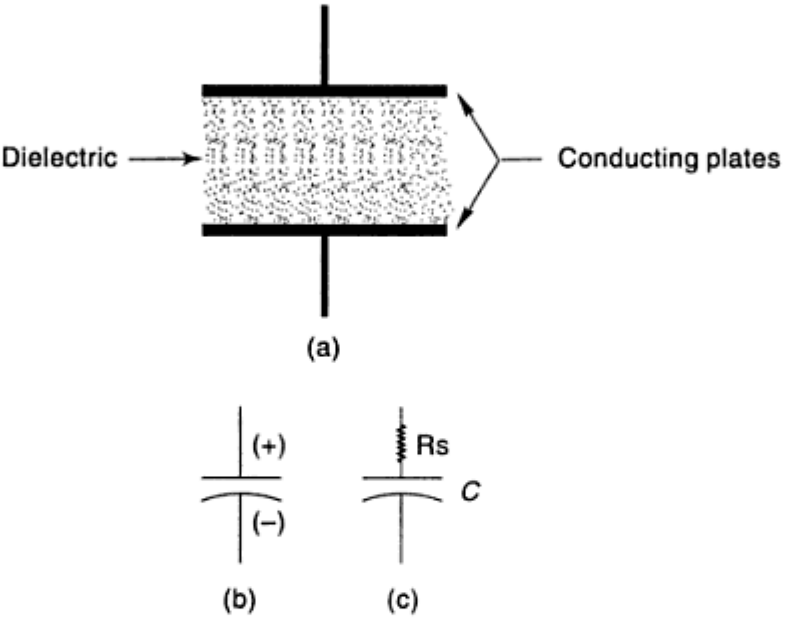


Fig. 2.12 Capacitor (a) Basic structure (b) Symbol and (c) Series equivalent circuit

where A = Area of each plate in m^2

d = distance between parallel plates in m

ε_o = dielectric constant (permittivity) of free space $= \frac{10^{-9}}{36\pi}$ F/m

$= 8.854 \times 10^{-12}$ F/m

ε_r = relative dielectric constant (permittivity).

For large capacitance, the area A and the value of the relative permittivity of the dielectric ε_r, must be large, while d must be very small. Larger capacitances can be obtained by using high permittivity dielectrics of smaller thickness.

When a voltage V is applied across the capacitor plates, the electrons accumulate on the side of the capacitor connected to the negative terminal of the voltage source. The plate connected to the positive terminal of the voltage source loses electrons. This accumulation of electrons produces a negative charge on one side of the capacitor, while the opposite side has a positive charge. Thus, with the application of voltage, the electrons are simply redistributed from one side of the capacitor to

2. Non-polarised electrolytic capacitor: The basic structure of a non-polarised electrolytic capacitor is shown in Fig. 2.13(b). This type has two oxide coated anodes. For an equal size polarized capacitor, the non-polarized type has one-half the capacitance for the same voltage rating. They are generally used in applications such as a.c. motor starting, crossover networks and large pulse signals.

(a) Aluminium type: This type of electrolytic capacitors are available in polarised or non-polarised form. The construction of this type of capacitors is similar to the construction of polarised or non-polarised capacitors, discussed in the previous section. This type of capacitors have high d.c. leakage and low insulation resistance. They are low in cost and have a high volumetric efficiency. The disadvantages are that their shelf-life is limited and their capacitance deteriorates with time and use.

(b) Tantalum type: Tantalum electrolytic capacitors have long shelf-life, stable operating characteristics, increased operating temperature range and greater volumetric efficiency. The disadvantages of tantalum type in comparison to the aluminium type are its greater cost and lower voltage rating.

Tantalum electrolytic capacitors can be divided into three types, namely, (i) Foil, (ii) Wet anode and (iii) Solid anode.

(i) Foil type The tantalum foil type is similar in construction to aluminium foil electrolytic type.

(ii) Wet anode type As shown in Fig. 2.14, the wet anode type capacitor is made by moulding a mixture of tantalum powder and binder into the shape of a pellet. Under high temperature and in a vacuum, the pellet mixture is welded together (sintered) and due to this sintering, the binder and impurities are driven off. The result is a porus pellet on which a layer of tantalum oxide is electrochemically formed. The volumetric efficiency of wet anode type is about three times that of the foil type.

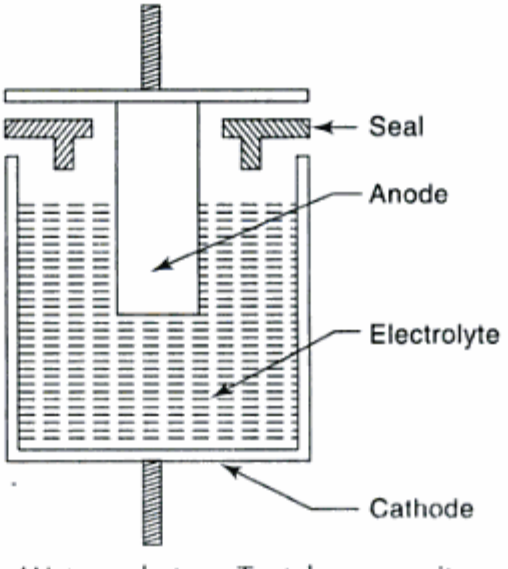


Fig. 2.14 Wet anode type Tantalum capacitor

(iii) Solid anode type As shown in Fig. 2.15, this type of capacitor is made by sintering an anode pellet on which a layer of tantalum oxide is formed. The pellet is then formed with a manganese dioxide layer which serves as a solid dielectric. The cathode connection is made by coating this pellet by carbon and silver paint. The

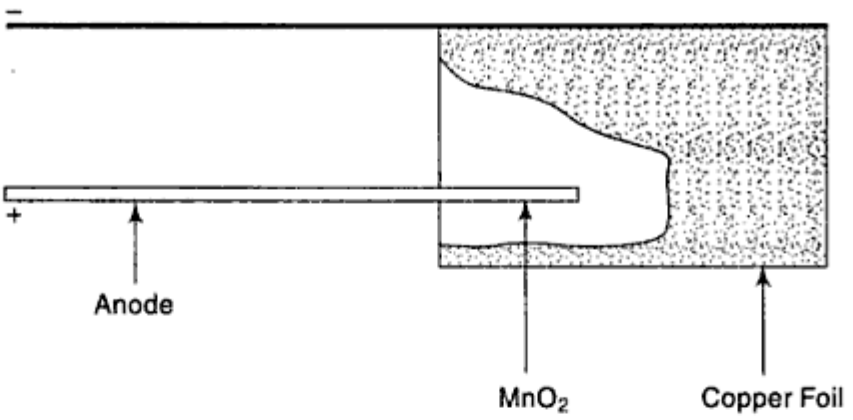


Fig. 2.15 Solid-anode type tantalum capacitor

solid-anode construction is most widely used. It has the longest life and lowest leakage current of the three types of tantalum capacitors.

Ceramic capacitors These capacitors use ceramic dielectric with thin metal films as electrodes bonded to the ceramic. There are three basic types of a ceramic dielectrics: (i) low permittivity, (ii) medium permittivity, and (iii) high permittivity types.

Low permittivity ceramic capacitors can be made to exhibit zero temperature coefficient and they find wide use in temperature compensation networks. The permittivity usually ranges from 6 to 400. It operates at voltages as high 6 KV and temperatures up to 125 °C. Its capacitance value is generally limited to 0.001 μF.

In the medium permittivity ceramic capacitors, the permittivity is in the range from 500 to about 4000 and is relatively stable over a wide temperature range of – 55 to + 125 °C with a maximum capacitance change of ±15%.

In the high permittivity ceramic capacitors, the permittivity ranges from 5000 to 30,000. They have capacitance values up to 2.2 μF and maximum working voltage of 100 V. However, these capacitors change their values appreciably with temperature, d.c. voltage and frequency.

The medium and high permittivity ceramic capacitors are used for bypass and decoupling applications or frequency discrimination where Q-factor and stability are not of major importance.

The ceramic capacitors are available as disc, tubular, monolithic and barrier type capacitors.

Disc capacitor: In the disc type capacitor, silver is fixed on to both sides of ceramic to form conductor plates. The sheets are then baked and cut to different sizes of 0.5 mm. The terminal leads are attached by pressure contact or soldering.

The disc capacitors have high capacitance per unit volume. Round disc types with axial leads are used for high voltage operation. The discs are lacquered or encapsulated in plastic or phenolic moulding. They are available up to a value of 0.01 μF. They have voltage rating of up to 750 V d.c. (350 V a.c.). Some disc type capacitors are shown in Fig. 2.16.

Fig. 2.16 Disc type ceramic capacitors

Tubular capacitor: The low loss type capacitors are normally made of steatite or similar material. Steatite has a permittivity of 8. These are usually made in tubular structures. The construction of tubular capacitor is shown in Fig. 2.17. The required materials are ground and mixed thoroughly. The mixture is compressed and fixed with suitable fluxes and moulded into tubes. Silver metal is deposited outside and inside the tube. Leads are formed and soldered. The outer surface may be protected by lacquer in the uninsulated types. The tubular capacitors are available in the range of 5 pF to 1000 pF in voltage ranges up to 5 kV, and up to 10,000 pF in the lower voltage ranges.

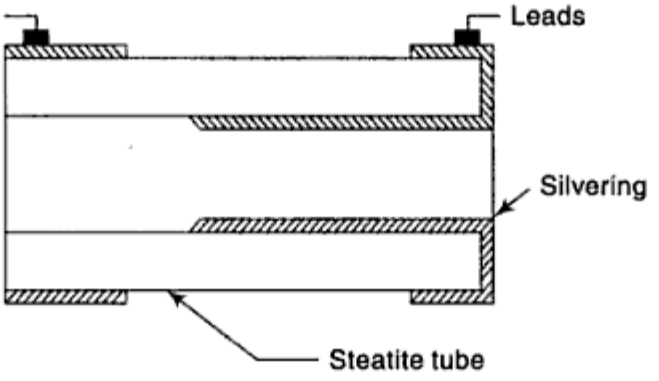


Fig. 2.17 Tubular ceramic capacitors

Monolithic capacitor: As shown in Fig. 2.18(a), the monolithic capacitors are formed by interleaving layers of ceramic and platinum electrodes. The first stage in the manufacturing process is the preparation of a slurry or slip in either aqueous or organic solution. Organic binders and plasticizers are added to the slurry to give strength and flexibility to the film. This is then cast on a carrier substrate in thicknesses of 0.025 mm to 0.75 mm and cut into strips. The next step is to apply screen, electrodes of platinum or palladium to the film. The electrode configuration is designed in such a way that each layer of film when blanked will carry an electrode. The deposition of metallic layer extends to alternate faces and accurate location of electrode pattern is required. After the required number of layers have been formed for a specific value of capacitance, the strips are subjected to consolidation under pressure. Then external electrodes are applied in order to pick up the edges of the

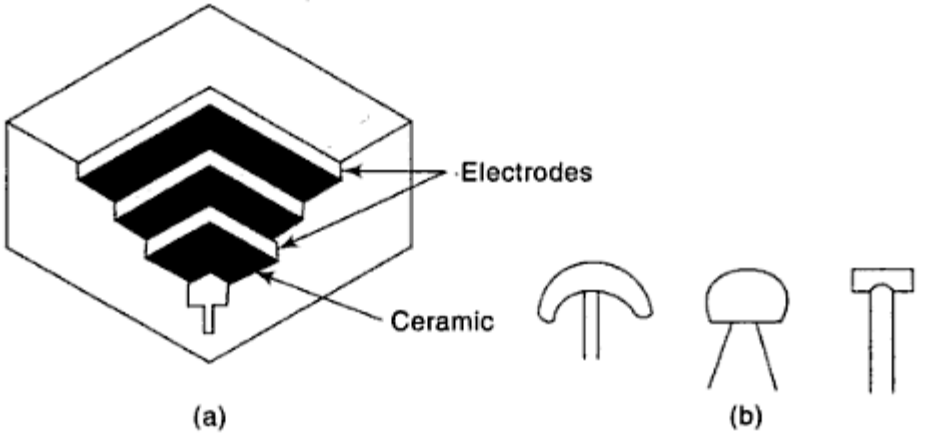


Fig. 2.18 Monolithic type ceramic capacitors

internal electrodes and form a parallel connection. This entire structure is fired to form a monolithic block. Sintering is done at high temperature (1300 °C) to reduce the volume of the block by 40 per cent. Some monolithic type ceramic capacitors are shown in Fig. 2.18(b).

Barrier layer capacitors: Barium titanate is used as the dielectric medium in these capacitors. If such a ceramic disc is reduced at 1300 °C, it becomes highly non-conducting medium. If reoxidised, a conducting layer of ceramic forms at the surface. This barrier of conducting to non-conducting layer possesses high values of dielectric constants. Suitable additives will improve the dielectric constant and the loss angle. Silver electrodes are fired on to this barrier layer and the entire assembly is reoxidised in hot air once again. This is encapsulated to prevent the entrance of moisture. The construction of a barrier type capacitor is shown in Fig. 2.19.

The capacitors have much lower insulation resistance and are used only for low voltage applications. They have temperature limitation and can operate from –40 °C to 80 °C. These are used in very high frequency circuit applications.

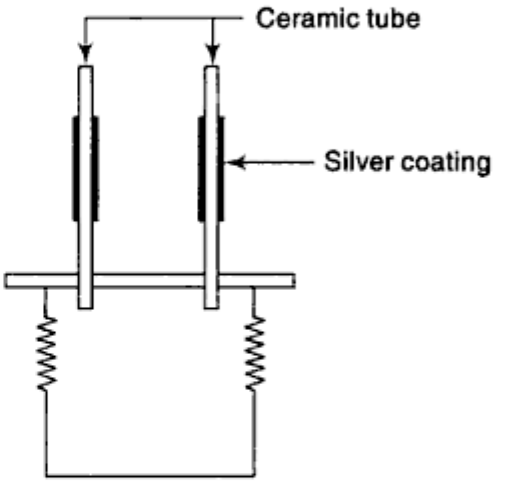


Fig. 2.19 Barrier type capacitor

Uses of ceramic capacitors: The ceramic capacitors are widely used as bypass capacitors and in decoupling and biasing applications. The major drawback is that the capacitance value varies with the temperature. Ceramic capacitors of the insulated type are used in TV receivers. Tubular capacitors are used to isolate the antennas in receivers. Capacitors with little temperature–capaci-tance variations are used to compensate for impedance–temperature changes in circuits.

Plastic capacitors These capacitors use a plastic film as a dielectric. The plastics used include polystyrene, polycarbonate, PTFE (Poly Tetra Fluora Ethylene) and polyester (mylar). Plastic capacitors are available in typical ranges of 500 pF to 10 μF. When sealed properly, these plastics exhibit good mechanical strength, resistance to heat and chemical inertness. The chief characteristics of these capacitors are their very high insulation resistance, high reliability, small physical size and low dielectric absorption.

Polystyrene capacitors: Polystyrene is a hydrocarbon material and has a lower permittivity than polyester and polycarbonate. These capacitors are made by rolling the polystyrene film with aluminium foils. These are compactly wound and the leads are soldered at the ends. The entire assembly is carefully heat treated so that the plastic softens (softening temperature, 90 °C) and makes good contact with the metal foil. The winding requires particular attention because it may accumulate electrostatic charges and attract dust. Polystyrene is not suitable for metallising because of lack of available oxygen for the self-healing action and low tensile strength.

They exhibit a low dissipation factor, small capacitance change with temperature and very good stability. They tend to be large in size, and their maximum operating temperature is 85 °C. Polystyrene capacitors are normally used in coupling, resonant and measuring circuits. As these capacitors have low dissipation factor, they can be used for a.c. applications.

Polyester capacitors: They are made of polyesters like Mylar, Melinex, Terelyne etc. Usually, this is metallised to give better results. Sometimes impregnation is done with silicone or mineral oil to improve the dielectric characteristics. A thin film of silver is deposited on to this plastic by vacuum evaporation. These capacitors are capable of withstanding temperatures up to 150 °C and voltages up to 400 V. These capacitors are available in the range from 100 pF to 2 μF. As the dielectric is prone to moisture, good encasement is needed. The capacitors are usually encased in glass or resin moulded or ceramic containers.

Polyester is a tough polymer with high tensile strength, free from pinholes and with good insulating properties over a reasonably wide temperature range. Its combination of good insulation properties over a wide temperature range and high mechanical strength are the outstanding characteristics. It also has a low moisture absorption. The high mechanical strength (up to approximately 210 °C) make polyester ideally suited for metallising by vacuum evaporation. It consists of sufficient oxygen to produce good self-healing.

Due to the good all-round mechanical and electrical properties and the high dielectric constant of polyester film, these capacitors are the most widely used of any film or paper capacitors. Polyester capacitors are normally used for coupling/decoupling applications at low and medium frequencies but find many other applications in general, and in power electronics as well as in certain a.c. applications. Polyester is ideal for d.c. applications because of its high dielectric strength and resistivity.

Polycarbonate capacitors: Polycarbonate is a polyester of carbonic acid bisphenols. It combines good physical properties with a lower dissipation factor than polyester. The metallised polycarbonate film is rolled and sealed in a metal case with epoxy resin. It is insulated with a polyester sleeve. As it does not have as much available oxygen, its self-healing properties are not quite as good as for polyester. It has a higher insulation resistance. It can operate at temperatures as high as 125 °C. Polycarbonate allows the manufacture of close-tolerance (typically ±1%), stable, small physical size, long life capacitors and is mainly used for d.c. applications, though it can also be used for a.c. applications. They are available in the range of 0.1 to 10 μF. They are not capable of withstanding higher voltages.

A comparison chart of some of the main characteristics of the three types of plastic dielectric capacitors is given in Table 2.4.

Mica capacitors The dielectric used in mica capacitors is muscovite or white mica, ruby or rose coloured mica and amber mica. It has a dielectric constant between 3 and 8, and a dielectric strength of 600–1500 V/mil. Its breakdown voltage widely varies from 500 V to 20 kV. It has low losses and the power factor varies depending on the dryness from 0.01 to 2% and can respond to very high frequencies. It can withstand temperatures up to 400 °C. It has better temperature

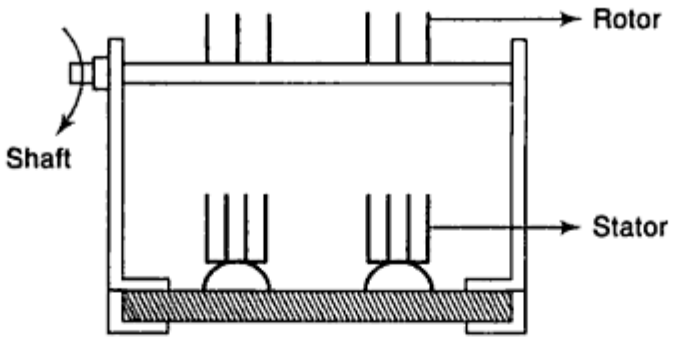


Fig. 2.22 Variable tuning capacitor

When the shaft is rotated, the effective area that exists between the rotor and stator plates varies, thereby varying the capacitance. The shape of the plates determines the manner in which capacitance varies with shaft rotation. The variation will be linear if the moving plates are semicircular. For tuning purposes, a uniform scale for frequency is required. A simple relation between the capacitance and the angular rotation is given by

$$\frac{1}{C} = (a\theta + b)^2$$

where C is the capacitance, a and b are constants, and θ is the angular rotation.

Capacitance values for air variable capacitors range from a few pico farads up to 500 pF with a maximum voltage rating of 9 kV. For higher operating voltages up to 60 KV; a vacuum variable capacitor is used.

Miniature type of variable capacitors use thin plastic film dielectrics between the plates. The entire assembly is enclosed in a moulded transparent plastic case to prevent entry of dust. These capacitors are typically low voltage units and they have a limited type.

2.3.3 Dissipation Factor

An ideal capacitor when discharged gives up all the electrical energy that was supplied to it in charging. But practical capacitors dissipate some of the energy delivered to them, i.e. they suffer from losses due to (i) dielectric losses—particularly at higher frequencies; (ii) insulation or leakage resistance; (iii) resistance of the leads and plates. The insulation resistance and dielectric strength are lowered if the capacitor is operated at high temperature, and high voltage and humidity. The leakage losses increase with frequency, applied voltage and temperature. The voltage rating of capacitors has to be reduced if operated at higher temperatures and humidity conditions.

The power factor (PF) of a capacitor expresses the ratio of the power dissipated per cycle to the power stored per cycle in a capacitor. The power factor is the ratio of the equivalent series resistance to the total impedance of the capacitor, i.e.

$$PF = \frac{R_s}{Z} = \frac{R}{\sqrt{R_s^2 + X_c^2}}$$

The dissipation factor (DF) of a capacitor is the ratio of the equivalent series resistance to the capacitive reactance, i.e. DF = R_s/X_C. For low losses (i.e., $R_s < X_c$), the power factor is equal to dissipation factor. The equivalent series resistance is due to capacitor plates with leads and insulation or leakage resistance. For an ideal capacitor $R_s = 0$ and hence PF = 0. Due to the losses, a practical capacitor has a non-zero power factor. For a practical capacitor the power factor angle Φ is not 90 ° but slightly differs given by 90 ° – θ. This angle is called the loss angle of a capacitor and tan θ is usually specified as the *dissipation factor* (DF). Dissipation factor is another parameter used to describe the quality of a capacitor. The reciprocal of the dissipation factor can be considered as the Q of the capacitor and is the ratio of the capacitor reactance (X_c) to the equivalent series resistance (R_s).

$$Q = \frac{1}{\text{DF}} = \frac{X_c}{R_s} = \frac{1}{\omega C R_s} = \tan \Phi = \tan (90^\circ - \theta) = \cot \theta = \frac{1}{\tan \theta}$$

Thus, higher Q means better quality of the capacitor.

2.3.4 Connecting Capacitors

Capacitors in parallel Figure 2.23 shows n capacitors connected in parallel. Here, as the effective area of the plates increases, the total equivalent capacitance of the parallel circuit C_{TP} increases with plate area (more charge is stored in the capacitor). Hence, C_{TP} is equal to the sum of the individual capacitances, i.e.,

$$C_{TP} = C_1 + C_2 + ... + C_n$$

When all capacitors are equal, then $C_{TP} = nC$. The relationship for C_{TP} is similar to that for resistors in series.

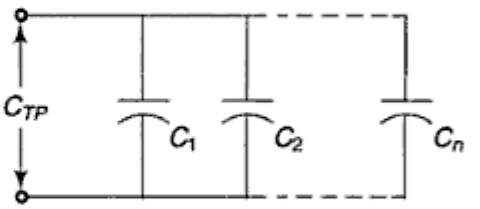


Fig. 2.23 n-capacitors connected in parallel

Capacitors in series Figure 2.24 shows n capacitors connected in series. Here, as the separation of the outer plates of the combination increases, the total equivalent capacitance of the series circuit C_{TS} will be less than the smallest capacitance of the individual capacitors. The value of C_{TS} can be calculated by

$$\frac{1}{C_{TS}} = \frac{1}{C_1} + \frac{1}{C_2} + \ldots + \frac{1}{C_n}$$

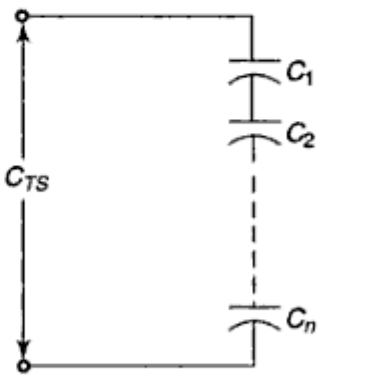


Fig. 2.24 n-capacitors connected in series

The relationship for C_{TS} is similar to that for resistors in parallel. Considering two capacitors, C_1 and C_2, connected in series, the total equivalent capacitance of the series circuit C_{TS} is equal to their product divided by their sum, i.e.

$$C_{TS} = \frac{C_1 C_2}{C_1 + C_2}$$

When equal valued capacitors are connected in series, then $C_{TS} = \frac{C}{n}$.

2.3.5 Characteristics and Applications

The characteristics and applications of various types of capacitors are summarised in Table 2.5.

2.4 INDUCTORS

This is the third passive component used in electronic circuits. It stores energy in the form of magnetic field and delivers it as and when required.

Whenever current passes through a conductor, lines of magnetic flux are generated around it. This magnetic flux opposes any change in current due to the induced e.m.f. This opposition to the change in current is known as inductance and the component producing inductance is known as inductor. The unit of inductance is Henry (symbol H). The induced e.m.f. is actually given by

$$e = -L\frac{di}{dt}$$

where e = induced e.m.f. in volts at any instant, L = inductance in Henry, and di/dt = rate of change of current.

The negative sign in the above equation indicates that the induced e.m.f. opposes the cause for the change in current.

An inductor is usually a coil of copper wire wound around a core made up of a ferromagnetic material, as shown in Fig. 2.25. The inductance (L) of the coil is given by

$$L = \frac{\mu_o \mu_r A N^2}{l} \text{ H}$$

where μ_o = permeability of free space = $4\pi \times 10^{-7}$ H/m = 1.257×10^{-6} H/m
μ_r = relative permeability of the core material
A = area of cross-section of the core
N = number of turns of the coil
l = length of the core.

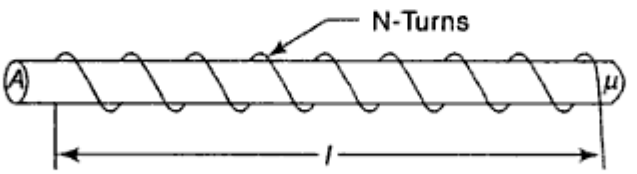


Fig. 2.25 Inductance of a coil

Hence, the value of the inductor depends upon the following factors: (i) number of turns, (ii) permeability of the core material, and (iii) size of core.

Inductors can be divided into two categories: (i) fixed inductors and (ii) variable inductors.

2.4.1 Fixed Inductors

Fixed type inductors can be further divided into three categories depending on the type of core used. They are (i) air core inductors, (ii) iron core inductors, and (iii) ferrite core inductors.

Table 2.5 Characteristics and applications of capacitors

Type	*Approximate capacitance range and tolerance*	*Power factor*	*Maximum working voltage (in volts)*	*Operating temperature*	*Typical applications*
1. Aluminium electrolytic	(i) 2 μF – 100 μF – 15% (ii) 50 μF – 1000 μF – 15%	0.02 – 0.2	600 50 – 200	85°C – 125°C -do-	Rectifier filters and smoothing Decoupling and bypassing in a.f.
2. Tantalum electrolytic	1 μF – 2000 μF – 2.5%	-do-	6 – 150	-do-	Coupling, decoupling, bypassing in amplifiers
3. Silvered ceramic	2 pF – 0.001 μF 1 – 20%	0.001	350	85°C – 150°C	r.f. amplifiers, r.f. bypass, decoupling and resonant circuit
4. Ceramic with high dieletric constant	50 pF – 0.01 μF 20%	0.001	350	-do-	-do-
5. Polyesterene	100 pF – 0.5 μF – 1/2 – 5%	0.002	125 – 500	65°C – 85°C	Resonant, coupling, measuring and a.c. applications
6. Polyester	0.005 – 2 μF – 10 – 20%	0.005	500	65°C – 125°C	Coupling, decoupling, smoothing and d.c. applications
7. Stacked mica	50 pF – 0.01 μF – 2 – 20%	0.001 – 0.005	2000	85°C – 120°C	r.f. coupling, bypassing circuits
8. Silvered mica	10 pF – 0.1 μF – 1/2 – 20%	0.001 – 0.005	1000	85°C – 150°C	r.f. resonant and measuring circuits
9. Tubular rolled paper	0.005 μF – 10 – 20%	0.004 – 0.01	1000	85°C – 125°C	a.f. coupling and decoupling, bypass, filter, motor-start
10. Metallised paper	0.005 μF – 25%	0.005 – 0.015	400	-do-	-do-
11. Air/polystyrene variable capacitance	15 pF – 500 pF	1000.0	1000	-do-	Tuning circuits in receivers and transmitters
12. Mica trimmers and padders	4 pF – 70 pF 100pF – 600pF	—	500	80°C	Tracking and alignment of receivers
13. Air trimmers	8 pF – 100 pF	—	-do-	-do-	-do-

These three types of inductors are shown in Fig. 2.26.

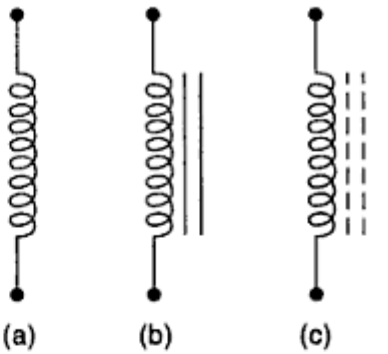


Fig. 2.26 (a) Air core inductors, (b) iron core inductors and (c) Ferrite core inductor

Air core inductors In radio frequency applications where very low values of inductance (from a fraction of a μH to a few μH) is required, air core inductors are generally used. Air core inductors consist of a few turns of wire wound on a hollow former.

Iron core inductors They have a coil containing a number of turns of copper wire wound on a hallow former and the core material passes through the former in such a way that it forms a closed magnetic path for the magnetic flux. The former is made up of paper or plastic material. The core is generally made up of silicon steel (a ferromagnetic material having high permeability) in the form of thin laminated sheets. Laminated sheets are used instead of solid mass to reduce the hysteresis and eddy current losses. Iron core transformers are used in low frequency applications such as filter circuits in power supplies, chokes in fluorescent tubes or as a reactive element in a.c. circuits. The value of inductors are generally in the order of a few Henries.

Ferrite core inductors Iron core inductors are not suitable for high frequency applications (because of enormous increase in hysteresis and eddy current losses). This difficulty is overcome by the use of ferrite materials as the core. A ferrite is basically an insulator having very high permeability.

Ferrite is made up of non-metallic compounds consisting mainly of ferric oxide in combination with one or two bivalent metal oxides. They are hard, dense ceramics and because of their high resistivity they can be used in the form of solid cores. The values of such inductors is in the range of few μH to few mH.

The typical applications of ferrite core inductors are in (i) r.f. chokes for supply decoupling purposes, (ii) switching regulated type d.c. power supplies, and (iii) various types of filters used in communication equipment.

2.4.2 Variable Inductors

In certain applications such as tuned circuits, it is required to vary the inductance from a minimum value to a maximum value. Ferrite core variable inductors are generally used for this purpose. In such inductors the hollow former on which the coil is wound has screw threads in the inner hollow portion. Similar matching threads are provided on the ferrite core which can be screwed in or out of the former. Because of the change of the position of the ferrite core the value of the inductance changes. It is maximum when the ferrite core is fully in. The variable ferrite core inductor is shown in Fig. 2.27.

Fig. 2.27 Variable inductor

very much less but increases in transformers with iron-power and ferrite cores. Air-core coils have k value of 0.05 to 0.3.

2.4.7 Connecting Inductors

Inductors in series The combined inductance of any two coils connected in series aiding and having self-inductances of L_1 and L_2, respectively, will be given by

$$L = L_1 + L_2 + 2M$$

But if the coils are connected in series opposing, their combined inductance will be

$$L = L_1 + L_2 - 2M$$

Inductors in parallel If two coils are connected in parallel with fields aiding, then

$$\frac{1}{L} = \frac{1}{(L_1 + M)} + \frac{1}{(L_2 + M)}$$

And for two coils in parallel with fields opposing, then

$$\frac{1}{L} = \frac{1}{(L_1 - M)} + \frac{1}{(L_2 - M)}$$

Review Questions

1. What is meant by resistance? Describe the various types of resistors.
2. What are the uses of resistors in electronic circuits?
3. Write short notes on wire wound resistors.
4. Describe different types of potentiometers.
5. Explain the tapering of potentiometers.
6. Why is colour coding used in resistors? Briefly explain its significance.
7. What is meant by tolerance in resistors?
8. Write short notes on: (a) Film resistors (b) Rheostat, and (c) Trimmer.
9. Explain the action of capacitor for d.c. and a.c. voltages.
10. Define the dielectric strength referred to a capacitor.
11. List the various capacitors used in electronic circuits and state uses of each type.
12. Compare the constructional features and characteristics of (a) Paper, (b) Mica, (c) Ceramic, and (d) Plastic capacitors.
13. Desctibe briefly polarized and non-polarized electrolytic capacitors.
14. Compare the constructional features and characteristics of aluminum and tantalum electrolytic capacitors.
15. What are the specifications of a capacitor? State the factors affecting the capacitance of a capacitor.
16. Explain the various losses of a capacitor.
17. What is meant by dissipation factor of a capacitor? Explain.
18. Write short notes on variable capacitors.
19. What are the factors that influence the value of inductance in an inductor?
20. Explain fixed inductors in detail.
21. Write down the applications of various types of inductors.
22. What are the losses that occur in an inductor?
23. Explain the quality factor of an inductor.
24. Classify the inductors and explain briefly.
25. Inductors and capacitors are energy storage elements—justify.

3 Electron Ballistics

3.1 INTRODUCTION

The operation of an electronic device depends upon the motion of electrons under the influence of electric and magnetic fields. The behaviour of an electron under the influence of these fields is termed as electron ballistics. This chapter deals with the study of basic properties of matter. This chapter starts with simple definitions, motion of particles in simple paths under uniform fields to complex paths under varying fields.

3.2 CHARGED PARTICLES

Electron forms the most fundamental charged particle. An electron is negatively charged; its magnitude is 1.602×10^{-19} Coulomb. The number of electrons per coulomb is 6×10^{18} and 1 A of current represents the motion of 6×10^{18} electrons per second. The mass of an electron is 9.107×10^{-31} kg. A direct measurement of mass of an electron is not possible but the ratio of electron's charge to mass can be determined by a number of experiments and found to be 1.759×10^{11} C/kg. The mass of an electron can be deduced from this value.

The positive ion possesses a charge that is an integral multiple of the charge of an electron and is positive. If a positive ion is singly ionized, it has a charge equal to that of an electron and if it is doubly ionized, its charge is twice that of an electron. The mass of a hypothetical atom of atomic weight unity is, as per definition, taken as one-sixteenth of the mass of monoatomic oxygen and has been calculated to be 1.660×10^{-27} kg. Hence, the mass of any atom, in kilogram, can be calculated by multiplying the atomic weight of the atom by 1.660×10^{-27} kg.

The radius of an electron is very small. It is estimated to be much closer to 3.8×10^{-10} m. The radius of an atom is estimated as 10^{-10} m. These dimensions being very small, all charges are considered to be mass points.

3.3 FORCE, FIELD INTENSITY, POTENTIAL, AND ENERGY

Force Charles Coulomb stated that the force between two very small objects separated by a distance which is large compared to their size is proportional to the charge on each and inversely proportional to the square of the distance between them, or

$$F = k\frac{Q_1 Q_2}{d^2} \tag{3.1}$$

where Q_1 and Q_2 are the charges, d the separation, and k the proportionality constant. In SI units, Q is measured in *Coulombs* (C), d in *metres*, and the force should be in *Newtons* (N). This will be achieved if the constant of proportionality k is written as

$$k = \frac{1}{4\pi\varepsilon_o} \tag{3.2}$$

where ε_o is the permittivity of free space and has a value of 8.854×10^{-12} F/m.

Coulomb's law shows that the force between two charges of 1 C each, separated by one metre, is 8.9×10^9 N and that between two electrons is 230×10^{-13} N.

Field intensity The electric field intensity is defined as the force on a unit positive charge, or conversely, the force on a unit positive charge at any point in an electric field is the electric field intensity E at that point. The unit of electric field intensity is *volts per metre* (V/m).

If a charge is moved against the electric field, a force equal to and opposite to the force exerted by the field is needed, and this requires expending energy or doing work. On the other hand, if the charge is moved in the direction of the field, the energy expenditure turns out to be negative; we do not do the work, the field does. When a charge q is moved in an electric field E, the force on q due to the electric field is

$$F_q = qE \tag{3.3}$$

The path of a charged particle in an electric field can be calculated by relating the force, given by Eqn. (3.3), to the mass and the acceleration of the particle by Newton's second law of motion. Thus,

$$F_q = q\,E = ma = m\frac{\mathrm{d}\,V}{\mathrm{d}\,t} \tag{3.4}$$

where m = mass, kg; a = acceleration, m/s^2; and V = velocity, m/s.

The solution of this equation with appropriate initial conditions gives the path of the particle resulting from the action of the electric forces.

Potential Consider a parallel-plate capacitor as shown in Fig. 3.1. A difference of potential is applied between the two plates as shown. The separation between the two plates is much smaller in comparison with the dimensions of the plates. This ensures uniform electric field between the plates, the lines of force pointing along negative Z -direction. At time $t = 0$, the initial velocity v_z is equal to zero. Since there is no force along the X or Y directions, the acceleration along these directions is zero and since the initial velocity is zero, the particle will not move in these axes.

By definition, the potential V (in volts) of point Z with respect to point Z_o is the work done against the field in taking a unit positive charge from Z_o to Z.

Thus,

$$V = -\int_{z_o}^{z} E \cdot \mathrm{d}z \tag{3.5}$$

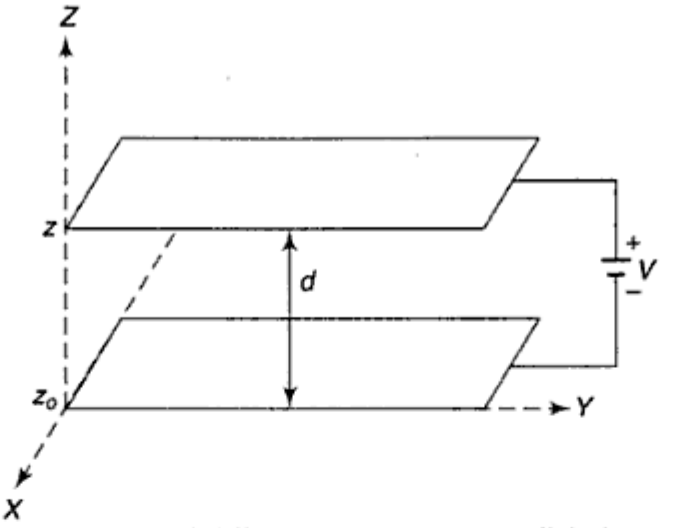


Fig. 3.1 Potential difference across two parallel plates

The integral in Eqn. (3.5) can be evaluated to the form

$$-\int_{z_o}^{z} E_z \cdot \mathrm{d}z = -E_z (z - z_o) = V \tag{3.6}$$

$$E_z = -\frac{V}{z - z_o} = -\frac{V}{d} \tag{3.7}$$

Thus, from Eqn. (3.7), the electric field intensity is negative of the ratio of applied potential difference between the two plates to the separation between the two plates. Equation (3.7) is valid for uniform electric fields; however, for fields varying with distance between the plates the field is given by

$$E_z = -\frac{\mathrm{d}V}{\mathrm{d}z} \tag{3.8}$$

The negative sign in Eqn. (3.8) signifies that the orientation of the electric field is from higher potential point to lower potential point.

Potential energy For an electron between two parallel plates, the Newton's second law can be applied. Then,

$$-q E_z = ma \tag{3.9}$$

$$-\frac{q}{m} E_z = \frac{\mathrm{d}v_z}{\mathrm{d}t} \tag{3.10}$$

The velocity of the particle in Z direction is defined as the rate of change of displacement of the particle in Z direction, i.e.,

$$v_z = \frac{\mathrm{d}z}{\mathrm{d}t}$$

$$\mathrm{d}z = v_z \, \mathrm{d}t \tag{3.11}$$

Multiplying Eqn. (3.10) by Eqn. (3.11) and integrating,

$$-\frac{q}{m} \int_{z_o}^{z} E_z \, \mathrm{d}z = \int_{-v_{oz}}^{v_z} \frac{\mathrm{d}v_z}{\mathrm{d}t} \cdot v_z \cdot \mathrm{d}t$$

$$-\frac{q}{m}\int_{z_o}^{z} E_z \, dz = \int_{v_{oz}}^{v_z} v_z \cdot dv_z \tag{3.12}$$

By virtue of Eqn. (3.5), Eqn. (3.12) integrates to

$$\frac{q}{m}V = \frac{1}{2}\left(v_z^2 - v_{oz}^2\right)$$

$$qV = \frac{1}{2}m\left(v_z^2 - v_{oz}^2\right) \tag{3.13}$$

where the energy qV is expressed in Joules.

Considering any two points A and B in space, with B at a higher potential than A by V_{BA}, Eqn. (3.13) can be stated in a more general form,

$$q\,V_{BA} = \frac{1}{2}m\,V_A^2 - \frac{1}{2}m\,V_B^2 \tag{3.14}$$

where q is the charge in *Coulombs*, qV_{BA} in *Joules*, and V_A and V_B are the initial and final velocities in *metres per second.*

By definition, the potential energy between two points equals the potential multiplied by the charge. Equation (3.14) is a statement of Law of Conservation of Energy. The left-hand side of Eqn. (3.14) is the rise in potential energy from A to B and the right-hand side is drop in kinetic energy from A to B. Thus, the rise in potential energy equals the drop in kinetic energy. Equation (3.14) is not valid for time-varying fields.

Example 3.1 Find the speed and the kinetic energy of an electron after it has moved through a potential difference of 5000 V.

Solution:

The speed of the electron, $v = \sqrt{\frac{2qV}{m}} = \sqrt{\frac{2\times1.6\times10^{-19}\times5000}{9.1\times10^{-31}}} = 4.2\times10^{7}$ m/s

The kinetic energy $= q \times V = 1.6\times10^{-19}\times5000 = 8\times10^{-16}$ J $= 5000$ eV

Example 3.2 A charged particle having mass equal to 1000 times of an electron and a charge same as that of an electron is accelerated through a potential difference of 1000 V. Calculate the velocity attained by the charged particle and kinetic energy in terms of electron volts and joules.

Solution:

The mass of the charged particle = 1000 times the mass of an electron

$= 1000\times9.1\times10^{-31} = 9.1\times10^{-28}$ kg

The charge of the particle $= 1.6\times10^{-19}$ C

Therefore, the velocity, $v = \sqrt{\frac{2qV}{m}} = \sqrt{\frac{2\times1.6\times10^{-19}\times1000}{9.1\times10^{-28}}} = 0.596\times10^{6}$ m s^{-1}

Kinetic energy $= q \times V$

$=1.6\times10^{-19}\times1000 = 1.6\times10^{-16}$ J $= 1000$ eV

Therefore, $v_x = v_{ox} + a_x t = 3 \times 10^6 + 1.026 \times 10^{16} (5.264 \times 10^{-10})$

$= 8.4 \times 10^6$ m/s

(ii) Consider $x = 6 \times 10^{-6}$ m

$$t^2 + 5.85 \times 10^{-10} t - 1.17 \times 10^{-18} = 0$$

Solving this equation,

$$t = 8.28 \times 10^{-10} \text{ s}$$

Therefore, $v_x = 3 \times 10^6 + 8.28 \times 10^{-10} (1.026 \times 10^{16}) = 11.5 \times 10^6$ m/s

---

Example 3.4 Two plane parallel plates A and B are placed 3 mm apart and potential of B is made 200 V positive with respect to plate A. An electron starts from rest from plate A. Calculate (i) the velocity of the electron on reaching plate B, (ii) time taken by the electron to travel from plate A to plate B, and (iii) kinetic energy of the electron on reaching the plate B.

Solution:

(i) The electron starts from rest at plate A, therefore, the initial velocity is zero. The velocity of the electron on reaching plate B is,

$$v = \sqrt{\frac{2qv}{m}} = \sqrt{\frac{2(1.6 \times 10^{-19})200}{(9.1 \times 10^{-31})}} = 8.38 \times 10^6 \text{ m/s}$$

(ii) Time taken by the electron to travel from plate A to plate B can be calculated from the average velocity of the electron in transit. The average velocity is

$$v_{\text{average}} = \frac{\text{Initial velocity} + \text{Final velocity}}{2} = \frac{0 + 8.38 \times 10^6}{2}$$

$= 4.19 \times 10^6$ m/s

Therefore, the time taken for travel is

$$\text{Time} = \frac{\text{Separation between the plates}}{\text{Average time}} = \frac{3 \times 10^{-3}}{4.19 \times 10^6}$$

$= 0.71 \times 10^{-9}$ s.

(iii) Kinetic energy of the electron on reaching the plate B is

Kinetic energy $= q\,V = (1.6 \times 10^{-19})\ 200 = 3.2 \times 10^{-17}$ J

Example 3.5 Two plane parallel plates A and B are placed 8 mm apart and plate B is 300 V more positive than plate A. The electron travels from plate A to plate B with an initial velocity of 1×10^6 m/s. Calculate the time of travel.

Solution:

The speed acquired by the electron due to the applied voltage is,

$$v = \sqrt{v_{\text{initial}}^2 + \frac{2qV}{m}}$$

$$= \sqrt{(1 \times 10^6)^2 + \frac{2(1.6 \times 10^{-19})300}{9.1 \times 10^{-31}}}$$

$= 10.33 \times 10^6$ m/s

The average velocity,

$$v_{average} = \frac{v_{initial} + v_{final}}{2} = \frac{1\times10^6 + 10.33\times10^6}{2} = 5.665\times10^6 \text{ m/s}$$

Therefore, the time for travel

$$= \frac{\text{separation between plates}}{v_{average}} = \frac{8\times10^{-3}}{5.665\times10^6} = 1.4\times10^{-9} \text{ s}$$

Example 3.6 An infinitely large parallel plane plates are spaced 0.8 cm apart. The voltage at one of the plates is raised from 0 to 5 V in 1 ns at a uniform rate, with respect to the other. After duration, the potential difference between the plates is suddenly dropped to 0 V and remains the same thereafter. Find (i) the position of the electron, which started with zero initial velocity from the negative plate, when the potential difference drops to zero volt, (ii) the total time of transit of the electron from the cathode to the anode.

Solution: The electric field intensity,

$$E = -\frac{5t}{10^{-9}\,d} = -\frac{5t}{10^{-9}\times1\times10^{-2}} = 5\times10^{11}\,t \quad (\text{for } 0 < t < t_1)$$

$$= 0 \quad (\text{for } t_1 < t < \infty)$$

The velocity of electron, $v_x = -\int \frac{q E_x}{m} dt + C$ (C is the initial velocity)

$$= \int 5\times10^{11}\,\frac{q}{m}\,t\,dt$$

$$= 5\times10^{11}\,\frac{q}{m}\cdot\frac{t^2}{2}$$

The distance travelled by the electron,

$$d = \int v_x\,dt = \int 5\times10^{11}\,\frac{q}{m}\cdot\frac{t^2}{2}\,dt + C \qquad (C = 0)$$

$$= 5\times10^{11}\,\frac{q}{m}\cdot\frac{t^3}{6}$$

(i) The position of the electron after 1 ns,

$$d = (5\times10^{11})\cdot(1.76\times10^{11})\cdot\frac{(1\times10^{-9})^3}{6}$$

$$= 14.7\times10^{-6} \text{ m} = 14.7\ \mu\text{m}$$

(ii) The rest of the distance to be covered by the electron = 0.8 cm – 14.7 μm

= 0.799 cm

Since, the potential difference drops to zero volt, after 1 ns, the electron will travel the distance of 0.799 cm with a constant velocity of

$$v_x = 5\times10^{11}\,\frac{q}{m}\cdot\frac{t^2}{2} = (5\times10^{11})\cdot(1.76\times10^{11})\cdot\frac{(1\times10^{-9})^2}{2}$$

$$= 44\times10^3 \text{ m/s}$$

Therefore, the time $t_2 = \frac{d}{v_x} = \frac{0.799 \times 10^{-2}}{44 \times 10^3} = 1.816 \times 10^{-7}$ s

The total time of transit of electron from cathode to anode = $1 \times 10^{-9} + 1.816 \times 10^{-7}$
$= 1.826 \times 10^{-7}$ s

3.5 FORCE IN MAGNETIC FIELD

The force acting on a conductor kept in a magnetic field is given by

$$f_m = B\,I\,L \tag{3.19}$$

where

f_m: force acting on the conductor, N
B: magnetic flux density, Wb/m^2
I: current in the conductor, A
L: length of the conductor, m.

The direction of force is perpendicular to the plane consisting of the components of B and I that are mutually perpendicular, and is directed along the advance of a right-handed screw as illustrated in Fig. 3.3.

