



**WE TRIP THE LIGHT
FANTASTIC**

INSTITUTE OF PHYSICS

SERIES IN ASTRONOMY AND ASTROPHYSICS

DARK SKY, DARK MATTER

J M OVERDUIN
P S WESSON

IoP

INSTITUTE OF PHYSICS

SERIES IN ASTRONOMY AND ASTROPHYSICS

**SERIES EDITORS: M BIRKINSHAW,
M ELVIS AND J SILK**

The *Series in Astronomy and Astrophysics* includes books on all aspects of theoretical and experimental astronomy and astrophysics. Books in the series range in level from textbooks and handbooks to more advanced expositions of current research.

Olbers' paradox states that in an unbounded Universe governed by the standard laws of physics and populated by light sources of constant luminosity, the night sky should be ablaze with light. Obviously this is not so. However, the paradox does not lie in nature but in our understanding of physics. A Universe with a finite age, such as follows from big-bang theory, necessarily has galaxies of finite age. This means we can only see some of the galaxies in the Universe, which is the main reason why the night sky is dark. Just how dark can be calculated using the astrophysics of galaxies and stars, and the dynamics of relativistic cosmology.

We know from the dynamics of individual galaxies and clusters of galaxies that the majority of the matter which exerts gravitational forces is not detectable by conventional telescopes. This dark matter could have many forms, and candidates include various types of elementary particles as well as vacuum fluctuations, black holes and others. Most of these candidates are unstable to decay and produce photons. So dark matter does not only affect the dynamics of the Universe, but the intensity of intergalactic radiation as well. Conversely, we can use observations of background radiation to constrain the nature and density of dark matter. Thus does Olbers' problem gain new importance.

Drawing on the latest observational data as well as current theoretical ideas in cosmology, general relativity and particle physics, this book reviews the astrophysics of the night sky and the nature of the dark matter that makes up most of the Universe.

Contents:

The dark night sky; The modern resolution and energy; The modern resolution and spectra; The dark matter; The vacuum; Axions; Neutrinos; Supersymmetric weakly interacting particles; Black holes; Conclusions; Appendices: Bolometric intensity integrals; Dynamics with a decaying vacuum; Absorption by galactic hydrogen.

ISBN 0-7503-0684-X



9 780750 306843 >

Institute of Physics Publishing
Bristol and Philadelphia
www.iop.org

IoP

Dark Sky, Dark Matter

Series in Astronomy and Astrophysics

Series Editors: **M Birkinshaw**, University of Bristol, UK
M Elvis, Harvard–Smithsonian Center for Astrophysics, USA
J Silk, University of Oxford, UK

The Series in Astronomy and Astrophysics includes books on all aspects of theoretical and experimental astronomy and astrophysics. Books in the series range in level from textbooks and handbooks to more advanced expositions of current research.

Other books in the series

Dust in the Galactic Environment, 2nd Edition
D C B Whittet

An Introduction to the Science of Cosmology
D J Raine and E G Thomas

The Origin and Evolution of the Solar System
M M Woolfson

The Physics of the Interstellar Medium
J E Dyson and D A Williams

Dust and Chemistry in Astronomy
T J Millar and D A Williams (eds)

Observational Astrophysics
R E White (ed)

Stellar Astrophysics
R J Tayler (ed)

Forthcoming titles

The Physics of Interstellar Dust
E Krügel

Very High Energy Gamma Ray Astronomy
T Weekes

Series in Astronomy and Astrophysics

Dark Sky, Dark Matter

J M Overduin

*Universität Bonn, Bonn, Germany and
Waseda University, Tokyo, Japan*

P S Wesson

University of Waterloo, Waterloo, Canada

IOP

Institute of Physics Publishing
Bristol and Philadelphia

© IOP Publishing Ltd 2003

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publisher. Multiple copying is permitted in accordance with the terms of licences issued by the Copyright Licensing Agency under the terms of its agreement with Universities UK (UUK).

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

ISBN 0 7503 0684 X

Library of Congress Cataloging-in-Publication Data are available

Series Editors: **M Birkinshaw**, University of Bristol, UK
M Elvis, Harvard–Smithsonian Center for Astrophysics, USA
J Silk, University of Oxford, UK

Commissioning Editor: John Navas
Production Editor: Simon Laurenson
Production Control: Sarah Plenty
Cover Design: Victoria Le Billon
Marketing: Nicola Newey and Verity Cooke

Published by Institute of Physics Publishing, wholly owned by The Institute of Physics, London

Institute of Physics Publishing, Dirac House, Temple Back, Bristol BS1 6BE, UK

US Office: Institute of Physics Publishing, The Public Ledger Building, Suite 929, 150 South Independence Mall West, Philadelphia, PA 19106, USA

Typeset in L^AT_EX 2_ε by Text 2 Text, Torquay, Devon
Printed in the UK by MPG Books Ltd, Bodmin, Cornwall

Contents

Preface	ix
1 The dark night sky	1
1.1 Olbers' paradox	1
1.2 A short history of Olbers' paradox	2
1.3 The paradox now: stars, galaxies and Universe	5
1.4 The resolution: age versus expansion	9
1.5 The data: optical and otherwise	10
1.6 Conclusion	13
2 The modern resolution and energy	16
2.1 Big-bang cosmology	16
2.2 The bolometric background	16
2.3 From time to redshift	20
2.4 Matter and energy	22
2.5 The expansion rate	24
2.6 The static analogue	27
2.7 A quantitative resolution	29
2.8 Light at the end of the Universe?	35
3 The modern resolution and spectra	41
3.1 The spectral background	41
3.2 From bolometric to spectral intensity	42
3.3 The delta-function spectrum	44
3.4 Gaussian spectra	50
3.5 Blackbody spectra	51
3.6 Normal and starburst galaxies	54
3.7 Back to Olbers	61
4 The dark matter	66
4.1 From light to dark matter	66
4.2 The four elements of modern cosmology	67
4.3 Baryons	68
4.4 Cold dark matter	73
4.5 Neutrinos	76

4.6	Vacuum energy	78
4.7	The coincidental Universe	85
5	The vacuum	90
5.1	Vacuum decay	90
5.2	The variable cosmological ‘constant’	91
5.3	Energy density	95
5.4	Source regions and luminosity	101
5.5	Bolometric intensity	105
5.6	Spectral energy distribution	106
5.7	The microwave background	108
6	Axions	113
6.1	Light axions	113
6.2	Rest mass	114
6.3	Axion halos	117
6.4	Intensity	119
6.5	The infrared and optical backgrounds	120
7	Neutrinos	127
7.1	The decaying-neutrino hypothesis	127
7.2	Bound neutrinos	128
7.3	Luminosity	130
7.4	Free-streaming neutrinos	133
7.5	Intergalactic absorption	135
7.6	The ultraviolet background	139
8	Supersymmetric weakly interacting particles	146
8.1	The lightest supersymmetric particle	146
8.2	Neutralinos	147
8.3	Pair annihilation	149
8.4	One-loop decays	154
8.5	Tree-level decays	157
8.6	Gravitinos	161
8.7	The x-ray and γ -ray backgrounds	163
9	Black holes	170
9.1	Primordial black holes	170
9.2	Initial mass distribution	171
9.3	Evolution and number density	173
9.4	Cosmological density	175
9.5	Spectral energy distribution	177
9.6	Luminosity	179
9.7	Bolometric intensity	179
9.8	Spectral intensity	183
9.9	Higher-dimensional ‘black holes’	186

10 Conclusions	192
A Bolometric intensity integrals	196
A.1 Radiation-dominated models	196
A.2 Matter-dominated models	197
A.3 Vacuum-dominated models	199
B Dynamics with a decaying vacuum	202
B.1 Radiation-dominated regime	202
B.2 Matter-dominated regime	203
B.3 Vacuum-dominated regime	204
C Absorption by galactic hydrogen	206
Index	209

Preface

The darkness of the night sky awed our ancestors and fascinates modern astronomers. It is a fundamental problem: *why* is the night sky dark, and just *how* dark is it? This book examines these overlapping problems from the viewpoints of both history and modern cosmology.

Olbers' paradox is a vintage conundrum that was known to other thinkers prior to its formulation by the Prussian astronomer in 1823. Given that the Universe is unbounded, governed by the standard laws of physics, and populated by light sources of constant intensity, the simple cube law of volumes and numbers implies that the sky should be ablaze with light. Obviously this is not so. However, the paradox does not lie in nature but in our understanding of physics. A Universe with a finite age, such as follows from big-bang theory, necessarily has galaxies of finite age. There is therefore a kind of imaginary spherical surface around us which depends on the age and the speed of light. That is, we can only see some of the galaxies in the Universe, and this is the main reason *why* the night sky is dark. Just *how* dark can be calculated using the astrophysics of galaxies and their stars, and the dynamics of relativistic cosmology.

We know from the dynamics of individual galaxies and clusters of galaxies that the majority of the matter which exerts gravitational forces is not detectable by conventional telescopes. This dark matter could, in principle, have many forms, and candidates include various types of elementary particles as well as vacuum fluctuations, black holes and others. Most of these candidates are unstable to decay and produce photons. So dark matter does not only affect the dynamics of the Universe, but the intensity of intergalactic radiation as well. Conversely, we can use observations of background radiation to constrain the nature and density of dark matter. Thus does Olbers' problem gain new importance.

Modern cosmology is a cooperative endeavour, and we would like to acknowledge some of the people who have helped to shape our ideas about the Universe and what it may contain. It is a pleasure to thank H-J Fahr and K-I Maeda for hospitality at the Institut für Astrophysik und Extraterrestrische Forschung in Bonn and at Waseda University in Tokyo, where much of this book was written with support from the Alexander von Humboldt Foundation and the Japan Society for the Promotion of Science. Many of the results we will present on astrophysics, general relativity and particle physics have been

derived in collaboration with colleagues over a number of years. For sharing their expertise we are particularly indebted to S Bowyer, T Fukui, W Priester, S Seahra and R Stabell. Finally, our motivation for carrying out these calculations is based simply on a desire to understand the Universe. This motivation is infectious, and in this regard we thank especially the late F Hoyle and his colleague D Clayton. We hope, in turn, that this book will inspire and educate others who wonder about the dark night sky.