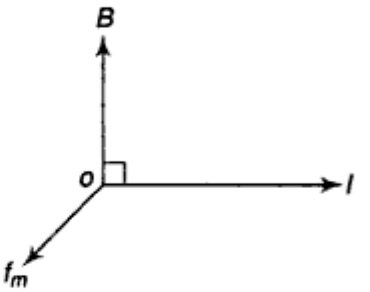


Fig. 3.3 Illustration of right-handed screw rule

If an electron takes T seconds to travel a distance of L metres in the conductor, the total number of electrons passing through any cross-section of conductor in unit time is N/T, where N is the total number of electrons contained in the conductor. Thus the total charge per second crossing any point, i.e. the current is

$$I = \frac{Nq}{T} \tag{3.20}$$

$$f_m = BIL = \frac{BNqL}{T} \tag{3.21}$$

The factor L/T is the drift speed v m/s of the electrons. Hence, the force per electron is

$$f_m = qBv \tag{3.22}$$

If the particle is a positive ion, the direction of current and drift velocity are the same. If the particle is electron, the direction of the current is opposite to that of the drift velocity.

3.6 MOTION IN MAGNETIC FIELD

The magnetic force acting on a charged particle in a uniform magnetic field can be expressed as

$$\bar{f}m = q\,\bar{B} \times \bar{v} \tag{3.23}$$

$$f_m = qBv \sin \varphi \tag{3.24}$$

where φ is the angle between the direction of the magnetic field and the direction of motion of the particle. From Eqn. (3.24), we see that the magnitude of the magnetic force is proportional to the charge of the particle, the magnetic flux density, the speed of the charged particle and the angle between the directions of motion of the charged particle and the magnetic flux density.

If an electron is placed in a uniform magnetic field with zero initial velocity, then the magnetic force on the electron is zero, in accordance with Eqn. (3.24). If the electron moves along the direction of magnetic flux density, then the angle is zero and the magnetic force is zero. A particle whose initial velocity has no component normal to a uniform magnetic field will continue to move with constant velocity along the lines of flux since the magnetic force on the particle is zero.

The magnetic force acting on an electron moving perpendicularly to the direction of the magnetic flux density is illustrated in Fig. 3.4. The magnetic field is perpendicular to the plane of paper and directed towards the reader. The electron is shown to enter the magnetic field from no-field region with an initial velocity v_o. The direction of the current is opposite to that of the motion. Hence, the magnetic force is upwards and the electron will have a change in its direction of motion. At every point in the magnetic field, force acts on the electron and the resultant direction will be perpendicular to both the magnetic field and the direction of motion of electron at that point. At every instant the force f_m is perpendicular to the direction of the particle and no work is done on the electron. This means that the kinetic energy is unaltered and the speed remains the same.

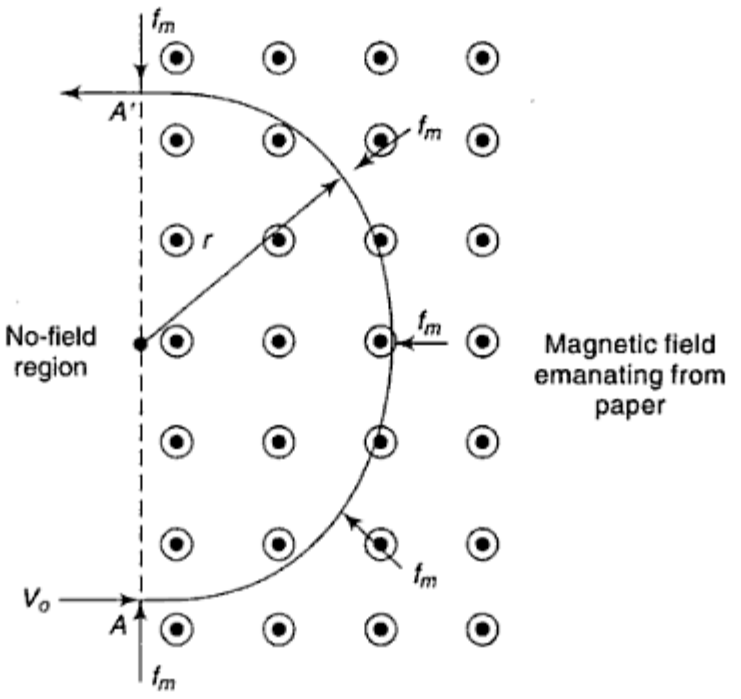


Fig. 3.4 Motion of an electron in a magnetic field

This type of force makes an electron to move in a circular path with uniform speed. As shown in the figure, the direction of the magnetic force is always towards the centre, O, of the circle. This force is same as the centripetal force which always tries to push the electron towards the centre. Then

$$\frac{mv^2}{r} = qBv \tag{3.25}$$

where 'r' is the radius of circular path of the electron.

From Eqn. (3.25),

$$r = \frac{mv^2}{qBv} = \frac{mv}{qB} \tag{3.26}$$

The angular velocity of the electron in radians per second is given by

$$\omega = \frac{v}{r} = \frac{qB}{m} \tag{3.27}$$

The time for one revolution is

$$T = \frac{2\pi}{\omega} = \frac{2\pi m}{qB}, \text{ seconds} \tag{3.28}$$

The time taken by the particle to complete one revolution is also called the period.

We see from Eqn. (3.26) that the radius is directly proportional to the velocity of the particle. The particles that move faster will traverse in larger circles and if the velocity is less the particles will be bent sharply in smaller circular paths.

Whenever a charged particle enters a uniform magnetic field, with an initial velocity, v, not perpendicular to the magnetic field but makes an angle with the direction of field, as shown in Fig. 3.5, then the velocity can be resolved into two components, $v \sin \varphi$, the velocity component perpendicular to the field and $v \cos \varphi$, parallel to the field. These two velocity components make the particle to move simultaneously in two directions.

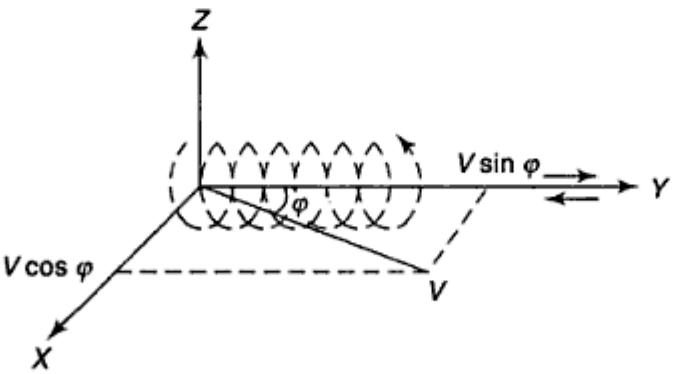


Fig. 3.5 Motion of an electron emerging in field at an angle φ

The velocity component parallel to the direction of flux makes the particle to move linearly along the direction of flux with no change in velocity. The perpendicular component of velocity makes the particle to move in a circular path as discussed above, with no change in the speed. The net effect, due to these concurrent motion, is that the electron takes a helical path as shown.

The pitch of the helix is defined as the distance travelled by an electron along the direction of the magnetic field in one revolution, and is given by

$$p = v_{oy} T \tag{3.29}$$

be a component of this force which is proportional to the X component of velocity and will be directed along the $+Z$ axis. The path will thus bend away from the $+X$ direction towards the $+Z$ direction. The electric and magnetic forces interact with one another and the net force will finally make the electron to travel in a cycloidal path.

3.9 ELECTROSTATIC DEFLECTION IN CATHODE RAY TUBE

The electrostatic deflection system uses a pair of deflection plates as shown in Fig. 3.6. The deflecting voltages are applied between the two plates. For deflecting the beam in the horizontal and vertical directions, two sets of plates are required. The horizontal deflection plates are kept vertically and the vertical deflection plates are kept horizontally, however, the plane of the plates are kept parallel to the axis of the Cathode ray tube. The electrons are attracted towards the positive plates and when they leave the region below the plates they travel in straight line, at an angle with the axis. Then the electron beam strikes the screen at a point.

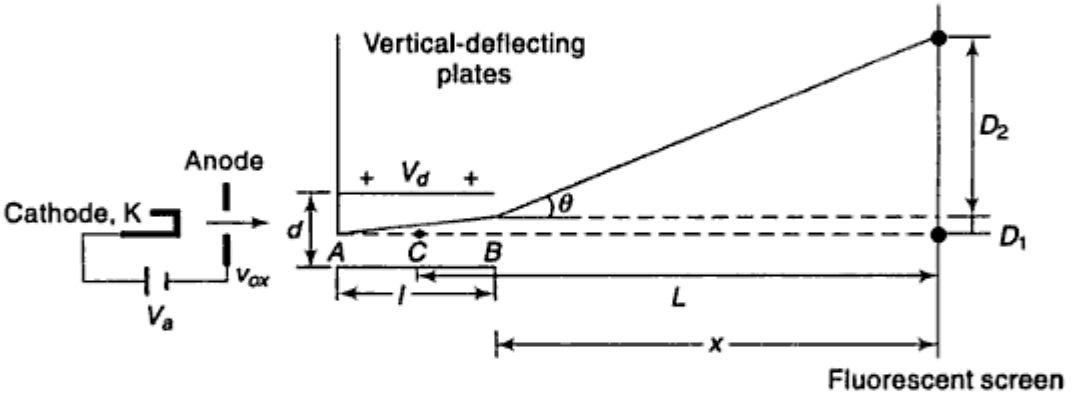


Fig. 3.6 Electrostatic deflection system in a CRT

Electrostatic deflection sensitivity of a pair of deflection plates of a CRT is defined as the amount of deflection (in mm or inch) of electron spot produced when a voltage of 1V d.c. is applied between the corresponding deflection plates. The deflection sensitivity may be different for X-X and Y-Y plates. The amount of deflection produced may be calculated knowing (i) the dimensions of the tube components and (ii) deflecting voltage.

Figure 3.6 shows the configuration of the electrostatic deflection system in a CRT. Two plates of length l and spacing d are kept at a distance x from the screen. Let the voltage applied between the plates be V_d volts, V_{ox} be the velocity of an electron on entering the field of deflection plates.

Then,

$$V_{ox} = \sqrt{\frac{2q}{m} \cdot V_a} = 5.94 \times 10^5 \sqrt{V_a} \text{ m/s}$$

where V_a is the final anode voltage (in volts), q is the charge of an electron in coulombs and m is the mass of an electron in kilograms.

As the beam passes through the field of the deflection plates, the electrons are attracted towards the positive plate by a force equal to

$$F = \frac{V_d \cdot q}{d}$$

The force produces an acceleration of 'a' m/s^2 and is equal to the force divided by mass m of an electron. This mass may be supposed to be the mass at rest because the deflecting voltage is hardly even greater than 2000 V.

$$a = \frac{V_d \cdot q}{dm}$$

The forward motion however continues at velocity V_{ox}. The time taken by an electron to traverse the field of the plates is then l/V_{ox} s. The upward velocity attained by an electron in this time interval l/V_{ox} is V_y. Then,

$$V_y = \frac{1}{V_{ox}} \cdot a = \frac{1}{V_{ox}} \cdot \frac{V_d \cdot q}{dm}$$

Then the ratio of upward velocity to forward velocity at the instant the electron leaves the field is

$$\frac{V_y}{V_{ox}} = \frac{V_d \cdot q \cdot l}{dm V_{ox}^2}$$

The electron follows a curved path from A, the point of entrance, to point B, the point leaving the field. Let the vertical displacement during this be D_1. Then D_1 is given by

$$D_1 = \frac{1}{2} a \left(\frac{l}{V_{ox}}\right)^2 = \frac{1}{2} \cdot \frac{V_a \cdot q \cdot l^2}{dm V_{ox}^2}$$

If θ is the angle with the axis that the electron beam makes after emerging out from the field of the deflection plates, then

$$\tan \theta = \frac{V_r}{V_{ox}} = \frac{D_2}{x}$$

(or)

$$D_2 = L \cdot \frac{V_r}{V_{ox}} = \frac{V_a}{d} \cdot \frac{q}{m} \cdot \frac{1}{V_{ox}^2} \cdot x$$

$$\text{Total deflection } D = D_1 + D_2 = \frac{V_x \cdot q\, l\,[l/2 + x]}{dm V_{ox}^2} = \frac{V_a \cdot qlL}{dm V_{ox}^2}$$

The total deflection of the spot will be same as like electron travelling straight line path at angle θ from point C instead of B. Since $L = \frac{1}{2} + x$, $D = \frac{V_d\, ql\, L}{dm V_{ox}^2}$.

But, $V_{ox}^2 = \dfrac{2qV_a}{m}$

where V_a is the final anode voltage.

$$D = \frac{V_d}{V_a} \times \frac{lL}{2d}$$

For a given CRT, the quantities l, L and d are fixed. Hence the deflection, 'D' of the spot can be changed only by changing the ratio $\dfrac{V_m}{V_a}$. The electrostatic deflection sensitivity of a CRT is defined as the deflection in metres on the screen per volt of deflecting voltage.

$$\text{Deflection sensitivity, } S = \frac{D}{V_d} = \frac{lL}{2dV_a}, \text{ cm /V} \tag{3.33}$$

where l = length of the deflecting plates, m

L = distance from the centre of the plate to the screen, m

V_a = accelerating potential, V

d = separation between the two plates, m

The deflection sensitivity is (i) directly proportional to the length of the deflection plates l, (ii) directly proportional to the distance L between the screen and the centre of the deflection plates, (iii) inversely proportional to the spacing d between the deflection plate and inversely proportional to the final anode voltage V_a.

Example 3.10 An electrostatic cathode ray tube has a final anode voltage of 600 V. The deflection plates are 3.5 cm long and 0.8 cm apart. The screen is at a distance of 20 cm from the centre of plates. A voltage of 20 V is applied to the deflection plates. Calculate (i) velocity of electron on reaching the field, (ii) acceleration due to deflection field, (iii) deflection produced on the screen in cm, and (iv) deflection sensitivity in cm/V.

Solution:

Given: $V_a = 600$ V, $l = 3.5$ cm, $d = 0.8$ cm, $L = 20$ cm, $V_d = 20$ V

(i) The velocity of the elctron, $v = \sqrt{\dfrac{2qV_a}{m}} = 14.5 \times 10^6$ ms^{-1}

(ii) $ma = qE$

Thus, acceleration, $a = qE/m = (q/m)(V_d/d) = 43.95 \times 10^{13}$ ms^{-2}

(iii) The deflection on the screen, $D = lLV_d/2V_ad = 1.45$ cm

(iv) Deflection sensitivity = $D/V_d = 0.0725$ cm/V.

Example 3.11 In a CRT, the deflection plates are 2 cm long and are spaced 0.5 cm apart. The screen is 20 cm away from the centre of the deflecting plates. The final anode voltage is 800 V. Calculate (i) the velocity of the beam on emerging from the field and (ii) the voltage that must be applied to the deflecting plates to have a displacement of 1 cm.

Solution:

Given: $V_a = 800$ V, $l = 2$ cm, $d = 0.5$ cm, $L = 20$ cm, $D = 1$ cm

(i) The velocity of the beam, $v = \sqrt{\frac{2qV_a}{m}} = \sqrt{\frac{2\times1.6\times10^{-19}\times800}{9.1\times10^{-31}}}$

$= 16.8 \times 10^6$ m/s

(ii) The deflection of the beam, $D = \frac{lLV_d}{2dV_a}$

i.e., $1 \times 10^{-2} = \frac{2\times10^{-2}\times20\times10^{-2}\times V_d}{2\times0.5\times10^{-2}\times800}$

Therefore, the voltage that must be applied to the plates,

$$V_d = 20 \text{ V}$$

Example 3.12 In an electrostatic deflecting CRT, the length of the deflection plate is 2 cm and spacing between deflecting plates is 0.5 cm. The distance from the centre of the deflecting plate to the screen is 20 cm and the deflecting voltage is 25 V. Find the deflection sensitivity, the angle of deflection and velocity of the beam. Assume final anode potential is 1000 V.

Solution:

Given, in an electrostatic deflecting CRT,

Length of deflection plate, $l = 2$ cm

Spacing between deflecting plates, $d = 0.5$ cm

Distance from center of deflecting plate to screen, $L = 20$ cm

Deflecting potential, $V_d = 25$ V

Anode potential, $V_a = 1000$ V

(i) Velocity of beam, $v = \sqrt{\frac{2qV_a}{m}} = \sqrt{\frac{2\times1.6\times10^{-19}\times1000}{9.1\times10^{-31}}}$

$= 18.75 \times 10^6$ ms^{-1}

(ii) Deflection sensitivity $= \frac{D}{V_d}$

where $D = \frac{ILV_d}{2V_a d} = \frac{2\times10^{-2}\times20\times10^{-2}\times25}{2\times1000\times0.5\times10^{-2}} = 10^{-2}$ cm

Therefore, the deflection sensitivity $= \frac{10^{-2}}{25} = 0.0004$ cm/V

(iii) To find the angle of deflection, θ:

$$\tan\theta = \frac{D}{(L-l)} = \frac{10^{-2}}{20-2} = \frac{10^{-2}}{18}$$

Therefore, $\theta = \tan^{-1}\left(\frac{1}{1800}\right) = 0.0318°$

Example 3.13 An electron with a velocity of 3×10^5 ms^{-1} enters an electric field of 910 V/m making an angle of 60° with the positive direction. The direction of the electric field is in the positive y direction. Calculate the time required to reach its maximum height.

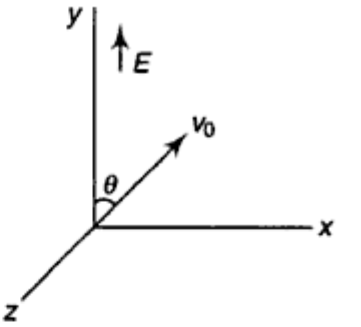


Fig. 3.7

Solution:
Given:

$$v_0 = 3 \times 10^5 \text{ m/s}$$
$$E = 910 \text{ V/m}$$
$$\theta = 60°$$

The electron starts moving in the +y direction, but, since acceleration is along the –y direction, its velocity is reduced to zero at time $t = t'$.

$$v_{0y} = v_0 \cos \theta = 3 \times 10^5 \times \cos 60° = 0.15 \times 10^6 \text{ m/s}$$

$$a_y = \frac{qE}{m} = \frac{1.602 \times 10^{-19}}{9.109 \times 10^{-31}} \times 910 = 1.5984 \times 10^{14} \text{ m/s}^2$$

$$t' = \frac{v_{0y}}{a_y} = \frac{0.15 \times 10^6}{1.5984 \times 10^{14}} = 9.384 \times 10^{-10} \text{ s} = 0.938 \text{ ns}$$

3.10 MAGNETIC DEFLECTION IN CATHODE RAY TUBE

The applied magnetic field is perpendicular to the direction of the electron beam. The force exerted on the electron beam by the magnetic field bends the electron beam in a direction perpendicular to both the field and the direction of electron movement.

Figure 3.8 shows the magnetic deflection system in a CRT. In the region of the uniform magnetic field, the electron experiences a force qBv, where v is the speed. The path OM is the arc of a circle whose center is at Q. The speed of the particles remains constant, which is given by

$$v = v_{cx} = \sqrt{\frac{2qv_a}{m}} \tag{3.34}$$

The small angle of deflection φ is equal to the length of the arc OM divided by R, the radius of the circle. Then,

$$\varphi \approx \frac{l}{R} \tag{3.35}$$

The magnetic field continues to deflect the electron beam at right angles to its movement and hence the path taken by the electron beam within the magnetic field is a part of a circle whose radius, R, is given by

$$R = \frac{mv}{qb} \tag{3.36}$$

system, the deflection is inversely proportional to electron velocity. Hence, the electrostatic deflection suffers deflection defocusing when the beam is bent through a large angle. For these reasons, the magnetic deflection is preferred in cathode ray tube used in television because large permissible angle of deflection reduces the length of the tube for a given diameter.

Defects of deflection When an electron beam is deflected from the axial direction, the spot on the fluorescent screen tends to distort and enlarge. This phenomenon is commonly referred to as *deflection defocusing*. This may be due to (i) non-uniform distance between the electron gun and the different parts of the screen, (ii) non-uniform electric and magnetic deflection fields, and (iii) unequal velocities of electrons in the beam. With non-uniform deflecting fields, the part of the beam which passes through region of weaker field is less deflected than the part of the beam which passes through region of stronger field. If the electrons are emitted from cathode with unequal initial velocities, as the degree of deflection is also a function of the electron velocity, electrons with different velocities are deflected differently which forms a blurred elliptical spot on the screen.

Review Questions

1. Define the terms (i) electric field intensity (ii) electric potential (iii) potential energy.
2. State the law of conservation of energy.
3. What are the assumptions made while analysing the motion of an electron in an electric field?
4. Discuss the motion of an electron between two parallel plates under the influence of applied potential.
5. Two plane parallel plates A and B are placed 5 mm apart and potential of B is made 300 V positive with respect to plate A. An electron starts from rest from plate A. Calculate (i) kinetic energy of the electron on reaching plate B, (ii) the velocity of the electron on reaching plate B, and (iii) time taken for the electron to reach plate B.
   [*Ans.* (i) 4.8×10^{-17} joule, (ii) 10.28×10^{6} m/s, (iii) 0.973×10^{-9} s]
6. Two parallel plates are kept 8 mm apart. One plate is kept 250 V positive with respect to the other. An electron starts from rest from the negative plate. Calculate the velocity, kinetic energy and distance travelled after time 0.4×10^{-19} s.
   [*Ans.* 2.2×10^{6} m/s; 22.84×10^{-19} joule; 0.44 mm]
7. In a vacuum diode, the spacing between the parallel plane plates of cathode and anode is 5 mm and the potential difference is 250 V. Calculate the time taken by the electron, with an initial velocity of 1×10^{6} m/s, to travel from cathode to anode.
   [*Ans.* 0.96 ns]
8. Two parallel plates are kept a distance 12 mm apart. One plate is 800 V positive with respect to the other. An electron starts from rest from the negative plate. Find the distance travelled, time taken and the kinetic energy of the electron when it has acquired a speed of 5×10^{6} m/s. At this instant, the potential across the plates is suddenly removed. Find the total time of travel of the electron from the negative plate to the positive plate.
   [*Ans.* 1.065×10^{-3} m, 0.426×10^{9} s, 113.8×10^{-19} joule, 2.613×10^{-9} s]
9. Discuss the motion of an electron under the influence of applied magnetic field.
10. An electron moving with intial velocity of 10^{6} m/s enters an uniform magnetic field at an angle of 30° with it. Calculate the magnetic flux density required in order that the

radius of helical path be 1 m. Also, calculate the time taken by the electron for one revolution and the pitch of the helix.

[*Ans.* 0.284×10^{-4} Wb/m^2; 12.57×10^{-6} s; 10.88 m.]

11. What happens to an electron when it is exposed to parallel electric and magnetic fields?
12. Why does an electron take a cycloidal path when it is exposed to perpendicular electric and magnetic fields?
13. Define electrostatic deflection sensitivity.
14. An electrostatic cathode ray tube has a final anode voltage of 400 V. The deflection plates are 2 cm long and 1 cm apart. The screen is at a distance of 10 cm from the centre of the plates. A voltage of 20 V is applied to the deflection plates. Calculate (i) velocity of electron on reaching the field, (ii) acceleration due to deflection field, (iii) deflection produced on the screen and (iv) deflection sensitivity.

[*Ans.* (i) 11.9×10^6 ms^{-1}, (ii) 35.1×10^{10} ms^{-2}, (iii) 0.5 cm, (iv) 0.025 cm/V]

15. In a cathode ray tube having electric deflection system, the deflection plates are 2 cm long and have a uniform spacing of 4 mm between them. The fluorescent screen is 25 cm away from the centre of the deflection plates. Calculate the deflection sensitivity, if the potential of the final anode is (i) 1000 V, (ii) 2000 V and (iii) 3500 V.

[*Ans.* (i) 0.0625 cm/V, (ii) 0.03125 cm/V, (iii) 0.01785 cm/V]

16. Define magnetic deflection sensitivity.
17. In a CRT, the distance of the screen from the centre of the magnetic field is 22 cm. The deflecting magnetic field of flux density 2×10^{-4} Wb/m^2 extends for a length of 2.5 cm along the tube axis. The final anode voltage is 1250 V. Calculate the deflection of the spot in cm. [*Ans.* 0.92 cm]
18. Compare electrostatic deflection with magnetostatic deflection.
19. Why is magnetic deflection preferred to electrostatic deflection in the CRT used in television?

4 Semiconductor Diodes

4.1 INTRODUCTION

The PN junction diode is one of the semiconductor devices with two semiconductor materials in physical contact, one with excess of holes (P-type) and other with excess of electrons (N-type). A PN junction diode may be found from a single-crystal intrinsic semiconductor by doping part of it with acceptor impurities and the remainder with donors. Such junctions can form the basis of very efficient rectifiers. The most important characteristic of a PN junction is its ability to allow the flow of current in only one direction. In the opposite direction, it offers very high resistance. The high-vacuum diode has largely been replaced by silicon and selenium rectifiers. Semiconductor diodes find wide applications in all phases of electronics, viz. radio and TV, optoelectronics, power supplies, industrial electronics, instrumentation, computers, etc.

4.2 CLASSIFICATION OF SEMICONDUCTORS

Semiconductors are classified as (i) intrinsic (pure) and (ii) extrinsic (impure) types. The extrinsic semiconductors are of N-type and P-type.

Intrinsic semiconductor A pure semiconductor is called intrinsic semiconductor. As already explained in the first chapter, even at the room temperature, some of the valence electrons may acquire sufficient energy to enter the conduction band to form free electrons. Under the influence of electric field, these electrons constitute electric current. A missing electron in the valence band leaves a vacant space there, which is known as a *hole*, as shown in Fig. 4.1. Holes also contribute to electric current.

In an intrinsic semiconductor, even at room temperature, electron-hole pairs are created. When electric field is applied across an intrinsic semiconductor, the current conduction takes place by two processes, namely, free electrons and holes. Under the influence of electric field, total current through the semiconductor is the sum of currents due to free electrons and holes.

Though the total current inside the semiconductor is due to free electrons and holes, the current in the external wire is fully by electrons. In Fig. 4.2, holes being positively charged move towards the negative terminal of the battery. As the holes reach the negative terminal of the battery, electrons enter the semiconductor near

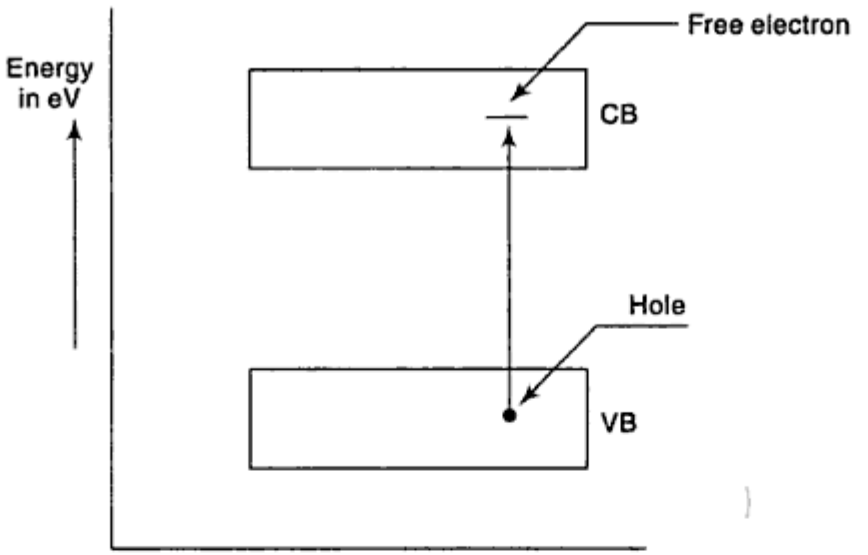


Fig. 4.1 Creation of electron-hole pair in a semiconductor

the terminal (X) and combine with the holes. At the same time, the loosely held electrons near the positive terminal (Y) are attracted away from their atoms into the positive terminal. This creates new holes near the positive terminal which again drift towards the negative terminal.

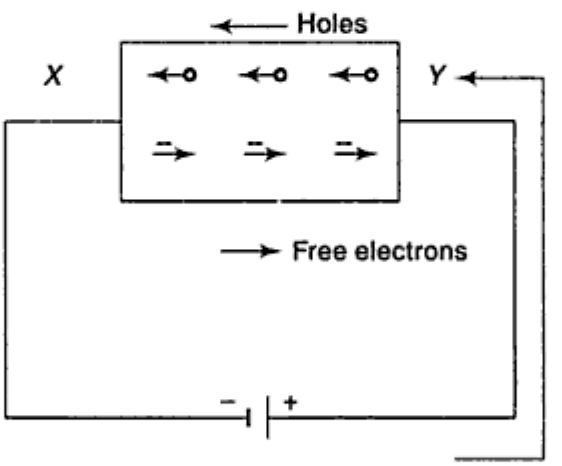


Fig. 4.2 Current conduction in semiconductor

Extrinsic semiconductor Due to the poor conduction at room temperature, the intrinsic seconductor as such, is not useful in the electronic devices. Hence, the current conduction capability of the intrinsic semiconductor should be increased. This can be achieved by adding a small amount of impurity to the intrinsic semiconductor, so that it becomes impure or extrinsic semiconductor. This process of adding impurity is known as *doping*.

The amount of impurity added is extremely small, say 1 to 2 atoms of impurity for 10^6 intrinsic atoms.

N-type semiconductor A small amount of pentavalent impurities such as arsenic, antimony or phosphorus is added to the pure semiconductor (germanium or silicon crystal) to get N-type semiconductor.

Germanium atom has four valence electrons and antimony has five valence electrons. As shown in Fig. 4.3, each antimony atom forms a covalent bond with surrounding four germanium atoms. Thus, four valence electrons of antimony atom form covalent bond with four valence electrons of individual germanium atom and fifth valence electron is left free which is loosely bound to the antimony atom.

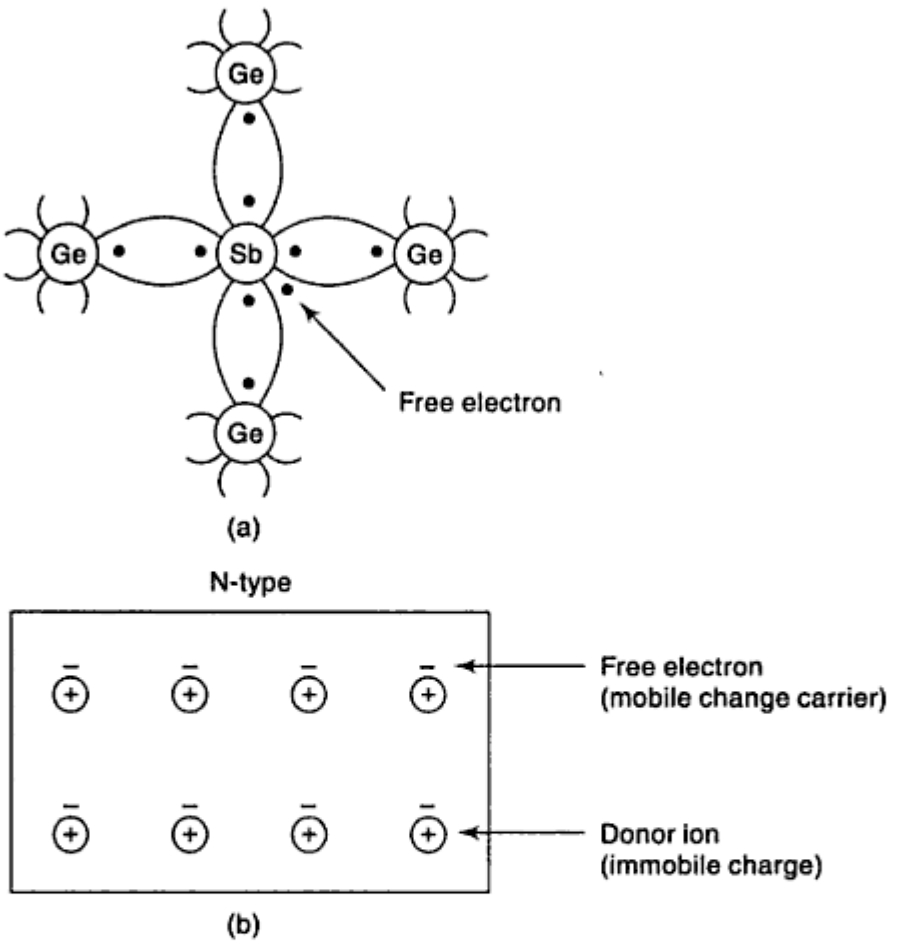


Fig. 4.3 N-type semiconductor: (a) Formation of covalent bonds, and (b) Charged carriers

This loosely bound electron can be easily excited from the valence band to the conduction band by the application of electric field or increasing the thermal energy. Thus every antimony atom contributes one conduction electron without creating a hole. Such pentavalent impurities are called donor impurities because it donates one electron for conduction. On giving an electron for conduction, the donor atom becomes positively charged ion because it loses one electron. But it cannot take part in conduction because it is firmly fixed in the crystal lattice.

Thus, the addition of pentavalent impurity (antimony) increases the number of electrons in the conduction band thereby increasing the conductivity of N-type semiconductor. As a result of doping, the number of free electrons far exceeds the number of holes in an N-type semiconductor. So electrons are called majority carriers and holes are called minority carriers.

P-type semiconductor A small amount of trivalent impurities such as aluminium or boron is added to the pure semiconductor to get the *p*-type semiconductor.

4.4 CARRIER CONCENTRATION IN INTRINSIC SEMICONDUCTOR

To calculate the conductivity of a semiconductor, the concentration of free electrons n and the concentration of free holes p must be known.

$$d_n = N(E)\, f(E)\, dE$$

where d_n represents the number of conduction electrons per cubic meter whose energies lie between E and $E + dE$ and $N(E)$ is the density of states. In a semiconductor, the lowest energy in the conduction band is E_C and hence,

$$N(E) = \gamma (E - E_C)^{1/2}$$

The Fermi Dirac probability function $f(E)$ is given by

$$f(E) = \frac{1}{1 + e^{(E - E_F)/kT}}$$

where E_F is the Fermi level or characteristic energy for the crystal in eV.

The concentration of electrons in the conduction band is,

$$n = \int_{E_c}^{\infty} N(E)\, f(E)\, dE$$

For $E \geq E_C$, $E - E_F >> kT$,

$$f(E) = e^{-(E - E_F)/kT}$$

and

$$n = \int_{E_c}^{\infty} \gamma (E - E_C)^{\frac{1}{2}}\, e^{-(E - E_F)/kT}\, dE$$

This integral evaluates to

$$n = N_C\, e^{-(E_c - E_F)/kT}$$

where $N_C = 2\left(\frac{2\pi m_n kT}{h^2}\right)^{3/2}$ $(1.60 \times 10^{-19})^{3/2}$, where m_n is the effective mass of an electron.

When the maximum energy in the valence band is E_v, the density of states is given by

$$N(E) = \gamma (E_V - E)^{1/2}$$

The Fermi function of a hole is $[1 - f(E)]$ and is given by

$$1 - f(E) = \frac{e^{(E - E_F)/kT}}{1 + e^{(E - E_F)/kT}} = e^{-(E_F - E)/kT}$$

where $E_F - E >> kT$ for $E \leq E_V$.

The concentration of holes in the valence band is,

$$p = \int_{-\infty}^{E_v} \gamma (E_V - E)^{1/2}\, e^{-(E_F - E)/kT}\, dE$$

This integral evaluates to

$$p = NV\, e^{(E_F - E_v)/kT}$$

where

$N_V = 2\left(\frac{2\pi m_p kT}{h^2}\right)^{3/2}$ $(1.60 \times 10^{-19})^{3/2}$, where m_p is the effective mass of a hole.

Fermi level in an intrinsic semiconductor In the case of intrinsic material, the crystal must be electrically neutral.

$$n_i = p_i$$

Therefore, $N_C\, e^{-(E_c - E_f)/kT} = N_V\, e^{-(E_F - E_v)/kT}$

Taking the logarithm on both sides,

$$\ln \frac{N_C}{N_V} = \frac{E_C + E_V - 2E_F}{kT}$$

$$E_F = \frac{E_C + E_V}{2} - \frac{kT}{2} \ln \frac{N_C}{N_V}$$

If the effective masses of a free electron and hole are the same,

$$N_C = N_V$$

Then, $$E_F = \frac{E_C + E_V}{2}$$

From the above equation, at the centre of the forbidden energy band, Fermi level is present.

Donor and acceptor impurities If a pentavalent substance (antimony, phosphorous or arsenic) is added as an impurity to a pure germanium, four of the five valance electrons of the impurity atoms will occupy covalent bonds and the fifth electron will be available as a carrier of current. These impurities donate excess electron carriers and hence called *donor* or N-type impurities.

If a trivalent impurity (boron, gallium or indium) is added to an intrinsic semiconductor, only three covalent bonds are filled, and the vacancy in the fourth bond constitutes a hole. These impurities are known as *acceptor* or P-type impurities.

Fermi level in a semiconductor having impurities The Fermi level in an N-type material is given by

$$E_F = E_C + kT \ln \frac{N_C}{N_D}$$

where $N_D = N_C\, e^{-(E_C - E_F)/kT}$, the concentration of donor atoms.

The Fermi level in a P-type material is given by

$$E_F = E_V + kT \ln \frac{N_V}{N_A}$$

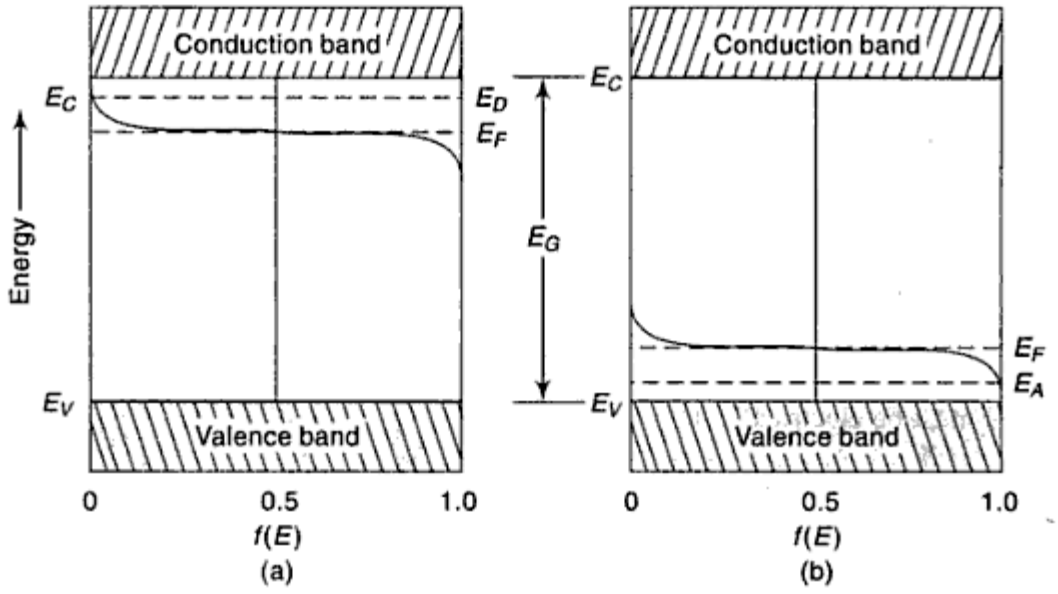


Fig. 4.5 Positions of Fermilevel in (a) N-type and (b) P-type semiconductors.

where $N_A = NVe^{-(E_F - E_v)/KT}$, the concentration of acceptor atoms. The change in th epositions of Fermilevel in N- and P-type semiconductors is shown in Fig. 4.5.

Movement of E_f with temperature In an N-type semiconductor, as temperature T increases, more number of electron-hole pairs are formed. At very high temperature T, the concentration of thermally generated electrons in the conduction band will be far greater than the concentration of donor electrons. In such a case, as concentration of electrons and holes become equal, the semiconductor becomes essentially intrinsic and E_F returns to the middle of the forbidden energy gap. Hence, it is concluded that as the temperature of the P-type and N-type semiconductor increases, E_F progressively moves towards the middle of the forbidden energy gap.

Example 4.2 In an N-type semiconductor, the Fermi level is 0.3 eV below the conduction level at a room temperature of 300 K. If the temperature is increased to 360 °K, determine the new position of the Fermi level.

Solution: The Fermi level in an N-type material is given by

$$E_F = E_C - kT \ln \frac{N_C}{N_D}$$

Therefore, $(E_C - E_F) = kT \ln \frac{N_C}{N_D}$

$$\text{At } T = 300 \text{ K}, 0.3 = 300 \text{ °K} \ln \frac{N_C}{N_D} \quad (1)$$

$$\text{Similarly, } E_C - E_{F1} = 360\, k \ln \frac{N_C}{N_D} \quad (2)$$

Eqn. (2) divided by Eqn. (1) gives

$$\frac{E_C - E_{F1}}{0.3} = \frac{360}{300}$$

Therefore, $E_C - E_{F1} = \frac{360}{300} \times 0.3 = 0.36$ eV

Hence, the new position of the Fermi level lies 0.36 eV below the conduction level.

Example 4.3 In a P-type semiconductor, the Fermi level is 0.3 eV above the valance band at a room temperature of 300 °K. Determine the new position of the Fermi level for temperatures of (a) 350 °K and (b) 400 °K.

The Fermi level in a P-type material is given by

$$E_F = E_V + kT \ln \frac{N_V}{N_A}$$

Therefore, $(E_F - E_V) = kT \ln \frac{N_V}{N_A}$

At $T = 300$ °K, $0.3 = 300\, k \ln \frac{N_V}{N_A}$

(a) At $T = 350$ °K, $(E_{F1} - E_V) = 350k \ln \frac{N_V}{N_A}$

Hence, from the above equation

$$\frac{E_{F1} - E_V}{0.3} = \frac{350}{300}$$

Therefore, $E_{F1} - E_V = \frac{350}{300} \times 0.3 = 0.35$ eV

(b) At $T = 400$ °K, $(E_{F2} - E_V) = 400\, k \ln \frac{N_V}{N_A}$

Hence, from the above equation,

$$\frac{E_{F2} - E_V}{0.3} = \frac{400}{300}$$

Therefore, $E_{F2} - E_V = \frac{400}{300} \times 0.3 = 0.4$ eV

Example 4.4 In an N-type semiconductor, the Fermi level lies 0.2 eV below the conduction band. Find the new position of Fermi level if the concentration of donor atoms is increased by a factor to (a) 4 and (b) 8. Assume $kT = 0.025$ eV.

Solution: In an N-type material, the concentration of donor atoms is given by

$$N_D = N_C\, e^{-(E_C - E_F)/KT}$$

Let initially $N_D = N_{DO}$, $E_F = E_{FO}$ and $E_C - E_{FO} = 0.2$ eV

Therefore, $N_{DO} = N_C\, e^{-0.2/0.025} = N_C\, e^{-8}$

(a) When $N_D = 4N_{DO}$ and $E_F = E_{F1}$, then

$$4N_{DO} = N_C\, e^{-(E_C - E_{F_1})/0.025} = N_C\, e^{-40(E_C - E_{F_1})}$$

Therefore, $4 \times N_C e^{-8} = N_C e^{-40(E_C - E_{F1})}$

Therefore, $4 = e^{-40(E_C - E_{F1}) + 8}$

Taking natural logarithm on both sides, we get

$$\ln 4 = -40 (E_C - E_{F1}) + 8$$

$$1.386 = -40(E_C - E_{F1}) + 8$$

Therefore, $E_C - E_{F1} = 0.165$ eV

(b) When $N_D = 8N_{DO}$ and $E_F = E_{F2}$, then

$$\ln 8 = -40 (E_C - E_{F2}) + 8$$

$$2.08 = -40 (E_C - E_{F2}) + 8$$

Therefore, $E_C - E_{F2} = 0.148$ eV

Example 4.5 In a P-type semiconductor, the Fermi level lies 0.4 eV above the valence band. Determine the new position of Fermi level if the concentration of acceptor atoms is multiplied by a factor of (a) 0.5 and (b) 4. Assume $kT = 0.025$ eV.

Solution: In a P-type material, the concentration of acceptor atoms is given by

$$N_A = N_V e^{-(E_F - E_V)/KT}$$

Let initially $N_A = N_{AO}$, $E_F = E_{FO}$ and $E_{FO} - E_V = 0.4$eV

Therefore, $N_{AO} = N_V e^{-0.4/0.025} = N_V e^{-16}$

(a) When $N_A = 0.5$, N_{AO} and $E_F = E_{F1}$, then

$$0.5N_{AO} = N_V e^{-(E_{F1} - E_V)/0.025} = N_V e^{-40(E_{F1} - E_V)}$$

Therefore, $0.5 \times N_V e^{-16} = N_V e^{-40(E_{F1} - E_V)}$

Therefore, $0.5 = e^{-40(E_{F1} - E_V) + 16}$

Taking natural logarithm on both sides, we get

$$\ln (0.5) = -40(E_{F1} - E_V) + 16$$

Therefore, $E_{F1} - E_V = 0.417$ eV

(b) When $N_A = 4N_{AO}$ and $E_F = E_{F2}$, then

$$\ln 4 = -40(E_{F2} - E_V) + 16$$

Therefore, $E_{F2} - E_V = 0.365$ eV

4.5 MASS-ACTION LAW

If a pure semiconductor is doped with N-type impurities, the number of electrons in the conduction band increases above a level and the number of holels in the valence band decreases below a level, which would be available in the intrinsic (pure) semiconductor. Similarly, the addition of P-type impurities to a pure semiconductor increases the number of holes in the valence band above a level and decreases the number of electrons in the conduction band below a level, which would have been available in the intrinsic semiconductor. This is because the rate of recombination increases due to the presence of a large number of free electrons (or holes).

Further, the experimental results state that under thermal equilibrium for any semiconductor, the product of the number of holes and the number of electrons is

constant and is independent of the amount of donor and acceptor impurity doping. This relation is known as *mass-action law* and is given by

$$n.p = n_i^2 \tag{4.3}$$

where n is the number of free electrons per unit volume, p the number of holes per unit volume and n_i the intrinsic concentration.

While considering the conductivity of the doped semiconductors, only the dominant majority charge carriers have to be considered.

Charge densities in N-type and P-type semiconductors The law of mass-action has given the relationship between free electron concentration and hole concentration. These concentrations are further related by the law of Electrical Neutrality as explained below.

Let N_D be the concentration of donor atoms in an N-type semiconductor. In order to maintain the electric neutrality of the crystal, we have

$$\begin{aligned} n_N &= N_D + p_N \\ &\approx N_D \end{aligned}$$

where n_N and p_N are the electron and hole concentration in the N-type semiconductor. The value of p_N is obtained from the relations of mass-action law as

$$p_N = \frac{n_i^2}{n_N}$$

$$\approx \frac{n_i^2}{N_D}, \text{ which is } << n_N \quad \text{or} \quad N_D.$$

Similarly, in a P-type semiconductor we have

$$\begin{aligned} p_P &= N_A + n_P \\ &\approx N_A \end{aligned}$$

From mass-action law, $n_P = \dfrac{n_i^2}{p_p}$

Therefore, $n_P = \dfrac{n_i^2}{N_A}$, which is $<< p_P$ or N_A

where N_A, p_P and n_P are the concentrations of acceptor impurities, holes and electrons respectively in a P-type semiconductor.

Extrinsic conductivity The conductivity of an N-type semiconductor is given by

$$\sigma_N = q n_N \mu_n \approx q N_D \mu_n, \text{ since } n_N \approx N_D.$$

The conductivity of a P-type semiconductor is given by

$$\sigma_P = q p_P \mu_P \approx q N_A \mu_P, \text{ since } p_P \approx N_A.$$

The doping of intrinsic semiconductor considerably increases its conductivity.

If the concentration of donor atoms added to a P-type semiconductor exceeds the concentration of acceptor atoms, i.e. $N_D >> N_A$, then the semiconductor is converted

from a P-type to N-type. Similarly, a large number of acceptor atoms added to an N-type semiconductor can convert it to a P-type semiconductor if $N_A >> N_D$. This concept is precisely used in the fabrication of PN junction, which is an essential part of semiconductor devices and integrated circuits.

Example 4.6 Find the conductivity of silicon (a) in intrinsic condition at a room temperature of 300 °K, (b) with donor impurity of 1 in 10^8, (c) with acceptor impurity of 1 in 5×10^7 and (d) with both the above impurities present simultaneously. Given that n_i for silicon at 300 °K is 1.5×10^{10} cm^{-3}, $\mu_n = 1300$ cm^2/V-s, $\mu_p = 500$ cm^2/V-s, number of Si atoms per cm^3 = 5×10^{22}.