J M Overduin and P S Wesson

10 June 2002

Chapter 1

The dark night sky

1.1 Olbers' paradox

It is a fundamental observation, which almost anyone can make, that the night sky is dark. But why?

It might be thought that to a person standing on the side of the Earth that faces away from the Sun the sky must necessarily be dark. However, on even a moonless night there is a faint but perceptible light from stars. Consider a situation in which there are no stars in the line of sight; for example, the hypothetical case of someone living on a planet at the edge of the Milky Way, who looks out into intergalactic space. Even in this situation, the night sky would not be perfectly dark because there are many other luminous galaxies in the Universe. It is at this point that the fundamental nature of the problem of the dark night sky begins to become apparent. For modern cosmology implies that the population of galaxies is uniform in density (on average) and unbounded. Given these conditions, and assuming that conventional geometry holds, the number of galaxies visible increases as the cube of the distance. This means that along the line of sight, the galaxies become increasingly crowded as the distance increases. In fact, under the conditions formulated, any given line of sight must end on a remote galaxy. Thus, the view in all directions must ultimately be blocked by overlapping galaxies; and the light from these, propagating towards us from the distant reaches of the Universe, should make the night sky bright.

Clearly this does not happen in reality, and the reason why has been a subject of controversy among astronomers for centuries. In 1823, Olbers published a paper that drew attention to this problem [1], which has since come to be known as Olbers' paradox. However, the concept of a paradox within modern physics implies a failure in terms of one or more of reasoning, theory or observational data. Below, we will see how a scrutiny of all three of these has recently enabled us to resolve Olbers' paradox.



Figure 1.1. H W M Olbers (1758–1840).

1.2 A short history of Olbers' paradox

H W M Olbers was a German astronomer who was born in 1758 and died in 1840. During his lifetime he was mainly notable for his research on the solar system. Thus, a comet that returned in 1815 (and again in 1887) was called Olbers' comet. He was also involved in the search for a hypothetical missing planet in the gap between the orbits of Mars and Jupiter. An object (Ceres) was discovered on 1 January 1801 by G Piazzi, and its orbit was worked out by the great mathematician K F Gauss and found to lie in the expected region. However, an unexpected second object (Pallas) was discovered in March of 1802 by Olbers with an orbit at almost the same distance from the Sun. This was followed in later years by the finding of yet more objects, including a second (Vesta) by Olbers in March of 1807. In this way Olbers was associated with the discovery of what later came to be known as the **asteroid belt**.

The work for which Olbers is mainly remembered today was first published in English in 1826 under the title 'On the Transparency of Space' [2]. It was then apparently forgotten for a good many years, only to be recalled in more recent

times following the proposal of the **steady-state cosmology** by H Bondi, T Gold and F Hoyle in 1948. The steady-state theory has an infinite past, and so is very different from the big-bang theory, which has become accepted as the standard cosmology because observations support its basic premise of a finite past (see later). These two theories actually involve quite different approaches to Olbers' paradox, and it is now the general opinion that any cosmology that hopes to be taken seriously must have a satisfactory solution to the problem of the dark night sky. This attitude, and the reawakening of interest in Olbers' paper and its central problem, appears to date from a discussion in Bondi's classic book *Cosmology*, which first came out in 1952 [3].

Bondi's discussion listed several conditions that result in the paradox of a bright night sky. These included the assumptions that the average density and luminosity of sources do not vary in space or time, that the sources do not have large systematic motions, that space is Euclidean, and that the known laws of physics apply. This list has been criticized by S L Jaki [4] and others, who have objected that it is not equivalent to what Olbers wrote in his original article. This may be true, and it is certainly the case that what is called Olbers' paradox today exists in many different versions. There is also confusion of other types: some people call the author of the paradox Wilhelm Olbers and some Heinrich Olbers; while some date the original formulation of the problem to 1826 rather than 1823. Also, it has been pointed out by various historians that the problem was, in fact, appreciated in the two centuries preceding Olbers, notably by J Kepler, N Hartsoeker, E Halley and P L de Chéseaux [5, 6]. But whatever confused path it has taken through history, it has arrived in modern times as a list of conditions that leads to a contradiction with observation and is associated primarily with the name of Olbers.

Currently, there are several slightly different lists of assumptions in use that lead to the paradox of a bright night sky. But they have in common an implicit dependence on modern astrophysical knowledge. For example, there is seldom much discussion today of a subject implied by the title of Olbers' original article, namely absorption of light in space. This is primarily because observations of remote sources such as quasars place low limits on intergalactic matter. However, it should also be pointed out that even if there were large amounts of absorbing matter in space, the energy absorbed at one wavelength would eventually be re-emitted at another wavelength of what we nowadays know to be a continuous electromagnetic spectrum, so the problem would not be eliminated but merely shifted. Also, today we know that the sources of light in the paradox should not be taken to be single stars as assumed in historical times, but galaxies—'island universes' as they were once termed—each containing billions of stars. Nowadays it is rare to find any mention of avoiding the paradox by taking the sources to be non-uniformly distributed, since modern observations imply that the galaxies may be assumed to be uniformly distributed, at least if averaged over large enough distances. Furthermore, while space over large distances may be curved in accordance with Einstein's theory of general relativity, the curvature

actually does not figure directly in the expression for the intensity of intergalactic light in the case where the Universe is uniform; and even if this were not the case, space could hardly be so far from Euclidean or flat as to offer a gravitational kind of resolution of Olbers' paradox. In the same vein, no compelling evidence has emerged that the laws of physics are significantly different in remote parts of the Universe from those nearby, so that way of solving the problem is hardly ever mentioned now.

Thus, the most up-to-date version of Olbers' paradox, slimmed down to avoid mentioning what modern astrophysics tells us can be taken for granted, is as follows. In a Universe consisting of galaxies that are of *infinite age* and *static*, the accumulated light would be so intense as to make the night sky bright and not dark as observed.

In this form, it is obvious that one or both of the remaining assumptions are wrong. It is actually the case, as will be seen later, that both are wrong. The galaxies have a **finite age**, which means that the amount of light which they have emitted into space is limited. And the galaxies are receding from each other due to **cosmic expansion**, which means that their light has been reduced in intensity. But this is merely to scratch the surface of a solution which, to be understood in depth, requires a critical look at several areas of modern astrophysics and cosmology. In particular, it is superficial to know only that the two assumptions in this version of the paradox are wrong. As in other scientific studies, this just leads to the important question of quantification: which assumption is more wrong? Or more precisely: *what is the exact relative importance of the two assumptions for the darkness of the night sky?*

The person often associated with this question is E R Harrison. In 1964 he discussed the timescales associated with stars and the non-static nature of the Universe, and concluded that the latter was not of major importance for the resolution of Olbers' paradox [7]. He published several more articles thereafter, and in 1981 in a popular book he restated his view that, in a Universe consisting of luminous sources, the fact that they have finite ages is far more important than whether they are static or not [8]. But while Harrison's view was supported by a few notable cosmologists, it was not universally accepted. This was mainly because, following Bondi's earlier discussion, the motions of the galaxies had somehow become accepted as *the* factor in resolving Olbers' paradox. Also, Harrison's technical papers on the subject were not easy to follow, and this delayed a proper evaluation of the age factor. Thus, while some headway was made in the 1960s, 1970s and early 1980s towards an objective treatment of the paradox, controversy and confusion continued to prevail. A survey of modern astronomy textbooks carried out in 1987 showed how divided opinion was. About 30% agreed with Harrison in saying that the failure of the infinite-age assumption was most important. While about 50% attributed the resolution of the paradox to the failure of both the infinite-age assumption and the static assumption, though without a proper assessment of which was the more important. And about 20% of the textbooks surveyed mentioned only the failure of the assumption that galaxies

are static. This state of confusion prompted the appearance in 1987 of an article by Wesson *et al* [9] that introduced a new method, to be discussed later, of tackling the problem of the dark night sky. Using this, they gave an exact quantitative assessment of the relative importance of the age and motion factors, and thereby resolved Olbers' paradox in a definitive fashion.

To complete this section, it is interesting to note that while Olbers' paradox has through history been a subject of mainly theoretical attention, in recent years it has also entered an observational phase that is likely to become of increasing importance in the future. Observational attempts to detect the faint intergalactic light have in the past produced useful but non-definitive data. However, recent advances in telescope technology have brought us to the threshold of detection [10]. This prospect is very interesting, since actual measurement would provide an opportunity to check theoretical estimates of the light's intensity, and thereby the underlying principles of astrophysics and cosmology.

1.3 The paradox now: stars, galaxies and Universe

We saw in the preceding section that Olbers' paradox has evolved in content and likely resolution since it was propounded by its author in 1823. In recent years, attention has focused on two ways in which a bright night sky can be avoided. First, the ages of galaxies may not be infinite as originally assumed but finite. Second, galaxies may not be static but instead have large systematic motions. Modern astrophysics tells us that both of these things are the case. But to understand how they affect the darkness of the night sky we need to look more closely at our current picture of the Universe.

Stars are the main radiation-producing bodies in the Universe. An average star converts elements of low atomic number (primarily hydrogen) to elements of higher atomic number by means of nuclear fusion reactions. The process by which there is a gradual change in the chemical makeup of a star during its life, from 'light' to 'heavy' elements, is called **nucleosynthesis**. Combining theoretical data on nucleosynthesis with observational data on the element abundances in a given star (gathered mainly from spectroscopy) allows its age to be determined. A by-product of the nuclear reactions in a star is energy, which is emitted largely in the form of optical radiation. The energy output of a typical star is colossal by conventional standards due to its very large mass, but the rate of energy production per unit mass is a parameter with a more reasonable value. As a rough average value for the energy output per unit time per unit mass of luminous matter in the Universe, we shall take $0.5 \text{ erg s}^{-1} \text{ g}^{-1}$. This is about one-quarter of the value for the matter in the Sun, which reflects the fact that most stars (as seen for instance on a Hertzsprung–Russell diagram of nearby stars) are intrinsically less luminous than ours. As we will show later, it is the relative smallness of this number that is one of the reasons the night sky appears dark.

Galaxies are the main mass concentrations in the Universe, each consisting

primarily of a clump of numerous stars bound together by their own gravity. There is a considerable range in the masses, luminosities and shapes of galaxies. The shapes are used to classify galaxies into spirals, ellipticals and irregulars. The masses and sizes are uncertain, because some galaxies may have halos of unseen **dark matter** that are considerably larger and more massive than the parts of the galaxies we can see. This is especially likely to be true of spirals like the Milky Way. However, the study of the light due to galaxies involves primarily luminous as opposed to dark matter, so we will defer discussion of dark matter to later chapters. An average value for the density of luminous matter in the Universe can be obtained if we assume that the galaxies are distributed uniformly in space. Galaxies actually tend to clump together because of their mutual gravitational interactions. But such clusters appear to have only limited sizes, with no sure observational evidence for very large ones. So if we consider large distances only, the distribution of galaxies can indeed be assumed to be uniform. Given this, we can describe the distribution of *luminous* matter in the Universe in terms of one parameter, namely the average density. Observational evidence indicates that this quantity takes a value of about $4 \times 10^{-32} \text{ g cm}^{-3}$. Like the stellar energy output per unit mass, this is a very small number by everyday standards, and is one of the main reasons the night sky appears dark.

Cosmology is the study of the Universe in the large; and because it has more observational support than its rivals, we will presume the validity of the **big-bang theory** of cosmology. This theory is based primarily on two assumptions. The first, which was discussed earlier, is that the matter distribution over sufficiently large distances can be taken to be uniform. It should be noted that the word ‘uniform’ implies that there is no identifiable centre or boundary to the matter distribution, so models of the Universe that emerge from this theory must in some sense be unbounded or infinite in extent. The second assumption on which the big-bang theory is based is that the Universe is expanding. This means that the galaxies are receding from each other (remember: no special or central one), and in particular they are receding from the Milky Way. We infer that this is happening because the light we receive from remote galaxies is redshifted, and the most natural way to account for this is in terms of a Doppler shift from a receding source. Thus, just as the sound waves from the whistle of a receding train are received with longer wavelengths, producing a lower pitch, so are the light waves from the stars of a receding galaxy received with longer wavelengths, producing a redshift. (Redshift in the big-bang theory is really more subtle than this, and involves expansion as distinct from velocity, but the two words are commonly taken to imply each other and this will be done here also.) Redshifts and recession speeds increase with distance, and it can be shown that the precise relationship—called **Hubble’s law**—is consistent with the assumption of uniformity. Insofar as both of the prime assumptions on which the big-bang theory is based are derived from observations, the theory is fairly well-founded.

If these assumptions are plugged into Einstein’s theory of general relativity, which is the appropriate theory of gravity and accelerations for large distances

and masses, then the evolution of the Universe can be worked out in detail. There are actually several different kinds of evolution allowed by the observational input data, and these are often called standard cosmological models. For each of these, the distances between galaxies change somewhat differently with time. But all of the standard models start in a big bang or state of infinite density at time zero. Note that the big bang is not localized at a special point in space, a concept that would violate the principle that the standard models are uniform with no preferred spot. Rather, *the big bang is like an explosion that occurs everywhere at the same time*. According to the theory, the real Universe just after the big bang must have been very hot, and we find evidence of this today in the cooled-down radiation of the 3 K **cosmic microwave background** or CMB (which should not be confused with galactic radiation). Galaxies are thought to have condensed out of the cooling matter not long after the big bang. Thus, the time that has elapsed since the big bang and the age of the galaxies are expected to be approximately equal. Insofar as the big-bang theory implies that galaxies condensed out at one fairly well-defined epoch, it agrees with the observation that many galaxies have nearly the same age. The age of the galaxies inferred indirectly (by matching their motions to the standard models) agrees reasonably well with the age of the galaxies inferred directly (by dating their oldest stars using nucleosynthesis). A typical nucleosynthesis value is 16×10^9 years, or 16 Gyr [11]. This number must clearly play a crucial role in the subject of Olbers' paradox. But while it is large, it will be seen later that it cannot offset the influence of the other two small numbers introduced earlier.

The contents of the three preceding paragraphs may be summed up thus: stars in galaxies produce light, but the galaxies have only been in existence for a time of order 16 Gyr and are also expanding away from each other. How do these last-mentioned factors, the finite age of the galaxies and the expansion of the Universe, reduce what would otherwise be a blaze of intergalactic light?

The finite age of the galaxies implies directly that they have not had time to populate the vast reaches of intergalactic space with enough photons to make it bright. (We will assume there were negligible numbers of optical photons around at the time galaxies and their stars formed, so the intensity of intergalactic light started at essentially zero.) This is the simple way to understand this factor. But there is another way which, while conceptually more difficult to appreciate, has the advantage of allowing a geometrical argument against the paradox of the bright night sky on a par with the one used originally by Olbers. The crux of this argument is **light-travel time**. To appreciate the importance of this, consider what happens when we look at a nearby galaxy, like the one in Andromeda. This is 'only' about 2 million light-years away, but even so the finite speed of light implies that it takes about 2 million years for photons from Andromeda to reach us, which means that we see it as it was about 2 million years ago. For galaxies further away, the light-travel time is even larger. Galaxies at very great distances can be seen, in principle, as they were around the time of their formation about 16 Gyr ago. So, there is an imaginary spherical surface about us at origin, with

a radius of order 16×10^9 light-years (or 16 Gly for short), where galaxies can in principle be seen as they were at formation. (In practice it is very difficult to observe such galaxies because they are very remote and thus appear very dim. Also the cosmological distance associated with a light-travel time of 16 Gyr is actually larger than 16 Gly, because the Universe has expanded in the meantime. These points do not alter the argument in a qualitative way.) The light of our night sky comes only from the limited number of galaxies within this imaginary surface, because if we could peer further out we would only see the presumably non-luminous material from which they formed.

Thus, Olbers' paradox is avoided because, contrary to what he thought, we do not see arbitrarily large numbers of sources as we look further away. Instead, we effectively run out of galaxies to see at a distance of order 16 Gly. This argument is logical, but conceptually difficult to accept for many people because it seems to imply that we are necessarily talking about a Universe that is finite in extent. However, this is not so. If we could by some miracle take an instantaneous snapshot of all of the Universe as it is 'now', without taking into account the lag associated with the light-travel time, we would see many more galaxies; and depending on the kind of Universe we are living in, these could be arbitrarily large in number and stretch to arbitrarily great distances. Thus, if we could see all of the Universe as it is 'now', it might well be infinite in extent and contain an arbitrarily large number of galaxies, and it would certainly be unbounded. (It is possible that it could contain a large but finite number of galaxies, because space according to general relativity can curve around and close up on itself, but such a Universe would still have no boundaries.) However, we do not see to arbitrarily great distances because light travels with finite speed, and this causes our view of the Universe to be of things as they were in the past. Galaxies can only be seen out to a distance corresponding to the time when they formed. We can sum up the finite-age factor by saying that it directly limits the number of photons that have been pumped into space by galaxies, *or* it creates an imaginary surface that limits the portion of the Universe from which we receive photons emitted by galaxies. On either interpretation, it is clear that intergalactic space and the night sky will not be arbitrarily bright.

The expansion of the Universe implies two things that both reduce the intensity of intergalactic light and are relatively easy to understand. Firstly, as the galaxies recede from each other and the volume of intergalactic space increases, the number of photons per unit volume and therefore the intensity of the light decreases. Secondly, as the galaxies recede the photons they emit are redshifted, so by **Planck's law** their intrinsic energies decrease, as does the intensity of the light. These things combine to reduce the brightness of intergalactic space and the night sky.

The age and expansion factors described here have been much discussed. The first has often been misunderstood because of the conceptual difficulty involved in understanding the light-travel-time effect. The second, perhaps because of this, has often been quoted in the literature as if it was the major or

only factor in avoiding Olbers' paradox. But before this problem can be laid to rest, a quantitative answer has to be given to the following question: Which of the two factors, the finite lifetime of the galaxies or the expansion of the Universe, is the more important in accounting for the dark night sky?

1.4 The resolution: age versus expansion

Finding the relative sizes of these factors is not easy: a proper calculation involves the heavy mathematics of Einstein's theory of general relativity, wherein the two factors are intertwined and deeply embedded. Later, a new method will be outlined that can deal with this complexity. But for the present, to get some idea of the kinds of numbers involved in the problem, let us do a simpler calculation.

From our discussion of stars, galaxies and cosmology, we are in possession of three important numbers. These are: the average rate of energy production per unit mass of luminous matter ($0.5 \text{ erg s}^{-1} \text{ g}^{-1}$); the average density of luminous matter ($4 \times 10^{-32} \text{ g cm}^{-3}$); and the age of the galaxies ($16 \text{ Gyr} = 5 \times 10^{17} \text{ s}$ approximately). The product of these three numbers is $1 \times 10^{-14} \text{ erg cm}^{-3}$, which is an energy density. If we multiply this by the speed of light ($3 \times 10^{10} \text{ cm s}^{-1}$) we obtain $3 \times 10^{-4} \text{ erg s}^{-1} \text{ cm}^{-2}$, which is an intensity. Dimensional analysis suggests that this number must be related to the intensity of intergalactic light. It can be shown [12] that it is actually equal to the intensity of light in a static Universe where relativistic effects are negligible.

The preceding calculation is useful because it shows that even if the Universe were not expanding the intensity of intergalactic light would still be quite low, of order 10^{-4} or $10^{-3} \text{ erg s}^{-1} \text{ cm}^{-2}$. (This result is quasi-Newtonian, and inaccurate mainly because it neglects the expansion.) In other words, since expansion can only reduce the intensity, the preceding calculation provides us with an upper limit for the brightness of intergalactic space. This upper limit can be better appreciated via the following comparison. Consider a 100 W light bulb hanging in the centre of an average-sized room whose walls, floor and ceiling have a summed area of 100 m^2 (10^6 cm^2). The intensity of illumination for this room, assuming all of the 10^9 erg s^{-1} power of the bulb is converted to light, is $10^3 \text{ erg s}^{-1} \text{ cm}^{-2}$. By comparison with the level of illumination in the average domestic situation, the light due to galaxies is at least a million times fainter.

From this comparison we can return to the main question concerning the relative importance of the age and expansion factors. How can we find out which is the more important in our Universe, in a way that is accurate and takes into account the complexities of Einsteinian cosmology?