(a) In intrinsic condition, $n = p = n_i$

$$\text{Hence, } \sigma_i = qn_i(\mu_n + \mu_p)$$
$$= (1.6 \times 10^{-19})(1.5 \times 10^{10})(1300 + 500)$$
$$= 4.32 \times 10^{-6} \text{ S/cm}$$

(b) Number of silicon atoms/cm^3 = 5×10^{22}

$$\text{Hence, } N_D = \frac{5 \times 10^{22}}{10^8} = 5 \times 10^{14} \text{ cm}^{-3}$$

Further, $n \approx N_D$

$$\text{Therefore, } p = \frac{n_i^2}{n} \approx \frac{n_i^2}{N_D}$$
$$= \frac{(1.5 \times 10^{10})^2}{5 \times 10^{14}} = 0.46 \times 10^6 \text{ cm}^{-3}$$

Thus $p << n$. Hence p may be neglected while calculating the conductivity.

$$\text{Hence, } \sigma = nq\mu_n = N_D q\mu_n$$
$$= (5 \times 10^{14})(1.6 \times 10^{-19})(1300)$$
$$= 0.104 \text{ S/cm.}$$

(c) $N_A = \frac{5 \times 10^{22}}{5 \times 10^7} = 10^{15}$ cm^{-3}

Further, $p \approx N_A$

$$\text{Hence, } n = \frac{n_i^2}{p} \approx \frac{n_i^2}{N_A}$$
$$= \frac{(1.5 \times 10^{10})^2}{10^{15}} = 2.25 \times 10^5 \text{ cm}^{-3}$$

Thus, $p >> n$. Hence n may be neglected while calculating the conductivity.

$$\text{Hence, } \sigma = pq\mu_P = N_A q\mu_P$$
$$= (10^{15} \times 1.6 \times 10^{-19} \times 500)$$
$$= 0.08 \text{ S/cm.}$$

(d) With both types of impurities present simultaneously, the net acceptor impurity density is,

$$N_A' = N_A - N_D = 10^{15} - 5 \times 10^{14} = 5 \times 10^{14} \text{ cm}^{-3}$$

Hence, $\sigma = N_A' q \mu_p$

$= (5 \times 10^{14})\ (1.6 \times 10^{-19})\ (500)$

$= 0.04$ S/cm.

Example 4.7 Determine the resistivity of germanium (a) in instrinsic condition at 300 °K (b) with donor impurity of 1 in 10^7 (c) with acceptor impurity of 1 in 10^8 (d) with both the above impurities simultaneously. Given that for germanium at room temperature $n_i = 2.5 \times 10^{13}/\text{cm}^3$, $\mu_n = 3800$ cm^2/V–Vs, $\mu_p = 1800$ cm^2/V-Vs and a number of Germanium atoms/cm$^3 = 4.4 \times 10^{22}$.

Solution: (a) $n = p = n_i = 2.5 \times 10^{13}$ cm^{-3}

Therefore, conductivity, $\sigma = q n_i (\mu_n + \mu_p)$

$= (1.6 \times 10^{-19})(2.5 \times 10^{13})(3800 + 1800)$

$= 0.0224$ S/cm

Hence, resistivity, $\rho = \dfrac{1}{\sigma} = \dfrac{1}{0.0224} = 44.64\ \Omega\text{-cm}$

(b) $N_D = \dfrac{4.4 \times 10^{22}}{10^7} = 4.4 \times 10^{15}$ cm^{-3}

Also, $n = N_D$

Therefore, $p = \dfrac{n_i^2}{n} = \dfrac{n_i^2}{N_D}$

$= \dfrac{(2.5 \times 10^{13})^2}{4.4 \times 10^{15}} = 1.42 \times 10^{11}$ holes/cm^3

Here, as $n >> p$, p can be neglected.

Therefore, conductivity, $\sigma = nq\mu_n = N_D q \mu_n$

$= (4.4 \times 10^{15})\ (1.6 \times 10^{-19})\ (3800) = 2.675$ S/cm

Hence, resistivity, $\rho = \dfrac{1}{\sigma} = \dfrac{1}{2.675} = 0.374\ \Omega$-cm

(c) $N_A = \dfrac{4.4 \times 10^{22}}{10^8} = 4.4 \times 10^{14}$ cm^{-3}

Also, $p = N_A$

Therefore, $n = \dfrac{n_i^2}{p} = \dfrac{n_i^2}{N_A}$

$= \dfrac{(2.5 \times 10^{13})^2}{4.4 \times 10^{14}} = 1.42 \times 10^{12}$ electrons/cm^3

Here, as $p >> n$, n may be neglected. Then

Conductivity, $\sigma = pq\mu_p = N_A q\mu_p$

$= (4.4 \times 10^{14})\ (1.6 \times 10^{-19})\ (1800) = 0.1267$ S/cm

Hence resistivity, $\rho = \frac{1}{\sigma} = \frac{1}{0.1267} = 7.89\ \Omega$-cm

(d) with both p and n type impurities present,

$N_D = 4.4 \times 10^{15}\ \text{cm}^{-3}$ and $N_A = 4.4 \times 10^{14}\ \text{cm}^{-3}$

Therefore, the net donor density N_D' is

$N_D' = (N_D - N_A) = (4.4 \times 10^{15} - 4.4 \times 10^{14})$

$= 3.96 \times 10^{15}\ \text{cm}^{-3}$

Therefore, effective $n = N_D' = 3.96 \times 10^{15}\ \text{cm}^{-3}$

$$p = \frac{n_i^2}{N_D'} = \frac{(2.5 \times 10^{13})^2}{3.96 \times 10^{15}}$$

$$= 1.578 \times 10^{11}\ \text{cm}^{-3}$$

Here again $p\left(= \frac{n_i^2}{N_D'}\right)$ is very small compared with N_D' and may be neglected in calculating the effective conductivity.

Therefore, $\sigma = N_D'\, q\mu_n$

$= (3.96 \times 10^{15})\ (1.6 \times 10^{-19})\ (3800) = 2.408$ S/cm

Hence, resistivity $\rho = \frac{1}{\sigma} = \frac{1}{2.408} = 0.415\ \Omega$-cm

---

Example 4.8 A sample of silicon at a given temperature T in intrinsic condition has a resistivity of $25 \times 10^4\ \Omega$-cm. The sample is now doped to the extent of 4×10^{10} donor atoms/cm^3 and 10^{10} acceptor atoms/cm^3. Find the total conduction current density if an electric field of 4 V/cm is applied across the sample. Given that $\mu_n = 1250\ \text{cm}^2/\text{V-s}$, $\mu_p = 475\ \text{cm}^2/\text{V-s}$ at the given temperature.

Solution: $\sigma_i = qni(\mu_n + \mu_p) = \frac{1}{25 \times 10^4}$

Therefore, $n_i = \frac{\sigma_i}{q(\mu_n + \mu_p)} = \frac{1}{(25 \times 10^4)\,(1.6 \times 10^{-19})(1250 + 475)}$

$= 1.45 \times 10^{10}\ \text{cm}^{-3}$

Net donor density $N_D\ (= n) = (4 \times 10^{10} - 10^{10})$

$= 3 \times 10^{10}\ \text{cm}^{-3}$

Hence, $p = \frac{n_i^2}{N_D} = \frac{(1.45 \times 10^{10})^2}{3 \times 10^{10}} = 0.7 \times 10^{10}\ \text{cm}^{-3}$

Hence, $\sigma = q(n\mu_n + p\mu_p)$

$= (1.6 \times 10^{-19})\ (3 \times 10^{10} \times 1250 + 0.7 \times 10^{10} \times 475)$

$= 6.532 \times 10^{-6}$

Therefore, total conduction current density, $J = \sigma E = 6.532 \times 10^{-6} \times 4$

$= 26.128 \times 10^{-6}$ A/cm^2

Example 4.9 Find the concentration (densities) of holes and electrons in N-type Silicon at 300 °K, if the conductivity is 300 S/cm. Also find these values for P-type silicon. Given that for Silicon at 300 °K, $n_i = 1.5 \times 10^{10}$/cm^3, $\mu_n = 1300$ cm^2/V-s and $\mu_p = 500$ cm^2/V-s.

Solution: (a) *Concentration in N-type Silicon*

The conductivity of an N-type Silicon is $\sigma = q\, n\, \mu_n$

Concentration of electrons, $n = \dfrac{\sigma}{q\mu_n}$

$$= \frac{300}{(1.6 \times 10^{-19})(1300)} = 1.442 \times 10^{18}\ \text{cm}^{-3}$$

Hence concentration of holes, $p = \dfrac{n_i^2}{n} = \dfrac{(1.5 \times 10^{10})^2}{1.442 \times 10^{18}} = 1.56 \times 10^{2}\ \text{cm}^{-3}$

(b) *Concentration in P-type silicon*

The conductivity of a P-type Silicon is $\sigma = qp\mu_p$

Hence, concentration of holes $p = \dfrac{\sigma}{q\mu_p}$

$$= \frac{300}{(1.6 \times 10^{-19})(500)} = 3.75 \times 10^{18}\ \text{cm}^{-3}$$

and concentration of electrons, $n = \dfrac{n_i^2}{p} = \dfrac{(1.5 \times 10^{10})^2}{3.75 \times 10^{18}} = 0.6 \times 10^{2}\ \text{cm}^{-3}$

Example 4.10 A specimen of pure germanium at 300 °K has a density of charge carriers 2.5×10^{19}/m^3. It is doped with donor impurity atoms at the rate of one impurity atom every 10^6 atoms of germanium. All impurity atoms are supposed to be ionised. The density of germanium atom is 4.2×10^{28} atoms/m^3. Calculate the resistivity of the doped germanium if electron mobility is 0.38 m^2/V-s.

If the Germanium bar is 5×10^{-3} m long and has a cross sectional area of $(5 \times 10^{-6})^2$ m^2, determine its resistance and the voltage drop across the semiconductor bar for a current of 1 μA flowing through it.

Solution: Density of added impurity atoms is

$$N_D = \frac{4.2 \times 10^{28}}{10^6} = 4.2 \times 10^{22}\ \text{atoms/m}^3$$

Also, $n \approx N_D$

Therefore, $p = \dfrac{n_i^2}{n} = \dfrac{n_i^2}{N_D}$

$$= \frac{(2.5 \times 10^{19})^2}{4.2 \times 10^{22}} = 1.488 \times 10^6 \text{ m}^{-3}$$

Here, as $p << n$, p may be neglected.

Therefore, $\sigma = qN_D \mu_n$

$= (1.6 \times 10^{-19})\ (4.2 \times 10^{22})\ (0.38) = 2.554 \times 10^3$ S/m

Therefore, resistivity, $\rho = \frac{1}{\sigma} = \frac{1}{2.554 \times 10^3} = 0.392 \times 10^{-3}\ \Omega$-m

Resistance, $R = \frac{\rho L}{A} = \frac{0.392 \times 10^{-3} \times 5 \times 10^{-3}}{(5 \times 10^{-6})^2}$

$= 78.4\ \text{k}\Omega$

Voltage drop, $V = RI = 78.4 \times 10^3 \times 10^{-6} = 78.4$ mV

4.6 PROPERTIES OF INTRINSIC SEMICONDUCTORS

Table 4.1 gives the important properties of intrinsic silicon and germanium at room temperature (300 °K)

Table 4.1 Properties of some common semiconductors at room temperature

	Si	*Ge*	*GaAs*
Atomic number	14	32	–
Atomic weight	28.09	72.59	–
Atomic density (m^{-3})	5.02×10^{28}	4.42×10^{28}	–
Lattice constant, a (nm)	0.543	0.565	0.563
Relative permittivity, ε_r	11.8	16.0	13.5
Density, g/cm^3	2.33	5.32	–
Energy gap, E_G (eV)	1.08	0.66	1.58
Electron mobility, μ_n (m^2/V–s)	0.13	0.38	0.85
Hole mobility, μ_p (m^2/V–s)	0.05	0.18	0.04
Intrinsic concentration, n_i(m^{-3})	1.38×10^{16}	2.5×10^{19}	9×10^{12}
Electron diffusion constant, D_n (m^2/S) = $\mu_n V_T$	0.0034	0.0099	0.020
Hole diffusion constant, D_P (m^2/S) = $\mu_p V_T$	0.0013	0.0047	–
Density of states at conduction band edge, N_C (m^{-3})	2.8×10^{25}	1.0×10^{25}	4.7×10^{23}
Density of states at valance band edge, N_V (m^{-3})	1.0×10^{25}	6.0×10^{24}	7.0×10^{24}
Intrinsic resistivity, Ω–cm	23×10^4	45	–
Melting point	1420	936	1250

4.7 VARIATION IN SEMICONDUCTOR PARAMETERS WITH TEMPERATURE

As the semiconductors operate on a wide range of temperature, the variation of semiconductor parameters such as intrinsic concentration (n_i), mobility (μ), conduc-

tivity (σ) and energy gap (E_G) with temperature are important and are discussed below.

Intrinsic concentration(n_i) We know that the electron concentration,

$$n = N_C e^{-(E_C - E_F)/kT}$$

and hole concentration,

$$p = N_V e^{-(E_F - E_V)/kT}$$

Therefore, the product of electron-hole concentrations,

$$np = N_C N_V e^{-(E_C - E_V)/kT}$$

$$= N_C N_V e^{-(E_G)/kT}$$

According to mass-action law, $np = n_i^2$. Hence, it is clear that the intrinsic concentration, n_i, is independent of the Fermi level but depends on temperature and energy gap, $E_G = E_C - E_V$. The above equation is valid for either an extrinsic or intrinsic material.

Also, in an intrinsic semiconductor, as temperature rises the intrinsic concentration (n_i) increases and the conductivity (σ) also increases. The intrinsic concentration n_i is very sensitive to temperature and is given by

$$n_i^2 = A_o T^3 e^{-E_{Go}/kT}$$

where A_o = a constant, independent of temperature
T = temperature (in °K)
E_{GO} = forbidden energy gap at °K (in eV) and
k = Boltzmann constant (in eV/°K).

The increase in n_i^2 with temperature has an effect on the charge densities in extrinsic semiconductor. In an N-type semiconductor, the number of free electrons (n) does not change appreciably with the increase in temperature, while the number of holes increases. In a P-type semiconductor, the number of free electrons (n) increases with the increase in temperature, while the number of holes (p) remains constant.

Mobility (μ) The mobility (μ) of an intrinsic semiconductor varies as T^{-m} over a temperature (T) range of 100 to 400 °K. For silicon, $m = 2.5$ for electrons and 2.7 for holes. Similarly, for germanium $m = 1.66$ for electrons and 2.33 for holes. The mobility of an intrinsic semiconductor decreases with the rise in temperature. Also, the mobility is a function of electric field intensity (EV) and it remains constant only if $E < 10^3$ m in N-type silicon. For $10^3 < E < 10^4$ V/m, μ_n varies approximately as $E^{-1/2}$. For higher fields, $\mu_n = \frac{v_d}{E}$ and hence, μ_n varies inversely as E and the carrier speed approaches a constant value of about 10^5 m/s.

Conductivity (σ) The conductivity (σ) of an intrinsic semiconductor depends upon the number of hole–electron pairs and mobility. The number of hole-electron pairs increases with the rise in temperature, while its mobility decreases. However, the increase in hole–electron pairs is greater than the decrease in their mobility. Hence,

the conductivity of an intrinsic semiconductor increases with the increase in temperature. The conductivity at any temperature (T °K) is given by

$$\sigma = \sigma_o [1 + \alpha(T - T_o)]$$

where α is the temperature coefficient.

The conductivity of extrinsic semiconductor decreases with the rise in temperature, as the number of majority carriers is almost constant and mobility decreases.

Energy gap The energy gap decreases with the increase in temperature and is given by

$$E_G(T) = E_{Go} - \beta T$$

where β = a constant, whose value depends upon the nature of the material. Its value for silicon is 3.6×10^{-4} and for germanium 2.23×10^{-4}, and E_{Go} = energy gap at 0°K. Its value for silicon is 1.21 eV and for germanium, 0.785 eV.

4.8 DRIFT AND DIFFUSION CURRENTS

The flow of charge, i.e current, through a semiconductor material are of two types, namely drift and diffusion. The net current that flows through a PN junction diode also has two components, viz. (i) drift current and (ii) diffusion current.

Drift current When an electric field is applied across the semiconductor material, the charge carriers attain a certain drift velocity v_d, which is equal to the product of the mobility of the charge carriers and the applied electric field intensity, E. The holes move towards the negative terminal of the battery and electrons move towards the positive terminal. This combined effect of movement of the charge carriers constitutes a current known as the *drift current*. Thus the drift current is defined as the flow of electric current due to the motion of the charge carriers under the influence of an external electric field. The drift current density due to the charge carriers such as free electrons and holes are the current passing through a square centimeter perpendicular to the direction of flow. The equation for the drift current density, J_n, due to free electrons is given by

$$J_n = qn\mu_n E \text{ A/cm}^2$$

and the drift current density, J_p, due to holes is given by

$$J_p = qp\mu_p E \text{ A/cm}^2$$

where n = number of free electrons per cubic centimetre
p = number of holes per cubic centimetre
μ_n = mobility of electrons in cm^2/V-s
μ_p = mobility of holes in cm^2/V-s
E = applied electric field intensity in V/cm
q = charge of an electron = 1.6×10^{-19} coulomb.

Diffusion current It is possible for an electric current to flow in a semiconductor even in the absence of the applied voltage provided a concentration gradient exists in the material. A concentration gradient exists if the number of either electrons or holes is greater in one region of a semiconductor as compared to the rest of the region. In a semiconductor material, the charge carriers have the tendency to move

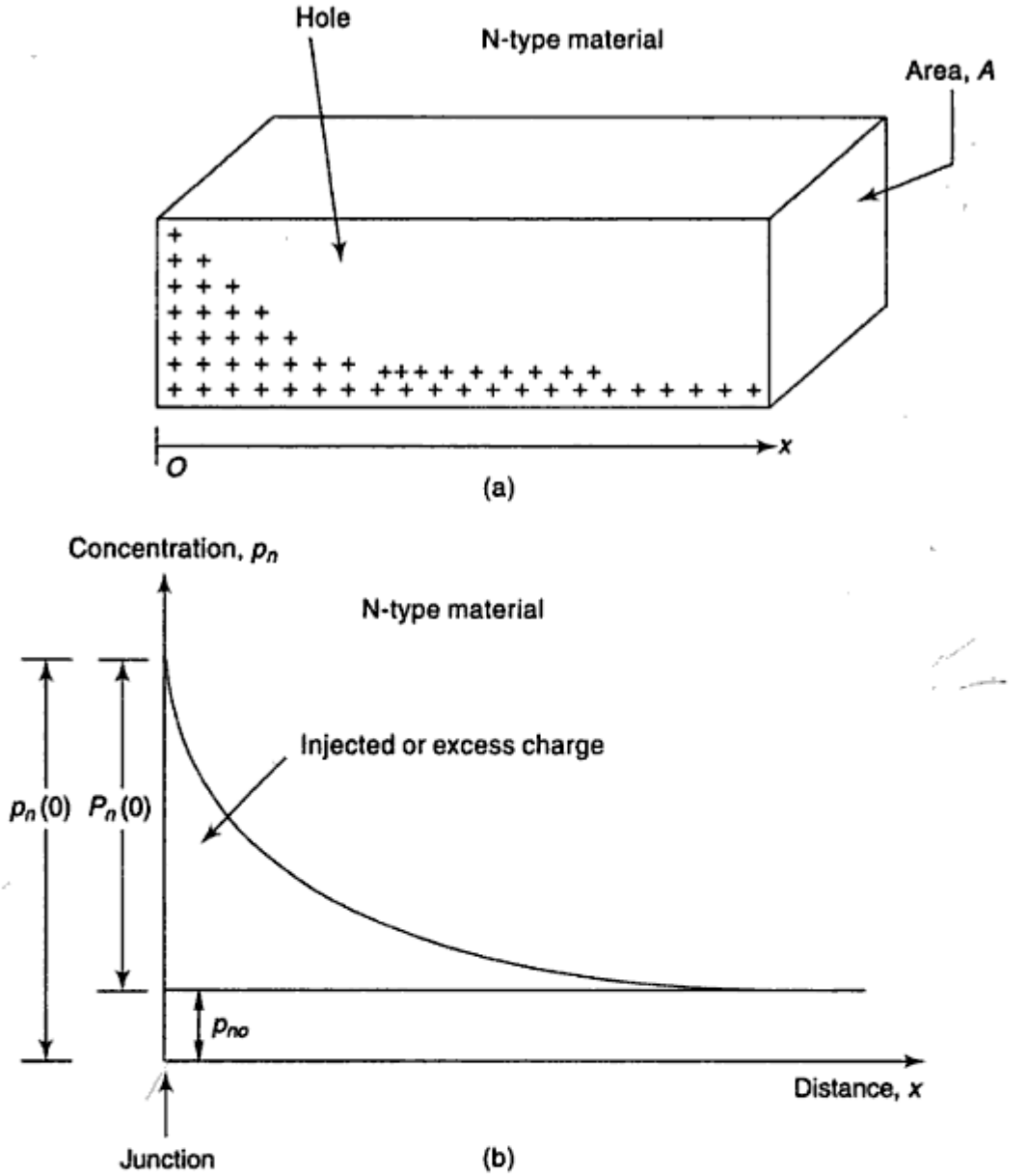


Fig. 4.6 (a) Excess hole concentration varying along the axis in an N-type semiconductor bar, and (b) The resulting diffusion current

from the region of higher concentration to that of lower concentration of the same type of charge carriers. Thus, the movement of charge carriers takes place resulting in a current called *diffusion current*. The diffusion current depends on the material of the semiconductor, type of charge carriers and the concentration gradient.

As indicated in Fig. 4.6(a), the hole concentration $p(x)$ in a semiconductor bar varies from a high value to a low value along the x-axis and is constant in the y- and z-directions.

Diffusion current density due to holes, J_p, is given by

$$J_p = - qD_p \frac{dp}{dx} \text{ A/cm}^2$$

Since the hole density $p(x)$ decreases with increasing x as shown in Fig. 4.6(b), dp/dx is negative and the minus sign in the above equation is needed in order that J_p has a positive sign in the positive x-direction.

Diffusion current density due to the free electrons, J_n, is given by

$$J_n = qD_n \frac{dn}{dx} \text{ A/cm}^2$$

where dn/dx and dp/dx are the concentration gradients for electrons and holes respectively, in the x-direction and D_n and D_p are the diffusion coefficients expressed in cm^2/s for electrons and holes, respectively.

Total current The total current in a semiconductor is the sum of drift current and diffusion current. Therefore, for a P-type semiconductor, the total current per unit area, i.e. the total current density is given by

$$J_p = qp\mu_p E - qD_p \frac{dp}{dx}$$

Similarly, the total current density for an N-type semiconductor is given by

$$J_n = qn\,\mu_n\,E + qD_n \frac{dn}{dx}$$

Einstein Relationship for Semiconductor

There exists a definite relationship between the mobility and diffusion coefficient of a particular type of charge carrier in the same semiconductor. The higher the value of mobility of a charge carrier, the greater will be its tendency to diffuse. The equation which relates the mobility μ and the diffusion coefficient D is known as the *Einstein Relationship*. The Einstein relationship is expressed as

$$\frac{D_p}{\mu_p} = \frac{D_n}{\mu_n} = \frac{kT}{q} = V_T \tag{4.5}$$

The importance of Einstein relationship is that it can be used to determine D_p (or D_n), if the mobility of holes (or electrons) is measured experimentally. For an intrinsic silicon, $D_p = 13$ cm^2/s and $D_n = 34$ cm^2/s. For an intrinsic germanium, $D_p = 47$ cm^2/s and $D_n = 99$ cm^2/s

Diffusion length (L) As shown in Fig. 4.6, the excess hole or electron densities fall off exponentially with distance as a result of the recombination of these excess minority carriers with the majority carriers of the semiconductor. Here, the excess charge carriers have a finite life time, τ, before they are totally destroyed by recombination. The average distance that on excess charge carrier can diffuse during its life time is called the diffusion length L, which is given by

$$L = \sqrt{D\tau}$$

where D is the diffusion coefficient that may be related to the drift mobility, μ, through the Einstein relation as

$$D = \mu \frac{kT}{q}$$

If the transverse length of the semiconductor is greater than the diffusion length L, then the terminal currents are the recombination currents arising out of the

recombination, as every electron lost by recombination is supplanted by the terminal electrode to maintain the charge neutrality.

4.9 CARRIER LIFE TIME

We know that in an intrinsic semiconductor the number of holes is equal to the number of free electrons. However, thermal agitation generates new electron – hole pairs. The electrons have a limited life time (τ_n) in the conduction band and periodically fall back to the valance band, so that a mobile electron – hole pair disappear in the recombination process with the energy of excitation appearing as heat energy. On the average, an electron will exist for τ_n seconds and a hole for τ_p seconds before recombination. The *mean life times* τ_n and τ_p of electrons and holes are very important parameters as they indicate the time required for the excessive electron and hole concentrations to return to their equilibrium values.

Thus, the carrier life time is defined as the time for which, on an average, a charge carrier will exist before recombination with a carrier of opposite charge. Its value varies from nanoseconds to hundreds of micro-seconds and depends on temperature and impurity concentration in the semiconductor material. Gold is extensively used as recombination agent by the manufacturers of semiconductor devices.

Consider an N-type semiconductor having thermal equilibrium concentration p_o and n_o of holes and electrons, respectively. When the specimen is illuminated by light or injection of carriers, additional electron–hole pairs are generated uniformly throughout the medium. This causes the concentration of holes and electrons to increase from p_o and n_o to new values. The rate of change of hole density is equal to their rate of generation minus the rate at which they recombine with electrons. Thus,

$$\frac{dp}{dt} = G - R$$

where p is the minority hole concentration at any time t, and G and R are the generation and recombination rates for the minority carriers.

The generation rate, G, is a function of temperature only since charge carriers are produced only by thermal excitation in the absence of current flow. Hence, at a constant temperature, $G(t)$ is constant.

Assuming that the mean life time of hole τ_p is independent of the magnitude of the hole concentration, we get

$$R = \frac{p}{\tau_p} = \text{decrease in hole concentration per second due to recombination.}$$

Also, from the definition of the generation rate, we get

G = increase in hole concentration per second due to thermal generation.

Since charge can neither be created nor destroyed, the rate dp/dt, at every instant of time, equals the algebraic sum of the rates given in the above equations of R and G. Thus,

$$\frac{dp}{dt} = G - \frac{p}{\tau_p}$$

under steady state conditions, $\frac{dp}{dt} = 0$ and with no radiation falling on the semiconductor, the hole concentration p reaches its equilibrium value p_o. Therefore, $G = \frac{p_o}{\tau_p}$

Hence, $\frac{dp}{dt} = \frac{p_o - p}{\tau_p}$.

A similar equation can be derived in an exactly similar manner for the change in excess electron density with time in the P-type semiconductor and is given by

$$\frac{dn}{dt} = \frac{n_o - n}{\tau_n}$$

4.10 CONTINUITY EQUATION

The fundamental law governing the flow of charge is called the *Continuity Equation*. The continuity equation as applied to semiconductors describes how the carrier concentration in a given elemental volume of the crystal varies with time and distance. The variation in density is attributable to two basic causes, viz. (i) the rate of generation and loss by recombination of carriers within the element, and (ii) drift of carriers into or out of the element. The continuity equations enable us to calculate the excess density of electrons or holes in time and space.

As shown Fig. 4.7, consider an infinitesimal N-type semiconductor bar of volume of area A and length dx and the average minority carrier (hole) concentration p, which is very small compared with the density of majority carriers. At time t, if minority carriers (holes) are injected, the minority current entering the volume at x is I_p and leaving at $x + dx$ is $I_p + d I_p$ which is predominantly due to diffusion. The minority carrier concentration injected into one end of the semiconductor bar decreases exponentially, with distance into the specimen, as a result of diffusion and recombination. Here, $d I_p$ is the decrease in number of coulombs per second within the volume.

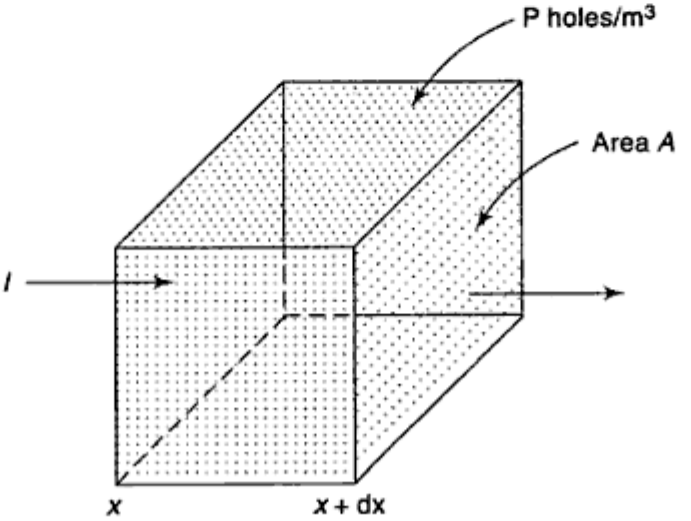


Fig. 4.7 Relating to continuity equation

Since the magnitude of the carrier charge is q, then $\frac{dI_p}{q}$ equals the decrease in the number of holes per second within the elemental volume $A \propto x$. As the current density $J_p = \frac{I_p}{A}$, we have

$$\frac{1}{q_A} \cdot \frac{dI_p}{dx} = \frac{1}{q} \cdot \frac{dJ_p}{dx} = \text{decrease in hole concentration per second, due to current } I_p.$$

We know that there is an increase of holes per unit volume per second given by $G = p_o/\tau_p$ due to thermal generation. Further, there is a decrease of holes per unit volume per second given by $R = p/\tau_p$ due to recombination but charge can neither be created nor destroyed. Hence, increase in holes per unit volume per second, dp/dt, must equal the algebraic sum of all the increases in hole concentration. Thus,

$$\frac{\partial p}{\partial t} = -\frac{p - p_o}{\tau p} - \frac{1}{q}\frac{\partial J_p}{\partial x}$$

where $$J_p = -qD_p \frac{dp}{dx} + qp\mu_p E$$

Therefore, $$\frac{\partial p}{\partial t} = -\frac{p - p_o}{\tau_p} + D_p \frac{d^2 p}{dx^2} - \mu_p \frac{d(pE)}{dx}$$

This is the *Continuity equation* or *equation of Conservation of Charge* for holes stating the condition of dynamic equilibrium for the density of mobile carrier holes. Here, partial derivatives have been used since both p and J_p are functions of both t and x.

Similarly, the continuity equation for electrons states the condition of dynamic equilibrium for the density of mobile carrier electrons and is given by

$$\frac{\partial n}{\partial t} = \frac{n_o - n}{\tau n} - \frac{1}{q}\frac{\partial J_n}{\partial x}$$

where $$J_n = -qD_n \frac{dn}{dx} + qn\mu_n E$$

Therefore, $$\frac{\partial n}{\partial t} = -\frac{n - n_o}{\tau_n} + D_n \frac{d^2 n}{dx^2} - \mu_n \frac{d(nE)}{dx}$$

We now consider three special cases of continuity equation.

(i) *Concentration independent of distance with zero electric field.* For this special case the continuity equation can be changed into

$$\frac{\partial p}{\partial t} = -\frac{p - p_o}{\tau_p}$$

Solving the above equation, we get

$$p - p_o = A_1 e^{-t/\tau_p} \text{ where } A \text{ is a constant}$$

(ii) *Concentration independent of time with zero electric field.* For this special case the continuity equation can be changed into

$$0 = -\frac{p - p_o}{\tau_p} + D_p \frac{d^2 p}{dx^2}$$

$$\frac{d^2 p}{dx^2} = \frac{p - p_o}{\tau_p D_p}$$

Solving the above equation, we get

$$p - p_o = A_1 e^{-x/L_p} + A_2 e^{x/L_p}$$

where A_1 and A_2 are constants.

$L_p = \sqrt{D_p \tau_p}$ = diffusion length for holes

(iii) *Concentration varies sinusoidally with time and with zero electric field.* Let $P(x, t) = P(x)\, e^{j\omega t}$

For this special case the continuity equation can be changed into

$$j\omega P(x) = -\frac{P(x)}{\tau_p} + D_p \frac{d^2 P(x)}{dx^2}$$

$$\frac{d^2 P}{dx^2} = \frac{(1 + j\omega \tau_p)}{L_p^2} P$$

At $\omega = 0$,

$$\frac{d^2 P}{dx^2} = \frac{P}{L_p^2}$$

The above equation is the same as that of the second special case.

(iv) *Minority carrier injection in homogeneous semiconductors* Consider a semiconductor bar shown in Fig. 4.6(a). This bar is uniformly doped with donar atoms so that the charge concentration $n = N_D$ is uniform throughout the bar on which radiation falls on one end of the bar at $x = 0$. Near the illuminated surface, the bound electrons in the covalent bonds capture some of the photons. This energy transfer results in breaking of covalent bonds and generation of hole-electron pairs.

The minority carrier (hole) concentration p is very small compared with the doping level, i.e., $p < n$. The condition $p = p + p_0 \ll n$ which states that the minority concentration is much smaller than the majority concentration is called the low-level injection. The controlling differential equation for p is

$$\frac{d^2 p}{dx^2} = \frac{p - p_0}{\tau_p D_p}$$

The diffusion length for holes L_p is given by

$$L_p = \sqrt{D_p \tau_p}$$

The differential equation for the injected concentration $p = p - p_0$ becomes

$$\frac{d^2 p}{dx^2} = \frac{P}{L_p^2}$$

The solution of the equation is

$$p(x) = A_1 e^{-x/L_p} + A_2 e^{x/L_p}$$

when $x \to \infty$, $A_2 = 0$. At $x = 0$, the injected concentration $P(0)$. To satisfy this boundary condition $A_1 = P(0)$. Therefore,

$$P(x) = P(0) e^{-x/L_p} = P(x) - p_0$$

Here the hole concentration decreases exponentially with distance as shown in Fig. 4.6(b).

(v) *Diffusion Current* The minority (hole) diffusion current is $I_p = AJ_p$, where A is the area of cross-section of the bar. Therefore,

$$I_p(x) = \frac{A_q D_p P(0)}{L_p} e^{-x/L_p}$$

$$= \frac{A_q D_p}{L_p} (P(0) - p_0) e^{-x/L_p}$$

This current decreases exponentially with distance x as that of minority carrier concentration.

The majority (electron) diffusion current is

$$A_q D_n \frac{dn}{dx} = A_q D_n \frac{dp}{dx} = -\frac{D_n}{D_p} I_p$$

where $I_p = -A_q D_n \frac{dp}{dx}$. The magnitude of ratio of majority to minority diffusion current is $D_n/D_p \approx 3$ for Si and 2 for Ge.

(vi) *Drift Current* For an open circuit semiconductor bar, the sum of the hole and electron currents should be zero everywhere. Therefore, a majority (electron) drift current I_{nd} exist such that

$$I_p + \left(I_{nd} - \frac{D_n I_p}{D_p} \right) = 0$$

Therefore,

$$I_{nd} = \left(\frac{D_n}{D_p} - 1 \right) I_p.$$

The hole drift current I_{pd} is given by

$$I_{pd} = Aqp\mu_p E = \frac{p}{n}\frac{\mu_p}{\mu_n}\left(\frac{D_n}{D_p} - 1\right) I_p$$

where the electric field, $$E = \frac{1}{Aqn\mu_n}\left(\frac{D_n}{D_p} - 1\right) I_p$$

Here as $p \ll n$, then $I_{pd} \ll I_p$, i.e., the hole drift current is negligible compared to the hole diffusion current.

4.11 THEORY OF PN JUNCTION DIODE

4.11.1 PN Junction Diode in Equilibrium with no Applied Voltage

In a piece of semiconductor material, if one half is doped by P-type impurity and the other half is doped by N-type impurity, a PN junction is formed. The plane dividing the two halves or zones is called PN junction. As shown in Fig. 4.8, the N-type material has high concentration of free electrons, while P-type material has high concentration of holes. Therefore, at the junction there is a tendency for the free electrons to diffuse over to the P-side and holes to the N-side. This process is called *diffusion.* As the free electrons move across the junction from N-type to P-type, the donor ions become positively charged. Hence a positive charge is built on the N-side of the junction. The free electrons that cross the junction uncover the negative acceptor ions by filling in the holes. Therefore, a net negative charge is established on the P-side of the junction. This net negative charge on the P-side prevents further diffusion of electrons into the P-side. Similarly, the net positive charge on the N-side repels the hole crossing from P-side to N-side. Thus a barrier is set-up near the junction which prevents further movement of charge carriers. i.e. electrons and holes. As a consequence of the induced electric field across the depletion layer, an electrostatic potential difference is established between P-and N-regions, which is called the potential barrier, junction barrier, diffusion potential, or contact potential, V_o. The magnitude of the contact potential V_o varies with doping levels and temperature. V_o is 0.3 V for germanium and 0.72 V for silicon.

The electrostatic field across the junction caused by the positively charged N-type region tends to drive the holes away from the junction and negatively charged P-type region tends to drive the electrons away from the junction. The majority holes diffusing out of the P-region leave behind negatively charged acceptor atoms bound to the lattice, thus exposing negative space charge in a previously neutral region. Similarly, electrons diffusing from the N-region expose positively ionised donor atoms, and a double space charge layer builds up at the junction as shown in Figs 4.8 (a) and (c).

It is noticed that the space-charge layers are of opposite sign to the majority carriers diffusing into them, which tends to reduce the diffusion rate. Thus, the double space of the layer causes an electric field to be set up across the junction directed from N- to P-regions, which is in such a direction to inhibit diffusion of majority electrons and holes, as illustrated in Figs 4.8 (a) and (d). The shape of the

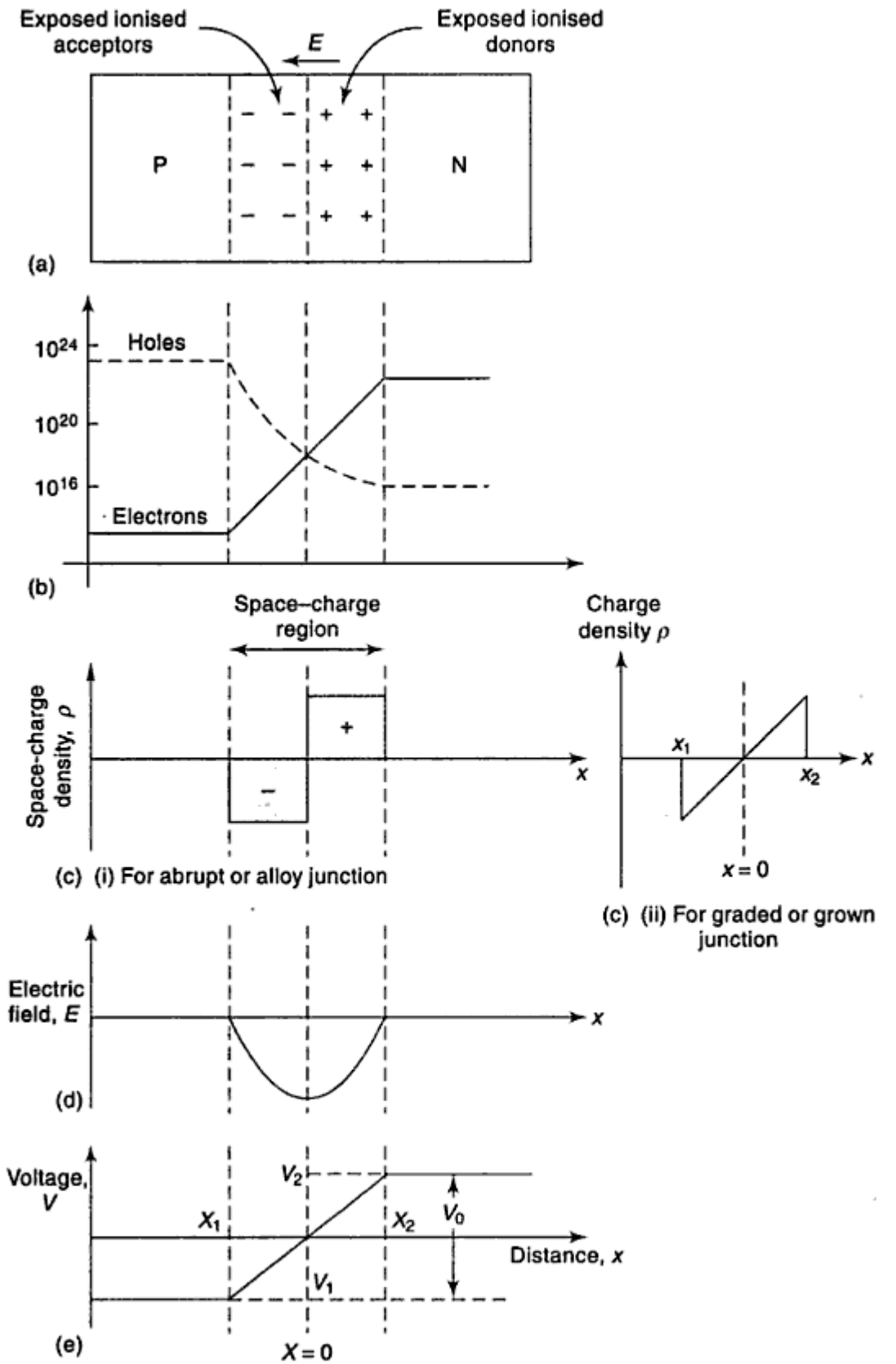


Fig. 4.8 Formation of PN junction

charge density, ρ, depends upon how the diode is doped, Thus, the junction region is depleted of mobile charge carriers. Hence, it is called the depletion region (layer), the space charge region, or the transition region. The depletion region is of order 0.5 μm thick. There are no mobile carriers in this very narrow depletion layer. Hence no current flows across the junction and the system is in equilibrium. To the left of this depletion layer, the carrier concentration is $p \approx N_A$, and to its right it is $n \approx N_D$.

Calculation of depletion width Let us now consider the width of the depletion region in the junction of Fig. 4.8. The region contains space charge due to the fact that, donors on the N-side and acceptors on the P-side have lost their accompanying electrons and holes. Hence, an electric field is established which, in turn, causes a difference in potential energy, qV_o, between the two parts of the specimen. Thus, a potential is built up across the junction and Fig. 4.8 (e) represents the variation in potential. Here, P-side of the junction is at a lower potential than the N-side which means that the electrons on the P-side have a great potential energy.

In this analysis, let us consider an *Alloy Junction* in which there is an abrupt change from acceptor ions on P-side to donor ions on N-side. Assume that the concentration of electrons and holes in the depletion region is negligible and that all of the donors and acceptors are ionised. Hence, the regions of space charge may be described as

$$\rho = -qN_A,\ 0 > x > X_1$$
$$\rho = +qN_D,\ X_2 > x > 0$$
$$\rho = 0, \text{ elsewhere}$$

where ρ is the space charge density, as indicated in Fig. 4.8c (i). The axes have been chosen in Fig. 4.8(e) in such a way that V_1 and X_1 have negative values. The potential variation in the space charge region can be calculated by using Poission's equation, which is given by

$$\nabla^2 V = -\frac{\rho(x, y, z)}{\varepsilon_o \varepsilon_r}$$

where ε_r is the relative permittivity. The relevant equation for the required one-dimensional problem is

$$\frac{d^2 V}{dx^2} = -\frac{\rho}{\varepsilon_o \varepsilon_r}$$

Applying the above equation to the P-side of the junction, we get

$$\frac{d^2 V}{dx^2} = \frac{qN_A}{\varepsilon_o \varepsilon_r}$$

Integrating twice, we get

$$V = \frac{qN_A x^2}{2\varepsilon_o \varepsilon r} + Cx + D$$

where C and D are the constants of integration.

From the Fig. 4.8(e), we have $V = 0$ at $x = 0$, and hence $D = 0$. When $x < X_1$ on the P-side, the potential is constant, so that $\frac{dV}{dx} = 0$ at $x = X_1$ Hence,

$$C = -\frac{qN_A}{\varepsilon_o \varepsilon_r} \cdot X_1$$

Therefore, $$V = \frac{qN_A x^2}{2\varepsilon_o \varepsilon_r} - \frac{qN_A}{\varepsilon_o \varepsilon_r} \cdot X_1 \cdot x$$

i.e. $$V = \frac{qN_A}{\varepsilon_o \varepsilon_r}\left(\frac{x^2}{2} - X_1 \cdot x\right)$$

As $V = V_1$ at $x = X_1$, we have

$$V_1 = -\frac{qN_A}{2\varepsilon_o \varepsilon_r} \cdot X_1^2$$

If we apply the same procedure to the N-side, we get

$$V_2 = \frac{qN_D}{2\varepsilon_o \varepsilon_r} \cdot X_2^2$$

Therefore, the total built-in potential or the contact potential is V_o, where

$$V_o = V_2 - V_1 = \frac{q}{2\varepsilon_o \varepsilon_r}\left(N_A X_1^2 + N_D X_2^2\right)$$

We know the fact that the positive charge on the N-side must be equal in magnitude to the negative charge on the P-side for the neutral specimen. Hence,

$$N_A X_1 = -N_D X_2$$

and substituting this relationship in the above equation and using the fact that X_1 is a negative quantity, we get

$$X_1 = -\left[\frac{2\varepsilon_o \varepsilon_r V_o}{qN_A\left(1 + \frac{N_A}{N_D}\right)}\right]^{1/2}$$

Similarly,

$$X_2 = \left[\frac{2\varepsilon_o \varepsilon_r V_o}{qN_D\left(1 + \frac{N_D}{N_A}\right)}\right]^{1/2}$$

The total depletion width, $W = X_2 - X_1$ and hence, $W^2 = X_1^2 + X_2^2 - 2X_1 X_2$, and then substituting for X_1 and X_2 from the above equations, we find

$$W = \left[\frac{2\varepsilon_o \varepsilon_r V_o}{q}\left(\frac{N_A + N_D}{N_A N_D}\right)\right]^{1/2}.$$

Here, in an *alloy junction*, the depletion width W is proportional to $(V_o)^{1/2}$.

In a *Grown Junction*, the charge density (ρ) varies linearly with distance (x) as shown in Fig. 4.8 c(ii). If a similar analysis is carried for this junction, it is found that W varies as $(V_o)^{1/3}$ instead of $(V_o)^{1/2}$.

4.11.2 Under Forward Bias Condition

When positive terminal of the battery is connected to the P-type and negative terminal to the N-type of the PN junction diode, the bias applied is known as forward bias.

Operation As shown in Fig. 4.9, the applied potential with external battery acts in opposition to the internal potential barrier and disturbs the equilibrium. As soon as equilibrium is disturbed by the application of an external voltage, the Fermi level is no longer continuous across the junction. Under the forward bias condition, the applied positive potential repels the holes in P-type region so that the holes move towards the junction and the applied negative potential repels the electrons in the N-type region and the electrons move towards the junction. Eventually, when the applied potential is more than the internal barrier potential the depletion region and internal potential barrier disappear.

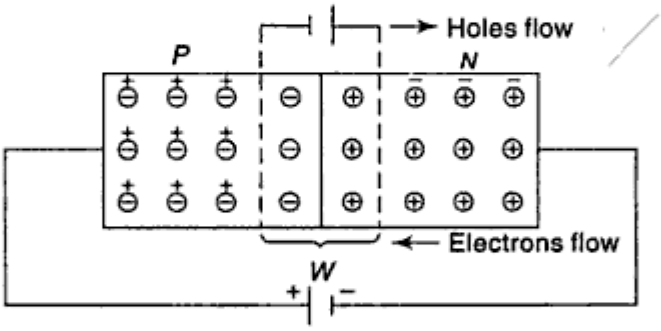


Fig. 4.9 PN junction under forward bias

V–I Characteristics of a diode under forward bias Under forward bias condition, the *V–I* characteristics of a PN unction diode are shown in Fig. 4.10. As the forward voltage (V_F) is increased, for $V_F < V_O$, the forward current I_F is almost zero (region *OA*) because the potential barrier prevents the holes from P-region and electrons from N-region to flow across the depletion region in the opposite direction.

For $V_F > V_O$, the potential barrier at the junction completely disappears and hence, the holes cross the junction from P-type to N-type and the electrons cross the junction in the opposite direction, resulting in relatively large current flow in the external circuit.

A feature worth to be noted in the forward characteristics shown in Fig. 4.10 is the cut in or threshold voltage (V_r) below which the

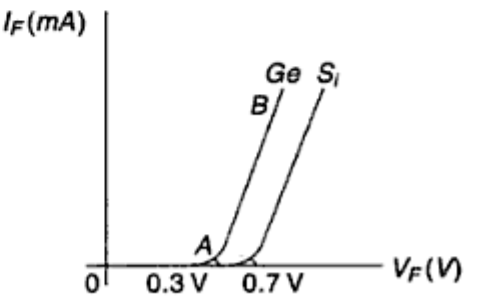


Fig. 4.10 *V–I* characteristics of a diode under forward bias condition

current is very small. It is 0.3 V and 0.7 V for germanium and silicon, respectively. At the cut in voltage, the potential barrier is overcome and the current through the junction starts to increase rapidly.

4.11.3 Under Reverse Bias Condition

When the negative terminal of the battery is connected to the P-type and positive terminal of the battery is connected to the N-type of the PN junction, the bias applied is known as reverse bias.

Operation Under applied reverse bias as shown in Fig. 4.11, holes which form the majority carriers of the P-side move towards the negative terminal of the battery and electrons which form the majority carrier of the N-side are attracted towards the positive terminal of the battery. Hence, the width of the depletion region which is depleted of mobile charge carriers increases. Thus, the electric field produced by applied reverse bias, is in the same direction as the electric field of the potential barrier. Hence, the resultant potential barrier is increased which prevents the flow of majority carriers in both directions; the depletion width, W, is proportional to $\sqrt{V_o}$ under reverse bias. Therefore, theoretically no current should flow in the external circuit. But in practice, a very small current of the order of a few microamperes

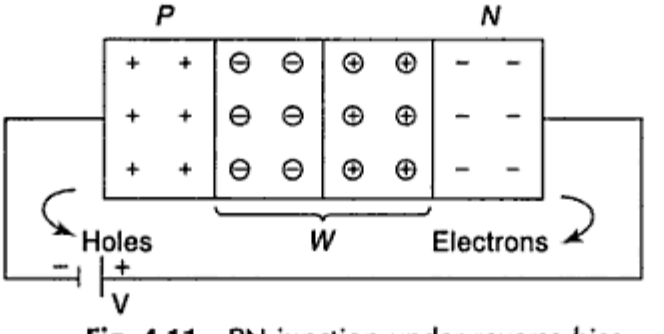


Fig. 4.11 PN junction under reverse bias

flows under reverse bias as shown in Fig. 4.12. Electrons forming covalent bonds of the semiconductor atoms in the P- and N-type regions may absorb sufficient energy from heat and light to cause breaking of some covalent bonds. Hence electron–hole pairs are continually produced in both the regions. Under the reverse bias condition, the thermally generated holes in the P-region are attracted towards the negative terminal of the battery and the electrons in the N-region are attracted towards the positive terminal of the battery. Consequently, the minority carriers, electrons in the P-region and holes in the N-region, wander over to the junction and flow towards their majority carrier side giving rise to a small reverse current. This current is known as *reverse saturation current, I_o*. The magnitude of reverse saturation current mainly depends upon junction temperature because the major source of minority carriers is thermally broken covalent bonds.

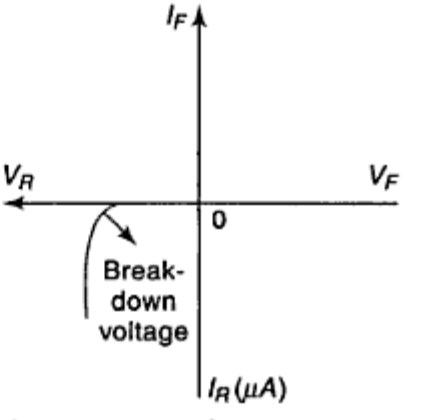


Fig. 4.12 *V–I* characteristics under reverse bias.

We know that

$$np = N_C N_V e^{-E_G/kT} \text{ and}$$

$$np = n_i^2 \text{ (Mass-action law)}$$

From the above equations, we get

$$E_G = kT \ln \frac{N_c N_v}{n_i^2} \tag{4.9}$$

We know that for N-type material $E_F = E_C - kT \ln \frac{N_C}{N_D}$. Therefore, from this equation, we get

$$E_{cn} - E_F = kT \ln \frac{N_C}{n_n} = kT \ln \frac{N_C}{N_D} \tag{4.10}$$

Similarly for P-type material $E_F = E_V + kT \ln \frac{N_V}{N_A}$. Therefore, from this equation, we get

$$E_F - E_{vp} = kT \ln \frac{N_V}{p_p} = kT \ln \frac{N_V}{N_A} \tag{4.11}$$

Substituting from Eqs (4.9), (4.10) and (4.11) into Eqn. (4.8), we get

$$E_0 = kT \left[\ln \frac{N_C N_V}{n_i^2} - \ln \frac{N_C}{N_D} - \ln \frac{N_V}{N_A} \right]$$

$$= kT \ln \left[\frac{N_C N_V}{n_i^2} \times \frac{N_D}{N_C} - \frac{N_D}{N_V} \right]$$

$$= kT \ln \frac{N_D N_A}{n_i^2} \tag{4.12}$$

As $E_0 = qV_o$, then the contact difference of potential or barrier voltage is given by

$$V_o = \frac{kT}{q} \ln \frac{N_D N_A}{n_i^2}$$

In the above equations, Es in electron volts and k is in electron volt per degree Kelvin. The contact difference of potential V_o is expressed in volt and is numerically equal to E_0. From Eqn. (4.12) we note that E_o (hence V_o) depends upon the equilibrium concentrations and not on the charge density in the transition region.

An alternative expression for E_0 may be obtained by substituting the equations of $n_n \approx N_D$, $p_n = \frac{n_i^2}{N_D}$, $n_n\, p_p = n_i^2$, $p_p \approx N_A$ and $n_p = \frac{n_i^2}{N_A}$ into Eqn. (4.12). Then we get

$$E_0 = kT \ln \frac{p_{po}}{p_{no}} = kT \ln \frac{n_{no}}{n_{po}} \tag{4.13}$$

where subscript 0 represents the thermal equilibrium condition.

Example 4.11 (a) The resistivities of the P-region and N-region of a germanium diode are 6 Ω-cm and 4Ω-cm, respectively. Calculate the contact potential V_o and potential energy barrier E_o. (b) If the doping densities of both P and N-regions are doubled, determine V_o and E_o. Given that $q = 1.6 \times 10^{-19}$ C, $n_i = 2.5 \times 10^{13}/\text{cm}^3$, $\mu_p = 1800$ cm^2/V-s, $\mu n = 3800$ cm^2/V-s and $V_T = 0.026$ V at 300 °K.

Solution: (a)Resistivity, $\rho = \frac{1}{\sigma} = \frac{1}{N_A q \mu_p} 6\Omega - \text{cm}$

Therefore, $N_A = \frac{1}{6q\mu_p} = \frac{1}{6 \times 1.6 \times 10^{-19} \times 1800} = 0.579 \times 10^{15}/\text{cm}^3$

Similarly, $N_D = \frac{1}{4q\mu_n} = \frac{1}{4 \times 1.6 \times 10^{-19} \times 3800} = 0.579 \times 10^{15}/\text{cm}^3$

Therefore, $V_o = V_T \ln \frac{N_D N_A}{n_i^2} = 0.026 \ln \frac{0.579 \times 0.411 \times 10^{30}}{(2.5 \times 10^{13})^2} = 0.1545$ V

Hence $E_0 = 0.1545$ eV

(b) $V_o = 0.026 \ln \frac{2 \times 0.579 \times 10^{15} \times 2 \times 0.411 \times 10^{15}}{(2.5 \times 10^{13})^2} = 0.1906$ V

Therefore, $E_0 = 0.1906$ eV

4.13 QUANTITATIVE THEORY OF PN DIODE CURRENTS

Let us now derive the expression for the total current as a function of applied voltage assuming that the width of the depletion region is zero. When a forward bias is applied to a diode, holes are injected from the P-side into the N-side. Due to this, the concentration of holes in the N-side (p_n) is increased from its thermal equilibrium value (p_{no}) and injected hole concentration [$P_n(x)$] decreases exponentially with respect to distance (x).

$$P_n(x) = p_n - p_{no} = P_n(0)\, e^{-x/L_p}$$

where L_p is the diffusion length for holes in the N-material.

$$p_n(x) = p_{no} + P_n(0) e^{-x/L_p} \tag{4.14}$$

Injected hole concentration at $x = 0$ is

$$P_n(0) = p_n(0) - p_{no} \tag{4.15}$$

These several components of hole concentration in the N-side of a forward biased diode are shown in Fig. 4.6, in which the density $p_n(x)$ decreases exponentially with distance (x).

Let p_p and p_n be the hole concentration at the edges of the space charge in the P-and N-sides, respectively. Let V_B (= $V_o - V$) be the effective barrier potential across the depletion layer. Then

$$p_p = p_n \, e^{\frac{V_B}{V_T}} \tag{4.16}$$

where V_T is the volt-equivalent of temperature.

This is the Boltzmann's relation of kinetic gas theory. This equation is valid as long as the hole current is small compared with diffusion or drift current. This condition is called Low level injection.

Under open circuit condition (i.e. V = 0), $p_p = p_{po}$, $p_n = p_{no}$ and $V_B = V_o$. Equation (4.11) can be changed into

$$p_{po} = p_{no} \, e^{V_o / V_T} \tag{4.17}$$

Under forward bias condition let V be the applied voltage, then the effective barrier voltage

$$V_B = V_o - V$$

The hole concentration throughout the P-side is constant and equal to the thermal equilibrium value ($p_p = p_{po}$). The hole concentration varies exponentially with distance into the N-side.

$$\text{At } x = 0, \; p_n = p_n \, (0)$$

Equation (4.3) can be changed into

$$p_{po} = pn \, (0) \; e^{(V_o - V)/V_T} \tag{4.18}$$

Comparing Eqs (4.17) and (4.18),

$$p_n \, (0) = p_{no} \, e^{V/V_T}$$

This boundary condition is called the *law of the junction*. Substituting this into Eq. (4.15), we get

$$p_n \, (0) = p_{no} \left(e^{V/V_T} - 1 \right) \in \tag{4.19}$$

The diffusion hole current in the N-side is

$$I_{pn}(x) = -Ae \, D_p \, \frac{dp_n \, (x)}{dx}$$

$$= -A \, eD_p \, \frac{d}{dx} \left[pn_o + P_n \, (0) \, e^{-x/L_p} \right]$$

$$= \frac{AeD_p P_n \, (0)}{L_p} \, e^{-x/L_p}$$

From this equation, it is evident that the injected hole current decreases exponentially with distance.

Forward currents The hole current crossing the junction into the N-side with $x = 0$ is

$$I_{pn}(0) = \frac{Ae D_p P_n(0)}{L_p} = \frac{AeD_p p_{no}}{L_p}(e^{V/V_T} - 1)$$

The electron current crossing the junction into the P-side with $x = 0$ is

$$I_{np}(0) = \frac{Ae D_n N_p(0)}{L_n} = \frac{AeD_n n_{po}}{L_n}(e^{V/V_T} - 1)$$

The total diode current,

$$I = I_{pn}(0) + I_{np}(0) = I_o\left(e^{V/V_T} - 1\right)$$

where $I_o = \frac{Ae D_p p_{no}}{L_p} + \frac{Ae D_n n_{po}}{L_n}$ = reverse saturation current.

If we consider carrier generation and recombination in the space-charge region, the general equation of the diode current is approximately given by

$$I = I_o\left[e^{(V/\eta V_T)} - 1\right]$$

where V = external voltage applied to the diode, and η = a constant, 1 for germanium and 2 for silicon.

Reverse saturation currents We know that $p_n = \frac{n_i^2}{N_D}$ and $n_p = \frac{n_i^2}{N_A}$. Applying these relationships in the above equation of reverse saturation current, I_o, we get

$$I_o = Ae\left[\frac{D_p}{L_p N_D} + \frac{D_n}{L_n N_A}\right] n_i^2$$

where $n_i^2 = A_o T^3 e^{\frac{-E_{G_o}}{kT}} = A_o T^3 e^{\frac{-V_{G_o}}{V_T}}$, where V_{Go} is a voltage which is numerically equal to the forbidden gap energy E_{Go} in electron volts.

For a germanium diode, the diffusion constants D_p and D_n vary approximately inversely proportional to T. Hence, the temperature dependence of I_o is

$$I_o = K_1 T^2 e^{\frac{-V_{Go}}{V_T}}$$

where K_1 is a constant independent of temperature.

For a silicon diode, I_o is proportional to n_i instead of n_i^2. Hence,

$$I_o = K_2 T^{\frac{3}{2}} e^{\frac{-V_{Go}}{2V_T}}$$

where K_2 is a constant independent of temperature.

4.14 DIODE CURRENT EQUATION

The diode current equation relating the voltage V and current I is given by

$$I = I_o [e^{(V/\eta V_T)} - 1]$$

where I = diode current
I_o = diode reverse saturation current at room temperature
V = external voltage applied to the diode
η = a constant, 1 for germanium and 2 for silicon
$V_T = kT/q = T/11600$, volt-equivalent of temperature, i.e., thermal voltage,
where k = Boltzmann's constant (1.38066×10^{-23} J/K
q = charge of the electron (1.60219×10^{-19} C)
T = temperature of the diode junction (K) = (°C + 273°)

At room temperature, (T = 300 °K), V_T = 26 mV. Substituting this value in the current equation, we get

$$I = I_o[e^{(40\,V/\eta)} - 1]$$

Therefore, for germanium diode, $I = I_o [e^{40V} - 1]$, since $\eta = 1$ for germanium
for silicon diode, $I = I_o[e^{20V} - 1]$, since $\eta = 2$ for silicon.

If the value of applied voltage is greater than unity, then the equation of diode current for germanium,

$$I = I_o (e^{40V})$$

and for silicon, $I = I_o (e^{20V})$

When the diode is reverse biased, its current equation may be obtained by changing the sign of the applied voltage V. Thus, the diode current with reverse bias is

$$I = I_o [e^{(-V/\eta V_T)} - 1]$$

If $V >> V_T$, then the term $e^{(-V/\eta V_T)} << 1$, therefore $I \approx -I_o$, termed as reverse saturation current, which is valid as long as the external voltage is below the breakdown value.

Example 4.12 When a reverse bias is applied to a germanium PN junction diode, the reverse saturation current at room temperature is 0.3 μA. Determine the current flowing in the diode when 0.15 V forward bias is applied at room temperature.

Solution: Given $I_o = 0.3 \times 10^{-6}$ A and $V_F = 0.15$ V

The current flowing through the PN diode under forward bias is

$$I = I_o (e^{40\,V_F} - 1)$$
$$= 0.3 \times 10^{-6} (e^{40 \times 0.15} - 1)$$
$$= 120.73\ \mu A$$

Example 4.13 The reverse saturation current of a silicon PN junction diode is 10 µA. Calculate the diode current for the forward-bias voltage of 0.6 V at 25 °C.

Given $V_F = 0.6$ V, $T = 273 + 25 = 298$ K

Example 4.16 Determine the forward resistance of a PN junction diode, when the forward current is 5 mA at $T = 300°$ K. Assume Silicon diode.

Solution:
Given: For a silicon diode, the forward current, $I = 5$ mA, $T = 300$ K

Forward resistance of a PN junction diode, $r_f = \dfrac{\eta V_T}{I}$ where $V_T = \dfrac{T}{11,600}$ and $\eta = 2$ for silicon

Therefore, $r_f = \dfrac{2 \times \dfrac{T}{11,600}}{5 \times 10^{-3}} = \dfrac{2 \times 300}{11,600 \times 5 \times 10^{-3}} = 10.34\ \Omega$

4.16 TRANSITION OR SPACE CHARGE (OR DEPLETION REGION) CAPACITANCE (C_T)

Under reverse bias condition, the majority carriers move away from the junction, thereby uncovering more immobile charges. Hence the width of the space–charge layer at the junction increases with reverse voltage. This increase in uncovered charge with applied voltage may be considered a capacitive effect. The parallel layers of oppositely charged immobile ions on the two sides of the junction form the capacitance, C_T, which is expressed as

$$C_T = \left|\frac{dQ}{dV}\right|$$

where dQ is the increase in charge caused by a change in voltage dV. A change in voltage dV in a time dt will result in a current $I = dQ/dt$ given by

$$I = C_T \frac{dV}{dt} m$$

Therefore C_T is important while considering a diode or a transistor as a circuit element. The quantity C_T is called the transition, space–charge, barrier or depletion region capacitance.

As shown in Fig. 4.14, consider a PN diode which is asymmetrically doped at the junction. Since the net charge is zero, then $eN_A W_p = eN_D W_n$.
If $N_A \gg N_D$, then $W_p \ll W_n \approx W$. The relationship between potential and charge density is given by the Poisson's equation, $\dfrac{d^2V}{dx^2} = \dfrac{eN_A}{\varepsilon}$

Integrating the above equation twice,

$$\int\int d^2V = \int\int \frac{eN_A}{\varepsilon} dx^2$$

Therefore, $$V = \frac{eN_A x^2}{2\varepsilon}$$

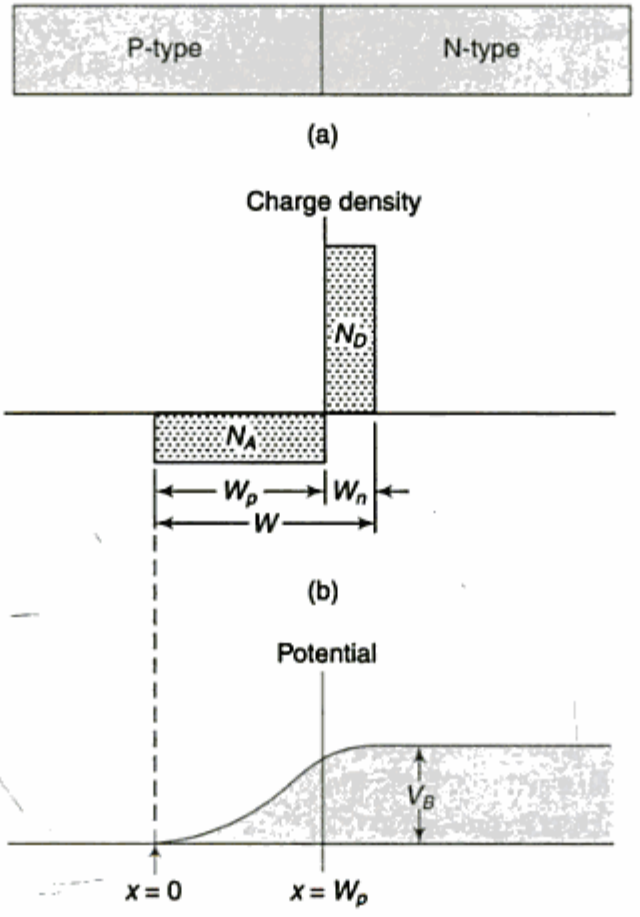


Fig. 4.14 Charge density and potential variation at an alloy PN junction

At $x = W_p \approx W$, $V = V_B$, the barrier potential that appears across the uncovered acceptor ions. Thus

$$V_B = \frac{eN_A W^2}{2\varepsilon} \tag{4.20}$$

Here $V_B = V_0 - V$, where V is a negative number for an applied reverse bias and V_0 is the contact potential. Hence, the width of the depletion layer increases with applied reverse voltage, i.e., $V_B \propto W^2$. Therefore, $W \propto \sqrt{V_B}$.

The total charge density of a p-type material with area of the junction A is given by

$$Q = en_A WA$$

Differentiating the above equation w.r.t. V, we get

$$C_T = \left|\frac{dQ}{dV}\right| = AeN_A \left|\frac{dW}{dV}\right| \tag{4.21}$$

Differentiating Eq. (2.20) w.r.t. V, we get

$$1 = \frac{eN_A\, 2W}{2\varepsilon}\left|\frac{dW}{dV}\right|$$

Therefore, $$\left|\frac{dW}{dV}\right| = \frac{\varepsilon}{eN_A W} \tag{4.22}$$

Substituting eqn. (4.22) in eqn. (4.21), we get

$$C_T = \left|\frac{dQ}{dV}\right| = AeN_A \frac{\varepsilon}{eN_A W}$$

Therefore, $C_T = \frac{\varepsilon A}{W}$.

Here ε is the permittivity of the material, A the cross-sectional area of the junction and W is the width of the depletion layer over which the ions are uncovered. The depletion width, W, is given by

$$W = \left[\frac{2\varepsilon_o \varepsilon_r (V_o - V)}{q}\left(\frac{N_A + N_D}{N_A N_D}\right)\right]^{1/2}$$

where V is the applied voltage and V_o is the barrier potential, or the contact potential.

When no external voltage is applied, i.e. $V = 0$, the width of the depletion region of a PN junction diode is of the order of 0.5 microns. The movement of majority carriers across the junction causes opposite charges to be stored at this distance W apart. This depletion region acts as a dielectric between the two conducting P- and N-regions. Therefore, these regions act as a parallel plate capacitor whose transition capacitance C_T is approximately 20 pF with no external bias.

When forward bias $+V$ is applied, the effective barrier potential, $V_B = [V_o - (+V)]$, is lowered and hence the width of the depletion region W decreases and C_T increases. Under reverse bias condition, the majority carriers move away from the junction, thereby uncovering more immobile charges. Now the effective barrier potential, $V_B = [V_o - (-V)]$, is increased and hence, W increases with reverse voltage and C_T decreases correspondingly. The values of C_T range from 5 to 200 pF, the larger values being for the high power diodes. This property of voltage variable capacitance with the reverse bias appears in varactors, vari-caps or volta-caps.

4.17 DIFFUSION (OR STORAGE) CAPACITANCE (C_D)

The capacitance that exists in a forward biased junction is called a diffusion or storage capacitance (C_D), whose value is usually much larger than C_T, which exists in a reverse-biased junction. This is also defined as the rate of change of injected charge with applied voltage, i.e., $C_D = \frac{dQ}{dV}$, where dQ represents the change in the number of minority carriers stored outside the depletion region when a change in voltage across the diode, dV, is applied.

Calculation of C_D

Let us assume that the P material in one side of the diode is heavily doped in comparison with the N side. Since the holes move from the P to the N side, the hole current $I \approx I_{pn}(0)$.

The excess minority charge Q existing on the N side is given by

$$Q = \int_0^\infty AeP_n(0)\,e^{-x/L_p}\,dx = \left[\frac{AeP_n(0)\,e^{-x/L_p}}{1/L_p}\right]_0^\infty = L_p\,AeP_n(0)$$

Differentiating the above equation, we get

$$C_D = \frac{dQ}{dV} = AeL_p\,\frac{d[P_n(0)]}{dV} \tag{4.23}$$

We know that the diffusion hole current in the N-side is $I_{pn}(x) = \dfrac{AeD_p P_n(0)}{L_p} e^{-x/L_p}$.

The hole current crossing the junction into the N-side with $x = 0$ is $I_{pn}(0) = \dfrac{AeD_p P_n(0)}{L_p}$.

Therefore, $$I = \frac{AeD_p P_n(0)}{L_p}$$

$$P_n(0) = \frac{IL_p}{AeD_p}$$

Differentiating the above equation w.r.t. "V", we get

$$\frac{d[P_n(0)]}{dV} = \frac{dI}{dV}\,\frac{L_p}{AeD_p}$$

Upon substituting in eqn.(2.23), we have

$$C_D = \frac{dQ}{dV} = \frac{dI}{dV}\,\frac{L_p^2}{D_p}$$

Therefore, $C_D = g\tau$, where $g = \dfrac{dI}{dV}$ is the diode conductance and $\tau = \dfrac{L_p^2}{D_p}$ is the mean life time of holes in the N-region.

From diode current equation, $g = \dfrac{I}{\eta V_T}$.

Therefore, $$C_D = \frac{\tau I}{\eta V_T}.$$

where τ is the mean life time for holes and electrons.

Diffusion capacitance C_D increases exponentially with forward bias or, alternatively, that it is proportional to diode forward current, I. The values of C_D range from 10 to 1000 pF, the larger values being associated with the diode carrying a larger anode current, I.

The effect of C_D is negligible for a reverse-biased PN junction. As the value of C_D is inversely proportional to frequency, it is high at low frequencies and it decreases with the increase in frequency.

4.18 EFFECT OF TEMPERATURE ON PN JUNCTION DIODES

The rise in temperature increases the generation of electron–hole pairs in semiconductors and increases their conductivity. As a result, the current through the PN junction diode increases with temperature as given by the diode current equation,

$$I = I_o \left[e^{(V/\eta V_T)} - 1 \right]$$

The reverse saturation current I_o of diode increases approximately 7 per cent/°C for both germanium and silicon. Since $(1.07)^{10} \approx 2$, reverse saturation current approximately doubles for every 10 °C rise in temperature. Hence, if the temperature is increased at fixed voltage, the current I increases. To bring the current I to its original value, the voltage V has to be reduced. It is found that at room temperature for either germanium or silicon, $\frac{dV}{dT} \approx -2.5$ mV/ °C in order to maintain the current I to a constant value.

At room temperature, i.e at 300 K, the value of barrier voltage or cut-in voltage is about 0.3 V for germanium and 0.7 V for silicon. The barrier voltage is temperature dependent and it decreases by 2 mV/°C for both germanium and silicon. This fact may be expressed in mathematical form, which is given by

$$I_{O2} = I_{O1} \times 2^{(T_2 - T_1)/10}$$

where I_{O1} = saturation current of the diode at temperature (T_1), and I_{O2} = saturation current of the diode at temperature (T_2).

Figure 4.15 shows the effect of increased temperature on the characteristic curve of a PN junction diode. A germanium diode can be used up to a maximum of 75 °C and a silicon diode to a maximum of 175 °C.

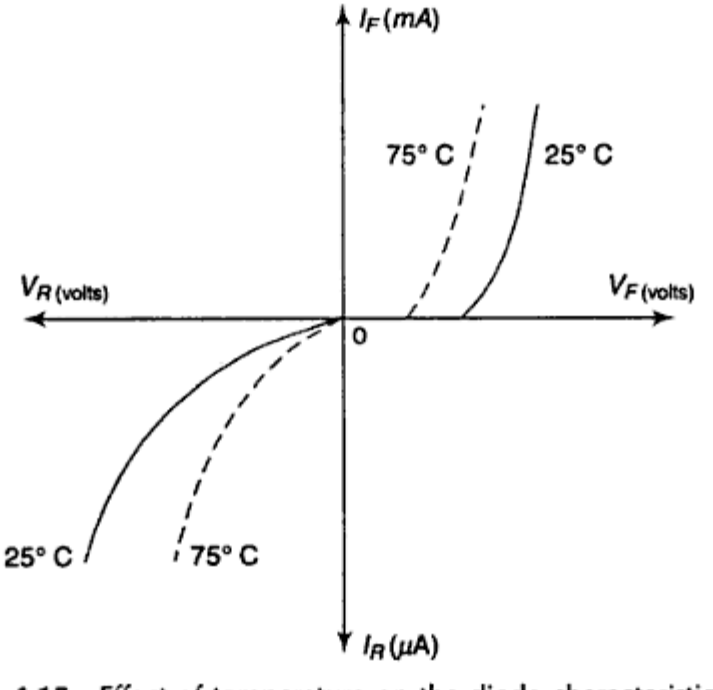


Fig. 4.15 Effect of temperature on the diode characteristics

completed (i) when even the minority carriers remote from the junction have diffused to the junction and crossed it, and (ii) when the junction transition capacitance, C_T, across the reverse-biased junction has got charged through the external resistor R_L to the voltage $-V_R$.

The reverse recovery time (or turn-off time) of a diode, t_{rr}, is the interval from the current reversal at $t = t_1$ until the diode has recovered to a specified extent in terms either of the diode current or of the diode resistance, i.e. $t_{rr} = t_s + t_t$.

For commercial switching type diodes the reverse recovery time, t_{rr}, ranges from less than 1 ns up to as high as 1 μs. This switching time obviously limits the maximum operating frequency of the device. If the time period of the input signal is such that $T = 2 \cdot t_{rr}$, then the diode conducts as much in reverse as in the forward direction. Hence it does not behave as a one way device. In order to minimise the effect of the reverse current, the time period of the operating frequency should be a minimum of approximately 10 times t_{rr}. For example, if a diode has t_{rr} of 2 ns, its maximum operating frequency is

$$f_{\max} = \frac{1}{T} = \frac{1}{10 \times t_{rr}} = \frac{1}{10 \times 2 \times 10^{-9}} = 50 \text{ MHz}$$

The t_{rr} can be reduced by shortening the length of the P-region in a PN junction diode. The stored charge and, consequently, the switching time can also be reduced by introduction of gold impurities into the junction diode by diffusion. The gold dopant, some times called a *life timekiller*, increases the recombination rate and removes the stored minority carriers. This technique is used to produce diodes and other active devices for high speed applications.

4.20 BREAKDOWN IN PN JUNCTION DIODES

The diode equation predicts that, under reverse bias conditions, a small constant current, the saturation current, I_o, flows due to minority carriers, which is independent of the magnitude of the bias voltage. But this prediction is not entirely true in practical diodes. There is a gradual increase of reverse current with increasing bias due to the ohmic leakage currents around the surface of the junction. Also, there is a sudden increase in reverse current due to some sort of *breakdown*, when the reverse bias voltage approaches a particular value called breakdown voltage, V_{BD}, as shown in Fig. 4.17. Once breakdown occurs, the diode is no longer blocking current and the diode current can be controlled only by the resistance of the external circuit.

The breakdown occurs due to avalanche effect in which thermally generated minority carriers cross the depletion region and acquire sufficient kinetic energy from the applied potential to produce new carriers by removing valence electrons from their bonds. These new carriers will in turn collide with other atoms and will increase the number of electrons and holes available for conduction. This multiplication effect of free carriers may be represented by the following equation:

$$M = \frac{1}{1 - \left(\frac{V}{V_{BD}}\right)^n}$$

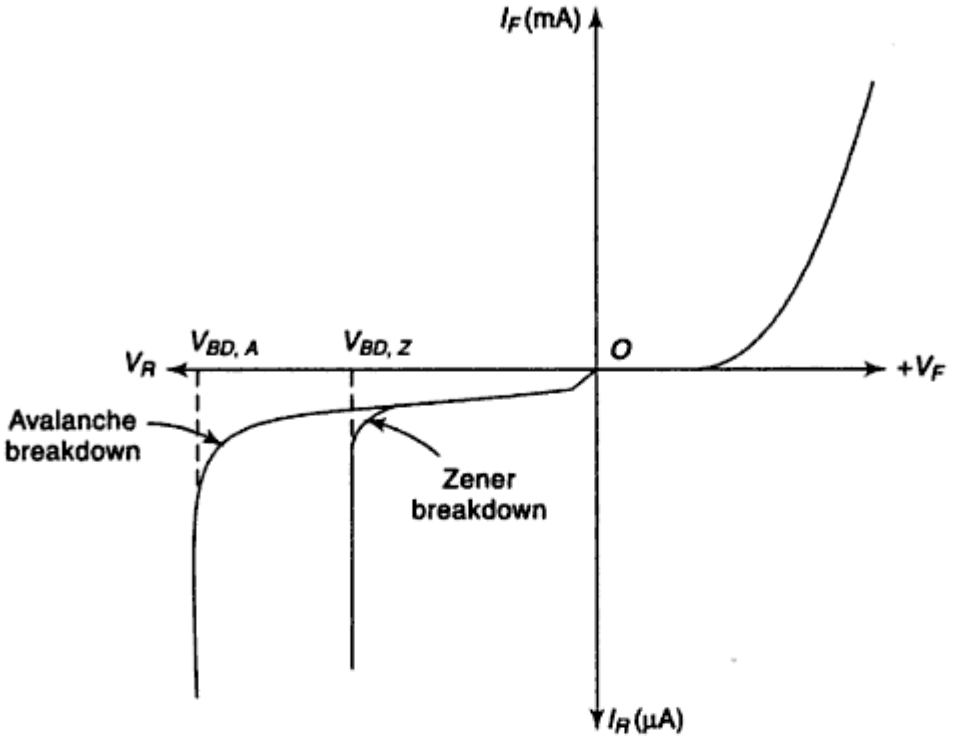


Fig. 4.17 Breakdown in PN junction diodes

where M = carrier multiplication factor, which is the ratio of the total number of electrons leaving the depletion region to the number entering the region

V = applied reverse voltage

V_{BD} = reverse breakdown voltage

n = empirical constant, which depends on the lattice material and the carrier type, for N-type silicon, $n \approx 4$ and for P-type, $n \approx 2$

It is evident from the above equation that M is being very small for $V = 0.9\ V_{BD}$. But when $V > 0.9\ V_{BD}$, M is very large and the resulting reverse current is given by $I_R = MI_o$, where I_o is the reverse saturation current before breakdown.

As V approaches the breakdown voltage V_{BD}, the value of M will become infinite and there is a rapid increase in carrier density and a corresponding increase in current. Because of the cumulative increase in carrier density after each collision, the process is known as *avalanche breakdown.*

Even if the initially available carriers do not gain enough energy to disrupt bonds, it is possible to initiate breakdown through a direct rupture of the bonds because of the existence of strong electric field. Under these circumstances the breakdown is referred to as *Zener breakdown.*

Voltage reference diodes that utilise the almost constant voltage characteristics in the breakdown region are also called variously *avalanche diodes* or sometimes *Zener diodes.* The Zener effect is in diodes with breakdown voltages below about 6V. The operating voltage in avalanche breakdown are from several volts to several hundred volts with power ratings up to 50 W.

True Zener diode action displays a negative temperature coefficient, i.e, breakdown voltage decreases with increasing temperature. True avalanche diode action

exhibits a positive temperature coefficient, i.e. breakdown voltage increases with increasing temperature.

It is clear that the breakdown voltage for a particular diode can be controlled during manufacture by altering the doping levels in the junction. The breakdown voltage for silicon diodes can be made to occur at a voltage as low as 5 V with 10^{17} impurity atoms per cubic cm or as high as 1000 V when doped to a level of only 10^{14} impurity atoms per cubic cm.

4.21 DIODE AS A CIRCUIT ELEMENT

The PN junction diode is considered as a circuit element. The basic diode circuit shown in Fig. 4.18 consists of a d.c. voltage V_S which is supplied across a resistor and a diode. In order to find the instantaneous diode voltage V and current I, the circuit can be analysed when the instantaneous source voltage is V_S. From Kirchhoff's voltage law (KVL), the instantaneous diode voltage is

$$V_S = IR + V \tag{4.24}$$

which can be expressed as

$$I = \frac{V_S}{R} - \frac{V}{R} \tag{4.25}$$

The ideal diode current equation relating the diode voltage V_D and the current I_D is

$$I = I_0 \left[e^{(V/\eta V_T)} - 1 \right] \tag{4.26}$$

where I_0 is the diode reverse saturation current at room temperature, η is a constant (1 for Ge and 2 for Si) and V_T is the thermal voltage, i.e., $V_T = \frac{T}{11,600}$ in which T is the temperature of the diode junction (300° K) or $V_T = 26$ mV at room temperature.

Substituting Eq. (4.26) into Eq. (4.24), we get

$$V_S = I_0 R \left[e^{(V/\eta V_T)} - 1 \right] + V \tag{4.27}$$

which has only one unknown variable V. Since Eq. (4.27) is a transcendental equation, this equation cannot be solved directly.

As these equations contain both linear and exponential terms, it is difficult to solve by hand. The iteration (trial and error) technique can be used to find a solution to this equation. The graphical analysis technique involves plotting two simultaneous equations and locating their point of intersection, which is called the quiescent point, or the Q-point.

4.21.1 Load Line Concept

The use of the load line construction allows the graphical analysis of many circuits including devices which are much more complicated than the PN diode. Two plots are needed to determine the operating point of the diode. Using Kirchhoff's current and voltage laws draws one plot and other is by plotting the volt–ampere characteristic of the diode. From KCL, $I = I_R$ and from KVL, $V + V_R = V_S$ as shown in Fig. 4.18. The relationship between diode current I and voltage V is obtained from

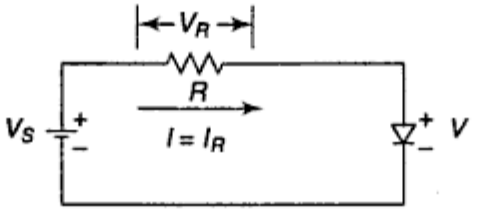


Fig. 4.18 Simple diode circuit

the diode characteristics curve as shown in Fig. 4.19(a). From Ohm's law, the current through the resistor and the voltage across the resistor are linearly related as shown in Fig. 4.19(b).

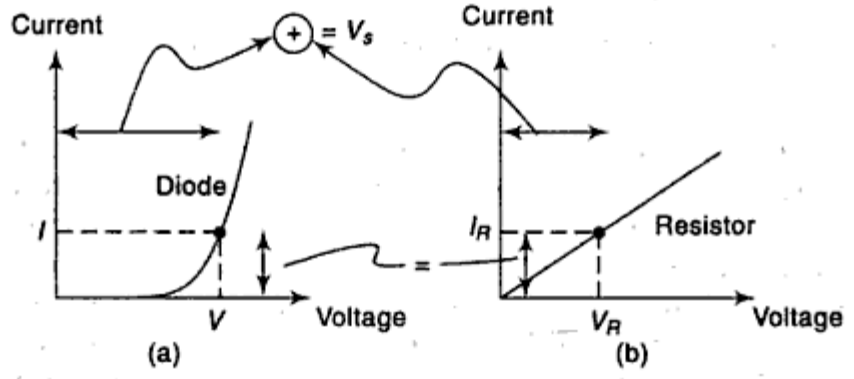


Fig. 4.19 (a) Diode characteristics for the circuit shown in Fig. 4.18 and (b) Output across resistor

To draw the load line, flip the resistor curve horizontally such that the slope of the curve is $-1/R$ and push the two curves, as shown in Figs 4.20(a) and 4.20(b), together horizontally until the y-axes are V_S apart.

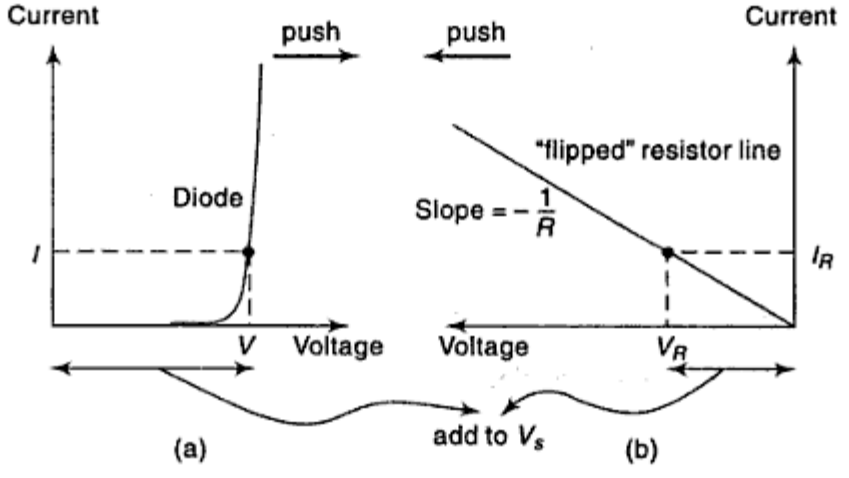


Fig. 4.20 (a) Forward voltage–current characteristic of a diode and (b) Flipped resistor line with slope $= -\frac{1}{R}$

The intersection point of the flipped resistor line called load line and the diode static characteristics curve is the operating point of the device, as shown in Fig. 4.21.

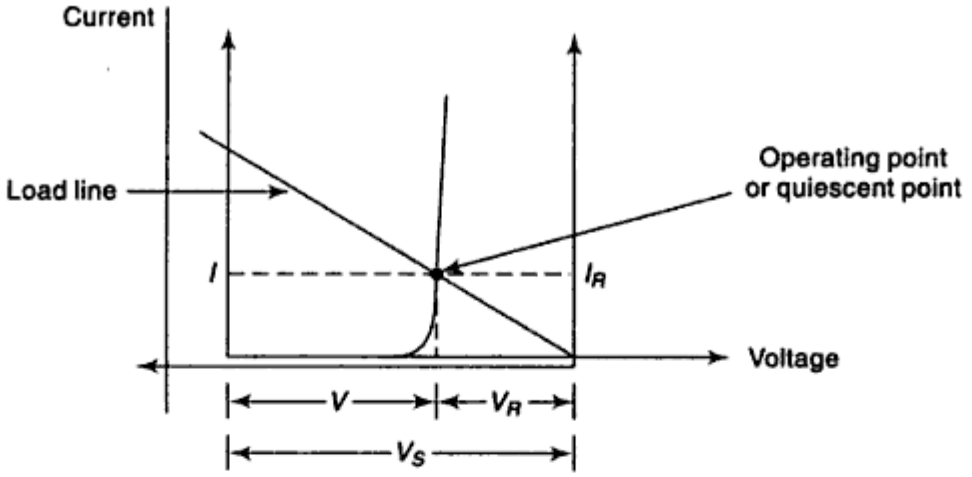


Fig. 4.21 Intersection point of the load line with the voltage-current characteristic of the diode

In the second method to draw the load line, I is determined when the device is short-circuited and V is determined when the device is open-circuited, so that the point B ($V = 0$, $I = V_S/R$) and $A(V = V_S, I = 0)$ lies somewhere on the y-axes and x-axes respectively of the diode characteristics curve as shown in Fig. 4.22. Thus a straight line drawn connecting the points A and B is called the load line. This load line intersects the diode characteristics at some point which is chosen as operating point for the device. The operating point provides the diode voltage V, appearing across the diode and the current I, flowing through the diode.

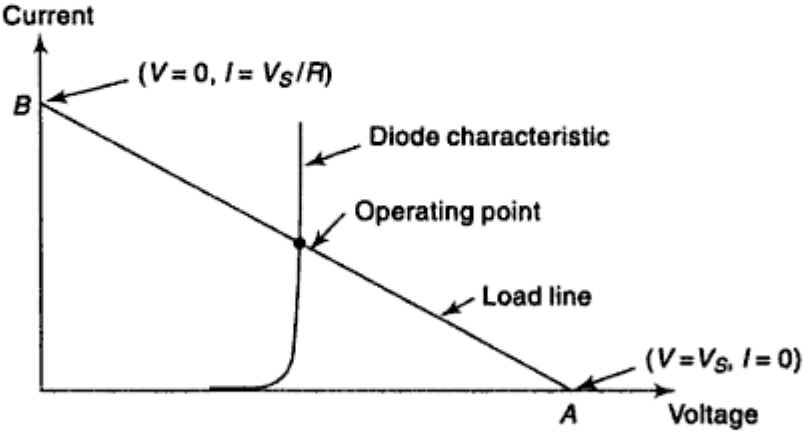


Fig. 4.22 Load line curve for fixing the operating point

4.22 PIECEWISE LINEAR DIODE MODEL

As the volt–ampere relationship of the diode is non-linear, the analysis of circuits containing diodes is difficult. With the help of piecewise linear approximation model of the diode, the results can easily be obtained. The particular regions of operation of the volt–ampere characteristics of the diode are broken into linear segments and the concept of a diode cut-in voltage is also used in this piecewise linear model.

If the reverse resistance R_r is included in the diode characteristics, then the piecewise linear and continuous volt–ampere characteristic is obtained. Piecewise

15. The mobility of electrons and holes in a sample of intrinsic germanium at room temperature are 0.36 m^2/V-s and 0.17 m^2/V-s, respectively. If the electron and hole densities are each equal to $2.5 \times 10^{19}/m^3$, calculate the conductivity.
(*Ans.* $\sigma_i = 2.12$ S/m)
16. Compute the conductivity of a silicon semiconductor which is doped with acceptor impurity to a density of 10^{22} atoms/m^3. Given that $n_i = 1.4 \times 10^{16}/m^3$, $\mu_n = 0.145$ m^2/V-s and $\mu_p = 0.05$ m^2/V-s. (Ans.: 80 S/m)
17. The conductivity of a pure silicon at room temperature is 5×10^{-4} S/m. How many aluminium atoms per m^3 are required so that a saturation conductivity of 200 S/m could be achieved in silicon using aluminium as an impurity? Given that the mobility of holes in Silicon is 0.05 m^2/V-s and the mobility of electrons is 0.13 m^2/V-s.
(*Ans.*: $2.5 \times 10^{22}/m^3$)
18. Calculate the conductivity of a pure silicon at room temperature of 300 °K. Given that $n_i = 1.5 \times 10^{16}/m^3$, $\mu_n = 0.13$ m^2/V-s, $\mu_p = 0.05$ m^2/V-s and $q = 1.6 \times 10^{-19}$ C. Now the silicon is doped 2 in 10^8 of a donor impurity. Calculate its conductivity if there are 5×10^{28} silicon atoms/m^3. By what factor has the conductivity increase?
(*Ans*: 4.32×10^{-4} s/m; 20.8 S/m; ≈ 48,000)
19. The mobility of free electrons and holes in pure silicon are 0.13 and 0.05 m^2/V-s and the corresponding values for pure germanium are 0.38 and 0.18 m^2/V-s respectively. Determine the values of intrinsic conductivity for both silicon and germanium. Given that $n_i = 2.5 \times 10^{19}/m^3$ for germanium and $n_i = 1.5 \times 10^{16}/m^3$ for silicon at room temperature. (*Ans*: 0.43 S/m; 2.24 S/m).
20. (a) A crystal of pure germanium has sufficient antimony (N-type or donor impurities) added to produce 1.5×10^{22} antimony atoms/m^3. The electron and hole mobility are 0.38 m^2/V-s and 0.18 m^2/V-s respectively, and the intrinsic charge carrier density is $2.5 \times 10^{19}/m^3$. Calculate (i) the density of electrons and holes in the crystal, and (ii) the conductivity.
    (b) A second Germanium crystal is produced which is doped with 2.5×10^{22} indium (P-types or acceptor impurities) atoms/m^3. Repeat the calculations listed in part (a).
    (c) A PN junction is made by joining the two crystals described above. Calculate its barrier voltage at 300 °K.
(*Ans.* (a) $n = 1.5 \times 10^{22}/m^3$, $p = 4.167 \times 10^{16}/m^3$, $\sigma = 912/\Omega$-m
(b) $n = 2.5 \times 10^{16}/m^3$, $p = 2.5 \times 10^{22}$ /m^3, $\sigma = 720/\Omega$-m (c) 0.335 V
21. Explain the drift and diffusion currents for a semiconductor.
22. State and explain Mass-action law.
23. What is Einstein relationship in a PN junction?
24. Derive the continuity equation from the first principle.
25. With reference to charge carriers in a semiconductor, define and explain the terms (i) mobility and (ii) life time.
26. What is a PN junction? How is it formed?
27. Explain the formation of depletion region in a PN junction.
28. Draw the energy band diagram of a PN junction and explain the working of a diode.
29. (a) The resistivities of the P-region and N-region of a silicon diode are 6 Ω-cm and 4Ω-cm respectively. Calculate the contact potential V_o and potential energy barrier E_o. (b) if the doping densities of both P-and N-regions are tripled, determine V_o and E_o. Given that $q = 1.6 \times 10^{-19}$ C, $n_i = 1.5 \times 10^{10}/cm^3$, $\mu_p = 500$ cm^2/V-s, $\mu_n = 1300$ cm^2/V-s and $V_T = 0.026$ V at 300 °K.
(*Ans.* (a) $V_o = 0.6$ *V*, $E_o = 0.6$ eV (b) $V_o = 0.66$V, $E_o = 0.66$ eV)
30. Show that in an intrinsic semiconductor, the Fermi level is located at the middle of the unallowable energy gap.

31. Sketch the conduction and valence bands before and after diffusion of carriers in a PN junction.
32. Explain how a barrier pointial is developed at the PN junction.
33. Describe the action of PN junction diode under forward bias and reverse bias.
34. Show that the PN diode works as a rectifier.
35. Explain how unidirectional current flow is possible through a PN junction diode.
36. Explain *V–I* characteristics of a PN junction diode.
37. Indicate the differences between the characteristics of silicon and germanium diodes and state approximately their cut-in voltages.
38. Explain the following terms in a PN junction diode:
    (a) Maximum forward current
    (b) Peak inverse voltage, and
    (c) Maximum power rating
39. Explain the terms: (i) Static resistance, (ii) dynamic resistance, (iii) junction resistance, and (iv) reverse resistance of a diode.
40. Write the volt–ampere equation for a PN diode. Give the meaning of each symbol.
41. What are the factors governing the reverse saturation current in a PN junction diode?
42. Determine the forward bias voltage applied to a silicon diode to cause a forward current of 10 mA and reverse saturation current, $I_o = 25 \times 10^{-7}$ A at room temperature. (*Ans.*: 0.4 V)
43. The reverse saturation current I_o in a germanium diode is 6 μA. Calculate the current flowing through the diode when the applied forward bias voltages are 0.2, 0.3 and 0.4 V at room temperature. (*Ans*: 13.15 mA, 21.5 mA, 28.8 mA)
44. Define the term transition capacitance C_T of a PN diode.
45. Explain the term diffusion capacitance C_D of a forward biased diode.
46. Explain the effect of temperature of a diode.
47. Distinguish between avalanche and Zener mechanisms.
48. Can an ordinary rectifier diode be used as a Zener diode? Explain.
49. Mention the applications of PN junction diode.
50. Give the ideal diode current-voltage relationship. Describe the meaning of I_0 and V_T.
51. What is iteration method of analysis?
52. Define a load line in a simple diode circuit.
53. Describe the piecewise linear approximation model of a diode.

5 Special Diodes

5.1 INTRODUCTION

In addition to the PN junction diode, other types of diodes are also manufactured for specific applications. These special diodes are two terminal devices with their doping levels carefully selected to give the desired characteristics.