If we could stop the expansion of the Universe, while keeping the properties of the galaxies (and especially their age) unchanged, we could work out the intensity of intergalactic light in such a static Universe, knowing that the finite age of the galaxies was the determining factor. Then we could allow the expansion to resume, knowing that the intensity of intergalactic light was determined by

both the age factor and the expansion factor. If we thereafter formed the ratio of intensities, with and without expansion, we would be able to state not only that it is less than unity, but also its exact value. If the ratio was small compared to unity, we would be able to conclude that expansion is the dominant factor. Whereas, if the ratio was of order unity, we would be able to conclude that expansion is not the dominant factor.

The procedure just outlined can be carried out by using models. If we were sure of the parameters that describe the real Universe, it would only be necessary to consider one model, and the calculation leading to the ratio of intensities with and without expansion could be carried out by hand. Unfortunately, we are not sure of the parameters that describe the real Universe, so it is necessary to consider many models and leave the calculations of the ratios of intensities to a computer. We will describe the detailed calculations—which have wide applicability—in the next chapter. But they have a consensus, which is that the ratio of intensities with and without expansion is approximately 0.5. In other words, the intensity of light due to galaxies in an expanding Universe is reduced from that in an equivalent static Universe by about a factor of two.

This conclusion effectively resolves Olbers' paradox once and for all. We previously saw that the intensity of intergalactic light in a static Universe, where only the finite age of the galaxies is significant, is of the order 10^{-4} or 10^{-3} $\text{erg s}^{-1} \text{cm}^{-2}$. We now see that the intensity in an expanding Universe with equivalent properties is merely reduced by half. If the order of magnitude of the intensity is determined by the finite age of the galaxies, and only modified by a number of order unity by the expansion of the Universe, then clearly we are justified in saying that the former factor is the more important. This agrees with previous arguments by a few respected cosmologists (notably E R Harrison), but it disagrees with statements in several textbooks. The method just discussed leaves no doubt about the true situation, however. The reason why the night sky is dark has mainly to do with the finite age of the galaxies, not the expansion of the Universe.

1.5 The data: optical and otherwise

Olbers' paradox has traditionally been discussed in terms of *light*. But of course this is only the optical, or human-eye-receiving, band of an infinitely wide electromagnetic spectrum. Due to absorption by the Earth's atmosphere, data are easy or difficult to obtain by ground-based observations, depending on the waveband.

In recent times, we have been able to obtain better data by using rocket-borne instruments and artificial satellites. However, technology progresses more slowly than insight. Thus we are in the position that observations provide mainly constraints rather than proofs of what we have discussed earlier. Specifically, we have good constraints in the optical band on intergalactic light. (This is

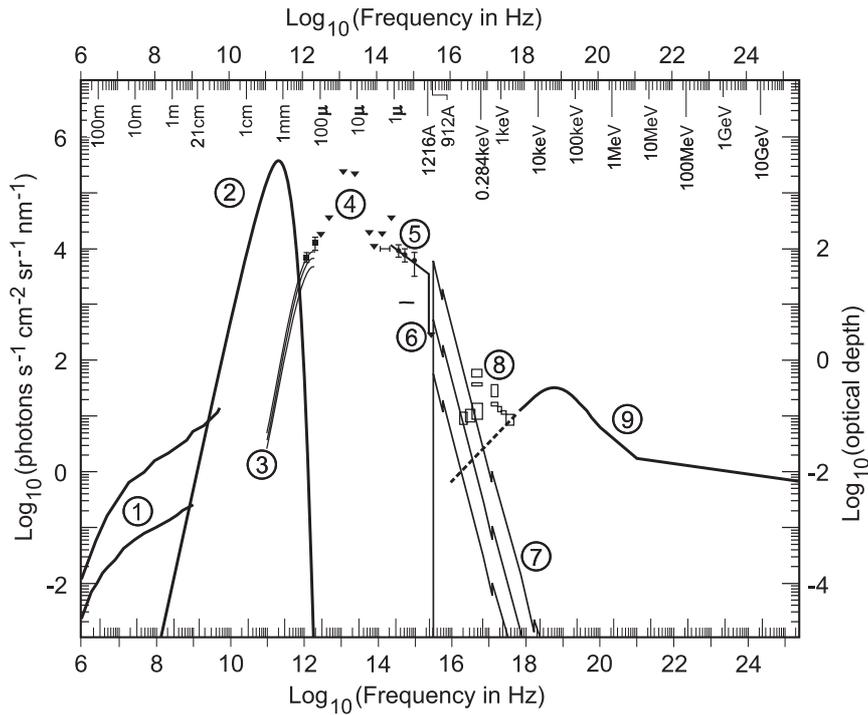


Figure 1.2. The full spectrum of the diffuse background radiation reaching our galaxy (figure courtesy of R C Henry [13]). This is a compilation of observational measurements and upper limits in wavebands from ① radio to ② microwave, ③④ infrared, ⑤ optical, ⑥⑦ ultraviolet, ⑧ x-ray and ⑨ γ -ray regions. Immense amounts of work, both theoretical and experimental, have gone into understanding each waveband. We will return to these in the chapters that follow, focusing on specific parts of the spectrum and comparing them to the predicted background intensities from galaxies and other sources of radiation.

commonly referred to in the literature as the **extragalactic background light** or EBL, meaning that we exclude primarily the light from nearby stars; and we have recently obtained improved data on this from the Hubble Space Telescope.) On the short-wavelength side of the optical, we have good data in the ultraviolet (UV), x-ray and γ -ray bands. On the long-wavelength side, we have usable data in the infrared (IR) band, and excellent data in the microwave band, beginning with the Cosmic Background Explorer (COBE). All this information is summarized graphically in figure 1.2. Here, we wish to briefly describe what techniques are used to study the optical band, and note how sharply the Olbers problem is constrained by data in other bands.

Several attempts have been made to detect the Olbers light or EBL using

large, ground-based telescopes. The idea behind such attempts is simple enough: if a telescope is pointed towards an area of the sky where there are no nearby sources, the light from the many remote galaxies in the field of view should combine to make a considerable flux, even if individual galaxies cannot be distinguished because their images are too faint. However, there is a major problem to contend with: the light from remote galaxies is swamped by other kinds of light that also enter the telescope. One of these is the light from the stars of the Milky Way. This has to be discounted, as does light from other non-galactic sources. The procedure to do this is complicated, but it is essential if one is to isolate the light from external galaxies. What was probably the most thorough ground-based search for the intergalactic light, both in terms of observations and the correction procedure just mentioned, was made by R R Dube, W C Wickes and D T Wilkinson, who used the Number 1 telescope at Kitt Peak in Arizona, USA [14]. Unfortunately, after their observations had been corrected for light from stars in the Milky Way and other places, they found that there was actually nothing measurable left. (This means that the integrated light from many remote but unresolved galaxies lying in the field of view of the telescope was too weak to be identified by the detection equipment used.) Such a null result can still be useful, though, because it provides an upper limit to the intensity of intergalactic light that can be employed to constrain and check cosmological theory. Numerically, the upper limit set by the three scientists is $3 \times 10^{-4} \text{ erg s}^{-1} \text{ cm}^{-2}$ approximately. It is doubtful if this can be improved on, or a positive identification made, using ground-based optical telescopes. Other techniques, however, provide more hope. One of these is to use the Hubble telescope in space, the data from which can be combined with other ground-based observations. This technique has been used by R Bernstein and co-workers [10]. As mentioned earlier, this method may provide the first concrete detection of the Olbers light in the optical band, but it is still in development.

We can also make observations from space in the *infrared* band. It is not difficult to understand the logic behind this. The intergalactic radiation field consists not only of optical photons from galaxies that are relatively nearby and at intermediate distances, but also includes photons from galaxies far away that have been redshifted to considerably longer infrared wavelengths by the expansion of the Universe. Just what part of the total energy of the field is at infrared wavelengths is hard to say, because it depends on various poorly known cosmological factors such as the formation epoch of galaxies. But it is probably significant. (It can be remarked in passing that theoretical estimates of the intensity of the radiation field due to galaxies are often bolometric ones, integrated over all wavelengths; whereas observations are made at one or a few wavelengths.) In addition to photons that have been redshifted into the infrared, there could also be photons of this type that were re-emitted from dust associated with young distant galaxies. Thus, while we must certainly receive a lot of optical photons from galaxies that are not too far away, we may also receive many infrared photons from galaxies that are remote. This situation, wherein the flux

of photons we receive at the Earth is a combination due to various mechanisms, is common to all wavebands.

The CMB is better understood than the infrared background. The intergalactic radiation field at wavelengths of order 1 cm is known from both ground and satellite observations to be highly isotropic and very cold. To be specific, direction-dependent fluctuations are known from the COBE satellite to be only of order 1 in 10^5 , and the mean temperature is close to 2.7 K (or about 3 degrees above absolute zero). These data are commonly taken to mean that this field was produced in the fireball that followed the big bang, and has been cooled by the expansion of the Universe. Irrespective of its origin, the intensity of the CMB provides a kind of cosmological baseline that severely restricts contributions from other sources.

On the short-wavelength side of the optical, we can use the data available to constrain the Olbers radiation and thus mechanisms that produce it. While we have yet to delve into the murky waters of dark matter, there is ample evidence from gravitational effects that we only see a minor fraction of the matter in the Universe. A well-known candidate for the rest is neutrinos. These particles are uncharged but ubiquitous, so even if they had a small mass they could, in principle, dictate the dynamical evolution of the cosmos. However, the most-discussed model for massive neutrinos, due to D W Sciama, predicts that they should decay. This would produce an Olbers field at ultraviolet wavelengths. Data on the intergalactic radiation field at wavelengths of order 1000 Å constrain this model, and enable us to work backward and set stringent constraints on neutrinos. The same argument, using data in other wavebands, applies to axions (neutral particles that some theories of particle physics predict should have been produced copiously in the big bang), WIMPs (weakly-interacting massive particles, for which there is evidence from particle accelerators), black holes (which decay by a process known as Hawking evaporation) and even the decay of the vacuum itself (allowed by quantum field theory). In short, dark matter is not black: while an individual particle may only produce an amount of energy that is small, the summed contribution to the intergalactic radiation field may be significant in a Universe that is vast.

1.6 Conclusion

The traditional paradox named after Olbers can be stated in modern language as follows: ‘in a Universe consisting of galaxies that are of infinite age and static, the accumulated light would be so intense as to make the night sky bright and not dark as observed’. We now know that both of the assumptions remaining in this formulation are wrong. However, the ways in which the age and expansion factors lead to a resolution of the problem are different.

- The age of the galaxies is finite, so the amount of light they have pumped into intergalactic space has been limited. This, in combination with the

relatively low rate of energy production and vast distances between galaxies, has resulted in a low intensity of intergalactic light. An alternative, though more subtle, way of regarding the age factor is as a geometrical limitation. The fact that the galaxies formed a finite time ago means that we can only see them out to a distance which is (roughly) of the order of their age multiplied by the speed of light. This implies that there is an imaginary surface about us which effectively delimits that part of the (unbounded) Universe from which we presently receive galactic light. Since we only receive light from a finite number of them, the light due to galaxies is limited in intensity.

- The Universe is expanding, which means that the galaxies are receding from each other. This causes the volume of intergalactic space to increase, and the energy density and intensity of intergalactic light to decrease. Also, the light emitted by galaxies is redshifted by the cosmological version of the Doppler effect. And since by Planck's law redder photons have less energy, the intensity of intergalactic light suffers a further decrease.

Of these two factors, the former is more important because it sets the order of magnitude of intergalactic light, while the second only reduces it further by about a half.

In later chapters, we will lay the calculational foundation for this conclusion, and then proceed to see how modern data on the darkness of the night sky can be applied to the problems of modern astrophysics. Of these problems, the largest is that we do not see all of the matter in the Universe and do not know its composition. However, the prime candidates involve a slow decay which feeds photons into intergalactic space. There is an intimate connection between the dark sky and dark matter.

References

- [1] Olbers W 1823 *Berliner Astronomisches Jahrbuch für das Jahr 1826* (Berlin: C F E Späthen) p 110
- [2] Olbers W 1826 *Edinburgh New Phil. J.* **1** 141
- [3] Bondi H 1952 *Cosmology* (Cambridge: Cambridge University Press)
- [4] Jaki S L 1967 *Am. J. Phys.* **35** 200
- [5] Hoskin M 1997 *Cambridge Illustrated History of Astronomy* (Cambridge: Cambridge University Press) pp 202–7
- [6] Jaki S L 2001 *The Paradox of Olbers' Paradox* 2nd edn (Pinckney, MI: Real View Books) pp 1–94
- [7] Harrison E R 1964 *Nature* **204** 271
- [8] Harrison E R 1981 *Cosmology, the Science of the Universe* (Cambridge: Cambridge University Press)
- [9] Wesson P S, Valle K and Stabell R 1987 *Astrophys. J.* **317** 601
- [10] Bernstein R A 1999 *The Low Surface Brightness Universe* (Astronomical Society of the Pacific Conference Series, Volume 170) ed J I Davies, C Impey and S Phillipps (San Francisco, CA: ASP) p 341

- [11] Vandenberg D A, Bolte M and Stetson P B 1996 *Ann. Rev. Astron. Astrophys.* **34** 461
- [12] Wesson P S 1986 *Sp. Sci. Rev.* **44** 169
- [13] Henry R C 1999 *Astrophys. J.* **516** L49
- [14] Dube R R, Wickes W C and Wilkinson D T 1979 *Astrophys. J.* **232** 333

Chapter 2

The modern resolution and energy

2.1 Big-bang cosmology

The best theory of the Universe we have is based on the assumptions that the galaxies are distributed approximately uniformly; that they are increasing their separations from one another; and that matter came out of a singularity. The best theory of gravity we have is Einstein's general relativity. It is justifiable to neglect the other interactions or forces known to modern physics in many astronomical problems: electromagnetism is long-range (like gravity) but there is no evidence of net electrical charges on astronomical bodies, so its effects are localized; the weak interaction, whereby a neutron changes to a proton and an electron, is the basis of nuclear fission but is not of astrophysical importance; the strong interaction figures crucially in the process whereby the Sun produces its energy by fusion, but the stars and galaxies can be treated as point sources of radiation for many purposes in cosmology. Thus, of the four known interactions, gravity rules in the Universe at large.

It would be philosophically pleasing, of course, to have a unified field theory of all the (≥ 4) forces. Such theories have been reviewed from an astronomical standpoint in a recent book [1], and we will have need in later chapters to 'import' the physics of particles into the physics of the Universe. Here, however, we will work with gravity alone, and specifically with general relativity, which has been amply verified by Solar System and other tests [2, 3]. Our aim in this chapter will be to show how Olbers' paradox is resolved in a modern context using Einstein's physics.

2.2 The bolometric background

At the heart of Einstein's general relativity lie the field equations

$$\mathcal{R}_{\mu\nu} - \frac{1}{2}\mathcal{R}g_{\mu\nu} - \Lambda g_{\mu\nu} = -\frac{8\pi G}{c^4}\mathcal{T}_{\mu\nu}. \quad (2.1)$$

The left-hand side of these equations describes the geometrical structure of spacetime, while the right-hand side describes the matter and energy content of the Universe. The **metric tensor** $g_{\mu\nu}$ relates distance and time via $ds^2 = -g_{\mu\nu} dx^\mu dx^\nu$. For the latter we will adopt spherical polar coordinates $x^\mu = (ct, r, \theta, \phi)$, where c is the speed of light, t is cosmic time and r is coordinate distance. The Ricci tensor $\mathcal{R}_{\mu\nu}$ and curvature scalar \mathcal{R} are functions of $g_{\mu\nu}$ and its derivatives. $\mathcal{T}_{\mu\nu}$ is known as the **energy–momentum tensor** while G , c and Λ are constants. About the value of Λ , in particular, we will say more in chapters 4 and 5. Our sign conventions throughout this book are the same as those of Weinberg [3].

Observations on the largest scales indicate that the Universe is both isotropic (similar in all directions) and homogeneous (similar at all places) along slices of constant time. If spatial isotropy and homogeneity are assumed, then the metric tensor is of **Robertson–Walker** form:

$$ds^2 = c^2 dt^2 - R^2(t) \left[\frac{dr^2}{(1 - kr^2)} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \right]. \quad (2.2)$$

Here $k(= \pm 1 \text{ or } 0)$ is the curvature constant and the **cosmological scale factor** $R = R(t)$ measures the change in distance between comoving objects due to expansion. Equation (2.2) tells us how to relate distances (integrals over $dr, d\theta, d\phi$) and times (integrals over dt) in a Universe whose geometry fluctuates in a manner which is driven by its matter and energy content, and described mathematically by the function $R(t)$.

Let us now consider the problem of adding up the light from all the galaxies in the Universe in such a way as to arrive at their combined energy density as received by us in the Milky Way. We begin by considering a single galaxy at coordinate distance r whose **luminosity**, or rate of energy emission per unit time, is given by $L(t)$. When it reaches us, this energy has been spread over an area

$$A = \int dA = \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} [R_0 r d\theta][R_0 r \sin\theta d\phi] = 4\pi R_0^2 r^2 \quad (2.3)$$

where we use the metric (2.2) to obtain the area element dA at the present time $t = t_0$ (figure 2.1). The subscript ‘0’ here and elsewhere denotes quantities taken at this time, so $R_0 \equiv R(t_0)$ is the present value of the cosmological scale factor.

The **intensity**, or energy flux per unit area, reaching us from this galaxy is given by

$$dQ_g = \left[\frac{R(t)}{R_0} \right]^2 \frac{L(t)}{A} = \frac{R^2(t)L(t)}{4\pi R_0^4 r^2}. \quad (2.4)$$

Here the subscript ‘g’ denotes a single galaxy, and the two factors of $R(t)/R_0$ are included to take *expansion* into account: this stretches the wavelength of the light in space (reducing its energy), and also spaces the photons more widely apart in time. These are sometimes known as Hubble’s ‘energy’ and ‘number’ effects [2].

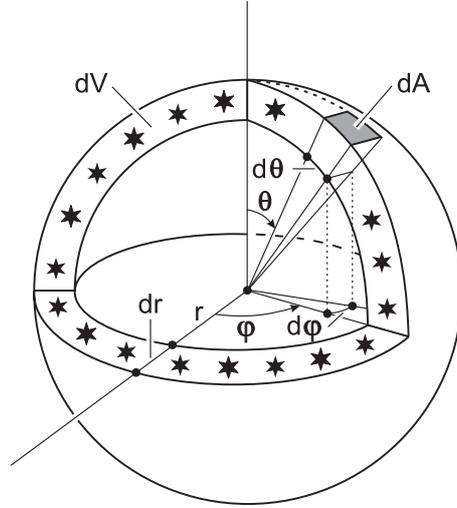


Figure 2.1. Surface area element dA and volume dV of a thin spherical shell, containing all the light sources at coordinate distance r .

We now consider a multitude of galaxies. Suppose that these are distributed through space with some physical number density $n_g(t)$. In the covariant language of general relativity, this quantity should be couched in four-dimensional terms. For this purpose one can imagine a **galaxy current** J_g^μ , analogous to the flow of charges in electromagnetic theory. In a homogeneous and isotropic Universe, this galaxy current may be shown [3] to have the simple form

$$J_g^\mu = n_g U^\mu \quad (2.5)$$

where $U^\mu \equiv (1, 0, 0, 0)$ is the galaxy four-velocity. The vanishing space components of this velocity indicate that galaxies in such a Universe are comoving with the expansion. Let us assume that they are also *conserved* (i.e. their rates of formation and destruction by merging or other processes are slow in comparison to the expansion rate of the Universe). Then J_g^μ obeys a conservation equation

$$\nabla_\mu J_g^\mu = 0 \quad (2.6)$$

where ∇_μ denotes the covariant derivative. Using the metric (2.2), one can put this into the form

$$\frac{1}{R^3} \frac{d}{dt} (R^3 n_g) = 0 \quad (2.7)$$

or

$$n_g = n \left(\frac{R}{R_0} \right)^{-3}. \quad (2.8)$$

We will henceforth reserve the symbol n to denote **comoving number density** (i.e. that measured in a frame which expands with the Universe). Under the assumption of galaxy conservation, which we shall make for the most part, this quantity is equal to its value at $z = 0$ ($n = n_0 = \text{constant}$). For cases where mergers or other galaxy-non-conserving processes are significant, the comoving number density becomes a function of redshift, $n = n(z)$. We shall consider one such case in chapter 3.