5.2 ZENER DIODE

When the reverse voltage reaches breakdown voltage in normal PN junction diode, the current through the junction and the power dissipated at the junction will be high. Such an operation is destructive and the diode gets damaged. Whereas diodes can be designed with adequate power dissipation capabilities to operate in the breakdown region. One such a diode is known as Zener diode. Zener diode is heavily doped than the ordinary diode.

From the *V–I* characteristics of the Zener diode, shown in Fig. 5.1, it is found that the operation of Zener diode is same as that of ordinary PN diode under forward-biased condition. Whereas under reverse-baised condition, breakdown of the junction occurs. The breakdown voltage depends upon the amount of doping. If the diode is heavily doped, depletion layer will be thin and, consequently, breakdown occurs at lower reverse voltage and further, the breakdown voltage is sharp. Whereas a lightly doped diode has a higher breakdown voltage. Thus breakdown voltage can be selected with the amount of doping.

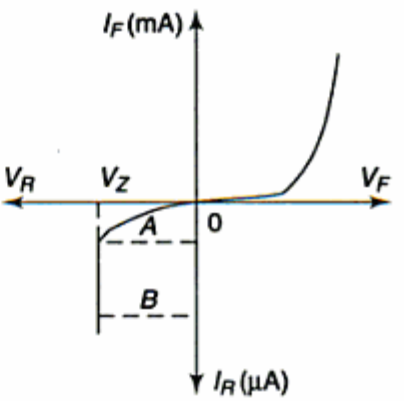


Fig. 5.1 *V–I* characteristics of a zener diode

The sharp increasing current under breakdown conditions are due to the following two mechanisms.

(1) Avalanche breakdown
(2) Zener breakdown.

5.2.1 Avalanche Breakdown

As the applied reverse bias increases, the field across the junction increases

correspondingly. Thermally generated carriers while traversing the junction acquire a large amount of kinetic energy from this field. As a result the velocity of these carriers increases. These electrons disrupt covalent bond by colliding with immobile ions and create new electron-hole pairs. These new carriers again acquire sufficient energy from the field and collide with other immobile ions thereby generating further electron–hole pairs. This process is cumulative in nature and results in generation of avalanche of charge carriers within a short time. This mechanism of carrier generation is known as Avalanche multiplication. This process results in flow of large amount of current at the same value of reverse bias.

5.2.2 Zener Breakdown

When the P and N regions are heavily doped, direct rupture of covalent bonds takes place because of the strong electric fields, at the junction of PN diode. The new electron–hole pairs so created increase the reverse current in a reverse biased PN diode. The increase in current takes place at a constant value of reverse bias typically below 6 V for heavily doped diodes. As a result of heavy doping of P and N regions, the depletion region width becomes very small and for an applied voltage of 6 V or less, the field across the depletion region becomes very high, of the order of 10^7 V/m, making conditions suitable for Zener breakdown. For lightly doped diodes, Zener breakdown voltage becomes high and breakdown is then predominantly by Avalanche multiplication. Though Zener breakdown occurs for lower breakdown voltage and Avalanche breakdown occurs for higher breakdown voltage, such diodes are normally called Zener diodes.

5.2.3 Applications

From the Zener characteristics shown in Fig. 5.1, under the reverse bias condition, the voltage across the diode remains almost constant although the current through the diode increases as shown in region *AB*. Thus, the voltage across the Zener diode serves as a reference voltage. Hence, the diode can be used as a voltage regulator.

In Fig. 5.2 it is required to provide constant voltage across load resistance R_L, whereas the input voltage may be varying over a range. As shown, Zener diode is reverse biased and as long as the input voltage does not fall below V_Z (Zener breakdown voltage), the voltage across the diode will be constant and hence the load voltage will also be constant.

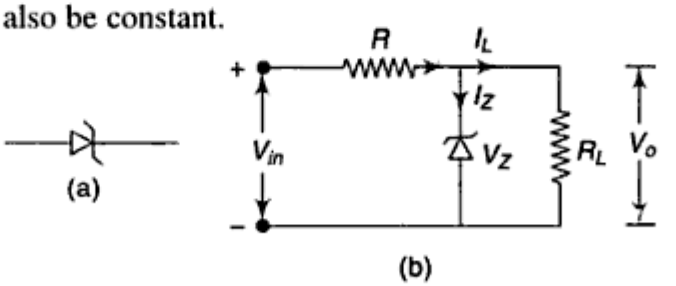


Fig. 5.2 Zener diode. (a) Circuit symbol and (b) as a voltage regulator

5.3 BACKWARD DIODE

Zener diodes are normally used under the reverse bias condition. Their breakdown voltages are normally greater than 2 V. Forward conduction occurs in the voltage around 0.7 V (Si). The breakdown, i.e. Zener effect under the reverse bias condition can also be obtained near zero (around 0.1 V).

Figure 5.3 shows the characteristics of a backward diode. It is called a backward diode because it conducts better in the reverse than in the forward direction.

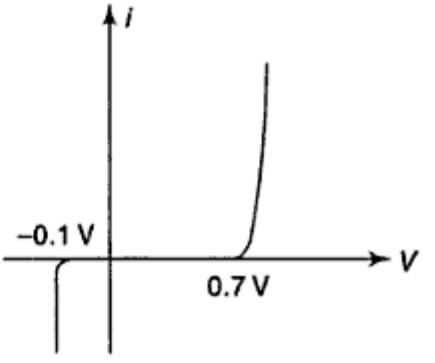


Fig. 5.3 Characteristics of backward diode

Figure 5.4 shows the low level rectifier circuit using a backward diode. Consider a sine wave with a peak of 0.5 V is given as the input to the backward diode. This voltage (0.5 V) is not enough to forward bias to diode into conduction, but it is enough to breakdown the diode. Hence, the output is a rectified half wave signal with a peak of 0.4 V (0.1 V is lost across the diode). Backward diodes are used to rectify weak signals whose peak amplitudes are between 0.1 and 0.7 V.

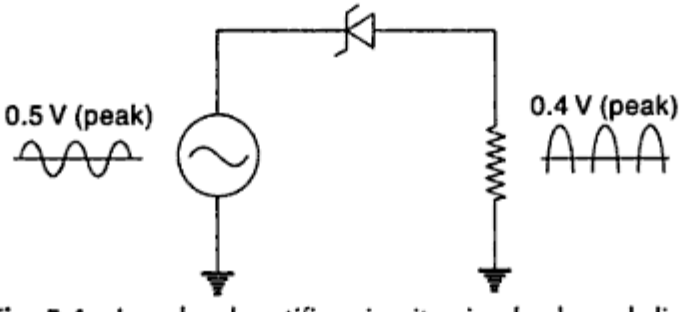


Fig. 5.4 Low level rectifier circuit using backward diode

5.4 VARACTOR DIODE

The varactor, also called a varicap, tuning or voltage variable capacitor diode, is a junction diode with a small impurity dose at its junction, which has the useful property that its junction or transition capacitance is easily varied electronically.

When any diode is reverse biased, a depletion region is formed, as seen in Fig. 5.5. The larger the reverse bias applied across the diode, the width of the depletion layer "*W*" becomes wider. Conversely, by decreasing the reverse bias voltage, the depletion region width "*W*" becomes narrower. This depletion region is devoid of majority carriers and acts like an insulator preventing conduction between the N and P regions of the diode, just like a dielectric, which separates the two plates of a capacitor. The varactor diode with its symbol is shown in Fig. 5.6(a).

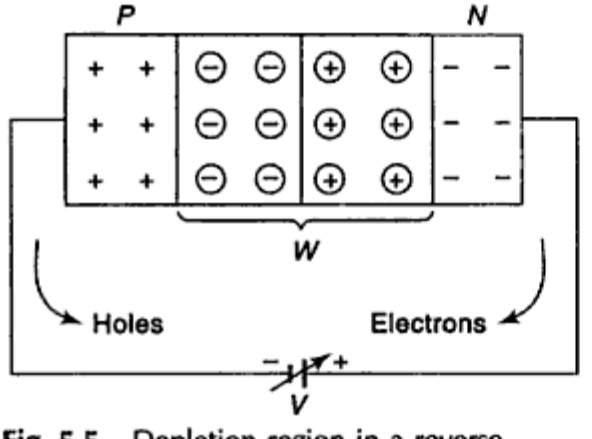


Fig. 5.5 Depletion region in a reverse biased PN junction

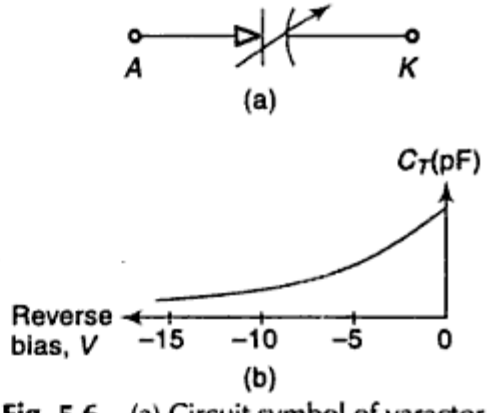


Fig. 5.6 (a) Circuit symbol of varactor diode (b) Characteristics of varactor diode

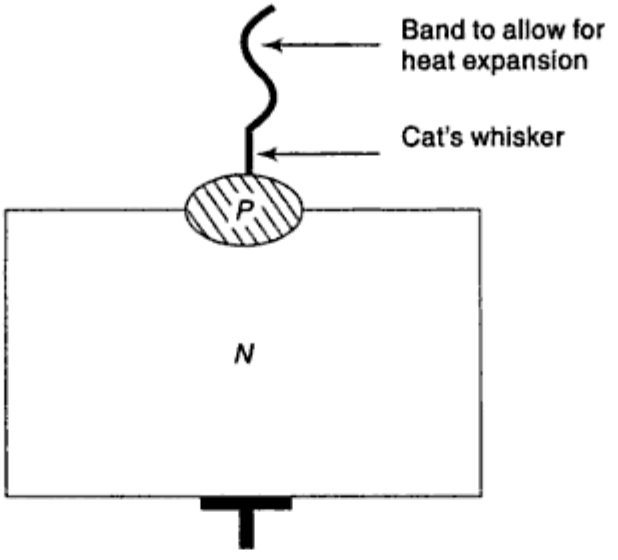


Fig. 5.10 Formation of PN junction in a point-contact diode

area of the PN junction results in very low junction capacitance (0.1 to 1pF) that makes it suitable for operation at frequencies as high as 10 GHz or more and for applications in pulse circuits.

5.7 METAL–SEMICONDUCTOR JUNCTIONS

Metal–semiconductor junctions are very common in all semiconductor devices and have very high importance. Depending upon the doping concentration, materials, and the characteristics of the interface, the metal–semiconductor junctions can act as either an ohmic contact or as a Schottky barrier. An analysis of metal–semiconductor junction is presented in this section.

5.7.1 Structure of Metal–Semiconductor Junction

A metal–semiconductor junction, as the name indicates, consists of a metal in contact with a piece of semiconductor. The structure of a typical metal–semiconductor junction is shown in Fig. 5.11. The active junction is the interface between the metal, which acts as an anode, and the semiconductor. The other interface between the semiconductor and the metal, which acts as a cathode, is an Ohmic contact and there is no potential drop at this junction.

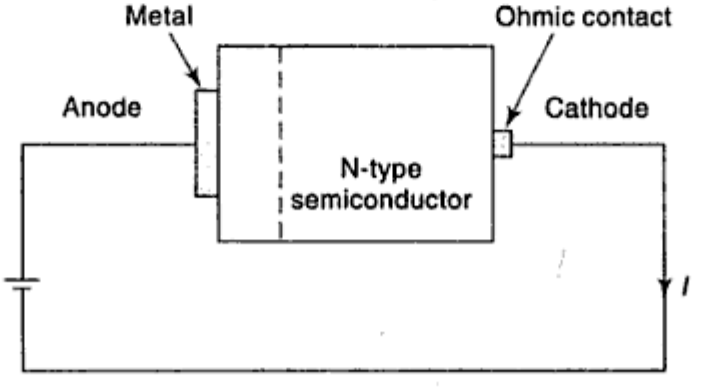


Fig. 5.11 Structure of a metal–semiconductor junction

5.7.2 Energy Band Diagram

The energy band diagram helps in identifying the barrier between the metal and the semiconductor. In order to understand the energy band structure at a metal–semiconductor junction, first let us consider the energy bands in metal and semiconductors separately, as shown in Fig. 5.12(a). The energy bands are aligned at the same vacuum level. When the metal and semiconductor are brought together, the Fermi levels do align themselves at thermal equilibrium. The condition that exists just before the thermal equilibrium is reached is depicted in Fig. 5.12(b).

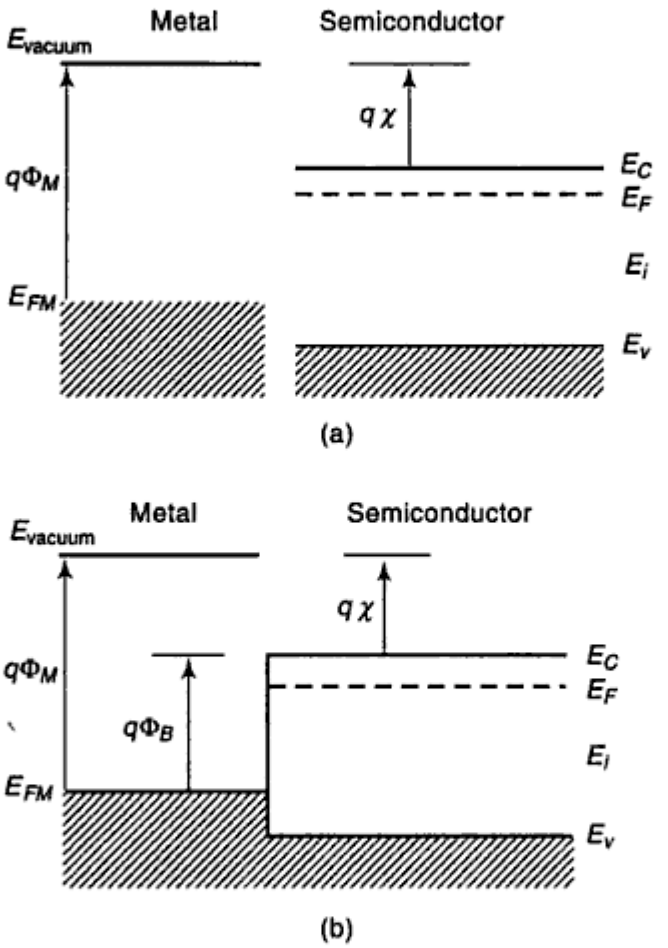


Fig. 5.12 Energy band diagram of metal and semiconductor (a) before and (b) after contact is made

Let us define, Φ_B, the barrier height as the potential difference between the Fermi level of the metal and the band edge where the majority carriers exist. For an N-type semiconductor, the barrier height is given by the difference between the metal work function (Φ_M) and the electron affinity (χ).

$$\Phi_{BN} = \Phi_M - \chi \tag{5.1}$$

The work function, Φ_M varies depending upon surface preparation. For P-type semiconductor, the barrier height is given by the difference between the valence band level and the Fermi level in the metal,

$$\Phi_{BP} = \chi + \frac{E_g}{q} - \Phi_M \tag{5.2}$$

where E_g is the energy gap between the conduction and valence bands. The sum of the barrier heights on N-type and P-type substrate is expected to be equal to the energy gap, E_g i.e., $(\Phi_{BN} + \Phi_{BP})\, q = E_g$.

In a metal-semiconductor junction, a barrier is formed if the Fermi level of the metal is somewhere between the valence and conduction band edges of the semiconductor, as shown in Fig. 5.12(b). Let us also define a built-in potential (Φ_1) as the difference between the Fermi level of the metal and the Fermi level of the semiconductor, given by,

For an N-type semiconductor, the barrier height is given by

$$\Phi_{BN} = \Phi_M - \chi$$

$$\Phi_{IN} = \Phi_{BN} - \frac{E_C - E_F}{q} = \Phi_M - \chi - \frac{E_C - E_F}{q} \tag{5.3}$$

For a P-type semiconductor, the Fermi level is closer to the valence band and the built-in potential is given by

$$\Phi_{IP} = \chi + \frac{E_F - E_V}{q} - \Phi_M \tag{5.4}$$

The Fermi level in an N-type semiconductor is given by

$$E_F = E_C - kT \ln \frac{N_C}{N_D} \tag{5.5}$$

and the Fermi level in a P-type semiconductor is given by

$$E_F = E_V + kT \ln \frac{N_V}{N_A} \tag{5.6}$$

Substitution Eq. (5.5) and Eq. (5.6) in Eq. (5.3) and Eq. (5.4), respectively, would give expressions for built in potentials in terms of the barrier height and doping concentration, as follows.

$$\Phi_{IN} = \Phi_{BN} - \frac{E_C - E_F}{q} = \Phi_{BN} - \frac{kT}{q} \ln \frac{N_C}{N_D} \text{ for N-type semiconductor} \tag{5.7}$$

and

$$\Phi_{IP} = \Phi_{BP} - \frac{E_F - E_V}{q} = \Phi_{BP} - \frac{kT}{q} \ln \frac{N_V}{N_A} \text{ for P-type semiconductor} \tag{5.8}$$

5.7.3 Thermal Equilibrium

After the metal and semiconductor have been brought into contact, electrons start to flow from the semiconductor into the metal, and as a result, a depletion region of width x_d, with uncompensated donors (positive charge) is formed. Electrons continue to flow into the metal until the Fermi energy levels of metal and semiconductor align with each other. In metal, the electron current forms a negative surface charge layer. This results in an electric field and the band edges are lowered in the semiconductor (Fig. 5.13).

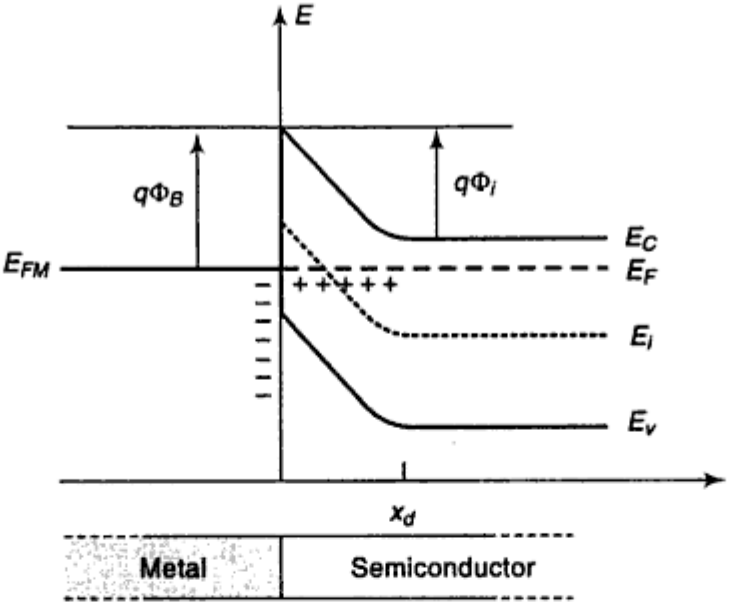


Fig. 5.13 Energy band diagram in thermal equilibrium

Table 5.1 gives the work function (F_M) and electron affinity (χ) of some commonly used metals and semiconductors.

Table 5.1

Element	*Work function, Φ_M (V)*
Aluminium, Al	4.28
Silver, Ag	4.26
Gold, Au	5.1
Chromium, Cr	4.5
Nickel, Ni	5.15
Platinum, Pt	5.65
Titanium, Ti	4.33
Tungsten, W	4.55
	Electron affinity, χ (Volt)
Silicon, Si	4.01
Germanium, Ge	4.13
Gallium Arsenide, GaAs	4.07

Example 5.1 A metal–semiconductor junction is made of silver and silicon with $N_D = 4 \times 10^{17}$ cm^{-3}. Calculate the barrier height and the built-in potential.

Solution:
The work function (Φ_M) and electron affinity (χ) for silver and silicon are 4.26 V and 4.01 V, respectively (from Table 5.1). The barrier height for N-type material, from Eq. (5.1), is given by

$$\Phi_{BN} = \Phi_M - \chi = 4.26 - 4.01 = 0.25 \text{ V}$$

The built-in potential is given by Eq. (5.7) i.e.,

$$\Phi_{IN} = \Phi_{BN} - \frac{kT}{q} \ln \frac{N_C}{N_D} = 0.25 - \frac{1.38 \times 10^{-23} \times 300}{1.6 \times 10^{-19}} \ln \frac{2.8 \times 10^{25}}{4 \times 10^{17}} = 0.47 \text{ V}$$

(The density of state at conduction band edge, $N_C = 2.8 \times 10^{25}$ from Table 4.1)

5.7.4 Forward and Reverse Bias

When an external bias is applied, the metal-to-semiconductor barrier remains unchanged, whereas, the semiconductor-to-metal barrier is either decreased (forward bias) or increased (reverse bias).

When the metal is connected to a positive bias with respect to the semiconductor. The Fermi energy level of the metal is lowered from its equilibrium level. The depletion region is narrowed, and the potential barrier in the semiconductor is reduced. The number of electrons that diffuse from semiconductor to metal is now more than the number of electrons that drift from metal into the semiconductor. Thus, there will be a positive current through the device. Figure 5.14 illustrates a metal–semiconductor junction under forward-bias condition.

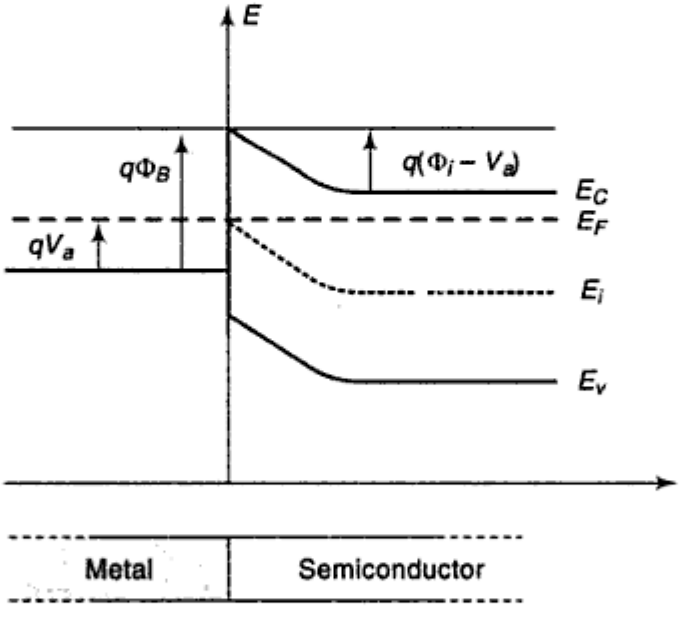


Fig. 5.14 Energy band diagram under forward bias

If the metal is connected to a negative bias with respect to the semiconductor, the metal is charged even more negatively than without any bias. The Fermi energy level of the metal is raised. The electrons in the semiconductor are repelled even more. The depletion region becomes wider and the potential barrier on semiconductor side is further increased, as shown in Fig. 5.15. However, the barrier on the metal side remains unchanged and limits the flow of electrons. A small current flows as a result of a few electrons in the metal acquiring enough thermal energy to overcome barrier.

That is,

$$\Phi_M \leq \chi \text{ for an N-type semiconductor}$$

or,

$$\Phi_M \geq \chi + E_g \text{ for a P-type semiconductor}$$

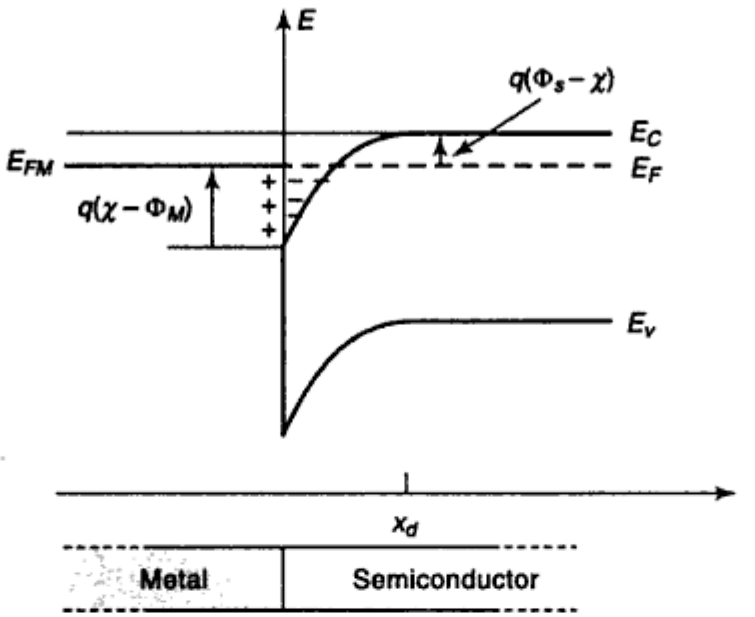


Fig. 5.18 Energy band diagram in an Ohmic contact

Figure 5.18 illustrates the energy band diagram when Q_M is less than the electron affinity, χ. In this case, the electrons flow from the metal into the semiconductor thus forming a positive surface charge layer in metal. The resultant electric field sets up an electric potential and the energy bands of the semiconductor bend downward. There is no barrier for the flow of electrons in both directions. The current is directly proportional to the potential across the junction and is symmetric about the origin, as shown in Fig. 5.19.

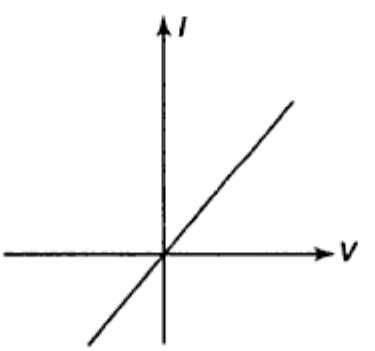


Fig. 5.19 V-I Characteristics of an Ohmic contact

A more practical method of providing contacts in semiconductor devices is to create a junction that consists of a thin barrier. Such contacts are also referred to as *tunnel contacts*. Such contacts have a positive barrier at the junction and a heavy doping in the semiconductor. This creates a very thin barrier separating the metal

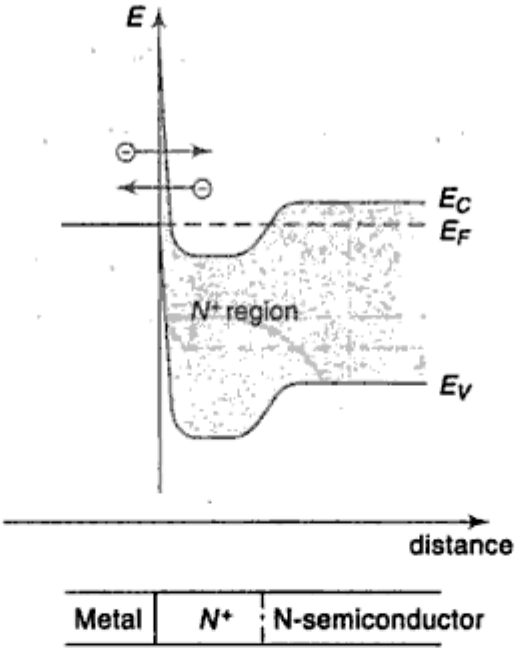


Fig. 5.20 Energy band levels in an Ohmic contact

from the semiconductor, through which the carriers can easily tunnel. Figure 5.20 shows the energy band in a heavily doped Ohmic contact.

5.8 TUNNEL DIODE

The Tunnel or Esaki diode is a thin-junction diode which exhibits negative resistance under low forward bias conditions.

An ordinary PN junction diode has an impurity concentration of about 1 part in 10^8. With this amount of doping the width of the depletion layer is of the order of 5 microns. This potential barrier restrains the flow of carriers from the majority carrier side to the minority carrier side. If the concentration of impurity atoms is greatly increased to the level of 1 part in 10^3. If the device characteristics are completely changed. The width of the junction barrier varies inversely as the square root of the impurity concentration and therefore, is reduced from 5 microns to less than 100 Å (10^{-8} m). This thickness is only about 1/50th of the wavelength of visible light. For such thin potential energy barriers, the electrons will penetrate through the junction rather than surmounting them. This quantum mechanical behavior is referred to as tunneling and hence, these high-impurity-density PN junction devices are called tunnel diodes.

The *V–I* characteristic for a typical germanium tunnel diode is shown in Fig. 5.21. It is seen that at first forward current rises sharply as applied voltage is increased, where it would have risen slowly for an ordinary PN junction diode (which is shown as dashed line for comparison). Also, reverse current is much larger for comparable back bias than in other diodes due to the thinness of the junction. The interesting portion of the characteristic starts at the point *A* on the curve, i.e. the peak voltage. As the forward bias is increased beyond this point, the forward current drops and continues to drop until point *B* is reached. This is the valley voltage. At *B*, the current starts to increase once again and does so very rapidly as bias is increased further. Beyond this point, characteristic resembles that

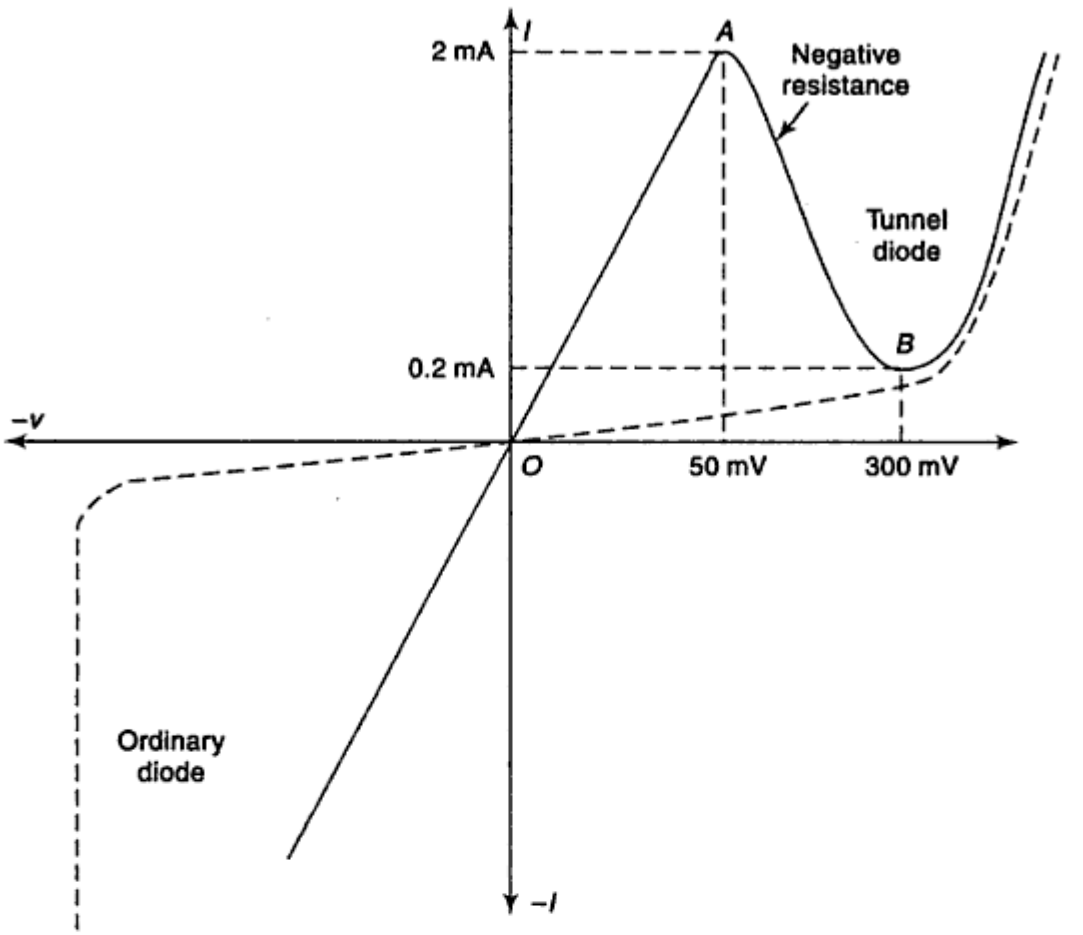


Fig. 5.21 *V–I* characteristic of tunnel diode

of an ordinary diode. Apart from the peak voltage and valley voltage, the other two parameters normally used to specify the diode behaviour are the peak current and the peak-to-valley current ratio, which are 2 mA and 10 respectively, as shown.

The *V–I* characteristic of the tunnel diode illustrates that it exhibits dynamic resistance between *A* and *B*. Figure 5.22 shows energy level diagrams of the tunnel diode for three interesting bias levels. The shaded areas show the energy states occupied by electrons in the valence band, whereas the cross hatched regions represent energy states in the conduction band occupied by the electrons. The levels to which the energy states are occupied by electrons on either side of the junctions are shown by dotted lines. When the bias is zero, these lines are at the same height. Unless energy is imparted to the electrons from some external source, the energy possessed by the electrons on the N–side of the junction is insufficient to permit the to climb over the junction barrier to reach the P-side. However, quantum mechanics show that there is a finite probability for the electrons to tunnel through the junction to reach the other side, provided there are allowed empty energy states in the P-side of the junction at the same energy level. Hence, the forward current is zero.

When a small forward bias is applied to the junction, the energy level of the P-side is lower as compared with the N-side. As shown in Fig. 5.22(b), electrons in the conduction band of the N-side see empty energy level on the P-side. Hence,

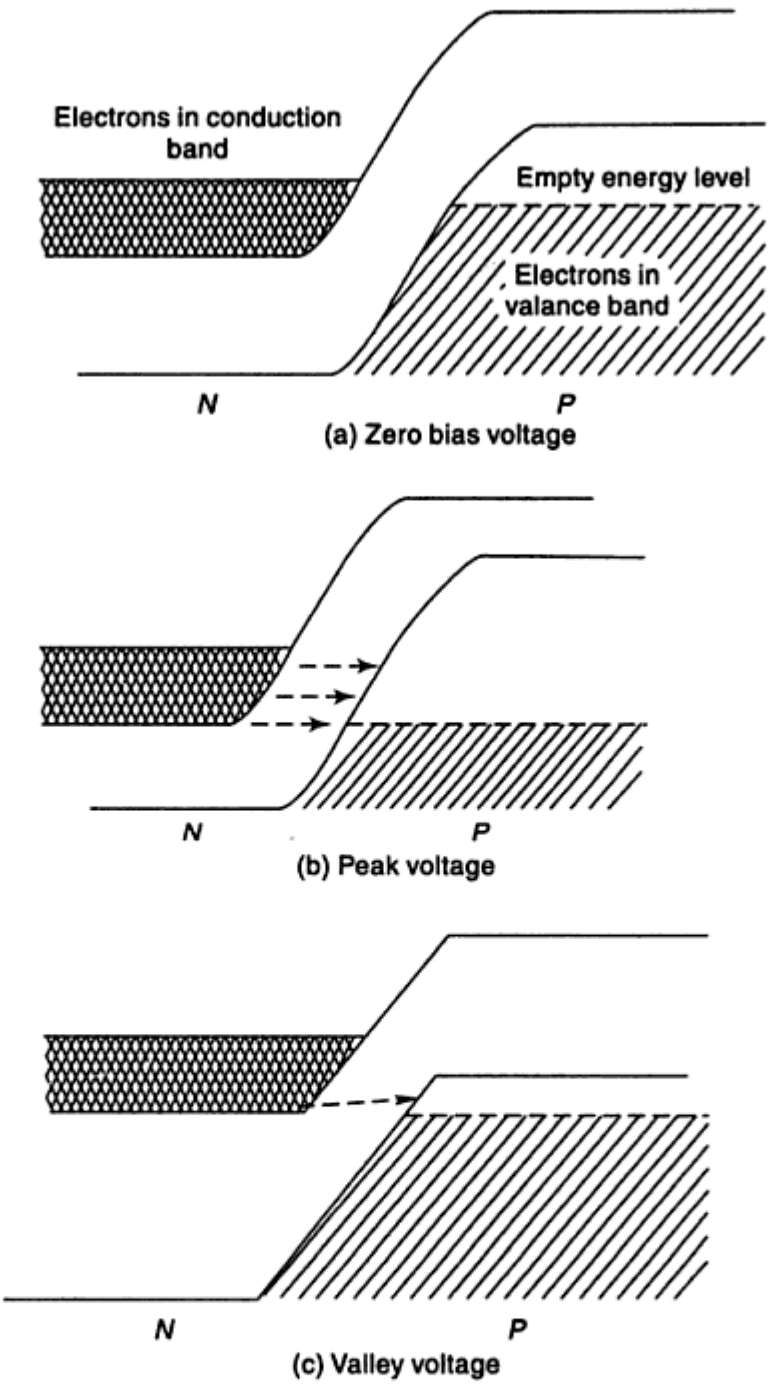


Fig. 5.22 Energy level diagrams of tunnel diode

tunneling from N-side to P-side takes place. Tunneling in other directions is not possible because the valence band electrons on the P-side are now opposite to the forbidden energy gap on the N-side. The energy band diagram shown in Fig. 5.22(b), is for the peak of the diode characteristic.

When the forward bias is raised beyond this point, tunneling will decrease as shown in Fig. 5.22(c). The energy of the P-side is now depressed further, with the result that fewer conduction band electrons on the N-side are opposite to the unoccupied P-side energy levels. As the bias is raised, forward current drops. This corresponds to the negative resistance region of the diode characteristic. As forward

bias is raised still further, tunneling stops altogether and it behaves as a normal PN junction diode.

5.8.1 Equivalent Circuit

The equivalent circuit of the tunnel diode when biased in the negative resistance region is as shown in Fig. 5.23(a). In the circuit, R_s is the series resistance and L_s is the series inductance which may be ignored except at highest frequencies. The resulting diode equivalent circuit is thus reduced to parallel combination of the junction capacitance C_j and the negative resistance $-R_n$. Typical values of the circuit components are $R_s = 6\ \Omega$, $L_s = 0.1$ nH, $C_j = 0.6$ pF and $R_n = 75\Omega$.

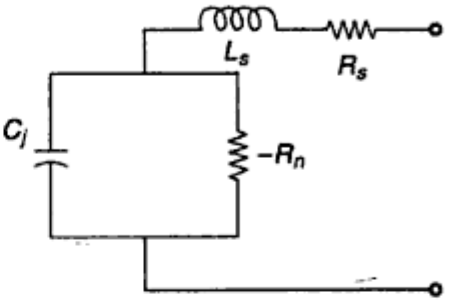


Fig. 5.23(a) Equivalent circuit of tunnel diode

Applications

1. Tunnel diode is used as an ultra-high speed switch with switching speed of the order of ns or ps
2. As logic memory storage device
3. As microwave oscillator
4. In relaxation oscillator circuit
5. As an amplifier.

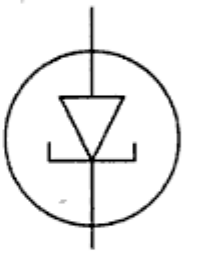


Fig. 5.23(b) Symbol of tunnel diode

Advantages

1. Low noise
2. Ease of operation
3. High speed
4. Low power

Disadvantages

1. Voltage range over which it can be operated is 1 V less
2. Being a two terminal device, there is no isolation between the input and output circuit.

5.9 GUNN DIODE

In 1963, Gunn discovered the transferred electron effect which is now commonly referred to a Gunn effect. This effect is effectively utilised in the Gunn diode for generation of microwave oscillations. It was the first instance of useful semiconductor device operation depending on the bulk properties of a material. Gunn effect is exhibited by semiconductor materials like gallium arsenide, indium phosphide, cadmium telluride and indium arsenide.

The structure of a practical Gunn diode is shown in Fig. 5.24.

5.9.1 Gunn Effect

If a relatively small d.c. voltage is applied across a thin slice of gallium arsenide, with thickness of the order of tens of micrometers, so that the voltage gradient across the slice is in excess of about 3300 V/cm, then negative resistance will

5.10 IMPATT DIODE

IMPATT diode stands for IMPact Avalanche and Transit Time diode. Any device that exhibits dynamic negative resistance for direct current will also exhibit it for alternate current. When alternating voltage is applied, current will rise when voltage falls. Hence, negative resistance may be redefined as that property of the device which causes the current through it to be 180° out of phase with the voltage across it.

IMPATT diode utilises a combination of delay involved in generating Avalanche current multiplication, together with the delay due to transit time through a drift space to provide the necessary 180° phase difference between applied voltage and the resulting current. The cross section of the active region of this device is shown in Fig. 5.27. From the figure, it should be noted that it is a diode, the junction being between the P^+ and N layers.

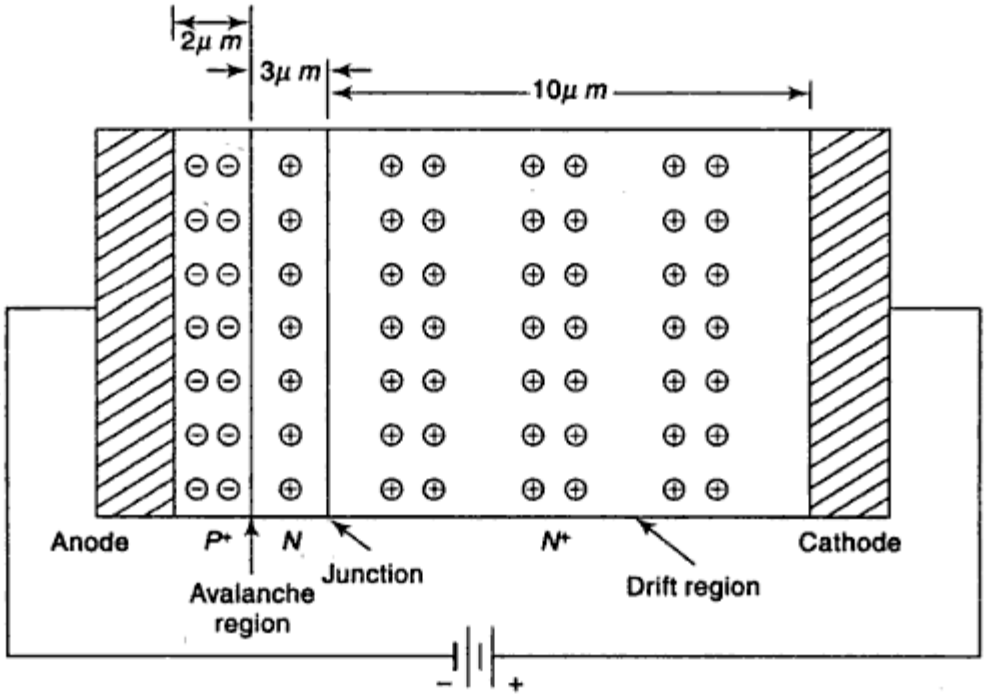


Fig. 5.27 Active region of IMPATT diode

An extremely high voltage gradient of the order of 400 kV/cm is applied to the IMPATT diode, eventually resulting in a very high current. The IMPATT diode is constructed to withstand such conditions repeatedly. Under such high potential gradient, applied in the reverse direction, results in flow of minority carriers across the junction. If it is assumed that oscillations exist, the effect of a positive swing of RF voltage superimposed on top of the high d.c. voltage should be considered. Electron and hole velocity now become so high, so that they form additional holes and electrons by knocking them out of the crystal structure by a process knows as impact ionization.

These additional carriers continue the process at the junction and turns out into an Avalanche. If the original d.c. field was just at the threshold of producing this Avalanche, then during the whole of the positive RF cycle Avalanche multiplication

will take place. As Avalanche is a multiplication process, the process takes a time, such that the resulting current pulse maximum at the junction, occurs at the instant when the RF voltage across the diode is zero and going negative, as shown in Fig. 5.28. Thus, a 90° phase difference between the voltage and current has been established.

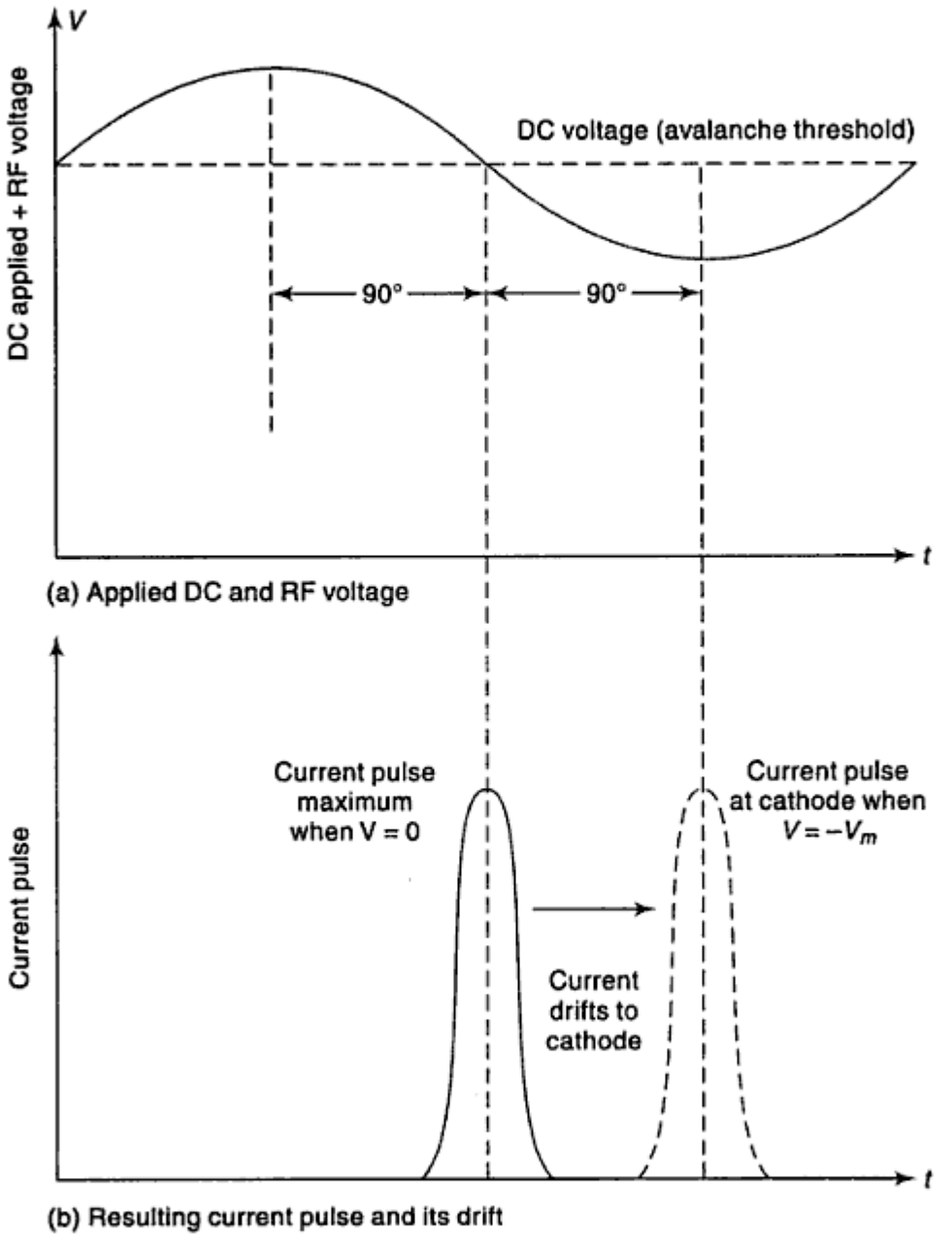


Fig. 5.28 Voltage–current waveform in an IMPATT diode

Because of the reverse bias that is applied, the current pulse flows to the cathode at a drift velocity dependent on the magnitude of the high d.c. field. The time taken by the pulse to reach the cathode depends on this velocity and on the thickness of the highly doped (N^+) layer. The thickness of the drift region is selected so that the time taken for the current pulse to arrive at the cathode corresponds to a further 90° phase difference. Thus, as shown in Fig. 5.28, when current pulse actually arrives at

the cathode terminal the RF voltage there is at its negative peak. As a result, voltage and current in the IMPATT diode is 180° out of phase, and a dynamic RF negative resistance has been proved to exist. Such a negative resistance finds extensive use in oscillators of amplifiers. Because of the short transit time involved they can be operated in the microwave frequency region.

IMPATT diodes are the most powerful CW solid state microwave power sources. A typical, commercial IMPATT diode for use below about 50 GHz is shown in Fig. 5.29.