Let us now shift our origin so that we are located at the centre of the spherical shell in figure 2.1, and consider those galaxies located inside the shell which extends from radial coordinate distance r to $r + dr$. The volume of this shell is given with the help of (2.2) by

$$dV = \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \left[\frac{R dr}{\sqrt{1-kr^2}} \right] [Rr d\theta][Rr \sin\theta d\phi] = \frac{4\pi R^3 r^2 dr}{\sqrt{1-kr^2}}. \quad (2.9)$$

We can simplify this by exploiting the fact that the only trajectories of interest are those of light rays striking our detectors at origin. By definition, these are radial ($d\theta = d\phi = 0$) null geodesics ($ds^2 = 0$), for which the metric (2.2) relates time t and coordinate distance r via

$$c dt = \frac{R dr}{\sqrt{1-kr^2}}. \quad (2.10)$$

The volume of the shell can then be written

$$dV = 4\pi R^2 r^2 c dt \quad (2.11)$$

and the latter may now be thought of as extending between **look-back times** $t_0 - t$ and $t_0 - (t + dt)$, rather than distances r and $r + dr$.

The total energy received at the origin from the galaxies in the shell is just the product of their individual intensities (2.4), their number per unit volume (2.8) and the volume of the shell (2.11):

$$dQ = dQ_g n_g dV = cn(t)\tilde{R}(t)L(t) dt. \quad (2.12)$$

Here we have defined the *relative* scale factor by

$$\tilde{R}(t) \equiv R(t)/R_0. \quad (2.13)$$

(We will use tildes throughout this book to denote dimensionless quantities taken relative to their value at the present time t_0 .) Integrating over all the spherical shells between t_0 and $t_0 - t_f$, where t_f is the source formation time, we obtain the result

$$Q = \int dQ = c \int_{t_f}^{t_0} n(t)L(t)\tilde{R}(t) dt. \quad (2.14)$$

Equation (2.14) defines the **bolometric intensity** of the extragalactic background light (EBL). This is the energy received by us (over all wavelengths of the

electromagnetic spectrum) per unit area, per unit time, from all the galaxies which have been shining since time t_f . In principle, if we let $t_f \rightarrow 0$, we will encompass the entire history of the Universe since the big bang. Although this sometimes provides a useful mathematical shortcut, we will see in later sections that it is physically more realistic to cut the integral off at a finite formation time. The quantity Q is a measure of the amount of light in the Universe, and Olbers' 'paradox' is merely another way of asking why it is low.

2.3 From time to redshift

While the cosmic time t is a useful independent variable for theoretical purposes, it is not directly observable. In studies aimed at making contact with eventual observation it is better to work in terms of **redshift** z , which is the relative shift in wavelength λ of a light between the time it is emitted and observed:

$$z \equiv \frac{\Delta\lambda}{\lambda} = \frac{R_0 - R(t)}{R(t)} = \tilde{R}^{-1} - 1. \quad (2.15)$$

Differentiation gives

$$dz = -\frac{R_0 dR}{R^2} = -\frac{R_0 \dot{R} dt}{R^2} \quad (2.16)$$

where an overdot signifies the time derivative. Inverting, we obtain

$$dt = -\frac{R^2 dz}{R_0 \dot{R}} = -\frac{dz}{(1+z)H(z)}. \quad (2.17)$$

Here we have introduced a new quantity, **Hubble's parameter**:

$$H \equiv \dot{R}/R \quad (2.18)$$

which is the expansion rate of the Universe. Its value at the present time is known as **Hubble's constant** (H_0). The numerical size of this latter quantity continues to be debated by observational cosmologists. For this reason it is usually written in the form

$$H_0 = 100h_0 \text{ km s}^{-1} \text{ Mpc}^{-1} = 0.102h_0 \text{ Gyr}^{-1}. \quad (2.19)$$

Here $1 \text{ Gyr} \equiv 10^9 \text{ yr}$ and the uncertainties have been absorbed into a dimensionless parameter h_0 whose value is currently thought to be in the range $0.6 \leq h_0 \leq 0.9$. We will have more to say about the observational status of h_0 in chapter 4.

Putting (2.17) into the bolometric EBL intensity integral (2.14), and using (2.15) to replace \tilde{R} with $(1+z)^{-1}$, we obtain

$$Q = c \int_0^{z_f} \frac{n(z)L(z) dz}{(1+z)^2 H(z)}. \quad (2.20)$$

Here z_f is the redshift of galaxy formation, and we have written the comoving number density and luminosity as functions of redshift rather than time. One may think of $n(z)$ and $L(z)$ as describing the physics of the sources, the two factors of $(1+z)$ as representing the dilution and stretching of the light signals and $H(z)$ as governing the evolution of the background spacetime. In practical terms, the transition from an integral over time (2.14) to one over redshift (2.20) is immensely valuable and might be likened to arming a traffic officer with a radar gun rather than a stopwatch.

For some problems, and for this chapter in particular, the physics of the sources themselves are of secondary importance, and it is reasonable to take $L(z) = L_0$ and $n(z) = n_0$ as constants over the range of redshifts of interest. In this case the integral (2.20) reads simply

$$Q = \frac{cn_0L_0}{H_0} \int_0^{z_f} \frac{dz}{(1+z)^2 \tilde{H}(z)}. \quad (2.21)$$

Here we have defined the **relative expansion rate** by

$$\tilde{H}(z) \equiv H(z)/H_0. \quad (2.22)$$

(This is merely Hubble's parameter, normalized to its present value.) The form of the function $\tilde{H}(z)$, which will be of central importance throughout this book, is derived in the section that follows.

All the dimensional information in (2.21) is now contained in the constant outside the integral. In the numerator of this constant we may identify the present **comoving luminosity density** of the Universe,

$$\mathcal{L}_0 \equiv n_0L_0. \quad (2.23)$$

This can be measured experimentally by counting galaxies down to some faint limiting apparent magnitude, and extrapolating to fainter ones based on assumptions about the true distribution of absolute magnitudes. A recent compilation of seven such studies over the past decade is that of Fukugita *et al* [5] in the B-band:

$$\begin{aligned} \mathcal{L}_0 &= (2.0 \pm 0.2) \times 10^8 h_0 L_\odot \text{ Mpc}^{-3} \\ &= (2.6 \pm 0.3) \times 10^{-32} h_0 \text{ erg s}^{-1} \text{ cm}^{-3}. \end{aligned} \quad (2.24)$$

(This waveband, centred near 4400 Å, is where galaxies emit most of their light, and is also close to the wavelength of peak sensitivity of human eyesight.) We will use this number throughout our book. Recent data from the Sloan Digital Sky Survey (SDSS) suggest a somewhat higher value with larger uncertainty, $(2.41 \pm 0.39) \times 10^8 h_0 L_\odot \text{ Mpc}^{-3}$ [6]; while a preliminary result from the Two Degree Field (2dF) team is slightly lower, $\mathcal{L}_0 = (1.82 \pm 0.17) \times 10^8 h_0 L_\odot \text{ Mpc}^{-3}$ [7]. If the final result inferred from large-scale galaxy surveys of this kind proves

to be significantly different from that in (2.24), then our EBL intensities (which are proportional to \mathcal{L}_0) would go up or down accordingly.

Numerically, the constant factor outside the integral (2.21) sets the order of magnitude of the integral itself, so it is of interest to see what value this takes. Denoting it by Q_* we find using (2.24) that

$$Q_* \equiv c\mathcal{L}_0/H_0 = (2.5 \pm 0.2) \times 10^{-4} \text{ erg s}^{-1} \text{ cm}^{-2}. \quad (2.25)$$

There are two important things to note about this quantity. First, because the factors of h_0 attached to both \mathcal{L}_0 and H_0 cancel each other out, it is *independent* of the uncertainty in Hubble's constant. This is not always appreciated but was first emphasized by Felten [8]. Second, the value of Q_* is *very small* by everyday standards: more than a million times fainter than the bolometric intensity produced by a 100 W bulb in an average-sized room (section 1.4). The smallness of this number already contains the essence of the resolution of Olbers' paradox. But to put the latter to rest in a definitive fashion, we need to evaluate the full integral (2.20). This, in turn, requires the Hubble expansion rate $\tilde{H}(z)$, which we proceed to derive next.

2.4 Matter and energy

The relative expansion rate $\tilde{H}(z)$ can be obtained from the field equations of general relativity, if the matter and energy content of the Universe are specified. Under the assumptions of isotropy and homogeneity, the latter two quantities can be described [3] by an energy–momentum tensor of **perfect fluid** form

$$\mathcal{T}_{\mu\nu} = (\rho + p/c^2)U_\mu U_\nu + pg_{\mu\nu}. \quad (2.26)$$

Here ρ is the density of the cosmic fluid and p its pressure. These two quantities, in turn, are related by an **equation of state**

$$p = (\gamma - 1)\rho c^2. \quad (2.27)$$

Three specific equations of state are of particular relevance in cosmology and will make regular appearances in the chapters that follow. They all have $\gamma = \text{constant}$, as follows:

$$\begin{aligned} \bullet \text{ radiation } (\gamma_r = \frac{4}{3}) : & \quad p_r = \frac{1}{3}\rho_r c^2 \\ \bullet \text{ dustlike matter } (\gamma_m = 1) : & \quad p_m = 0 \\ \bullet \text{ vacuum energy } (\gamma_v = 0) : & \quad p_v = -\rho_v c^2 \end{aligned} \quad (2.28)$$

The first of these is a good approximation to the *early* Universe, when conditions were so hot and dense that matter and radiation existed in nearly perfect thermodynamic equilibrium (the radiation era). The second has usually been used to model the *present* Universe, since we know that the energy density of

electromagnetic radiation now is far below that of dustlike (pressure-free) matter. The third may be a good description of the *future* state of the Universe, if recent measurements of the magnitudes of high-redshift of Type Ia supernovae are borne out. These indicate that the vacuum energy may already be more important than all other contributions to the density of the Universe, including those from any unseen dark-matter component.

A vacuum-like equation of state has also often been discussed as a possible description of conditions during an ‘inflationary’ phase in the early stages of (or perhaps even preceding) the radiation era. Various values have been proposed for the index γ in this context, from zero (de Sitter inflation) to $1/3$ (domain walls) and $2/3$ (cosmic strings). Inflation, or non-negative acceleration of the scale factor, occurs for any perfect fluid with $0 \leq \gamma \leq 2/3$, and cosmological matter with this property has come to be known as **quintessence** (a reference to the substance making up the realm beyond the planetary spheres in Aristotelian astronomy). Such a fluid might be realized physically, for instance, in an early Universe dominated by a minimally-coupled **scalar field** whose equation of state would have $\gamma_\varphi = \dot{\varphi}^2/[\dot{\varphi}^2/2 + V(\varphi)]$, where φ is the value of the scalar field and $V(\varphi)$ its potential energy. This is ‘quintessential’ ($\gamma_\varphi \leq 2/3$) whenever the field’s potential energy exceeds its kinetic energy, $V(\varphi) \geq \dot{\varphi}^2$. A pure vacuum ($\gamma_\varphi \approx 0$) is recovered in the ‘slow-roll’ limit, $V(\varphi) \gg \dot{\varphi}^2$. In the opposite limit, $V(\varphi) \approx 0$, one has $\gamma_\varphi \approx 2$ and the scalar fluid equation of state approximates that of **stiff matter**, $p_s = \rho_s c^2$. The index γ is not necessarily constant, especially for situations involving multiple interacting fluids. For our purposes, however, it will be sufficient to assume that a single component dominates the cosmic fluid during any given epoch, so that the Universe can be described by one of the cases in (2.28).

Assuming that energy and momentum are neither created nor destroyed, one can proceed exactly as with the galaxy current J_g^μ . The conservation equation in this case reads

$$\nabla^\mu \mathcal{T}_{\mu\nu} = 0. \quad (2.29)$$

This reduces to

$$\frac{1}{R^3} \frac{d}{dt} [R^3 (\rho c^2 + p)] = \frac{dp}{dt} \quad (2.30)$$

which may be compared with (2.7) for galaxies. Equation (2.30) is solved with the help of the equation of state (2.27) to yield

$$\rho(t) = \rho_0 [R(t)/R_0]^{-3\gamma}. \quad (2.31)$$

In particular, for the single-component fluids in (2.28):

$$\begin{aligned} \rho_r &= \rho_{r,0} (R/R_0)^{-4} && \text{(radiation)} \\ \rho_m &= \rho_{m,0} (R/R_0)^{-3} && \text{(dustlike matter)} \\ \rho_v &= \rho_{v,0} = \text{constant} && \text{(vacuum energy)} \end{aligned} \quad (2.32)$$

These expressions will frequently prove useful in later chapters. They are also applicable to cases in which several components are present, but only if these do not exchange energy at significant rates (i.e. relative to the expansion rate), so that each is in effect conserved separately.

2.5 The expansion rate

With geometry as described by the metric (2.2), and matter–energy as described by the energy–momentum tensor (2.26), we are in a position to solve the Einstein field equations (2.1) for cosmology. This results in two differential equations for the scale factor R and its time-derivatives. The equation for the expansion rate H in terms of R is

$$H^2 + \frac{kc^2}{R^2} - \frac{\Lambda c^2}{3} = \frac{8\pi G}{3}\rho. \quad (2.33)$$

We will take the cosmological fluid to consist of both radiation and dustlike matter components so that $\rho = \rho_r + \rho_m$. It then proves useful to define two new quantities. One is the **vacuum energy density** associated with the cosmological constant:

$$\rho_\Lambda c^2 \equiv \frac{\Lambda c^4}{8\pi G} = \text{constant}. \quad (2.34)$$

The corresponding pressure is given by (2.28) as $p_\Lambda = -\rho_\Lambda c^2$. The second useful quantity is known as the **critical density**, for reasons that will become clear momentarily. It reads:

$$\rho_{\text{crit}}(t) \equiv \frac{3H^2(t)}{8\pi G}. \quad (2.35)$$

In particular the *present* value of this quantity is

$$\begin{aligned} \rho_{\text{crit},0} &\equiv \frac{3H_0^2}{8\pi G} = (2.78 \times 10^{11})h_0^2 M_\odot \text{Mpc}^{-3} \\ &= (1.88 \times 10^{-29})h_0^2 \text{g cm}^{-3}. \end{aligned} \quad (2.36)$$

This works out to the equivalent of between 4.0 protons per cubic metre (if $h_0 = 0.6$) and 9.1 protons per cubic metre (if $h_0 = 0.9$). Using (2.34) and (2.36) we can evaluate (2.33) at the present time as follows:

$$\frac{kc^2}{R_0^2} = H_0^2 \left(\frac{\rho_{r,0} + \rho_{m,0} + \rho_\Lambda - \rho_{\text{crit},0}}{\rho_{\text{crit},0}} \right). \quad (2.37)$$

The physical significance of $\rho_{\text{crit},0}$ is now apparent: its value determines the *spatial curvature* of the Universe. If the sum of the densities $\rho_{r,0}$, $\rho_{m,0}$ and ρ_Λ is exactly equal to $\rho_{\text{crit},0}$ then $k = 0$. The Universe in this case is flat or *Euclidean*. Such a Universe is unbounded and infinite in extent. Alternatively, if the sum

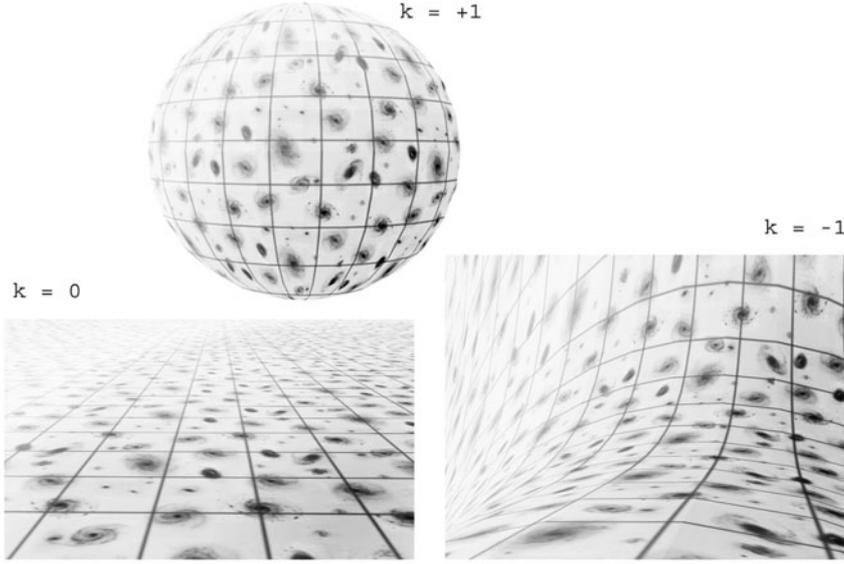


Figure 2.2. The spatial geometry of the Universe (here represented as a two-dimensional surface) is determined by its total density, expressed in units of the critical density $\rho_{\text{crit},0}$ by the parameter $\Omega_{\text{tot},0}$. If $\Omega_{\text{tot},0} = 1$ then the Universe is flat (‘Euclidean’) and the curvature parameter k can be set to zero in the metric (2.2). If $\Omega_{\text{tot},0} > 1$ then the Universe is positively curved (spherical) and k can be set to $+1$. If $\Omega_{\text{tot},0} < 1$ then it is negatively curved (hyperbolic) and $k = -1$. (Figure courtesy E C Eekels.)

of $\rho_{r,0}$, $\rho_{m,0}$ and ρ_{Λ} exceeds $\rho_{\text{crit},0}$, then $k > 0$ and the Universe is positively curved. This is often referred to as a ‘ $k = +1$ model’ because the magnitude of the curvature constant k can be normalized to unity by choice of units for R_0 . Spatial hypersurfaces in this model are spherical in shape, and the Universe is unbounded, but closed and finite. Finally, if $\rho_{r,0} + \rho_{m,0} + \rho_{\Lambda} < \rho_{\text{crit},0}$, then the Universe has negative curvature ($k = -1$) and hyperbolic spatial sections. It is open, unbounded and infinite in extent. (Strictly speaking one can also obtain flat and hyperbolic solutions which are *finite* by adopting a non-trivial topology and suitably ‘identifying’ pairs of points; we do not pursue this here.) All three possibilities are depicted in figure 2.2.

To simplify the notation we now rescale all our physical densities, defining three dimensionless **density parameters**:

$$\Omega_{r,0} \equiv \frac{\rho_{r,0}}{\rho_{\text{crit},0}} \quad \Omega_{m,0} \equiv \frac{\rho_{m,0}}{\rho_{\text{crit},0}} \quad \Omega_{\Lambda,0} \equiv \frac{\rho_{\Lambda}}{\rho_{\text{crit},0}} = \frac{\Lambda c^2}{3H_0^2}. \quad (2.38)$$

These parameters occupy central roles throughout the book, and we will usually refer to them for brevity as ‘densities’ (or ‘present densities’ where this is not

clear from the context). Defining $\Omega_{\text{tot},0} \equiv \Omega_{\text{r},0} + \Omega_{\text{m},0} + \Omega_{\Lambda,0}$, we find from (2.37) that the curvature constant is

$$k = (R_0 H_0 / c)^2 (\Omega_{\text{tot},0} - 1). \quad (2.39)$$

Thus flat models are just those with $\Omega_{\text{tot},0} = 1$, while closed and open models have $\Omega_{\text{tot},0} > 1$ and $\Omega_{\text{tot},0} < 1$ respectively. Substituting (2.39) back into (2.33), using the definitions (2.38) and recalling from (2.15) that $z = R_0/R - 1$, we obtain the expansion rate of the Universe:

$$\begin{aligned} \tilde{H}(z) = & [\Omega_{\text{r},0}(1+z)^4 + \Omega_{\text{m},0}(1+z)^3 + \Omega_{\Lambda,0} \\ & - (\Omega_{\text{tot},0} - 1)(1+z)^2]^{1/2}. \end{aligned} \quad (2.40)$$

When $\Omega_{\text{r},0} = 0$, this is sometimes referred to as the Friedmann equation after the Russian mathematician who obtained the first cosmological solutions of Einstein's equations under the assumption of zero pressure. This restriction was first lifted several years later by Lemaître, so we refer to (2.40) as the **Friedmann–Lemaître equation**.