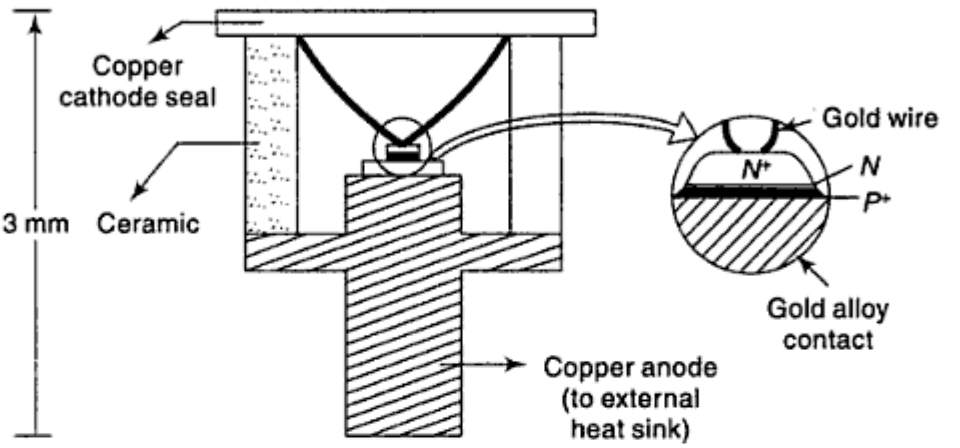


Fig. 5.29 Structure of IMPATT diode

5.11 PIN DIODE

It is composed of three regions. In addition to the usual N and P regions, an intrinsic layer (I region) is sandwiched between them, to form the PIN structure as shown in Fig. 5.30. Being intrinsic, the intermediate layer offers relatively high resistance which gives it two advantages compared to an ordinary PN diode. They are (1) decrease in capacitance between P and N regions as it is inversely proportional to the separation between these regions. It allows a faster response time for the diode. Hence, PIN diodes are used at high frequencies (more than 300 MHz). (2) possibility of greater electric field between the P and N junctions, so that the charge carriers drift towards their majority carrier side. This enhances faster response of the diode.

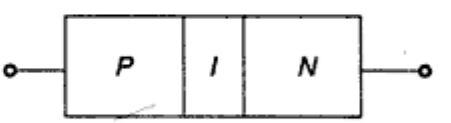


Fig. 5.30 Structure of PIN diode

It offers a variable resistance under forward bias condition as shown in Fig. 5.31. Forward resistance offered is given by $r_{ac} \propto 50/I$, where I is the d.c. current in mA.

Hence, for large d.c. currents, the diode will look like a short. In reverse biased condition it looks like an open, i.e. it offers an infinite resistance.

It is used as a switching diode for signal frequencies up to GHz range and as an AM modulator of very high frequency signals.

5.12 PIN PHOTODIODE

It is used for the detection of light at the receiving end in Optical communication. It is a three-region reverse biased junction diode. A layer of intrinsic silicon is sand-

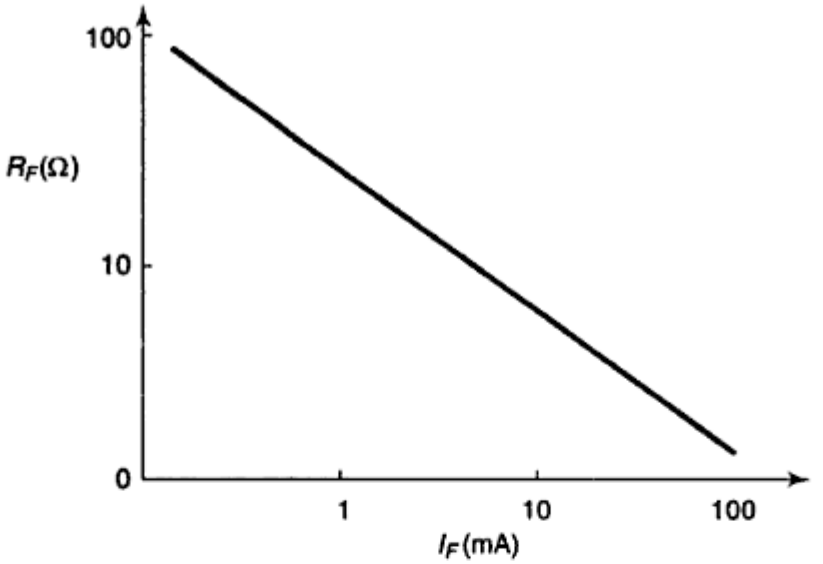


Fig. 5.31 Variation of forward resistance in a PIN diode

witched between heavily doped P and N type semiconductor materials. As shown in Fig. 5.32, the depletion region extends almost to the entire intrinsic layer where most of the absorption of light photons take place. The width of the intrinsic layer is large compared to the width of the other two layers. This ensures large absorption of light photons in the depletion region which also forms the absorption region. Light photons incident on the PIN photodiode are absorbed in the absorption region which leads to the generation of electron–hole pairs. These charge carriers present in the depletion region drift under the influence of the existing electric field that is set up due to the applied reverse bias. The reverse current flowing in the external circuit increases linearly with the level of illumination.

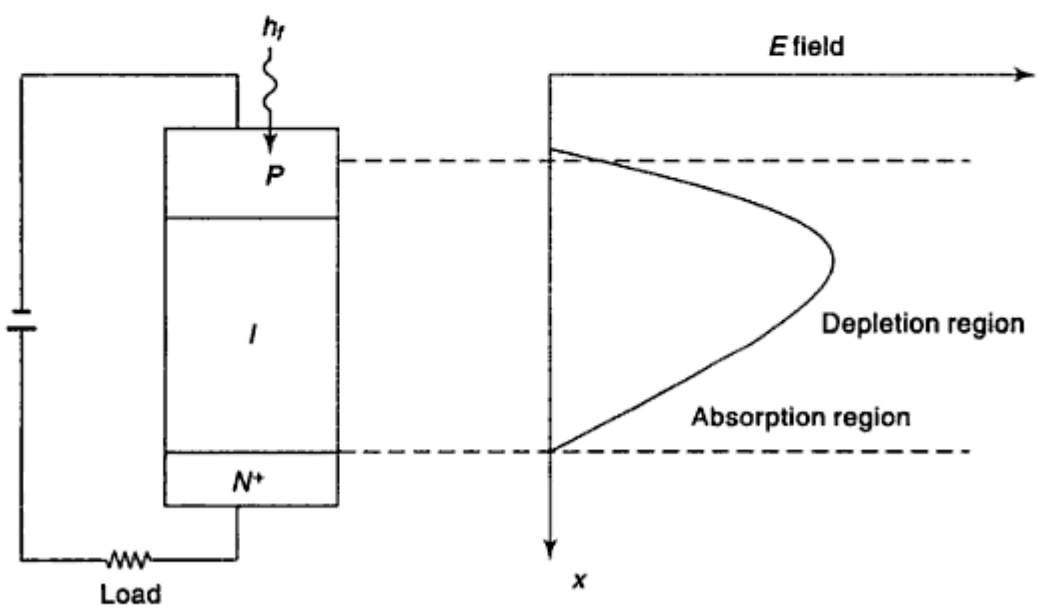


Fig. 5.32 Structure of PIN photodiode

As the process of drifting is quicker than diffusion, the transit time of the charge carriers is small so that the response time is considerably reduced. The large width of the depletion region results in achieving high quantum efficiency.

5.13 AVALANCHE PHOTO DIODE (APD)

APD is used in Optical communication for detection of light at the receiving end. It converts the input light signal into electrical signal. The structure of an APD is shown in Fig. 5.33. It essentially consists of reverse biased PN junction. The depletion region in this reverse biased PN junction is formed by immobile positively charged donor atoms in the N-type semiconductor material and immobile negatively charged acceptor atoms in the P-type material. The electric field in this depletion region is very high where most of the photons are absorbed and primary charge carriers (electron–hole pair) generated. These charge carriers acquire sufficient energy from the electric field to excite new electron-hole pairs by a process known as impact ionisation. These new carriers created by impact ionisation can themselves produce additional carriers by the same mechanism. For this process, APD requires a high reverse bias voltage in the order of 100–400 V. Carrier multiplication factors as great as 10^4 may be obtained using defect free materials. Electron-hole pairs thus generated separate and drift under the influence of the electric field in the depletion region and diffuse outside the depletion region so that they are finally collected in the detector terminals. This leads to a flow of current in the external circuit whose magnitude is proportional to the intensity of light incident on APD.

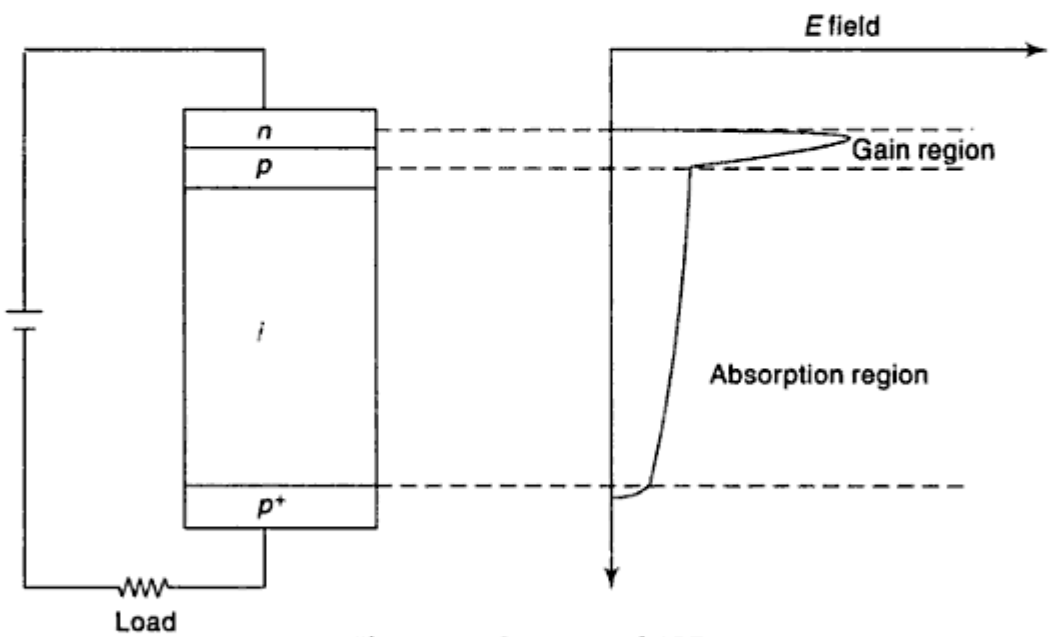


Fig. 5.33 Structure of APD

Due to the internal gain mechanism in an APD, a large electrical response is obtained even for a weak input light signal. Quantum efficiency closer to 100% in the working region can be obtained.

5.14 LASER DIODE

Similar to LED, Lasers are used to convert the electrical signal to light signal. In direct band gap materials where high recombination velocities exist, optical gain can be achieved by creating population inversion of carriers through high-level current injection and by forming a resonant cavity. This cavity is usually produced

6 Bipolar Junction Transistor

6.1 INTRODUCTION

A Bipolar Junction Transistor (BJT) is a three terminal semiconductor device in which the operation depends on the interaction of both majority and minority carriers and hence the name Bipolar. The BJT is analogous to a vacuum triode and is comparatively smaller in size. It is used in amplifier and oscillator circuits, and as a switch in digital circuits. It has wide applications in computers, satellites and other modern communication systems.

6.2 CONSTRUCTION

The BJT consists of a silicon (or germanium) crystal in which a thin layer of N-type Silicon is sandwiched between two layers of P-type silicon. This transistor is referred to as PNP. Alternatively, in a NPN transistor, a layer of P-type material is sandwiched between two layers of N-type material. The two types of the BJT are represented in Fig. 6.1.

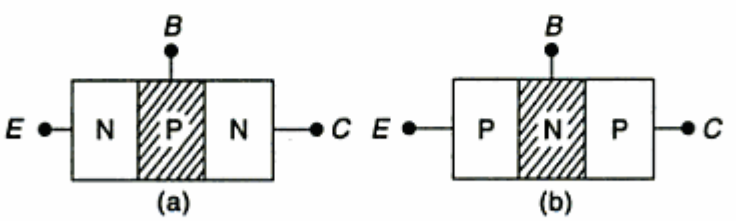


Fig. 6.1 Transistor (a) NPN and (b) PNP

The symbolic representation of the two types of the BJT is shown in Fig. 6.2. The three portions of the transistor are Emitter, Base and Collector, shown as *E*, *B* and *C*, respectively. The arrow on the emitter specifies the direction of current flow when the *EB* junction is forward biased.

Emitter is heavily doped so that it can inject a large number of charge carriers into the base. Base is lightly doped and very thin. It passes most of the injected charge carriers from the emitter into the collector. Collector is moderately doped.

6.3 TRANSISTOR BIASING

As shown in Fig. 6.3, usually the emitter-base junction is forward biased and collector-base junction is reverse biased. Due to the forward bias on the emitter-base

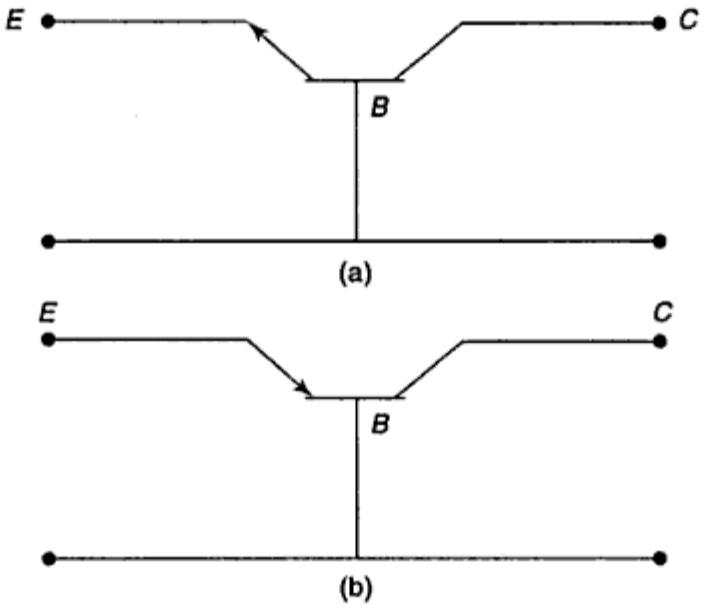


Fig. 6.2 Circuit symbol. (a) NPN transistor and (b) PNP transistor

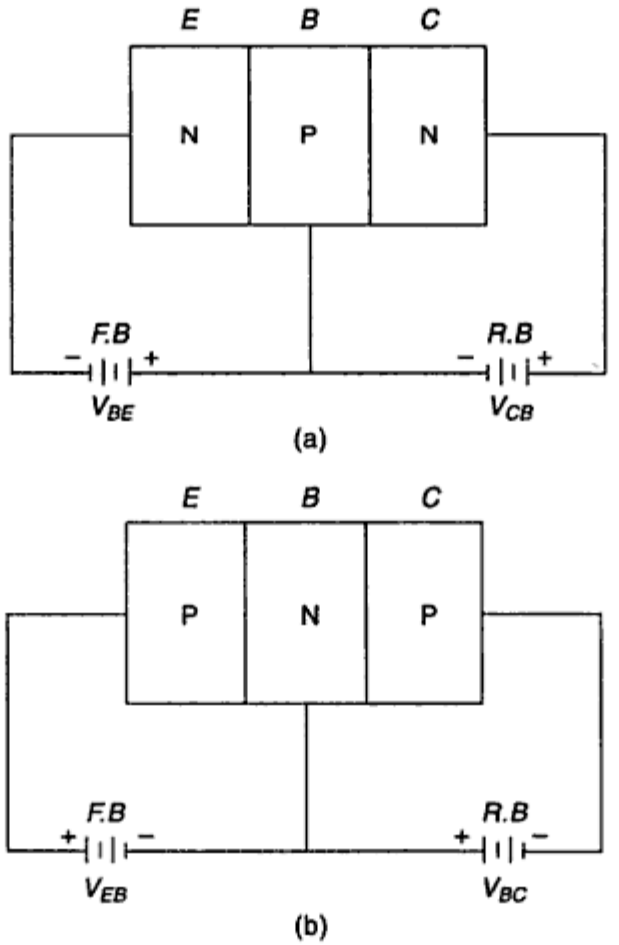


Fig. 6.3 Transistor biasing (a) NPN transistor and (b) PNP transistor

junction an emitter current flows through the base into the collector. Though the, collector-base junction is reverse biased, almost the entire emitter current flows through the collector circuit.

6.4 OPERATION OF NPN TRANSISTOR

As shown in Fig. 6.4, the forward bias applied to the emitter base junction of an NPN transistor causes a lot of electrons from the emitter region to crossover to the base region. As the base is lightly doped with P-type impurity, the number of holes in the base region is very small and hence the number of electrons that combine with holes in the P-type base region is also very small. Hence a few electrons combine with holes to constitute a base current I_B. The remaining electrons (more than 95%) crossover into the collector region to constitute a collector current I_C. Thus the base and collector current summed up gives the emitter current, i.e. $I_E = -(I_C + I_B)$.

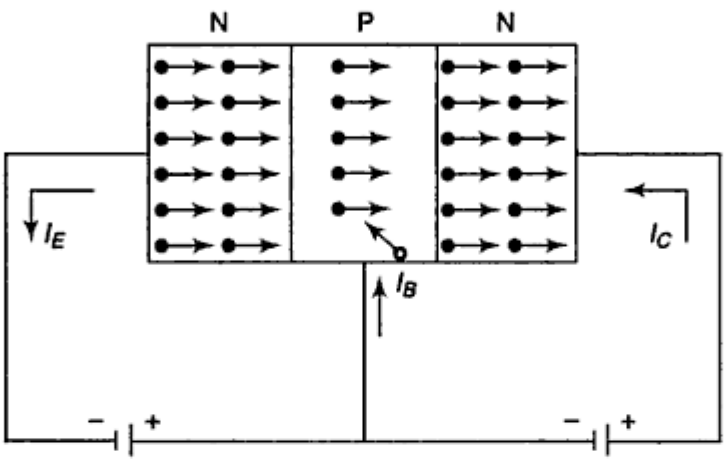


Fig. 6.4 Current in NPN transistor

In the external circuit of the NPN bipolar junction transistor, the magnitudes of the emitter current I_E, the base current I_B and the collector current I_C are related by $I_E = I_C + I_B$.

6.5 OPERATION OF PNP TRANSISTOR

As shown in Fig. 6.5, the forward bias applied to the emitter-base junction of a PNP transistor causes a lot of holes from the emitter region to crossover to the base region as the base is lightly doped with N-types impurity. The number of electrons in the base region is very small and hence the number of holes combined with

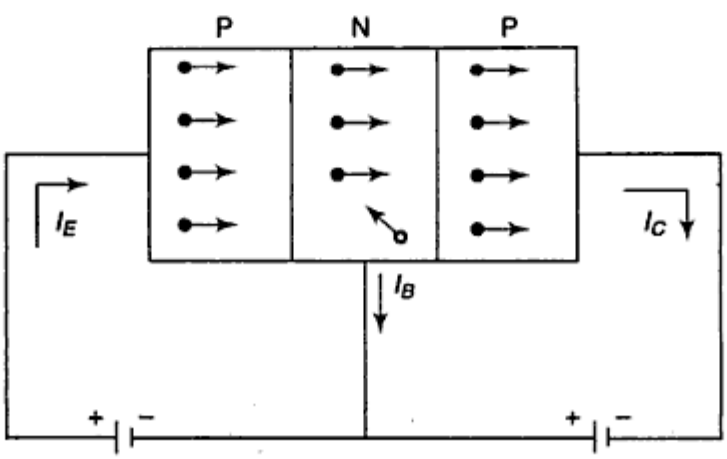


Fig. 6.5 Current in PNP transistor

electrons in the N-type base region is also very small. Hence a few holes combined with electrons to constitute a base current I_B. The remaining holes (more than 95%) crossover into the collector region to constitute a collector current I_C. Thus the collector and base current when summed up gives the emitter current, i.e. $I_E = -(I_C + I_B)$.

In the external circuit of the PNP bipolar junction transistor, the magnitudes of the emitter current I_E, the base current I_B and the collector current I_C are related by

$$I_E = I_C + I_B \tag{6.1}$$

This equation gives the fundamental relationship between the currents in a bipolar transistor circuit. Also, this fundamental equation shows that there are current amplification factors α and β in common base transistor configuration and common emitter transistor configuration respectively for the static (d.c.) currents, and for small changes in the currents.

Large-signal current gain(α) The large signal current gain of a common base transistor is defined as the ratio of the negative of the collector-current increment to the emitter-current change from cutoff ($I_E = 0$) to I_E, i.e.

$$\alpha = -\frac{(I_c - I_{CBO})}{I_E - 0}$$

where I_{CBO} (or I_{CO}) is the reverse saturation current flowing through the reverse biased collector-base junction, i.e. the collector to base leakage current with emitter open. As the magnitude of I_{CBO} is negligible when compared to I_E, the above expression can be written as

$$\alpha = \frac{I_C}{I_E} \tag{6.3}$$

Since I_C and I_E are flowing in opposite directions, α is always positive. Typical value of α ranges from 0.90 to 0.995. Also, α is not a constant but varies with emitter current I_E, collector voltage V_{CB} and temperature.

General Transistor Equation

In the active region of the transistor, the emitter is forward biased and the collector is reverse biased. The generalised expression for collector current I_C for collector junction voltage V_C and emitter current I_E is given by

$$I_C = -\alpha I_E + I_{CBO}(1 - e^{V_c/V_T}) \tag{6.4}$$

If V_C is negative and $|V_c|$ is very large compared with V_T, then the above equation reduces to

$$I_C = -\alpha I_E + I_{CBO} \tag{6.5}$$

If V_C, i.e. V_{CB}, is few volts, I_C is independent of V_C. Hence the collector current I_C is determined only by the fraction α of the current I_E flowing in the emitter.

Relation among I_C, I_B and I_{CBO} From Eqn. (6.5), We have

$$I_C = -\alpha I_E + I_{CBO}$$

Since I_C and I_E are flowing in opposite directions,

$$I_E = -(I_C + I_B)$$

Therefore,
$$I_C = -\alpha[-(I_C + I_B)] + I_{CBO}$$
$$I_C - \alpha I_C = \alpha I_B + I_{CBO}$$
$$I_C(1-\alpha) = \alpha I_B + I_{CBO}$$
$$I_C = \frac{\alpha}{1-\alpha} I_B + \frac{I_{CBO}}{1-\alpha}$$

Since
$$\beta = \frac{\alpha}{1-\alpha}, \tag{6.6}$$

the above expression becomes

$$I_C = (1 + \beta) I_{CBO} + \beta I_B \tag{6.7}$$

Relation among I_C, I_B and I_{CEO} In the common-emitter (CE) transistor circuit, I_B is the input current and I_C is the output current. If the base circuit is open, i.e, $I_B = 0$, then a small collector current flows from the collector to emitter. This is denoted as I_{CEO}, the collector-emitter current with base open. This current I_{CEO} is also called the collector to emitter leakage current.

In this CE configuration of the transistor, the emitter-base junction is forward-biased and collector-base junction is reverse-biased and hence the collector current I_C is the sum of the part of the emitter current I_E that reaches the collector, and the collector-emitter leakage current I_{CEO}. Therefore, the part of I_E, which reaches collector is equal to $(I_C - I_{CEO})$.

Hence, the *large-signal current gain* (β) is defined as,

$$\beta = \frac{(I_C - I_{CEO})}{I_B} \tag{6.8}$$

From the equation, we have

$$I_C = \beta I_B + I_{CEO} \tag{6.9}$$

Relation between I_{CBO} and I_{CEO} Comparing Eqs. (6.7) and (6.9), we get the relationship between the leakage currents of transistor common-base (CB) and common-emitter (CE) configurations as

$$I_{CEO} = (1 + \beta) I_{CBO} \tag{6.10}$$

From this equation, it is evident that the collector-emitter leakage current (I_{CEO}) in CE configuration is $(1 + \beta)$ times larger than that in CB configuration. As I_{CBO} is temperature-dependent, I_{CEO} varies by large amount when temperature of the junctions changes.

Expression for Emitter Current

The magnitude of emitter-current is

$$I_E = I_C + I_B$$

Substituting Eqn. (6.7) in the above equation, we get

$$I_E = (1 + \beta) I_{CBO} + (1 + \beta) I_B \tag{6.11}$$

Substituting Eqn. (6.6) into Eqn. (6.11), we have

$$I_E = \frac{1}{1-\alpha} I_{CBO} + \frac{1}{1-\alpha} I_B \tag{6.12}$$

DC current gain ($\beta_{d.c.}$ or h_{FE}) The d.c. current gain is defined as the ratio of the collector current I_C to the base current I_B. That is,

each value of I_E. This is repeated for different fixed values of I_E. Now the curves of I_C versus V_{CB} are plotted for constant values of I_E and the output characteristics thus obtained is shown in Fig. 6.9.

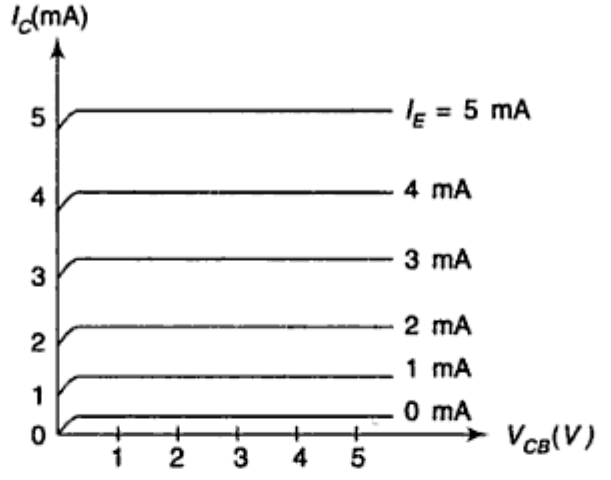


Fig. 6.9 CB output characteristics

From the characteristics, it is seen that for a constant value of I_E, I_C is independent of V_{CB} and the curves are parallel to the axis of V_{CB}. Further, I_C flows even when V_{CB} is equal to zero. As the emitter-base junction is forward biased, the majority carriers, i.e. electrons, from the emitter are injected into the base region. Due to the action of the internal potential barrier at the reverse biased collector-base junction, they flow to the collector region and give rise to I_C even when V_{CB} is equal to zero.

Early effect or base-width modulation As the collector voltage V_{CC} is made to increase the reverse bias, the space charge width between collector and base tends to increase, with the result that the effective width of the base decreases. This dependency of base-width on collector-to-emitter voltage is known as the *Early effect*. This decrease in effective base-width has three consequences:

(i) There is less chance for recombination within the base region. Hence, α increases with increasing $|V_{CB}|$.

(ii) The charge gradient is increased within the base, and consequently, the current of minority carriers injected across the emitter junction increases.

(iii) For extremely large voltages, the effective base-width may be reduced to zero, causing voltage breakdown in the transistor. This phenomenon is called the *punch through*.

For higher values of V_{CB}, due to Early effect, the value of α increases. For example, α changes, say from 0.98 to 0.985. Hence, there is a very small positive slope in the CB output characteristics and hence the output resistance is not zero.

Transistor parameters The slope of the CB characteristics will give the following four transistor parameters. Since these parameters have different dimensions, they are commonly known as common base *hybrid parameters* or *h-parameters*.

(i) Input impedance (h_{ib}): It is defined as the ratio of the change in (input) emitter voltage to the change in (input) emitter current with the (output) collector voltage V_{CB} kept constant. Therefore,

$$h_{ib} = \frac{\Delta V_{EB}}{\Delta I_E},\ V_{CB} \text{ constant} \tag{6.14}$$

It is the slope of CB input characteristics I_E versus V_{EB} as shown in Fig. 6.8. The typical value of h_{ib} ranges from 20Ω to 50Ω.

(ii) Output admittance (h_{ob}): It is defined as the ratio of change in the (output) collector current to the corresponding change in the (output) collector voltage with the (input) emitter current I_E kept constant. Therefore,

$$h_{ob} = \frac{\Delta I_C}{\Delta V_{CB}},\ I_E \text{ constant} \tag{6.15}$$

It is the slope of CB output characteristics I_C versus V_{CB} as shown in Fig. 6.9. The typical value of this parameter is of the order of 0.1 to 10 μ mhos.

(iii) Forward current gain (h_{fb}): It is defined as a ratio of the change in the (output) collector current to the corresponding change in the (input) emitter current keeping the (output) collector voltage V_{CB} constant. Hence,

$$h_{fb} = \frac{\Delta I_C}{\Delta I_E},\ V_{CB} \text{ constant.} \tag{6.16}$$

It is the slope of I_C versus I_E curve. Its typical value varies from 0.9 to 1.0.

(iv) Reverse voltage gain (h_{rb}): It is defined as the ratio of the change in the (input) emitter voltage and the corresponding change in (output) collector voltage with constant (input) emitter current, I_E. Hence,

$$h_{rb} = \frac{\Delta V_{EB}}{\Delta V_{CB}},\ I_E \text{ constants} \tag{6.17}$$

It is the slope of V_{EB} versus V_{CB} curve. Its typical value is of the order of 10^{-5} to 10^{-4}.

6.6.2 CE Configuration

Input characteristics To determine the input characteristics, the collector to emitter voltage is kept constant at zero volt and base current is increased from zero in equal steps by increasing V_{BE} in the circuit shown in Fig. 6.10.

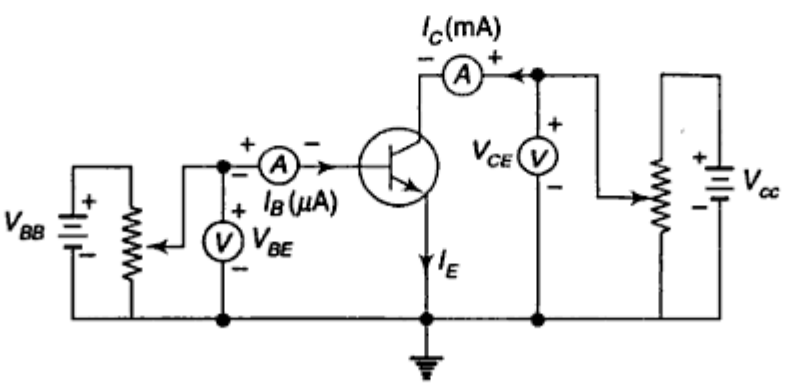


Fig. 6.10 Circuit to determine CE static characteristics

The value of V_{BE} is noted for each setting of I_B. This procedure is repeated for higher fixed values of V_{CE}, and the curves of I_B V_s. V_{BE} are drawn. The input characteristics thus obtained are shown in Fig. 6.11.

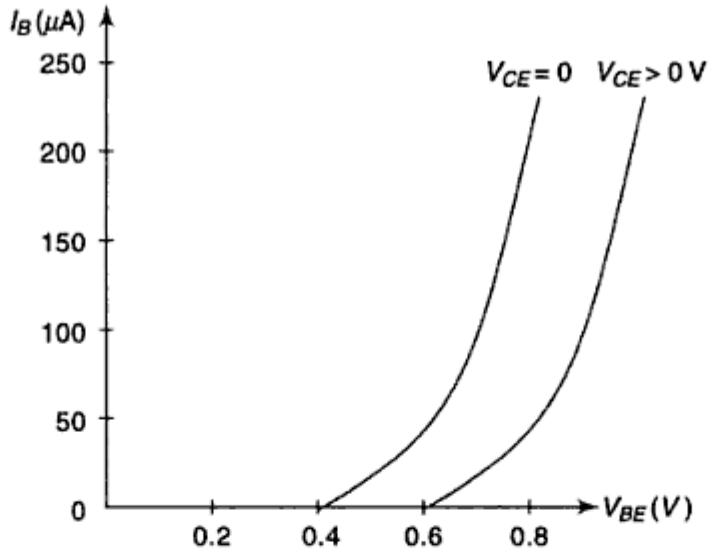


Fig. 6.11 CE input characteristics

When $V_{CE} = 0$, the emitter-base junction is forward biased and the junction behaves as a forward biased diode. Hence the input characteristic for $V_{CE} = 0$ is similar to that of a forward-biased diode. When V_{CE} is increased, the width of the depletion region at the reverse biased collector-base junction will increase. Hence the effective width of the base will decrease. This effect causes a decrease in the base current I_B. Hence, to get the same value of I_b as that for $V_{CE} = 0$, V_{BE} should be increased. Therefore, the curve shifts to the right as V_{CE} increases.

Output characteristics To determine the output characteristics, the base current I_B is kept constant at a suitable value by adjusting base-emitter voltage, V_{BE}. The magnitude of collector-emitter voltage V_{CE} is increased in suitable equal steps from zero and the collector current I_C is noted for each setting V_{CE}. Now the curves of I_C versus V_{CE} are plotted for different constant values of I_B. The output characteristics thus obtained are shown in Fig. 6.12. From Eqs. (6.6) and (6.7), we have

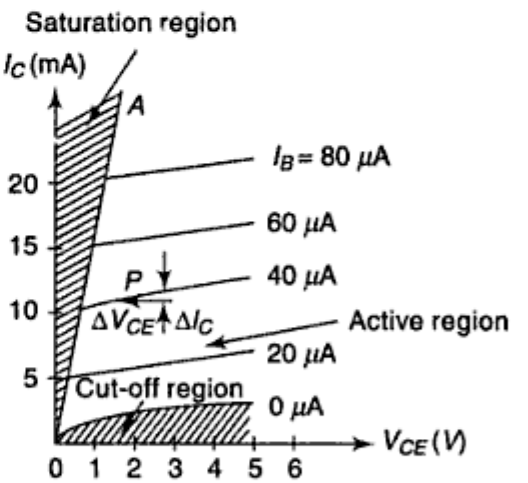


Fig. 6.12 CE output characteristics

$$\beta = \frac{\alpha}{1-\alpha} \qquad \text{and} \qquad I_C = (1+\beta)\, I_{CBO} + \beta I_B$$

For larger values of V_{CE}, due to Early effect, a very small change in α is reflected in a very large change in β. For example, when $\alpha = 0.98$, $\beta = \frac{0.98}{1-0.98} = 49$. If α increases to 0.985, then $\beta = \frac{0.985}{1-0.985} = 66$. Here, a slight increase in α by about 0.5% results in an increases in β by about 34%. Hence, the output characteristics of CE configuration show a larger slope when compared with CB configuration.

The output characteristics have three regions, namely, saturation region, cutoff region and active region. The region of curves to the left of the line *OA* is called the

saturation region (hatched), and the line *OA* is called the saturation line. In this region, both junctions are forward biased and an increase in the base current does not cause a corresponding large change in I_C. The ratio of $V_{CE(sat)}$ to I_C in this region is called saturation resistance.

The region below the curve for $I_B = 0$ is called the *cut-off region* (hatched). In this region, both junctions are reverse biased. When the operating point for the transistor enters the cut-off region, the transistor is OFF. Hence, the collector current becomes almost zero and the collector voltage almost equals V_{CC}, the collector supply voltage. The transistor is virtually an open circuit between collector and emitter.

The central region where the curves are uniform in spacing and slope is called the *active region* (unhatched). In this region, emitter-base junction is forward biased and the collector-base junction is reverse biased. If the transistor is to be used as a linear amplifier, it should be operated in the active region.

If the base current is subsequently driven large and positive, the transistor switches into the saturation region via the active region, which is traversed at a rate that is dependent on factors such as gain and frequency response. In this ON condition, large collector current flows and collector voltage falls to a very low value, called V_{CEsat}, typically around 0.2 V for a silicon transistor. The transistor is virtually a short circuit in this state.

High speed switching circuits are designed in such a way that transistors are not allowed to saturate, thus reducing switching times between ON and OFF times.

Transistor parameters The slope of the CE characteristics will give the following four transistor parameters. Since these parameters have different dimensions, they are commonly known as common emitter *hybrid parameters* or *h-parameters.*

(i) Input impedance (h_{ie}): It is defined as the ratio of the change in (input) base voltage to the change in (input) base current with the (output) collector voltage V_{CE} kept constant. Therefore,

$$h_{ie} = \frac{\Delta V_{BE}}{\Delta I_B}, \; V_{CE} \text{ constant} \tag{6.18}$$

It is the slope of CE input characteristics I_B versus V_{BE} as shown in Fig. 6.11. The typical value of h_{ie} ranges from 500 to 2000 Ω.

(ii) Output admittance (h_{oe}): It is defined as the ratio of change in the (output) collector current to the corresponding change in the (output) collector voltage with the (input) base current I_B kept constant. Therefore,

$$h_{oe} = \frac{\Delta I_C}{\Delta V_{CE}}, \; I_B \text{ constant} \tag{6.19}$$

It is the slope of CE output characteristic I_C versus V_{CE} as shown in Fig. 6.12. The typical value of this parameter is of the order of 0.1 to 10 μ mhos.

(iii) Forward current gain (h_{fe}): It is defined as a ratio of the change in the (output) collector current to the corresponding change in the (input) base current keeping the (output) collector voltage V_{CE} constant. Hence,

$$h_{fe} = \frac{\Delta I_C}{\Delta I_B}, \; V_{CE} \text{ constant} \tag{6.20}$$

It is the slope of I_C versus I_B curve. Its typical value varies from 20 to 200.

(iv) Reverse voltage gain (h_{re}): It is defined as the ratio of the change in the (input) base voltage and the corresponding change in (output) collector voltage with constant (input) base current, I_B. Hence,

$$h_{re} = \frac{\Delta V_{BE}}{\Delta V_{CE}}, \quad I_B \text{ constant} \tag{6.21}$$

It is the slope of V_{BE} versus V_{CE} curve. Its typical value is of the order of 10^{-5} to 10^{-4}.

6.6.3 CC Configuration

The circuit diagram for determining the static characteristics of an NPN transistor in the common collector configuration is shown in Fig. 6.13.

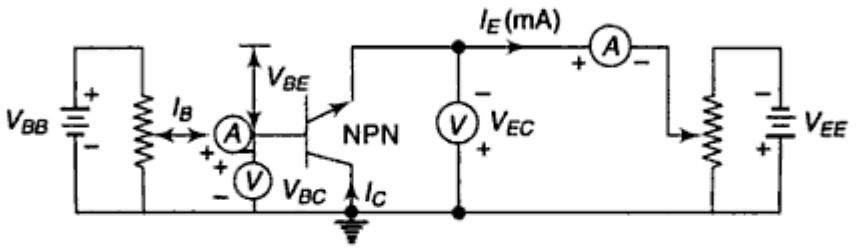


Fig. 6.13 Circuit to determine CC static characteristics

Input characteristics To determine the input characteristics, V_{EC} is kept at a suitable fixed value. The base-collector voltage V_{BC} is increased in equal steps and the corresponding increase in I_B is noted. This is repeated for different fixed values of V_{EC}. Plots of V_{BC} versus I_B for different values of V_{EC} shown in Fig. 6.14 are the input characteristics.

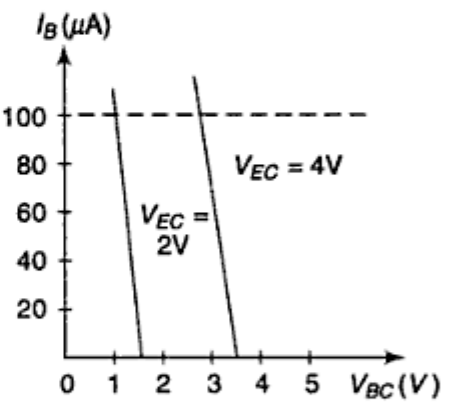


Fig. 6.14 CC input characteristics

Output characteristics The output characteristics shown in Fig. 6.15 are the same as those of the common emitter configuration.

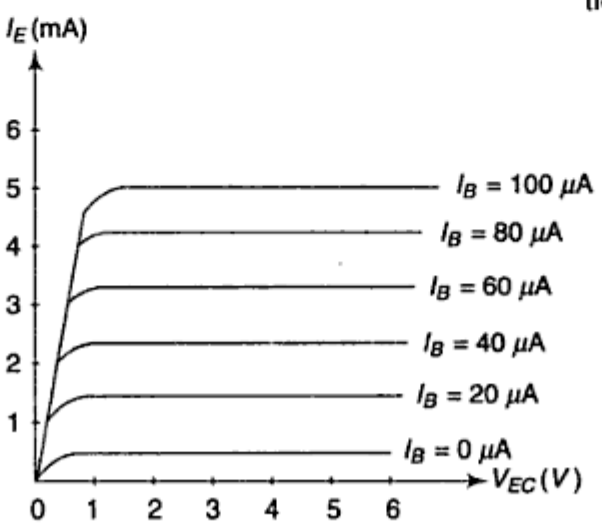


Fig. 6.15 CC output characteristics

and $$I_C = (1 + \beta)\, I_{CBO} + \beta I_B$$

The above equation can be expressed as

$$I_C - I_{CBO} = \beta I_{CBO} + \beta I_B$$

Therefore, $$\beta = \frac{I_C - I_{CBO}}{I_B - (-I_{CBO})}$$

The output characteristics of CE configuration show that in the *cut-off* region, the values $I_E = 0$, $I_C = I_{CBO}$ and $I_B = -I_{CBO}$. Therefore, the above equation gives the ratio of the collector-current increment to the base-current change from cut off to I_B, and hence *β is called the large-signal current gain of common-emitter transistor.* The d.c. current gain of the transistor is given by

$$\beta_{\text{d.c.}} \equiv h_{FE} \equiv \frac{I_C}{I_B}$$

Based on this h_{FE} value, we can determine whether the transistor is in saturation or not. For any transistor, in general, I_B is large compared to I_{CBO}. Under this condition, the value of $h_{FE} \approx \beta$.

The small-signal CE forward short-circuit gain β' is defined as the ratio of a collector-current increment ΔI_C for a small base-current change ΔI_B, at a fixed collector-to-emitter voltage V_{CE}.

i.e. $$\beta' \equiv \left.\frac{\partial I_C}{\partial I_B}\right|_{V_{CE}}$$

If β is independent of currents, then $\beta' = \beta \approx h_{FE}$. However, β is a function of current, then $\beta' = \beta + (I_{CBO} + I_B)\, \dfrac{\partial \beta}{\partial I_B}$. By using $\beta' = h_{fe}$ and $\beta \approx h_{FE}$. Therefore, the above equation becomes

$$h_{fe} = \frac{h_{FE}}{1 - (I_{CBO} + I_B)\dfrac{\partial h_{FE}}{\partial I_C}}$$

In Fig. 6.17, the h_{FE} versus I_C shows a maximum and hence $h_{fe} > h_{FE}$ for smaller currents, and $h_{fe} < h_{FE}$ for larger currents. Therefore, the above equation is valid only for the active region.

Example 6.1 In a common-base transistor circuit, the emitter current I_E is 10 mA and the collector current I_C is 9.8 mA. Find the value of the base current I_B.
Solution:

$$I_E = 10 \text{ mA and } I_C = 9.8 \text{ mA}$$

We know that emitter current is

$$I_E = I_B + I_C$$

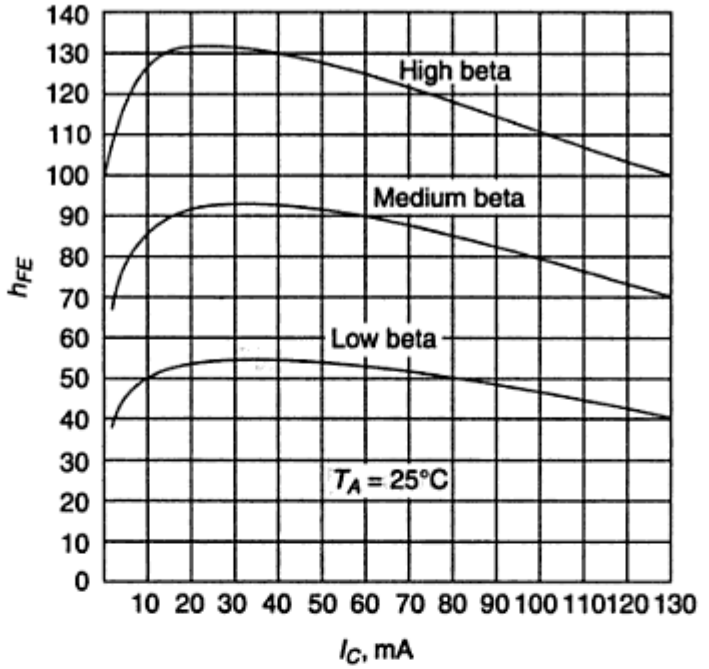


Fig. 6.17 Characteristic curves of d.c. current gain h_{FE} (at V_{CE} = – 0.25 V) versus collector current for low, medium and high beta values

i.e., $10 = I_B + 9.8$

Therefore, $I_B = 0.2$ mA

Example 6.2 In a common-base connection, the emitter current I_E is 6.28 mA and the collector current I_C is 6.20 mA. Determine the common-base d.c. current gain.
Solution:

Given: $I_E = 6.28$ mA and $I_C = 6.20$ mA

We know that common-base d.c. current gain,

$$\alpha = \frac{I_C}{I_E} = \frac{6.20 \times 10^{-3}}{6.28 \times 10^{-3}} = 0.987$$

Example 6.3 The common-base d.c. current gain of a transistor is 0.967. If the emitter current is 10 mA, what is the value of base current?
Solution:

Given: $\alpha = 0.967$ and $I_E = 10$ mA

The common-base d.c. current gain (α) is

$$\alpha = 0.967 = \frac{I_C}{I_E} = \frac{I_C}{10}$$

Therefore, $I_C = 0.967 \times 10 = 9.67$ mA

The emitter current $I_E = I_B + I_C$

i.e. $10 = I_B + 9.67$

Therefore, $I_B = 0.33$ mA

Example 6.4 The transistor has $I_E = 10$ mA and $\alpha = 0.98$. Determine the values of I_C and I_B.
Solution:

Given: $I_E = 10$ mA and $\alpha = 0.98$

The common-base d.c. current gain, $\alpha = \dfrac{I_C}{I_E}$

i.e. $0.98 = \dfrac{I_C}{10}$

Therefore $I_C = 0.98 \times 10 = 9.8$ mA

The emitter current $I_E = I_B + I_C$

i.e. $10 = I_B + 9.8$

Therefore, $I_B = 0.2$ mA

Example 6.5 If a transistor has a α of 0.97, find the value of β. If $\beta = 200$, find the value of α.
Solution:

If $\alpha = 0.97, \beta = \dfrac{\alpha}{1-\alpha} = \dfrac{0.97}{1-0.97} = 32.33$

If $\beta = 200, \alpha = \dfrac{\beta}{\beta+1} = \dfrac{200}{200+1} = 0.995$

Example 6.6 A transistor has $\beta = 100$. If the collector current is 40 mA, find the value of emitter current.
Solution:

Given: $\beta = 100$ and $I_C = 40$ mA

$$\beta = 100 = \frac{I_C}{I_B} = \frac{40}{I_B}$$

Therefore, $I_B = 40/100 = 0.4$ mA and

$$I_E = I_B + I_C = (0.4 + 40) \times 10^{-3} = 40.4 \text{ mA}$$

Example 6.7 A transistor has $\beta = 150$. Find the collector and base currents, if $I_E = 10$ mA.
Solution:

Given: $\beta = 150$ and $I_E = 10$ mA

The common-base current gain, $\alpha = \dfrac{\beta}{\beta+1} = \dfrac{150}{150+1} = 0.993$

Also, $$\alpha = \frac{I_C}{I_E}$$

i.e. $$0.993 = \frac{I_C}{10}$$

Therefore, $$I_C = 0.993 \times 10 = 9.93 \text{ mA}$$

The emitter current $$I_E = I_B + I_C$$

i.e., $$10 \times 10^{-3} = I_B + 9.93 \times 10^{-3}$$

Therefore, $$I_B = (10 - 9.93) \times 10^{-3} = 0.07 \text{ mA}$$

Example 6.8 Determine the values of I_B and I_E for the transistor circuit if I_C = 80 mA and $\beta = 170$.

Solution:

Given: $$\beta = 170 \text{ and } I_C = 80 \text{ mA}$$

We know that (β), $$\beta = 170 = \frac{I_C}{I_B} = \frac{80 \times 10^{-3}}{I_B}$$

Therefore, $$I_B = \frac{80 \times 10^{-3}}{170} = 0.47 \text{ mA}$$

and $$I_E = I_B + I_C = (0.47 + 80) \text{ mA} = 80.47 \text{ mA}$$

Example 6.9 Determine the values of I_C and I_E for the transistor circuit of $\beta = 200$ and $I_B = 0.125$ mA.

Solution:

Given: $$I_B = 0.125 \text{ mA and } \beta = 200$$

Therefore, $$\beta = 200 = \frac{I_C}{I_B} = \frac{I_C}{0.125 \times 10^{-3}}$$

Therefore, $$I_C = 200 \times 0.125 \times 10^{-3} = 25 \text{ mA}$$

and $$I_E = I_B + I_C = (0.125 + 25) \times 10^{-3} = 25.125 \text{ mA}$$

Example 6.10 Determine the values of I_C and I_B for the transistor circuit of I_E = 12 mA and $\beta = 100$.

Solution:

Given: $$I_E = 12 \text{ mA and } \beta = 100$$

We know that base current,

$$I_B = \frac{I_E}{1+\beta} = \frac{12 \times 10^{-3}}{1+100} = 0.1188 \text{ mA}$$

and collector current, $$I_C = I_E - I_B = (12 - 0.1188) \times 10^{-3} = 11.8812 \text{ mA}$$

Example 6.11 A transistor has $I_B = 100\ \mu A$ and $I_C = 2$ mA. Find (a) β of the transistor, (b) α of the transistor, (c) emitter current I_E, (d) if I_B changes by + 25 μA and I_C changes by + 0.6 mA, find the new value of β.
Solution:

Given: $I_B = 100\ \mu A = 100 \times 10^{-6}$ A and $I_C = 2$ mA $= 2 \times 10^{-3}$ A.

(a) To find β of the transistor

$$\beta = \frac{I_C}{I_B} = \frac{2 \times 10^{-3}}{100 \times 10^{-6}} = 20$$

(b) To find α of the transistor

$$\alpha = \frac{\beta}{\beta + 1} = \frac{20}{1 + 20} = 0.952$$

(c) To find emitter current, I_E

$$I_E = I_B + I_C = 100 \times 10^{-6} + 2 \times 10^{-3}\ \text{A}$$
$$= (0.01 + 2) \times 10^{-3} = 2.01 \times 10^{-3}\ \text{A} = 2.01\ \text{mA}$$

(d) To find the new value of β when $\Delta I_B = 25\ \mu A$ and $\Delta I_C = 0.6$ mA

Therefore, $I_B = (100 + 25)\ \mu A = 125\ \mu A$

$I_C = (2 + 0.6)$ mA $= 2.6$ mA

New value of β of the transistor,

$$\beta = \frac{I_C}{I_B} = \frac{2.6 \times 10^{-3}}{125 \times 10^{-6}} = 20.8$$

Example 6.12 For a transistor circuit having $\alpha = 0.98$, $I_{CBO} = I_{CO} = 5\ \mu A$ and $I_B = 100\ \mu A$, find I_C and I_E.
Solution:

Given: $\alpha = 0.98$, $I_{CBO} = I_{CO} = 5\ \mu A$ and $I_B = 100\ \mu A$

The collector current is

$$I_C = \frac{\alpha \cdot I_B}{1 - \alpha} + \frac{I_{CO}}{1 - \alpha} = \frac{0.98 \times 100 \times 10^{-6}}{1 - 0.98} + \frac{5 \times 10^{-6}}{1 - 0.98} = 5.15\ \text{mA}$$

The emitter current is

$$I_E = I_B + I_C = 100 \times 10^{-6} + 5.15 \times 10^{-3} = 5.25\ \text{mA}$$

Example 6.13 A germanium transistor used in a complementary symmetry amplifier has $I_{CBO} = 10\ \mu A$ at 27°C and $h_{FE} = 50$. (a) Find I_C when $I_B = 0.25$ mA and (b) assuming h_{FE} does not increase with temperature, find the value of new collector current, if the transistor's temperature rises to 50°C.
Solution:

Given: $I_{CBO} = 10\ \mu A$ and $h_{FE}\ (= \beta) = 50$

Example 6.20 Calculate the values of I_C and I_E for a transistor with $\alpha_{d.c.} = 0.99$ and $I_{CBO} = 5\ \mu A$. I_B is measured as 20 μA.
Solution:

Given, $\alpha_{d.c.} = 0.99$, $I_{CBO} = 5\ \mu A$ and $I_B = 20\ \mu A$

$$I_C = \frac{\alpha_{d.c.} I_B}{1-\alpha_{d.c.}} + \frac{I_{CBO}}{1-\alpha_{d.c.}} = \frac{0.99 \times 20 \times 10^{-6}}{1-0.99} + \frac{5 \times 10^{-6}}{1-0.99} = 2.48\ mA$$

Therefore, $I_E = I_B + I_C = 20 \times 10^{-6} + 2.48 \times 10^{-3} = 2.5\ mA$

Example 6.21 The reverse leakage current of the transistor when connected in CB configuration is 0.2 μA and it is 18 μA when the same transistor is connected in CE configuration. Calculate $\alpha_{d.c.}$ and $\beta_{d.c.}$ of the transistor.
Solution:

The leakage current $I_{CBO} = 0.2\ \mu A$

$I_{CEO} = 18\ \mu A$

Assume that $I_B = 30\ mA$

$I_E = I_B + I_C$

$I_C = I_E - I_B = \beta I_B + (1 + \beta)\ I_{CBO}$

We know that $I_{CEO} = \frac{I_{CEO}}{1-\alpha} = (1 + \beta) I_{CBO}$

$$\beta = \frac{I_{CEO}}{I_{CBO}} - 1 = \frac{18}{0.2} - 1 = 89$$

$$I_c = \beta I_B + (1 + \beta) I_{CBO}$$
$$= 89\ (30 \times 10^{-3}) + (1 + 89)\ (0.2 \times 10^{-6})$$
$$= 2.67\ A$$

$$\alpha_{d.c.} = 1 - \frac{I_{CBO}}{I_{CEO}} = 1 - \frac{0.2 \times 10^{-6}}{18 \times 10^{-6}} = 0.988$$

$$\beta_{d.c.} = \frac{I_C - I_{CBO}}{I_B - I_{CEO}}$$
$$= \frac{2.67 - 0.2 \times 10^{-6}}{30 \times 10^{-3} - 18 \times 10^{-6}} = 89$$

Example 6.22 If $\alpha_{d.c.} = 0.99$ and $I_{CBO} = 50\ \mu A$, find emitter current.
Solution:

Given: $\alpha_{d.c.} = 0.99$ and $I_{CBO} = 1\ \mu A$,

Assume that $I_B = 1\ mA$

$$I_C = \frac{\alpha_{d.c.} I_B}{1-\alpha_{d.c.}} + \frac{I_{CBO}}{1-\alpha_{d.c.}} = \frac{0.99\left(1\times10^{-3}\right)}{1-0.99} + \frac{5\times10^{-6}}{1-0.99}$$

$$= \frac{0.99\times10^{-3}}{0.01} + \frac{50\times10^{-6}}{0.01} = 99\text{ mA} + 5\text{ mA} = 104\text{ mA}$$

$$I_E = I_C + I_B = 104\text{ mA} + 1\text{ mA} = 105\text{ mA}$$

6.9 BREAKDOWN IN TRANSISTORS

There is a possibility of voltage breakdown in the transistor at high voltages even though the rated dissipation of the transistor is not exceeded. Therefore, there is an upper limit to the maximum allowable collector junction voltage. There are two types of breakdown, namely, ***Avalanche multiplication*** or ***Avalanche breakdown*** and ***reach-through*** or ***punch-through***.

6.9.1 Avalanche Breakdown and Multiplication

When a diode is reverse biased, there is a limit on the voltage that can be applied which is the Avalanche voltage. Similarly, in the transistor, the maximum reverse biasing voltage which may be applied before breakdown between the collector and base terminals with the emitter open is called breakdown voltage BV_{CBO}. Therefore, an upper limit is set on the collector voltage V_{CB} by avalanche breakdown in the reverse biased collector–base junction. Fields of order 10^6 V/m are required in the depletion layer for breakdown to occur, which usually limits V_{CB} to a maximum of several tens of volts.

Breakdown may occur because of Avalanche multiplication of the current I_{CO} that crosses the collector junction. As a result of this multiplication, the current becomes MI_{CO}, where M is the Avalanche multiplication factor. At the breakdown voltage BV_{CBO}, multiplication factor M becomes infinite and the current rises abruptly in the breakdown region, as shown in Fig. 6.19, there will be large changes in current with small changes in applied voltage.

The Avalanche multiplication factor depends on the voltage V_{CB} between collector and base, which has been found to be given empirically by

$$M = \frac{1}{1-\left(\frac{V_{CB}}{BV_{CBO}}\right)^n} \tag{6.27}$$

where the empirical constant, n, depends on the lattice material and the carrier type, which is usually in the range of about 2 to 10; for N-type silicon $n \approx 4$ and for P-type $n \approx 2$, and controls the sharpness of the onset breakdown. Taking Avalanche multiplication into account, I_C has the magnitude $M\,\alpha\,I_E$, where α is the common base current gain. As a result, in the presence of Avalanche multiplication, the current gain of CB transistor has become $M\mu$. As the effective current gain exceeds unity, the emitter circuit may then display negative resistance effects, which can lead to undesirable instabilities.

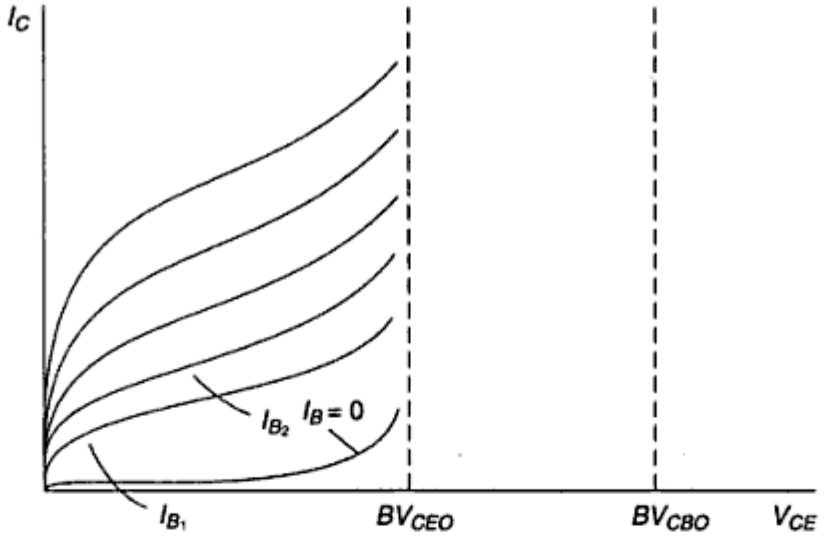


Fig. 6.19 CE characteristics in the breakdown region

For the CE configuration, the collector to emitter breakdown voltage BV_{CEO} with base open is

$$BV_{CEO} = BV_{CBO} \sqrt[n]{\frac{1}{h_{FE}}} \tag{6.28}$$

In general, BV_{CEO} is 40–50% of BV_{CBO}. This is the upper limit of V_{CE} that can be placed across the transistor without damaging it.

6.9.2 Reach-Through or Punch-Through

According to Early effect, the width of the collector-junction transition region increases with increased collector-junction voltage. As the voltage applied across the junction V_{CB} increases the transition region penetrates deeper into the base and will have spread completely across the base to reach the emitter junction, as the base is very thin. Thus, the collector voltage has reached through the base region. This effect, known as reach-through, also affects the output characteristics of a transistor since I_C versus V_{CB} curves are no longer horizontal but take on a positive slope indicating that the device has a finite output impedance that is voltage dependent. Since the input characteristics are also affected, the input impedance is also influenced by V_{CB}.

It is possible to raise the punch-through voltage by increasing the doping concentration in the base, but this automatically reduces the emitter efficiency.

Punch-through takes place at a fixed voltage between collector and base and is not dependent on circuit configuration, whereas Avalanche multiplication takes place at different voltages depending upon the circuit configuration. Therefore, the voltage limit of a particular transistor is determined by either of the two types of breakdown, whichever occurs at lower voltage.

6.10 EBERS–MOLL MODEL

The general expression for collector current I_C of a transistor for any voltage across collector junction V_C and emitter current I_E is

$$I_C = -\alpha_N I_E - I_{CO}\left(e^{\frac{V_C}{V_T}} - 1\right) \tag{6.29}$$

where α_N is the current gain in *normal* operation and I_{CO} is the collector junction reverse saturation current.

In inverted mode of operation, the above equation can be written as

$$I_E = -\alpha_I I_C - I_{EO}\left(e^{\frac{V_E}{V_T}} - 1\right) \tag{6.30}$$

where α_I is the inverted common-base current gain and I_{EO} is the emitter junction reverse saturation current.

The above four parameters are related by the condition

$$\alpha_I I_{CO} = \alpha_N I_{EO} \tag{6.31}$$

For many transistors I_{EO} lies in the range 0.5 I_{CO} to I_{CO}.

Figure 6.20 shows the Ebers–Moll model for a PNP transistor. Here, two separate ideal diodes are connected back to back with saturation currents $-I_{EO}$ and $-I_{CO}$ and there are two dependent current-controlled current sources shunting the ideal diodes. The current sources account for the minority carrier transport across the base. An application of Kirchoff's current law to the collector node of Fig. 6.20 gives

$$I_C = -\alpha_N I_E + I = -\alpha_N I_E + I_O\left(e^{\frac{V_C}{V_T}} - 1\right) \tag{6.32}$$

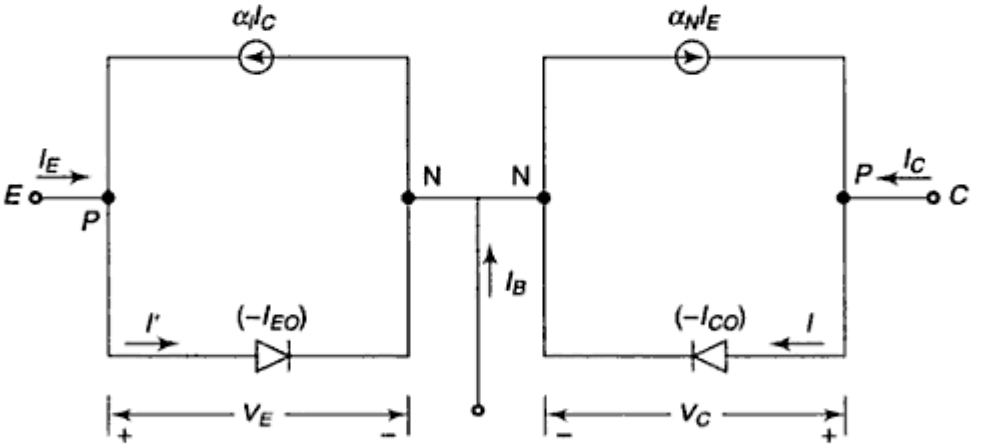


Fig. 6.20 Ebers–Moll model for a PNP transistor

where I is the diode current. As I_O is the magnitude of reverse saturation current, then $I_O = -I_{CO}$. Substituting this value of I_O in Eqn. (6.32), we get

$$I_C = -\alpha_N I_E - I_{CO}\left(e^{\frac{V_C}{V_T}} - 1\right)$$

which is nothing but the general expression for collector current of a transistor given in Eqn. (6.29). Hence, this model is valid for both forward and reverse static voltages applied across the transistor junctions. Here, base spreading resistance has been omitted and the difference between I_{CBO} and I_{CO} have been neglected.

The dependent current sources may be removed from the Fig. 6.20, provided $\alpha_N = \alpha_I = 0$. If the base-width is made much larger than the diffusion length of minority carriers in the base, all minority carriers will recombine in the base and no minority carrier will be available to reach the collector. Therefore, the transistor amplification factor α will become zero. As a result, transistor action ceases. Hence, it is not possible to construct a transistor by simply placing two isolated diodes back to back.

6.11 BIAS STABILITY

The quiescent operating point of a transistor amplifier should be established in the active region of its characteristics. Since the transistor parameters such as β, I_{CO} and V_{BE} are functions of temperature, the operating point shifts with changes in temperature. The stability of different methods of biasing transistor (BJT, FET and MOSFET) circuits and compensation techniques for stabilizing the operating point are discussed in this chapter.

Need for Biasing

In order to produce distortion-free output in amplifier circuits, the supply voltages and resistances in the circuit must be suitably chosen. These voltages and resistances establish a set of d.c. voltage V_{CEQ} and current I_{CQ} to operate the transistor in the active region. These voltages and currents are called *quiescent values* which determine the *operating point* or Q-point for the transistor. The process of giving proper supply voltages and resistances for obtaining the desired Q-point is called *biasing*. The circuits used for getting the desired and proper operating point are known as *biasing circuits*.

The collector current for common-emitter amplifier is expressed by

$$I_C = \beta I_B + I_{CEO} = \beta I_B + (1 + \beta) I_{CO}$$

Here the three variables h_{FE}, i.e., β, I_B and I_{CO} are found to increase with temperature. For every 10°C rise in temperature, I_{CO} doubles itself. When I_{CO} increases, I_C increases significantly. This causes power dissipation to increase and hence to make I_{CO} increase. This will cause I_C to increase further and the process becomes cumulative which will lead to thermal runaway that will destroy the transistor. In addition, the quiescent operating point can shift due to temperature changes and the transistor can be driven into the region of saturation. The effect of β on the Q-point is shown in Fig. 6.21. To establish the operating point in the active region, compensation techniques are needed.

DC Load Line Referring to the biasing circuit of Fig. 6.22(a), the values of V_{CC} and R_C are fixed and I_C and V_{CE} are dependent on R_B.

Applying Kirchhoff's voltage law to the collector circuit in Fig. 6.22(a), we get $V_{CC} = I_C R_C + V_{CE}$.

The straight line represented by AB in Fig. 6.22(b) is called the d.c. *load line*. The coordinates of the end point A are obtained by substituting $V_{CE} = 0$ in the above equation. Then $I_C = \dfrac{V_{CC}}{R_C}$. Therefore, the coordinates of A are $V_{CE} = 0$ and $I_C = \dfrac{V_{CC}}{R_C}$.

Therefore, the d.c. load line AB is drawn with the point B ($0B$ = 24 V) on the V_{CE} axis and the point A ($0A$ = 3 mA) on the I_C axis, as shown in Fig. 6.22(b).

(ii) For fixing the optimum operating point Q, mark the middle of the d.c. load line AB and the corresponding V_{CE} and I_C values can be found.

Here, $$V_{CEQ} = \frac{V_{CC}}{2} = 12 \text{ V} \quad \text{and} \quad I_{CQ} = 1.5 \text{ mA}$$

(iii) a.c. load line:
To draw an a.c. load line, two end points viz. maximum V_{CE} and maximum I_C when the signal is applied are required.

The A.C. load, $$R_{a.c.} = R_C \| R_L = \frac{8 \times 24}{8 + 24} = 6 \text{ k}\Omega$$

Maximum $$V_{CE} = V_{CEQ} + I_{CQ} R_{a.c.}$$
$$= 12 + 1.5 \times 10^{-3} \times 6 \times 10^3 = 21 \text{ V}$$

This locates the point D ($0D$ = 21 V) on the V_{CE} axis.

Maximum collector current $$= I_{CQ} + \frac{V_{CEQ}}{R_{a.c.}}$$

$$= 1.5 \times 10^{-3} + \frac{12}{6 \times 10^3} = 3.5 \text{ mA}$$

This locates the point C ($0C$ = 3.5 mA) on the I_C axis. By joining points C and D, a.c. load line CD is constructed.

Example 6.24 For the transistor amplifier shown in Fig. 6.23(a), V_{CC} = 12 V, R_1 = 8 kΩ, R_2 = 4 kΩ, R_C = 1 kΩ, R_E = 1 kΩ and R_L = 1.5 kΩ. Assume V_{BE} = 0.7 V. (i) Draw the d.c. load line, (ii) determine the operating point, and (iii) draw the a.c. load line.

Solution:

(i) *DC load line:*
Referring to Fig. 6.23(a), we have $V_{CC} = V_{CE} + I_C (R_C + R_E)$.
To draw the d.c. load line, we need two end points, viz. maximum V_{CE} point (at $I_C = 0$) and maximum I_C point (at $V_{CE} = 0$).
Maximum $V_{CE} = V_{CC}$ = 12 V, which locates the point B ($0B$ = 12 V) of the d.c. load line.

Maximum $$I_C = \frac{V_{CC}}{R_C + R_E} = \frac{12}{(1+1) \times 10^3} = 6 \text{ mA}$$

$$\text{Maximum } I_C = I_{CQ} + \frac{V_{CEQ}}{R_{a.c.}}$$

$$= 3.3 \times 10^{-3} + \frac{5.4}{0.6 \times 10^3} = 12.3 \text{ mA}$$

This locates the point D ($0D$ = 12.3 mA) on the I_C axis. By joining points C and D, a.c. load line CD is constructed.

Example 6.25 Design the circuit shown in Fig. 6.24, given Q-point values are to be I_{CQ} = 1 mA and V_{CEQ} = 6 V. Assume that V_{CC} = 10 V, β = 100 and V_{BE} (on) = 0.7 V.

Solution: The collector resistance is

$$R_C = \frac{V_{CC} - V_{CEQ}}{I_{CQ}} = \frac{10 - 6}{1 \times 10^{-3}} = 4 \text{ k}\Omega$$

The base current is

$$I_{BQ} = \frac{I_{CQ}}{\beta} = \frac{1 \times 10^{-3}}{100} = 10 \text{ } \mu\text{A}$$

The base resistance is

$$R_B = \frac{V_{CC} - V_{BE}\ (on)}{I_{BQ}} = \frac{10 - 0.7}{10 \times 10^{-6}} = 0.93 \text{ M}\Omega$$

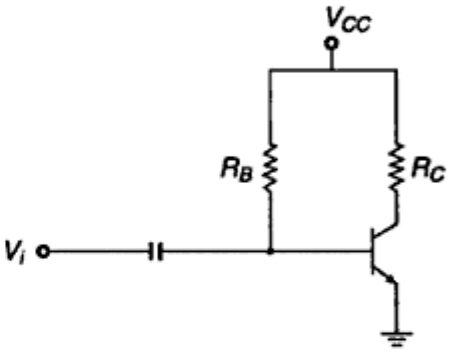


Fig. 6.24

Example 6.26 Determine the characteristics of a circuit shown in Fig. 6.25. Assume that β = 100 and V_{BE} (on) = 0.7 V.

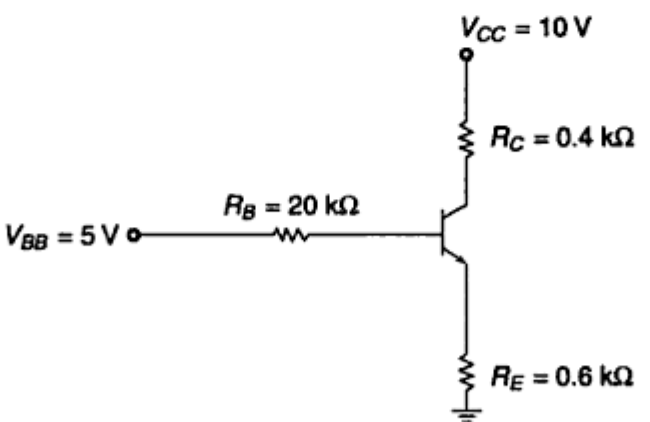


Fig. 6.25

Solution: Referring to Fig. 6.25, Kirchhoff's voltage law equation is

$$V_{BB} = I_B R_B + V_{BE}(on) + I_E R_E$$

We know that

$$I_E = I_B + I_C = I_B + \beta I_B = (1 + \beta) I_B$$

The base current $$I_B = \frac{V_{BB} - V_{BE}\,(on)}{R_B + (1+\beta)\,R_E} = \frac{5 - 0.7}{20 \times 10^3 + 101 \times 600} = 53.34\ \mu A$$

Therefore,
$$I_C = \beta I_B = 100 \times 53.34 \times 10^{-6} = 5.334\ mA$$
$$I_E = I_C + I_B = 5.334 \times 10^{-3} + 53.34 \times 10^{-6} = 5.38734\ mA$$
$$V_{CE} = V_{CC} - I_C R_C - I_E R_E = 10 - 5.334 \times 10^{-3} \times 400 - 5.38734 \times 10^{-3} \times 600 = 4.634\ V$$

The Q point is at $V_{CEQ} = 4.634$ V and $I_{CQ} = 5.334$ mA

6.11.1 Thermal Runaway

The collector current for the CE circuit of Fig. 6.22 is given by $I_C = \beta I_B + (1 + \beta)\, I_{CO}$. The three variables in the equation, β, I_B and I_{CO} increase with rise in temperature. In particular, the reverse saturation current or leakage current I_{CO} changes greatly with temperature. Specifically, it doubles for every 10 °C rise in temperature. The collector current I_C causes the collector–base junction temperature to rise which, in turn, increase I_{CO}, as a result I_C will increase still further, which will further rise the temperature at the collector–base junction. This process will become cumulative leading to "*thermal runaway*." Consequently, the ratings of the transistor are exceeded which may destroy the transistor itself.

The collector is normally made larger in size than the emitter in order to help dissipate the heat developed at the collector junction.

However, if the circuit is designed such that the base current I_B is made to decrease automatically with rise in temperature, then the decrease in βI_B will compensate for the increase in $(1 + \beta)I_{CO}$, keeping I_C almost constant.

In power transistors, the heat developed at the collector junction may be removed by the use of heat sink, which is a metal sheet fitted to the collector and whose surface radiates heat quickly.

6.11.2 Stability Factor (*S*)

The extent to which the collector current I_C is stabilised with varying I_{CO} is measured by a stability factor S. It is defined as the rate of change of collector current I_C with respect to the collector–base leakage current I_{CO}, keeping both the current I_B and the current gain β constant

$$S = \frac{\partial I_C}{\partial I_{CO}} \approx \frac{dI_C}{dI_{CO}} \approx \frac{\Delta I_C}{\Delta I_{CO}},\ \beta \text{ and } I_B \text{ constant} \qquad (6.33)$$

The collector current for a CE amplifier is given by

$$I_C = \beta I_B + (\beta + 1) I_{CO} \qquad (6.34)$$

Differentiating the above equation with respect to I_C, we get

$$1 = \beta \frac{dI_B}{dI_C} + (\beta + 1) \frac{dI_{CO}}{dI_C}$$

Therefore, $$\left(1-\beta\frac{dI_B}{dI_C}\right)=\frac{(\beta+1)}{S}$$

$$S=\frac{1+\beta}{1-\beta\left(\frac{dI_B}{dI_C}\right)} \tag{6.35}$$

From this equation, it is clear that this factor S should be as small as possible to have better thermal stability.

Stability Factors S′ and S″ The stability factor S' is defined as the rate of change of I_C with V_{BE}, keeping I_{CO} and β constant

$$S'=\frac{\partial I_C}{\partial V_{BE}}\approx\frac{\Delta I_C}{\Delta V_{BE}}$$

The stability factor S'' is defined as the rate of change of I_C with respect to β, keeping I_{CO} and V_{BE} constant

$$S''=\frac{\partial I_C}{\partial \beta}\approx\frac{\Delta I_C}{\Delta \beta}$$

6.12 METHODS OF TRANSISTOR BIASING

The staibility factor for some commonly used biasing circuits are discussed below.

6.12.1 Fixed Bias or Base Resistor Method

A common emitter amplifier using fixed bias circuit is shown in Fig. 6.26. The d.c. analysis of the circuit yields the following equation:

$$V_{CC}=I_BR_B+V_{BE} \tag{6.36}$$

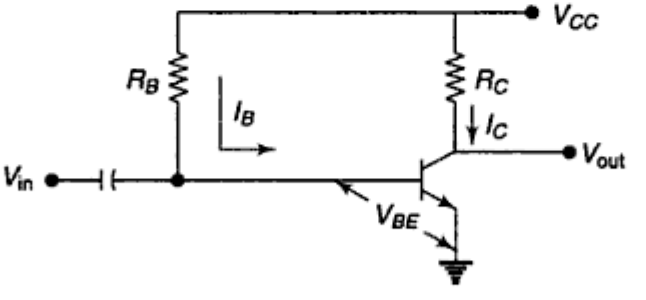


Fig. 6.26 Fixed bias circuit

Therefore, $$I_B=\frac{V_{CC}-V_{BE}}{R_B}$$

Since this equation is independent of current I_C, $dI_B/dI_C = 0$ and the stability factor given in Eq. (6.35) reduces to

$$S=1+\beta$$

Since β is a large quantity, this is a very poor bias stable circuit. Therefore, in practice, this circuit is not used for biasing the base.

The advantage of this method are: (a) simplicity, (b) small number of components required, and (c) if the supply voltage is very large as compared to V_{BE} of the transistor, then the base current becomes largely independent of the voltage V_{BE}.

Example 6.27 In the fixed bias compensation method (refer to Fig. 6.26), a silicon transistor with $\beta = 100$ is used. $V_{CC} = 6$ V, $R_C = 3$ kΩ, $R_B = 530$ kΩ. Draw the d.c. load line and determine the operating point. What is the stability factor?

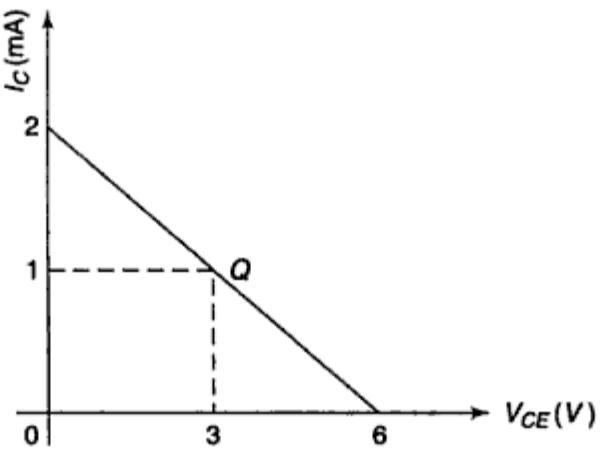


Fig. 6.27 DC load line

Solution:

(i) *DC load line:*

$$V_{CE} = V_{CC} - I_C R_C$$

When $I_C = 0$, $V_{CE} = V_{CC} = 6$ V

When $V_{CE} = 0$, $I_C = \dfrac{V_{CC}}{R_C} = \dfrac{6}{3 \times 10^3} = 2$ mA

(ii) *Operating point Q:*

For silicon transistor, $V_{BE} = 0.7$ V

$$V_{CC} = I_B R_B + V_{BE}$$

Therefore, $$I_B = \frac{V_{CC} - V_{BE}}{R_B} = \frac{6 - 0.7}{530 \times 10^3} = 10\ \mu\text{A}$$

Therefore, $$I_C = \beta I_B = 100 \times 10 \times 10^{-6} = 1 \text{ mA}$$

$$V_{CE} = V_{CC} - I_C R_C = 6 - 1 \times 10^{-3} \times 3 \times 10^3 = 3 \text{ V}$$

Therefore operating point is $V_{CEQ} = 3$ V and $I_{CQ} = 1$ mA.

(iii) *Stability factor:* $S = 1 + \beta = 1 + 100 = 101$

6.12.2 Collector-to-Base Bias or Biasing with Feedback Resistor

A common emitter amplifier using collector-to-base bias circuit is shown in Fig. 6.28. This circuit is the simplest way to provide some degree of stabilisation to the amplifier operating point.

Solution: Refer to Fig. 6.26. We know that for a silicon transistor, $V_{BE} = 0.7$ V.

(a) *To determine R_B:*

The operating point is at $V_{CE} = 7$ V and $I_C = 1$ mA

Here, $$R_C = \frac{V_{CC} - V_{CE}}{I_C} = \frac{12 - 7}{1 \times 10^{-3}} = 5 \text{ k}\Omega$$

$$I_B = \frac{I_C}{\beta} = \frac{1 \times 10^{-3}}{100} = 10 \text{ μA}$$

Using the relation,

$$R_B = \frac{V_{CC} - V_{BE} - I_C R_C}{I_B} = \frac{12 - 0.7 - 1 \times 10^{-3} \times 5 \times 10^3}{10 \times 10^{-6}} = 630 \text{ k}\Omega$$

(b) *Stability factor,* $$S = \frac{1 + \beta}{1 + \beta\left[\frac{R_C}{R_C + R_B}\right]}$$

$$= \frac{1 + 100}{1 + 100\left[\frac{5 \times 10^3}{(5 + 630) \times 10^3}\right]} = 56.5$$

(c) *To determine new operating point when $\beta = 50$*

$$V_{CC} = \beta I_B R_C + I_B R_B + V_{BE}$$

$$= I_B (\beta R_C + R_B) + V_{BE}$$

i.e $$12 = I_B (50 \times 5 \times 10^3 + 630 \times 10^3) + 0.7$$

$$I_B = \frac{11.3}{880 \times 10^3} = 12.84 \text{ μA}$$

Therefore, $$I_C = \beta I_B = 50 \times 12.84 \times 10^{-6} = 0.642 \text{ mA}$$

$$V_{CE} = V_{CC} - I_C R_C = 12 - 0.642 \times 10^{-3} \times 5 \times 10^3 = 8.79 \text{ V}$$

Therefore the coordinates of the new operating point are $V_{CEQ} = 8.79$ V and $I_{CQ} = 0.642$ mA.

Example 6.29 In a collector-to-base CE amplifier circuit of Fig. 6.26 having $V_{CC} = 12$ V, $R_C = 250\ \Omega$, $I_B = 0.25$ mA, $\beta = 100$ and $V_{CEQ} = 8$ V, calculate R_B and stability factor.

Solution: $$R_B = \frac{V_{CEQ}}{I_B} = \frac{8}{0.25 \times 10^{-3}} = 32 \text{ k}\Omega$$

Substituting this equation in Eqn. (6.35), we get

$$S = \frac{1+\beta}{1+\beta\left(\frac{R_E}{R_E + R_B}\right)}$$

Therefore,

$$S = (1+\beta)\frac{1+\frac{R_B}{R_E}}{1+\beta+\frac{R_B}{R_E}} \tag{6.41}$$

As can be seen, the value of S is equal to one if the ratio R_B/R_E is very small as compared to 1. As this ratio becomes comparable to unity, and beyond towards infinity, the value of the stability factor goes on increasing till $S = 1 + \beta$.

This improvement in the stability up to a factor equal to 1 is achieved at the cost of power dissipation. To improve the stability, the equivalent resistance R_B must be decreased, forcing more current in the voltage divider network of R_1 and R_2.

Often, to prevent the loss of gain due to the negative feedback, R_E is shunted by a capacitor C_E. The capacitive reactance X_{CE} must be equal to about one-tenth of the value of the resistance R_E at the lowest operating frequency.

To determine the stability factor S'

The stability factor S′ is defined as the rate of change of I_C with V_{BE}, keeping I_{CO} and β constant.

$$S' = \frac{\partial I_C}{\partial V_{BE}} \approx \frac{\Delta I_C}{\Delta V_{BE}}$$

From Fig. 6.29(b),

$$\begin{aligned} V_T &= I_B R_B + V_{BE} + I_E R_E \\ &= I_B[R_B + R_E] + I_C R_E + V_{BE} \quad \text{since } [I_E = I_B + I_C] \end{aligned} \tag{6.42}$$

From Eq. (6.34), we have

$$I_B = \frac{I_C - (1+\beta) I_{CO}}{\beta} \tag{6.43}$$

Substituting Eq. (6.43) in Eq. (6.42), we get

$$V_T = \frac{I_C}{\beta}(R_B + R_E) + V_{BE} + I_C R_E + \frac{I_{CO}}{\beta}(1+\beta)\{R_B + R_E\} \tag{6.44}$$

Differentiating the above Eq. w.r.t. V_{BE}, we get

$$0 = \frac{dI_C}{dV_{BE}}\left(\frac{R_B + R_E}{\beta}\right) + 1 + R_E\frac{dI_C}{dV_{BE}} + 0$$

$$-1 = \frac{dI_C}{dV_{BE}}\left[R_E + \frac{R_B + R_E}{\beta}\right]$$

$$-1 = \frac{dI_C}{dV_{BE}}\left[\frac{R_B + (1+\beta)R_E}{\beta}\right]$$

Therefore, $$S' = \frac{dI_C}{dV_{BE}} = \frac{-\beta}{R_B + (1+\beta)R_E} \tag{6.45}$$

To determine the stability factor S''

The stability factor S'' is defined as the rate of change of I_C w.r.t. to β, keeping I_{CO} and V_{BE} constant.

Rearranging Eqn. (6.44), we have

$$I_C = \frac{\beta(V_T - V_{BE})}{R_B + (1+\beta)R_E} + \frac{\beta\left(\frac{1+\beta}{\beta}\right)I_{CO}\ (R_B + R_E)}{R_B + (1+\beta)R_E} \tag{6.46}$$

Since $\beta \gg 1$, the numerator of the second term can be written as

$$(R_B + R_E)\left(\frac{1+\beta}{\beta}\right) I_{CO} \approx (R_B + R_E)\, I_{CO} \tag{6.47}$$

Substituting Eqn. (6.47) in Eqn. (6.46), we have

$$I_C = \frac{\beta(V_T - V_{BE})}{R_B + (1+\beta)R_E} + \frac{\beta(R_B + R_E)\, I_{CO}}{R_B + (1+\beta)R_E}$$

Therefore, $$I_C = \frac{\beta[V_T - V_{BE} + (R_B + R_E)I_{CO}]}{R_B + (1+\beta)R_E}$$

Let, $$V' = (R_B + R_E)\, I_{CO},$$

Therefore, $$I_C = \frac{\beta[V_T - V_{BE} + V']}{R_B + (1+\beta)R_E} \tag{6.48}$$

Differentiating the above equation w.r.t. β and simplifying, we obtain

$$S'' = \frac{dI_C}{d\beta} = \frac{I_C}{\beta\left[1 + \beta\left(\frac{R_E}{R_E + R_B}\right)\right]} = \frac{S\, I_C}{\beta(1+\beta)} \tag{6.49}$$

Example 6.30 In a CE germanium transistor amplifier circuit, the bias is provided by self bias, i.e., emitter resistor and potential divider arrangement (refer to Fig. 6.27). The various parameters are: $V_{CC} = 16$ V, $R_C = 3$ kΩ, $R_E = 2$ kΩ, $R_1 = 56$ kΩ, $R_2 = 20$ kΩ and $\alpha = 0.985$. Determine (a) the coordinates of the operating point, and (b) the stability factor S.

Solution: For a germanium transistor, $V_{BE} = 0.3$ V. As $\alpha = 0.985$,

$$\beta = \frac{\alpha}{1-\alpha} = \frac{0.985}{1-0.985} = 66$$

(a) *To find the coordinates of the operating point*
Referring to Fig. 6.29, we have

Thevenin's voltage, $$V_T = \frac{R_2}{R_1 + R_2} V_{CC}$$

$$= \frac{20 \times 10^3}{76 \times 10^3} \times 16 = 4.21 \text{ V}$$

Thevenin's resistance, $$R_B = \frac{R_1 R_2}{R_1 + R_2} = \frac{20 \times 10^3 \times 56 \times 10^3}{76 \times 10^3}$$

$$= 14.737 \text{ k}\Omega$$

The loop equation around the base circuit is

$$V_T = I_B R_B + V_{BE} + (I_B + I_C) R_E$$

$$= \frac{I_C}{\beta} R_B + V_{BE} + \left(\frac{I_C}{\beta} + I_C \right) R_E$$

$$4.21 = \frac{I_C}{66} \times 14.737 \times 10^3 + 0.3 + I_C \left(\frac{1}{66} + 1 \right) \times 2 \times 10^3$$

$$3.91 = I_C \, [0.223 + 2.03] \times 10^3$$

Therefore, $$I_C = \frac{3.91}{2.253 \times 10^3} = 1.73 \text{ mA}$$

Since I_B is very small, $I_C \approx I_E = 1.73$ mA

Therefore, $$V_{CE} = V_{CC} - I_C R_C - I_E R_E$$

$$= V_{CC} - I_C [R_C + R_E] = 16 - 1.73 \times 10^{-3} \times 5 \times 10^3$$

$$= 7.35 \text{ V}$$

Therefore, the coordinates of the operating point are $I_C = 1.73$ mA and $V_{CE} = 7.35$ V.

(b) *To find the stability factor S*

$$S = (1 + \beta) \frac{1 + \frac{R_B}{R_E}}{1 + \beta + \frac{R_B}{R_E}}$$

$$= (1 + 66) \frac{1 + \frac{14.737}{2}}{1 + 66 + \frac{14.737}{2}}$$

$$= 67 \times \frac{8.3685}{74.3685} = 7.537$$

Example 6.31 A CE transistor amplifier with voltage divider bias circuit of Fig. 6.29 is designed to establish the quiescent point at $V_{CE} = 12$ V, $I_C = 2$ mA and stability factor ≤ 5.1. If $V_{CC} = 24$ V, $V_{BE} = 0.7$ V, $\beta = 50$ and $R_C = 4.7$ kΩ, determine the values of resistors R_E, R_1 and R_2.

Solution:

(a) *To determine* R_E

$$V_{CE} = V_{CC} - I_C R_C - I_E R_E$$
$$= V_{CC} - I_C[R_C + R_E], \text{ since } I_C \approx I_E$$
$$12 = 24 - 2 \times 10^{-3}\,[4.7 \times 10^3 + R_E]$$

Therefore, $R_E = 1.3$ kΩ

(b) *To determine* R_1 *and* R_2

Stability factor, $$S = \frac{1+\beta}{1+\beta\left(\frac{R_E}{R_E + R_B}\right)}, \quad \text{where } R_B = \frac{R_1 R_2}{(R_1 + R_2)}$$

$$5.1 = \frac{51}{1+50\left(\frac{1.3\times10^3}{1.3\times10^3 + R_B}\right)}$$

i.e. $$1+50\left(\frac{1.3\times10^3}{1.3\times10^3 + R_B}\right) = \frac{51}{5.1} = 10$$

Therefore, $$\left(\frac{50\times1.3\times10^3}{1.3\times10^3 + R_B}\right) = 9$$

$$1.3 \times 10^3 + R_B = \frac{50\times1.3\times10^3}{9} = 7.2 \text{ kΩ}$$

$$R_B = 5.9 \text{ kΩ}$$

Also, we know that for a good voltage divider, the value of resistor $R_2 = 0.1\beta R_E$.

Therefore, $R_2 = 0.1 \times 50 \times 1.3 \times 10^3 = 6.5$ kΩ

$$R_B = \frac{R_1 R_2}{R_1 + R_2}$$

$$5.9 \times 10^3 = \frac{R_1 \times 6.5\times10^3}{R_1 + 6.5\times10^3}$$

Simplifying, we get $R_1 = 64$ kΩ

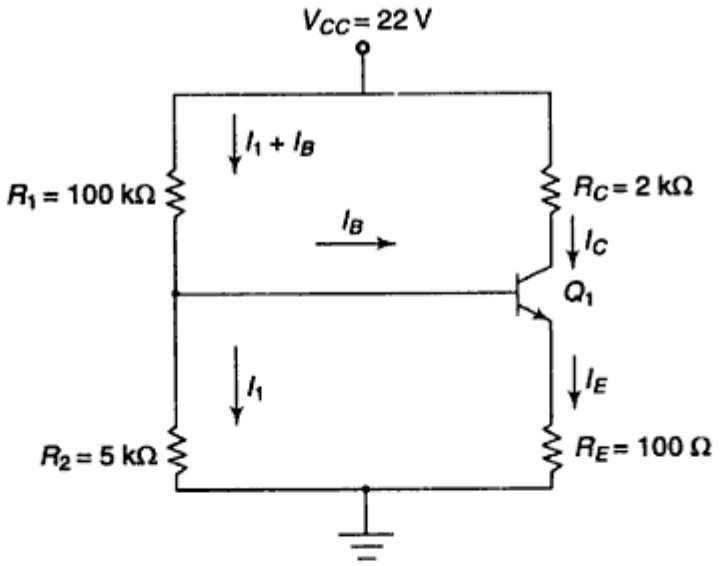


Fig. 6.31

As $I_E = I_C + I_B$

$= \beta I_B + I_B = (1 + \beta) I_B$

Hence $V_{CC} = R_1[I_1 + I_B] + V_{BE} + (1 + \beta)\, I_B R_E$

Substituting for I_1 from Eqn. (1), we get

$$V_{CC} = R_1 \left[\frac{V_{CC} - I_B R_1}{R_1 + R_2} + I_B \right] + V_{BE} + (1 + \beta)\, I_B R_E$$

$$V_{CC} = R_1 \left[\frac{V_{CC} + I_B R_2}{R_1 + R_2} \right] + V_{BE} + (1 + \beta)\, I_B R_E$$

Substituting for V_{CC}, R_1, R_2, V_{BE}, β and R_E

$$22 = 100 \times 10^3 \left[\frac{22 + I_B \times 5 \times 10^3}{(100 + 5) \times 10^3} \right] + 0.6 + (1 + 60)\, I_B \times 100$$

$$22 = 0.952\, [22 + I_B \times 5 \times 10^3] + 0.6 + 6100\, I_B$$

$$22 = 20.944 + 11{,}100\, I_B + 0.6$$

$$I_B = 41.08\ \mu A$$

$$I_C = \beta I_B = 60 \times 41.08 \times 10^{-6} = 2.46\ mA$$

Applying KVL to collector circuit, we get

$$V_{CC} = I_C R_C + V_{CE} + I_E R_E = I_C R_C + V_{CE} + (1 + \beta)\, I_B R_E$$

Hence $V_{CE} = V_{CC} - I_C R_C - (1 + \beta)\, I_B R_E$

$= 22 - (2.46) \times 10^{-3} \times 2 \times 10^3 - (61)\,(41.08 \times 10^{-6})\,(100)$

$= 22 - 4.92 - 0.25 = 16.83$ V

To find stability factor, (*S*):

Stability factor for voltage divider bias is

6.12.4 Common Base Stability

In a common base amplifier circuit, the equation for collector current I_C is given by

$$I_C = \alpha I_E + I_{CO}$$

$$S \approx dI_C/dI_{CO} = 1$$

Since this is the most stable possible, the common base amplifier circuit is not in need of bias stabilisation.

6.12.5 Advantage of Self Bias (Voltage Divider Bias) Over Other Types of Biasing

In fixed bias method discussed in Section 6.12.1, the stability factor is given by

$$S = 1 + \beta$$

Since β, is normally a large quantity, this circuit provides very poor stability. Therefore the fixed biasing technique is not preferred for biasing the base.
In collector-to-base bias method discussed in Section 6.12.2, when R_C is very small, $S \approx 1 + \beta$, which is equal to that of fixed bias. Hence, collector-to-base bias method is also not preferable. In self bias method discussed in Section 6.12.3, when R_B/R_E is very small, $S \approx 1$, which provides good stability. Hence, self bias method is the best one over other types of "biasing".

6.13 BIAS COMPENSATION

The various biasing circuits considered in the previous sections used some types of negative feedback to stabilise the operation point. Also, diodes, thermistors and sensistors can be used to compensate for variations in current.

6.13.1 Diode Compensation

Figure 6.32 shows a transistor amplifier with a diode D connected across the base–emitter junction for compensation of change in collector saturation current I_{CO}. The diode is of the same material as the transistor and it is reverse biased by the base–emitter junction voltage V_{BE}, allowing the diode reverse saturation current I_O to flow through diode D. The base current $I_B = I - I_O$.

As long as temperature is constant, diode D operates as a resistor. As the temperature increases, I_{CO} of the transistor increases. Hence, to compensate for this, the base current I_B should be decreased.

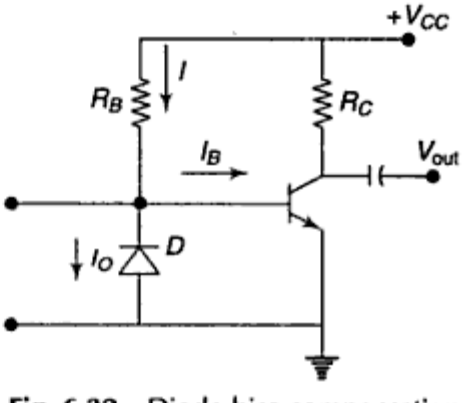


Fig. 6.32 Diode bias compensation

The increase in temperature will also cause the leakage current I_O through D to increase and thereby decreasing the base current I_B. This is the required action to keep I_C constant.

This method of bias compensation does not need a change in I_C to effect the change in I_B, as both I_O and I_{CO} can track almost equally according to the change in temperature.

6.13.2 Thermistor Compensation

In Fig. 6.33, a thermistor, R_T, having a negative temperature coefficient is connected in parallel with R_2. The resistance of thermistor decreases exponentially with increase of temperature. An increase in temperature will decrease the base voltage V_{BE}, reducing I_B and I_C. Bias stabilisation is also provided by R_E and C_E.

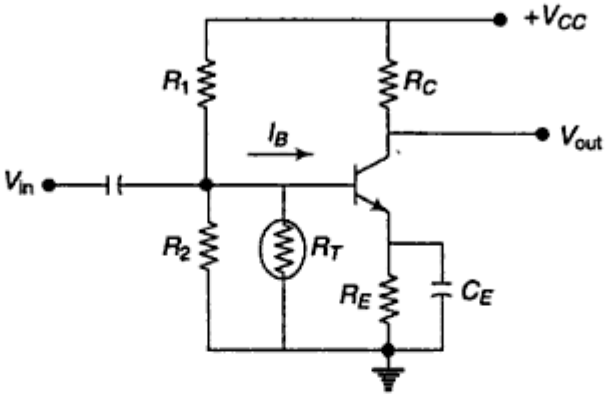


Fig. 6.33 Thermistor bias compensation

6.13.3 Sensistor Compensation

In Fig. 6.34, a sensistor, R_S, having a positive temperature coefficient is connected across R_1 (or R_E). R_S increases with temperature. As temperature increases, the equivalent resistance of the parallel combination of R_1 and R_S also increases and hence the base voltage V_{BE} decreases, reducing I_B and I_C. This reduced I_C compensates for the increased I_C caused by the increase in I_{CO}, V_{BE} and β due to temperature rise.

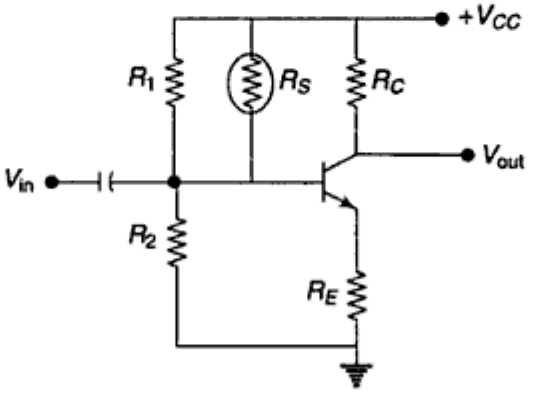


Fig. 6.34 Sensistor bias compensation

Review Questions

1. What is a Bipolar junction transistor? How are its terminals named?
2. Explain the operation of NPN and PNP transistors.
3. What are the different configurations of BJT?
4. Explain the input and output characteristics of a transistor in CB configuration.
5. Explain the early effect and its consequences.
6. Derive the relationship between α and β.

7. Why does the CE configuration provide large current amplification while the CB configuration does not?
8. Draw the circuit diagram of an NPN junction transistor CE configuration and describe the static input and output characteristics. Also, define active, saturation and cutoff regions, and saturation resistance of a CE transistor.
9. How will you determine h-parameters from the characteristics of CE configuration?
10. Determine the h-parameters from the characteristics of CB configuration.
11. What is the relation between I_B, I_E and I_C in CB configuration?
12. Explain the laboratory setup for obtaining the CC characteristics.
13. Compare the performance of a transistor in different configurations.
14. Define α, β and γ of a transistor. Show how are they related to each other?
15. Explain how a transistor is used as an amplifier.
16. From the characteristics of CE configuration, explain the large signal, d.c., and small signal CE values of current gain.
17. Calculate the values of I_C and I_E for a transistor with $\alpha_{d.c.} = 0.99$ and $I_{CBO} = 5\ \mu A$. I_B is measured as 20 μA. [*Ans.* $I_C = 2.48$ mA, $I_E = 2.5$ mA]
18. If $\alpha_{d.c.} = 0.99$ and $I_{CBO} = 50\ \mu A$, find emitter current. [*Ans.* $I_C = 104$ mA, $I_E = 105$ mA]
19. If I_C is 100 times larger than I_B, find the value of $\beta_{d.c.}$. [*Ans.* 100]
20. Find the value of $\alpha_{d.c.}$, if $\beta_{d.c.}$ is equal to 100. [*Ans.* 0.99]
21. Find the voltage gain of a transistor amplifier if its output is 5 V r.m.s. and the input is 100 mV r.m.s. [*Ans.* 50]
22. Find the value of $\alpha_{d.c.}$, when $I_C = 8.2$ mA and $I_E = 8.7$ mA. [*Ans.* 0.943]
23. If $\alpha_{d.c.}$ is 0.96 and $I_E = 9.35$ mA, determine I_C. [*Ans.* 8.98 mA]
24. Describe the two types of breakdown in transistors.
25. Draw the Ebers–Moll model for a PNP transistor and give the equations for emitter current and collector current.
26. What is meant by Q-point?
27. What is the need for biasing a transistor?
28. What factors are to be considered for selecting the operating point Q for an amplifier?
29. Distinguish between d.c. and a.c. load lines with suitable diagrams.
30. Briefly explain the reasons for keeping the operating point of a transistor as fixed.
31. What is thermal runaway? How can it be avoided?
32. What three factors contribute to thermal instability?
33. Define 'stability factor.' Why would it seem more reasonable to call this an instability factor?
34. Draw a fixed bias circuit and derive an expression for the stability factor.
35. If the coordinates of the operating point of a CE amplifier using fixed bias or base resistor method of biasing are $V_{CE} = 6$ V and $I_C = 1$ mA, determine the value of R_C and R_B. [*Ans.* $R_C = 3$ kΩ, $R_B = 300$ kΩ]
36. Consider a common emitter NPN transistor with fixed bias as shown in Fig. 6.20. If $\beta = 80$, $R_B = 390$ kΩ, $R_C = 1.5$ kΩ and $V_{CC} = 30$ V, find the coordinates of the Q-point. [*Ans.* 21 V, 6 mA]
37. A germanium transistor having $\beta = 100$ and $V_{BE} = 0.2$ V is used in a fixed bias amplifier circuit where $V_{CC} = 16$ V, $R_C = 5$ kΩ and $R_B = 790$ kΩ. Determine its operating point.
38. Derive an expression for the stability factor of a collector-to-base bias circuit.
39. Mention the disadvantages of collector-to-base bias. Can they be overcome?
40. In a germanium transistor CE amplifier biased by feedback resistor method, $V_{CC} = 20$ V, $V_{BE} = 0.2$ V, $\beta = 100$ and the operating point is chosen such that $V_{CE} = 10.4$ V and $I_C = 9.9$ mA. Determine the values of R_B and R_C. [*Ans.* 100 kΩ, 1 kΩ]

41. Draw a circuit diagram of CE transistor amplifier using emitter biasing. Describe qualitatively the stability action of the circuit.
42. Draw a voltage divider bias circuit and derive an expression for its stability factor.
43. Why does potential divider method of biasing become universal?
44. If the various parameters of a CE amplifier which uses the self bias method are V_{CC} = 12 V, R_1 = 10 kΩ, R_2 = 5 kΩ, R_C = 1 kΩ, R_E = 2 kΩ and β = 100, find (i) the coordinates of the operating point, and (ii) the stability factor, assuming the transistor to be of silicon. [*Ans.* V_{CE} = 7.05 V, I_C = 1.65 mA, S = 2.62]
45. In a CE germanium transistor amplifier using self bias circuit, R_C = 2.2 kΩ, β = 50, V_{CC} = 9 V and the operating point is required to be set at I_C = 2 mA and V_{CE} = 3V. Determine the values of R_1, R_2 and R_E.
[*Ans.* R_1 = 17.75 kΩ, R_2 = 4.75 kΩ, R_E = 800 Ω]
46. Determine the operating point for the circuit of a potential divider bias arrangement with $R_2 = R_C$ = 5 kΩ, R_E = 1 kΩ and R_1 = 40 kΩ. [*Ans.* V_{CE} = 6 V, I_C = 1 mA]
47. Calculate the values of R_1 and R_C in the voltage divider bias circuit so that Q-point is at V_{CE} = 6 V and I_C = 2 mA. Assume the transistor parameters are: α = 0.985, I_{CBO} = 4 μA and V_{BE} = 0.2 V. [*Ans.* R_C = 3 kΩ, R_1 = 5.54 kΩ]
48. Determine the stability factor for a CB amplifier circuit.
49. Draw a circuit which uses a diode to compensate for changes in I_{CO}. Explain how stabilisation is achieved in the circuit.
50. How will you provide temperature compensation for the variations of V_{BE} and stabilisation of the operating point?
51. What is the principle of providing thermal stabilisation by means of different methods of transistor biasing? How does this differ from the compensation techniques using a diode or thermistor or sensistor?

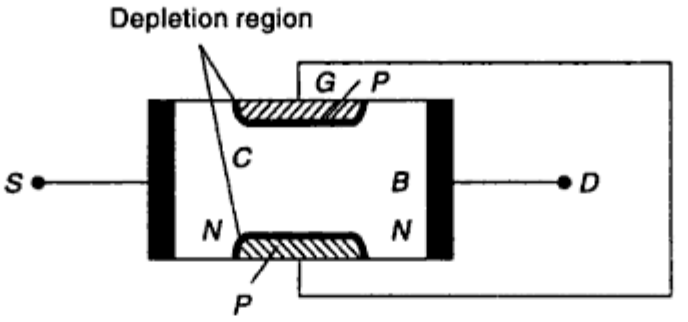


Fig. 7.1 JFET construction

When $V_{DS} = 0$ and V_{GS} is decreased from zero In this case, the PN junctions are reverse biased and hence the thickness of the depletion region increases. As V_{GS} is decreased from zero, the reverse bias voltage across the PN junction is increased and hence, the thickness of the depletion region in the channel until the two depletion regions make contact with each other. In this condition, the channel is said to be cut-off. The value of V_{GS} which is required to cut-off the channel is called the cut-off voltage V_C.

When $V_{GS} = 0$ and V_{DS} is increased from zero Drain is positive with respect to the source with $V_{GS} = 0$. Now the majority carriers (electrons) flow through the N-channel from source to drain. Therefore the conventional current I_D flows from drain to source. The magnitude of the current will depend upon the following factors:

1. The number of majority carriers (electrons) available in the channel, i.e. the conductivity of the channel.
2. The length L of the channel.
3. The cross-sectional area A of the channel at B.
4. The magnitude of the applied voltage V_{DS}. Thus the channel acts as a resistor of resistance R given by

$$R = \frac{\rho L}{A} \tag{7.1}$$

$$I_D = \frac{V_{DS}}{R} = \frac{AV_{DS}}{\rho L} \tag{7.2}$$

where ρ is the resistivity of the channel. Because of the resistance of the channel and the applied voltage V_{DS}, there is a gradual increase of positive potential along the channel from source to drain. Thus the reverse voltage across the PN junctions increases and hence the thickness of the depletion regions also increases. Therefore, the channel is wedge shaped as shown in Fig. 7.2.

As V_{DS} is increased, the cross-sectional area of the channel will be reduced. At a certain value V_P of V_{DS}, the cross-sectional area at B becomes minimum. At this voltage, the channel is said to be pinched off and the drain voltage V_P is called the pinch-off voltage.

As a result of the decreasing cross-section of the channel with the increase of V_{DS}, the following results are obtained.

the junction is further increased. Hence for a negative value of V_{GS}, the curve of I_D versus V_{DS} is similar to that for $V_{GS} = 0$, but the values of V_P and BV_{DGO} are lower, as shown in Fig. 7.3.

From the curves, it is seen that above the pinch-off voltage, at a constant value of V_{DS}, I_D increases with an increase of V_{GS}. Hence, a JFET is suitable for use as a voltage amplifier, similar to a transistor amplifier.

It can be seen from the curve that for voltage $V_{DS} = V_P$, the drain current is not reduced to zero. If the drain current is to be reduced to zero, the ohmic voltage drop along the channel should also be reduced to zero. Further, the reverse biasing to the gate-source PN junction essential for pinching off the channel would also be absent.

The drain current I_D is controlled by the electric field that extends into the channel due to reverse biased voltage applied to the gate; hence, this device has been given the name *Field Effect Transistor.*

In a bar of P-type semiconductor, the gate is formed due to N-type semiconductor. The working of the P-channel JFET will be similar to that of N-channel JFET with proper alterations in the biasing circuits; in this case holes will be the current carriers instead of electrons. The circuit symbols for N-channel and P-channel JFETs are shown in Fig. 7.4. It should be noted that the direction of the arrow points in the direction of conventional current which would flow into the gate if the PN junction was forward biased.

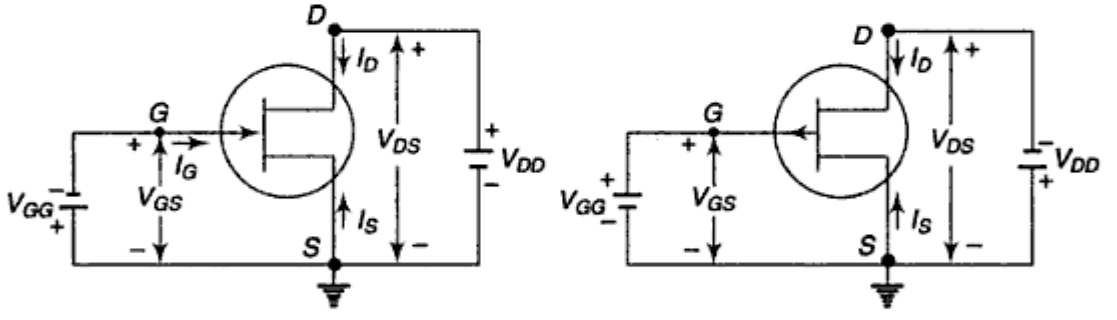


Fig. 7.4 Circuit symbols for N- and P-channel JFET

7.4 CHARACTERISTIC PARAMETERS OF THE JFET

In a JFET, the drain current I_D depends upon the drain voltage V_{DS} and the gate voltage V_{GS}. Any one of these variables may be fixed and the relation between the other two are determined. These relations are determined by the three parameters which are defined below.

(1) Mutual conductance or transconductance, g_m It is the slope of the transfer characteristic curves, and is defined by

$$g_m = \left(\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{DS}} = \frac{\Delta I_D}{\Delta V_{GS}}, V_{DS} \text{ held constant.}$$

It is the ratio of a small change in the drain current to the corresponding small change in the gate voltage at a constant drain voltage. The change in I_D and V_{GS}

should be taken on the straight part of the transfer characteristics. It has the unit of conductance in mho.

(2) Drain resistance, r_d It is the reciprocal of the slope of the drain characteristics and is defined by

$$r_d = \left(\frac{\partial V_{DS}}{\partial I_D}\right)_{V_{GS}} = \frac{\Delta V_{DS}}{\Delta I_D}, V_{GS} \text{ held constant.}$$

It is the ratio of a small change in the drain voltage to the corresponding small change in the drain current at a constant gate voltage. It has the unit of resistance in ohms.

The drain resistance at $V_{GS} = 0$ V, i.e. when the depletion regions of the channel are absent, is called as *drain-source ON resistance,* represented as R_{DS} or $R_{DS(ON)}$.

The reciprocal of r_d is called the drain conductance. It is denoted by g_d or g_{os}.

(3) Amplification factor, μ It is defined by

$$\mu = -\left(\frac{\partial V_{DS}}{\partial V_{GS}}\right) I_D = -\frac{\Delta V_{DS}}{\Delta V_{GS}}, I_D \text{ held constant.}$$

It is the ratio of a small change in the drain voltage to the corresponding small change in the gate voltage at a constant drain current. Here, the negative sign shows that when V_{GS} is increased, V_{DS} must be decreased for I_D to remain constant.

(4) Relationship among FET parameters As I_D depends on V_{DS} and V_{GS}, the functional equation can be expressed as

$$I_D = f(V_{DS}, V_{GS})$$

If the drain voltage is changed by a small amount from V_{DS} to $(V_{DS} + \Delta V_{DS})$ and the gate voltage is changed by a small amount from V_{GS} to $(V_{GS} + \Delta V_{GS})$, then the corresponding small change in I_D may be obtained by applying Taylor's theorem with neglecting higher order terms. Thus the small change ΔI_D is given by

$$\Delta I_D = \left(\frac{\partial I_D}{\partial V_{DS}}\right)_{V_{GS}} \Delta V_{DS} + \left(\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{DS}} \Delta V_{GS}$$

Dividing both the sides of this equation by ΔV_{GS}, we obtain

$$\frac{\Delta I_D}{\Delta V_{GS}} = \left(\frac{\partial I_D}{\partial V_{DS}}\right)_{V_{GS}} \left(\frac{\Delta V_{DS}}{\Delta V_{GS}}\right) + \left(\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{DS}}$$

If I_D is constant, then $\frac{\Delta I_D}{\Delta V_{GS}} = 0$

Therefore, we have

$$0 = \left(\frac{\partial I_D}{\partial V_{DS}}\right)_{V_{GS}} \left(\frac{\Delta V_{DS}}{\Delta V_{GS}}\right)_{I_D} + \left(\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{DS}}$$

Substituting the values of the partial differential coefficients, we get

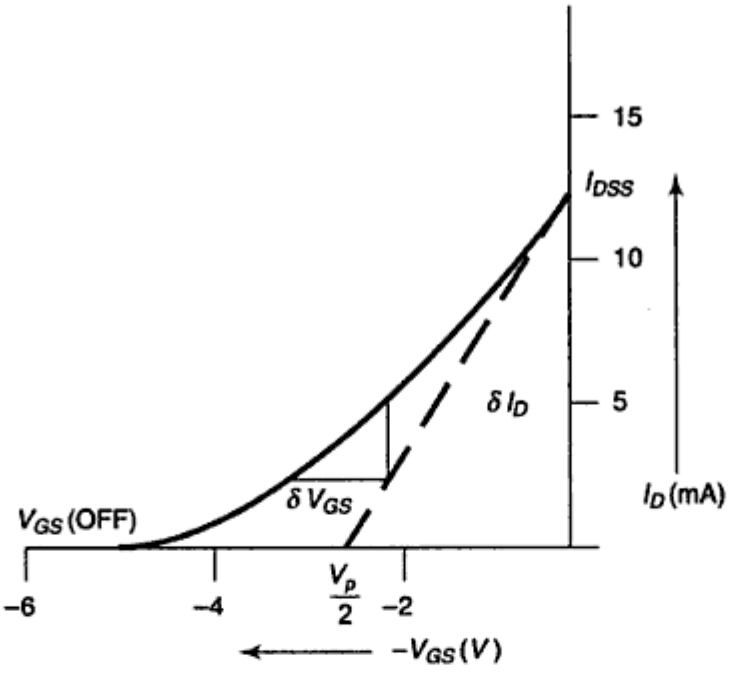


Fig. 7.5 Transfer characteristics of JFET

shape of the transfer characteristic is very nearly a parabola. It is found that the characteristic is approximately represented by the parabola,

$$I_{DS} = I_{DSS}\left(1-\frac{V_{GS}}{V_P}\right)^2 \tag{7.3}$$

where I_{DS} is the saturation drain current, I_{DSS} is the value of I_{DS} when $V_{GS} = 0$, and V_P is the pinch-off voltage.

Differentiating Eq. (7.3) with respect to V_{GS} we can obtain an expression for g_m.

$$\frac{\partial I_{DS}}{\partial V_{GS}} = I_{DSS} \times 2\left(1-\frac{V_{GS}}{V_P}\right)\left(-\frac{1}{V_P}\right)$$

We know that $g_m = \dfrac{\delta I_{DS}}{\delta V_{GS}}$, V_{DS} is constant.

Therefore, $g_m = \dfrac{-2\,I_{DSS}}{V_P}\left(1-\dfrac{V_{GS}}{V_P}\right)$ (7.4)

Now from Eqn. (7.3), we have

$$\left(1-\frac{V_{GS}}{V_P}\right) = \sqrt{\frac{I_{DS}}{I_{DSS}}} \tag{7.5}$$

Substituting this value in Eqn. (7.4), we get

$$g_m = \frac{2\sqrt{I_{DS}\ I_{DSS}}}{V_P}$$

Suppose $g_m = g_{mo}$, when $V_{GS} = 0$, then from Eqn. (7.4)

$$g_{mo} = -\frac{2I_{DSS}}{V_P} \tag{7.6}$$

Therefore, from Eqs (7.4) and (7.6)

$$g_m = g_{mo}\left(1 - \frac{V_{GS}}{V_P}\right) \tag{7.7}$$

Equation (7.5) shows that g_m varies as the square root of the saturation drain current I_{DS}, and Eqn. (7.7) shows that g_m decreases linearly with increase of V_{GS}.

7.6 SLOPE OF THE TRANSFER CHARACTERISTIC AT I_{DSS}

From Eq. (7.5), we have

$$g_m = \frac{2\sqrt{I_{DS}\ I_{DSS}}}{V_P}$$

or

$$\frac{\partial I_{DS}}{\partial V_{GS}} = \frac{2\sqrt{I_{DS}\ I_{DSS}}}{V_P}$$

Substituting $I_{DS} = I_{DSS}$,

$$\frac{\partial I_{DS}}{\partial V_{GS}} = -\frac{2I_{DSS}}{V_P} = \frac{I_{DSS}}{\frac{-V_P}{2}}$$

This equation shows that the tangent to the curve at $I_{DS} = I_{DSS}$, $V_{GS} = 0$, will have an intercept at $\frac{-V_P}{2}$ on the axis of V_{GS} as shown in Fig. 7.5. Therefore, the value of V_P can be found by drawing the tangent at $I_{DS} = I_{DSS}$, $V_{GS} = 0$.

The gate-source cut off voltage, $V_{GS(off)}$, on the transfer characteristic is equal to the pinch off voltage, V_P, on the drain characteristics, i.e. $V_P = |V_{GS(off)}|$.

Therefore, $I_D = I_{DSS}\left[1 - \frac{V_{GS}}{V_{GS\,(off)}}\right]^2$

Example 7.4 An N-channel JFET has $I_{DSS} = 8$ mA and $V_p = -5$ V. Determine the minimum value of V_{DS} for pinch-off region and the drain current I_{DS}, for $V_{GS} = -2$ V in the pinch-off region.

Solution: The minimum value of V_{DS} for pinch-off to occur for $V_{GS} = -2$ V is

$$V_{DSmin} = V_{GS} - V_P = -2 - (-5) = 3\text{V}$$

$$I_{DS} = I_{DSS}\left[1 - \frac{V_{GS}}{V_P}\right]^2$$

$$= 8 \times 10^{-3}\,[1 - (-2)/(-5)]^2 = 2.88\text{ mA}$$

7.7 COMPARISON OF JFET AND BJT

1. FET operation depends only on the flow of majority carriers-holes for P-channel FETs and electrons for N-channel FETs. Therefore, they are called Unipolar devices. Bipolar transistor (BJT) operation depends on both minority and majority current carriers.
2. As FET has no junctions and the conduction is through an N-type or P-type semiconductor material, FET is less noisy than BJT.
3. As the input circuit of FET is reverse biased, FET exhibits a much higher input impedance (in the order of 100 M Ω) and lower output impedance and there will be a high degree of isolation between input and output. So, FET can act as an excellent buffer amplifier but the BJT has low input impedance because its input circuit is forward biased.
4. FET is a voltage controlled device, i.e. voltage at the input terminal controls the output current, whereas BJT is a current controlled device, i.e. the input current controls the output current.
5. FETs are much easier to fabricate and are particularly suitable for ICs because they occupy less space than BJTs.
6. The performance of BJT is degraded by neutron radiation because of the reduction in minority-carrier lifetime, whereas FET can tolerate a much higher level of radiation since they do not rely on minority carriers for their operation.
7. The performance of FET is relatively unaffected by ambient temperature changes. As it has a negative temperature coefficient at high current levels, it prevents the FET from thermal breakdown. The BJT has a positive temperature coefficient at high current levels which leads to thermal breakdown.
8. Since FET does not suffer from minority carrier storage effects, it has higher switching speeds and cut-off frequencies. BJT suffers from minority carrier storage effects and therefore has lower switching speed and cut-off frequencies.
9. FET amplifiers have low gain bandwidth product due to the junction capacitive effects and produce more signal distortion except for small signal operation.
10. BJTs are cheaper to produce than FETs.

7.8 APPLICATIONS OF JFET

1. FET is used as a buffer in measuring instruments, receivers since it has high input impedance and low output impedance.
2. FETs are used in RF amplifiers in FM tuners and communication equipment for the low noise level.
3. Since the input capacitance is low, FETs are used in cascade amplifiers in measuring and test equipments.
4. Since the device is voltage controlled, it is used as a voltage variable resistor in operational amplifiers and tone controls.
5. FETs are used in mixer circuits in FM and TV receivers, and communication equipment because inter modulation distortion is low.
6. It is used in oscillator circuits because frequency drift is low.
7. As the coupling capacitor is small, FETs are used in low frequency amplifiers in hearing aids and inductive transducers.

continuous channel in an enhancement MOSFET, this condition is represented by the broken line in the symbols.

Two highly doped N^+ regions are diffused in a lightly doped substrate of P-type silicon substrate. One N^+ region is called the source *S* and the other one is called the drain *D*. They are separated by 1 mil (10^{-3} inch). A thin insulating layer of SiO_2 is grown over the surface of the structure and holes are cut into the oxide layer, allowing contact with source and drain. Then a thin layer of metal aluminium is formed over the layer of SiO_2. This metal layer covers the entire channel region and it forms the gate *G*.

The metal area of the gate, in conjunction with the insulating oxide layer of SiO_2 and the semiconductor channel forms a parallel plate capacitor. This device is called the insulated gate FET because of the insulating layer of SiO_2. This layer gives an extremely high input impedance for the MOSFET.

Operation If the substrate is grounded and a positive voltage is applied at the gate, the positive charge on *G* induces an equal negative charge on the substrats side between the source and drain regions. Thus, an electric field is produced between the source and drain regions. The direction of the electric field is perpendicular to the plates of the capacitor through the oxide. The negative charge of electrons which are minority carriers in the P-type substrate forms an inversion layer. As the positive voltage on the gate increases, the induced negative charge in the semiconductor increases. Hence, the conductivity increases and current flows from source to drain through the induced channel. Thus the drain current is enhanced by the positive gate voltage as shown in Fig. 7.7.

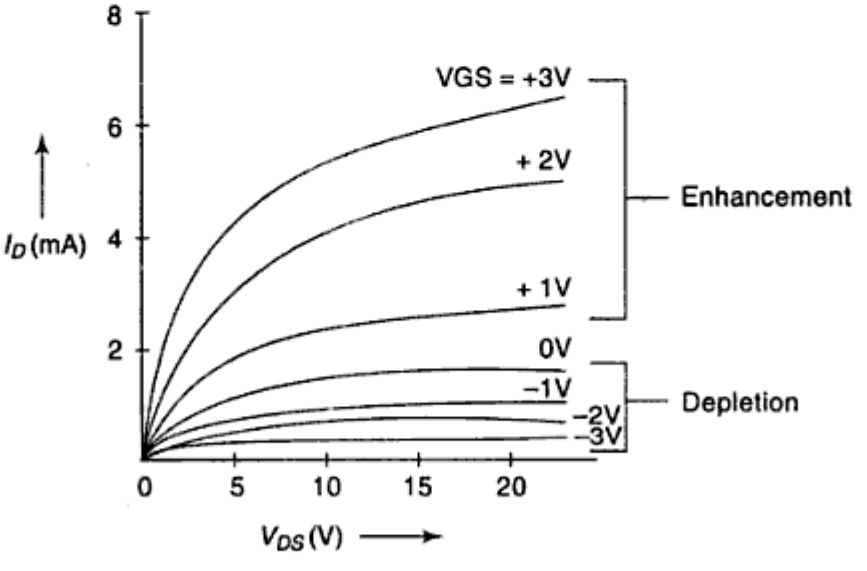


Fig. 7.7 Volt–ampere characteristics of MOSFET

7.11 DEPLETION MOSFET

The construction of an N-channel depletion MOSFET is shown in Fig. 7.8(a) where an N-channel is diffused between the source and drain to the basic structure of MOSFET. The circuit symbols for an N-channel and a P-channel depletion MOSFET are shown in Figs. 7.8(b) and (c), respectively.

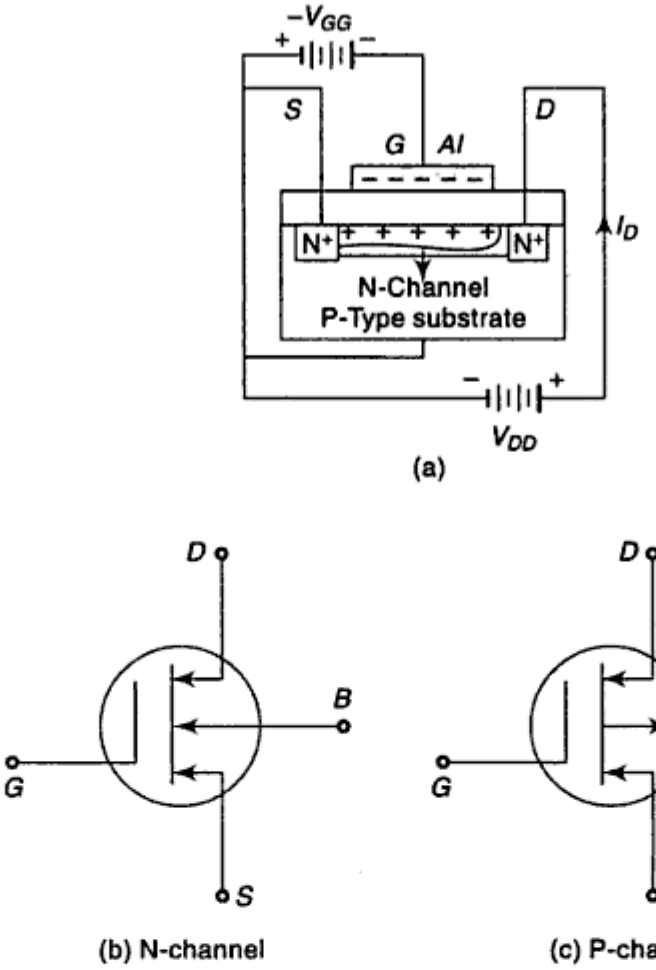


Fig. 7.8 (a) N-channel depletion MOSFET, (b) and (c) Circuit symbols for depletion MOSFETs.

With $V_{GS} = 0$ and the drain D at a positive potential with respect to the source, the electrons (majority carriers) flow through the N-channel from S to D. Therefore, the conventional current I_D flows through the channel D to S. If the gate voltage is made negative, positive charge consisting of holes is induced in the channel through SiO_2 of the gate-channel capacitor. The introduction of the positive charge causes depletion of mobile electrons in the channel. Thus a depletion region is produced in the channel. The shape of the depletion region depends on V_{GS} and V_{DS}. Hence the channel will be wedge shaped as shown in Fig. 7.8. When V_{DS} is increased, I_D increases and it becomes practically constant at a certain value of V_{DS}, called the pinch-off voltage. The drain current I_D almost gets saturated beyond the pinch-off voltage.

Since the current in an FET is due to majority carriers (electrons for an N-type material), the induced positive charges make the channel less conductive, and I_D drops as V_{GS} is made negative.

The depletion MOSFET may also be operated in an enhancement mode. It is only necessary to apply a positive gate voltage so that negative charges are induced into the N-type channel. Hence, the conductivity of the channel increases and I_D increases. As the depletion MOSFET can be operated with bipolar input signals irrespective of doping of the channel, it is also called as *dual mode MOSFET*. The volt-ampere characteristics are indicated in Fig. 7.7.

The output resistance can be determined at the Q-point by

$$r_0 = [\lambda K_N (V_{GSQ} - V_{TN})^2]^{-1}$$

$$= \left[\lambda I_{DSS} \left(1 - \frac{V_{GSQ}}{V_P} \right)^2 \right]^{-1}$$

i.e., $$r_0 \equiv \left[\lambda I_{DQ} \right]^{-1} = \frac{1}{\lambda I_{DQ}} = \frac{V_E}{I_{DQ}}$$

The output resistance is an important factor in the analysis of small signal equivalent circuit of MOSFET.

Temperature Effects

The threshold voltage V_{TN} or V_{TP} and conduction parameter K_N or K_P are functions of temperature. The magnitude of threshold voltage decreases with temperature and hence the drain current I_D increases with temperature at a given V_{GS}. The conduction parameter is directly proportional to mobility μ_N or μ_P of the carrier, which increases as the temperature decreases. Here the temperature dependent of mobility is larger than that of the threshold voltage. Hence the net effect of decreasing drain current at a given V_{GS} due to increase in temperature provides a negative feedback condition and hence the stability for a power MOSFET.

7.12 COMPARISON OF MOSFET WITH JFET

1. In enhancement and depletion types of MOSFET, the transverse electric field induced across an insulating layer deposited on the semiconductor material controls the conductivity of the channel. In the JFET the transverse electric field across the reverse biased PN junction controls the conductivity of the channel.
2. The gate leakage current in a MOSFET is of the order of 10^{-12} A. Hence the input resistance of a MOSFET is very high in the order of 10^{10} to 10^{15} Ω. The gate leakage current of a JFET is of the order of 10^{-9} A and its input resistance is of the order of 10^8 Ω.
3. The output characteristics of the JFET are flatter than those of the MOSFET and hence, the drain resistance of a JFET (0.1 to 1 MΩ) is much higher than that of a MOSFET (1 to 50 kΩ).
4. JFETs are operated only in the depletion mode. The depletion type MOSFET may be operated in both depletion and enhancement mode.
5. Comparing to JFET, MOSFETs are easier to fabricate.
6. MOSFET is very susceptible to overload voltage and needs special handling during installation. It gets damaged easily if it is not properly handled.
7. MOSFET has zero offset voltage. As it is a symmetrical device, the source and drain can be interchanged. These two properties are very useful in analog signal switching.
8. Special digital CMOS circuits are available which involve near-zero power dissipation and very low voltage and current requirements. This makes them most suitable for portable systems.

MOSFETs are widely used in digital VLSI circuits than JFETs because of their advantages.

7.13 HANDLING PRECAUTIONS FOR MOSFET

The MOSFET has the drawback of being very susceptible to overload voltage and may require special handling during installation. The MOSFET gets damaged easily if it is not properly handled. A very thin layer of SiO_2, between the gate and channel is damaged due to high voltage and even by static electricity. The static electricity may result from the sliding of a device in a plastic bag. If a person picks up the transistor by its case and brushes the gate against some grounded object, a large electrostatic discharge may result. In a relatively dry atmosphere, a static potential of 300 V is not uncommon on a person who has high resistance soles on his footwear.

MOSFETs are protected by a shorting ring that is wrapped around all four terminals during shipping and must remain in place until after the device is soldered into position. Prior to soldering, the technician should use a shorting strap to discharge his static electricity and make sure that the tip of the soldering iron is grounded. Once in circuit, there are usually low resistances present to prevent any excessive accumulation of electrostatic charge. However, the MOSFET should never be inserted into or removed from a circuit with the power ON. JFET is not subject to these restrictions, and even some MOSFETs have a built in gate protection known as "integral gate protection," a system built into the device to get around the problem of high voltage on the gate causing a puncturing of the oxide layer. The manner in which this is done is shown in the cross sectional view of Fig. 7.11. The symbol clearly shows that between each and the source is placed a back-to-back (or front-to-front) pair of diodes, which are built right into the P-type substrate.

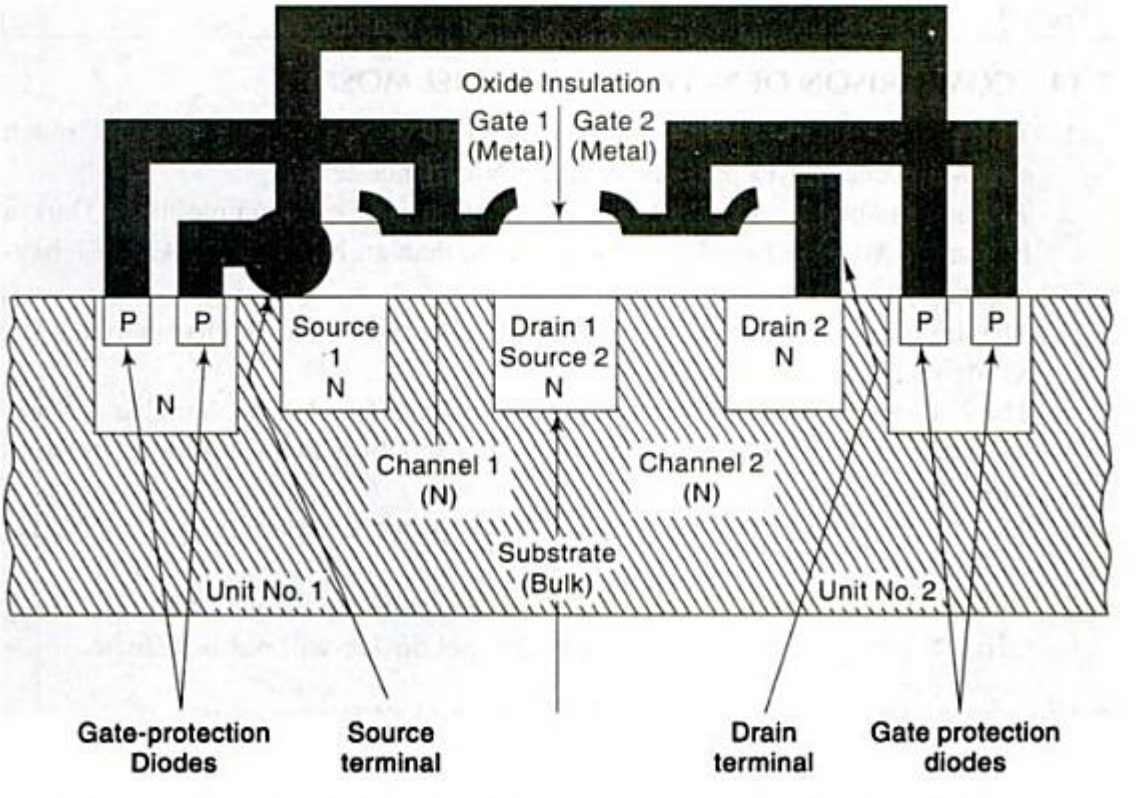


Fig. 7.11 Cross sectional view of the dual gate protected N-channel depletion type MOSFET

These diodes are designed so that if either gate exceeds +10 V typically, with respect to the source, the upper diode will conduct and the lower diode will breakdown by Zener effect, providing a shunt path for excessive charge from gate to source. Similarly, if the gate voltage exceeds –10 V, the lower diode conducts and the upper diode breaks down. The breakdown voltage from either gate to source is in a range such that normal application of signal will not be affected by the gate protection diodes. In addition to providing protection against preinstallation high static voltages, these diodes also guard against in-circuit transients.

Current–Voltage Relationships of the N-channel and P-channel MOSFETs

The current-voltage relationships of the N-channel and P-channel MOSFETs are given in Table 7.1.

Table 7.1 Current–voltage relationships of the N-MOSFET and P-MOSFET

N-Channel MOSFET	*P-Channel MOSFET*
Saturation region ($V_{DS} > V_{DS}$(sat)) $I_D = K_N(V_{GS} - V_{TN})^2$	Saturation region ($V_{SD} > V_{SD}$ (sat)) $I_D = K_P(V_{SG} + V_{TP})^2$
Non saturation region ($V_{DS} < V_{DS}$(sat)) $I_D = K_N[2(V_{GS} - V_{TN})V_{DS} - V_{DS}^2]$	Non saturation region ($V_{SD} < V_{SD}$(sat)) $I_D = K_P[2(V_{SG} + V_{TP})V_{SD} - V_{DS}^2]$
Transition point V_{DS}(sat) = $V_{GS} - V_{TN}$	Transition point V_{SD}(sat) = $V_{SG} + V_{TP}$
Enhancement mode $V_{TN} > 0$	Enhancement mode $V_{TP} < 0$
Depletion mode $V_{TN} < 0$	Depletion mode $V_{TP} > 0$

7.14 COMPARISON OF N- WITH P-CHANNEL MOSFETS

1. The P-channel enhancement MOSFET is very popular because it is much easier and cheaper to produce than the N-channel device.
2. The hole mobility is nearly 2.5 times lower than the electron mobility. Thus, a P-channel MOSFET occupies a larger area than an N-channel MOSFET having the same I_D rating.
3. The drain resistance of P-channel MOSFET is three times higher than that for an identical N-channel MOSFET.
4. The N-channel MOSFET has the higher packing density which makes it faster in switching applications due to the smaller junction areas and lower inherent capacitances.
5. The N-channel MOSFET is smaller for the same complexity than P-channel device.
6. Due to the positively charged contaminants, the N-channel MOSFET may turn ON prematurely, whereas the P-channel device will not be affected.

7.15 COMPARISON OF N- WITH P- CHANNEL FETS

1. In an N-channel JFET the current carriers are electrons, whereas the current carriers are holes in a P-channel JFET.
2. Mobility of electrons is large in N-channel JFET; mobility of holes is poor in P-channel JFET.

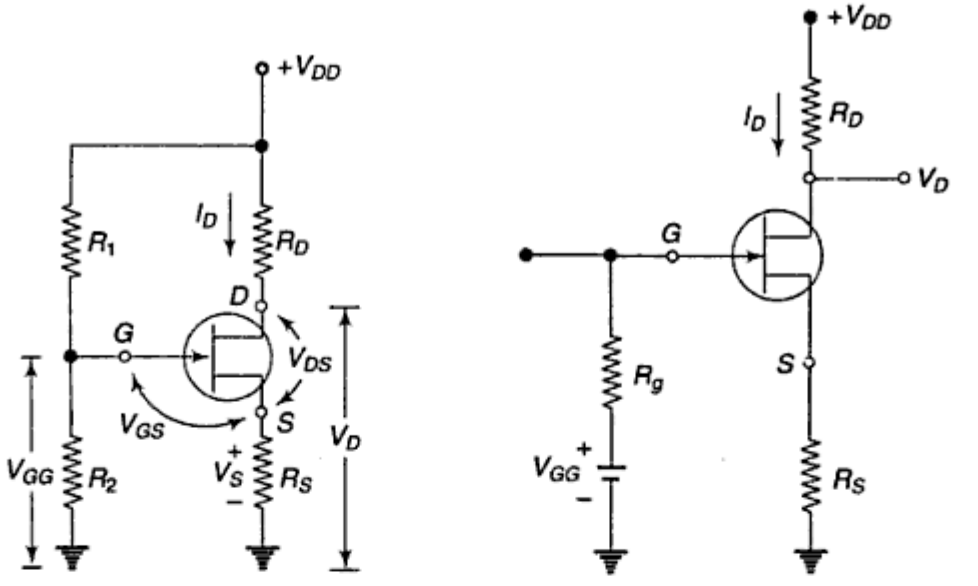


Fig. 7.13 (a) Voltage divider bias circuit, and (b) Thevenin's equivalent circuit

This is because, in BJT, $V_{BE} \approx 0.7$ V (silicon) with only minor variations from one transistor to another. But in a JFET, the V_{GS} can vary several volts from one JFET to another.

7.17 USE OF JFET AS VOLTAGE-VARIABLE RESISTOR

FET is operated in the constant-current portion of its output characteristics for the linear applications. In the region before pinch-off, where V_{DS} is small, the drain to source resistance r_d can be controlled by the bias voltage V_{GS}. The FET is useful as a *Voltage Variable Resistor (VVR)* or *Voltage Dependent Resistor (VDR).*

In JFET, the drain-to-source conductance $g_d = \dfrac{I_D}{V_{DS}}$ for small values of V_{DS}, which may also be expressed as,

$$g_d = g_{do}\left[1 - \left(\frac{V_{GS}}{V_P}\right)^{\frac{1}{2}}\right]$$

where g_{do} is the value of drain conductance when the bias voltage V_{GS} is zero.

The variation of the r_d with V_{GS} can be closely approximated by the empirical expression,

$$r_d = \frac{r_0}{1 - KV_{GS}}$$

where r_0 = drain resistance at zero gate bias, and K = a constant, dependent upon FET type.

Thus, small signal FET drain resistance r_d varies with applied gate voltage V_{GS} and FET acts like a variable passive resistor.

FET finds wide applications where VVR property is useful. For example, the VVR can be used in Automatic Gain Control (AGC) circuit of a multistage amplifier.

7.18 BIASING THE MOSFET

7.18.1 Biasing of Enhancement MOSFET

Figure 7.14 shows the drain-to-gate bias circuit for enhancement mode MOSFET. Here, the gate bias voltage is $V_{GS} = \left[\frac{R_1}{R_1 + R_f}\right] V_{DS}$. This circuit offers the d.c. stabilisation through the feedback resistor R_f. However, the input resistance is reduced because of Miller effect.

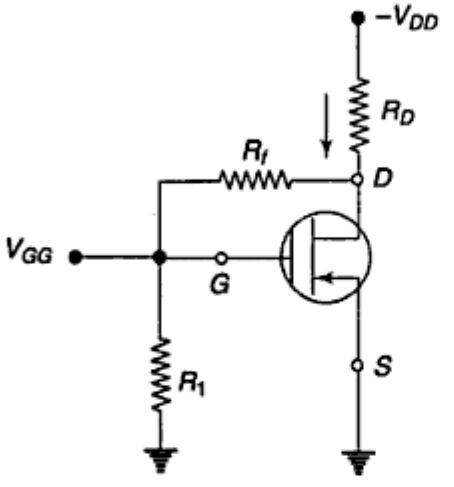


Fig. 7.14 Drain-to-gate bias circuit for enhancement MOSFET

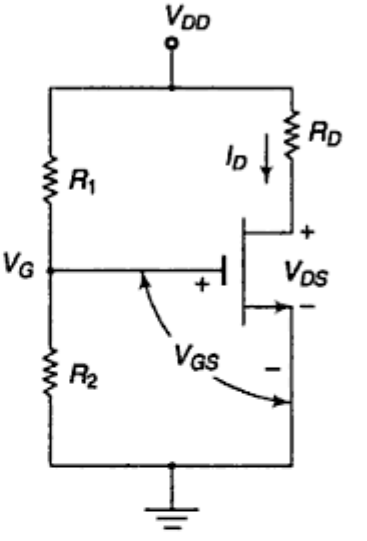


Fig. 7.15 N-channel enhancement mode MOSFET common source circuit with source resistor

Also, the voltage divider biasing technique given for JFET can be used for the enhancement MOSFET. Here, the d.c. stability is accomplished by the d.c. feedback through R_S.

But the self-bias technique given for JFET cannot be used for establishing an operating point for the enhancement MOSFET because the voltage drop across R_S is in a direction to reverse-bias the gate and it actually needs forward gate bias.

Figure 7.15 shows an N-channel enhancement mode MOSFET common source circuit with source resistor. The gate voltage is

$$V_G = V_{GS} = \left(\frac{R_2}{R_1 + R_2}\right)(V_{DD})$$

and the gate-to-source voltage is

$$V_{GS} = V_{DD} - V_G$$

Assuming that $V_{GS} > V_{TN}$ and the MOSFET is biased in the saturation region, the drain current is

$$I_D = K_N(V_{GS} - V_{TN})^2$$

Here the threshold voltage V_{TN} and conduction parameter K_N are functions of temperature

Substituting the given values, we get

$$I_D = 40 \times 10^{-3} \left[1 - \frac{(-5)}{(-10)}\right]^2 = 10 \text{ mA}$$

Therefore, $$R_S = \left|\frac{V_{DSQ}}{I_D}\right| = \frac{5}{10 \times 10^{-3}} = 500\ \Omega$$

Example 7.7 A JFET amplifier with a voltage divider biasing circuit, shown in Fig. 7.13, has the following parameters: $V_P = -2$ V, $I_{DSS} = 4$ mA, $R_D = 910\ \Omega$, $R_S = 3$ kΩ, $R_1 = 12$ MΩ, $R_2 = 8.57$ MΩ and $V_{DD} = 24$ V. Find the value of the drain current I_D at the operating point. Verify whether the FET will operate in the pinch-off region.

Solution: From Fig. 7.13, we obtain

$$V_{GG} = V_{DD} \frac{R_2}{R_1 + R_2} = 24 \frac{8.57 \times 10^6}{(12 + 8.57) \times 10^6} = 10 \text{ V}$$

We know that

$$I_D = I_{DSS} \left(1 - \frac{V_{GS}}{V_P}\right)^2$$

$$= I_{DSS} \left(1 - \frac{V_{GG} - I_D R_S}{V_P}\right), \text{ where } V_{GS} = V_{GG} - I_D R_S$$

Expressing I_D and I_{DSS} in mA, we have

$$I_D = 4 \times \left(1 - \frac{10 - I_D \times 3}{-2}\right)^2$$

i.e., $9I^2_D - 73I_D + 144 = 0$

Therefore, $I_D = 3.39$ mA or 4.72 mA

As $I_D = 4.72$ mA > 4 mA = I_{DSS}, this value is inappropriate. So, $I_{DQ} = 3.39$ mA is selected.

Therefore,

$$V_{GSQ} = V_{GG} - I_{DQ}R_S = 10 - (3.39 \times 10^{-3} \times 3 \times 10^3) = -0.17 \text{ V}$$

and $$V_{DSQ} = V_{DD} - I_{DQ}(R_D + R_S) = 24 - 3.39 \times 10^{-3} (0.91 + 3) \times 10^3$$
$$= 10.745 \text{ V}$$

Then $$V_{DGQ} = V_{DSQ} - V_{GQS} = 10.745 + 0.17 = 10.915 \text{ V}$$

which is greater than $|V_P| = 2$ V. Hence, the FET is in the pinch-off region.

Example 7.8 A voltage divider bias is provided to an N-channel JFET circuit as shown in Fig. 7.16. To establish $I_{DSS} = 10$ mA, $V_P = -3.5$ V, $R_1 + R_2 = 120$ kΩ, $I_D = 5$ mA and $V_{DS} = 5$ V, determine the values of R_1, R_2 and R_D.