The form of this equation already reveals a great deal about the contents and evolution of the Universe. For example, the term $\Omega_{\text{m},0}(1+z)^3$ shows that matter acts to increase the expansion rate $H(z)$ as one goes to higher z —that is, to *slow down* the expansion rate with time. This is the braking effect of matter's gravitational self-attraction, and (as we shall learn in chapter 4) it appears to be generated almost entirely by matter which we cannot see.

The term $\Omega_{\text{r},0}(1+z)^4$ shows that radiation has the same effect, but with a stronger dependence on redshift (this is related to the fact that pressure, as well as density, acts as a source of gravitation in general relativity). As one moves backward in time, photons (and relativistic particles) therefore become increasingly important compared to pressureless matter. In fact, the dynamics of the early Universe (at redshifts above $z \gtrsim 10^4$) must have been completely dominated by them. The total *present* radiation density $\Omega_{\text{r},0}$ is, however, several orders of magnitude below that of non-relativistic matter. (This is inferred, not only from measurements such as those of the COBE satellite, but also from the fact that too much pressure would have slowed expansion so much that the Universe could not have lived long enough to contain the oldest stars.) In later chapters we will usually be interested only in the matter-dominated era (or later), and we can safely neglect the radiation term in (2.40).

The term $\Omega_{\Lambda,0}$ is independent of redshift, which means that its influence is not diluted with time. Any Universe with $\Lambda > 0$ will therefore eventually be dominated by vacuum energy. In the limit $t \rightarrow \infty$, in fact, the other terms drop out of (2.40) altogether and the vacuum energy density reads simply $\Omega_{\Lambda,0} = (H_\infty/H_0)^2$ where H_∞ is the limiting value of $H(t)$ as $t \rightarrow \infty$ (assuming that this latter quantity exists; i.e. that the Universe does not recollapse). From (2.38) it then follows that

$$\Lambda c^2 = 3H_\infty^2. \quad (2.41)$$

This constitutes a link between Λ (a constant of nature in Einstein's theory) and the asymptotic expansion rate H_∞ (a dynamical parameter). If $\Lambda > 0$, then we will *necessarily* measure $\Omega_{\Lambda,0} \sim 1$ at late times, regardless of the microphysical origin of the vacuum energy.

The last term in equation (2.40), finally, shows that an excess of $\Omega_{\text{tot},0}$ over one (i.e. a positive curvature) acts to offset the contribution of the first three terms to the expansion rate, while a deficit (i.e. a negative curvature) enhances them. Open models, in other words, expand more quickly at any given redshift z (and therefore last longer) than closed ones. This curvature term, however, goes only as $(1+z)^2$, which means that its importance drops off relative to the matter and radiation terms at early times, and relative to the vacuum term at late ones.

It became commonplace during the 1980s to work with a simplified version of equation (2.40), in which not only the first (radiation) term on the right-hand side was neglected, but the third (vacuum) and fourth (curvature) terms as well. This **Einstein–de Sitter** (EdS) model appeared reasonable at the time, for four principal reasons. First, the four terms in question differ sharply from each other in their dependence on redshift z , and the probability that we should happen to find ourselves in an era when they have similar values would seem *a priori* very remote. By this argument, which goes back to Dicke [9], it was felt that only one term ought to dominate at any given time. Second, the vacuum term was regarded with particular suspicion for reasons to be discussed in chapter 4. Third, a period of **inflation** was asserted to have driven $\Omega_{\text{tot}}(t)$ to unity. (This is still widely believed, but depends on the initial conditions preceding inflation, and does not necessarily hold in all plausible models [10].) And finally, the EdS model was favoured on grounds of simplicity.

These arguments are no longer compelling today, and the determination of $\Omega_{r,0}$, $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$ (along with H_0) has shifted largely back into the domain of observation. We defer a fuller discussion of these issues to chapter 4, being content here to assume merely that radiation and matter densities are positive and not too large ($0 \leq \Omega_{r,0} \leq 1.5$ and $0 \leq \Omega_{m,0} \leq 1.5$), and that vacuum energy density is neither too large nor too negative ($-0.5 \leq \Omega_{\Lambda,0} \leq 1.5$).

2.6 The static analogue

This is a good place to pause and take stock of our results so far. In the foregoing sections we have obtained a simple integral for the bolometric intensity of the Universe, equation (2.21):

$$Q = Q_* \int_0^{z_f} \frac{dz}{(1+z)^2 \tilde{H}(z)}. \quad (2.42)$$

The value of Q_* is given by (2.25). We have also obtained a general expression for the expansion rate $\tilde{H}(z)$ in terms of the present densities of radiation, matter and vacuum energy, equation (2.40). Strictly speaking, (2.42) assumes constant

galaxy luminosity and comoving number density, as discussed in section 2.3. This is adequate for our purposes, because we are not primarily concerned here with the physics of the galaxies themselves. We wish to obtain a definitive answer to the question posed in chapter 1: is it cosmic expansion, or the finite age of the Universe, which is primarily responsible for the low value of Q ?

The relative importance of these two factors continues to be a subject of controversy and confusion (see [4] for a review). In particular there is a lingering perception that general relativity ‘solves’ Olbers’ paradox chiefly because the expansion of the Universe stretches and dims the light it contains.

There is a simple way to test this supposition using the formalism we have already laid out, and that is to ‘turn off’ expansion by setting the scale factor of the Universe equal to a constant value, $R(t) = R_0$. Then $\dot{R}(t) = 1$ from (2.13), and the bolometric intensity of the EBL is given by (2.14) as an integral over time:

$$Q_{\text{stat}} = Q_* H_0 \int_{t_f}^{t_0} dt. \quad (2.43)$$

Here we have taken $n = n_0$ and $L = L_0$ for convenience, as before, and used (2.25) for Q_* . The subscript ‘stat’ denotes the *static analogue* of Q ; that is, the bolometric EBL intensity that one would measure in a Universe which did not expand. Equation (2.43) shows that this is nothing more than the *length of time for which the galaxies have been shining*, measured in units of Hubble time (H_0^{-1}) and scaled by Q_* .

Now, we wish to compare (2.42) in the expanding Universe with its static analogue (2.43), while keeping all other factors the same. In particular, if the comparison is to be meaningful, the *lifetime* of the galaxies should be identical. This is just the integral over dt , which may—in an expanding Universe—be converted to one over dz by means of (2.17):

$$H_0 \int_{t_f}^{t_0} dt = \int_0^{z_f} \frac{dz}{(1+z)\tilde{H}(z)} \quad (2.44)$$

where we have used the definition (2.22) of $\tilde{H}(z)$. In a static Universe, of course, redshift does not carry its usual physical significance. But nothing prevents us from retaining z as an integration variable. Substitution of (2.44) into (2.43) then yields

$$Q_{\text{stat}} = Q_* \int_0^{z_f} \frac{dz}{(1+z)\tilde{H}(z)}. \quad (2.45)$$

We emphasize that z and \tilde{H} are to be seen here as algebraic parameters whose usefulness lies in the fact that they ensure consistency in *age* between the static and expanding pictures.

Equation (2.42) and its static analogue (2.45) allow us to make a meaningful comparison of the bolometric intensity of the EBL in models which are alike in all respects, except that one is expanding while the other stands still. The

two equations are almost identical, differing only by an extra factor of $(1+z)^{-1}$ attached to the integrand in the expanding case. Since this is less than unity for all z , we expect Q to be less than Q_{stat} , although the magnitude of the difference will depend on the shape of $\tilde{H}(z)$. The latter must have been fairly smooth over the lifetime of the galaxies in any plausible model, so that the difference between the two integrals cannot amount to much more than a few orders of magnitude. It follows immediately that the intensity of the EBL in expanding models (as in their static analogues) *must be determined primarily by the lifetime of the galaxies, so that the effects of expansion are secondary*. We proceed in the next section to make this conclusion quantitative.

2.7 A quantitative resolution

The most straightforward way to do this is to evaluate the ratio Q/Q_{stat} for a wide range of values of the cosmological parameters $\Omega_{r,0}$, $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$. If we find that $Q/Q_{\text{stat}} \ll 1$ over much of this phase space, then expansion would reduce Q quite significantly from what it would otherwise be in an equivalent static Universe. Conversely, values of $Q/Q_{\text{stat}} \approx 1$ would confirm that expansion has little effect.

This section, then, is a largely technical one, where we calculate Q/Q_* , Q_{stat}/Q_* and Q/Q_{stat} for the widest possible range of models and present the results in analytic or graphical form. For readers who are interested primarily in final results and prefer to skip the details of this exercise, we give here the conclusion, which is that the ratio Q/Q_{stat} turns out to lie in the range 0.6 ± 0.1 across nearly the entirety of the cosmological phase space. While expansion does reduce the bolometric intensity of the EBL, in other words, the reduction is slight. If we could freeze expansion without affecting any of the other relevant factors, the total amount of light reaching the Milky Way would go up by a factor of less than two in any plausible cosmological model.

Let us begin with the simplest examples, in which the Universe has one critical-density component or contains nothing at all (table 2.1). Consider first the **radiation model** with a critical density of radiation or ultrarelativistic particles ($\Omega_{r,0} = 1$) but $\Omega_{m,0} = \Omega_{\Lambda,0} = 0$. Bolometric EBL intensity in the expanding

Table 2.1. Simple flat and empty models

Model name	$\Omega_{r,0}$	$\Omega_{m,0}$	$\Omega_{\Lambda,0}$	$1 - \Omega_{\text{tot},0}$
Radiation	1	0	0	0
Einstein–de Sitter	0	1	0	0
De Sitter	0	0	1	0
Milne (empty)	0	0	0	1

Universe is, from (2.42),

$$\frac{Q}{Q_*} = \int_1^{1+z_f} \frac{dx}{x^4} = \begin{cases} 21/64 & (z_f = 3) \\ 1/3 & (z_f = \infty) \end{cases} \quad (2.46)$$

where $x \equiv 1 + z$. The corresponding result for a static model is given by (2.45) as

$$\frac{Q_{\text{stat}}}{Q_*} = \int_1^{1+z_f} \frac{dx}{x^3} = \begin{cases} 15/32 & (z_f = 3) \\ 1/2 & (z_f = \infty). \end{cases} \quad (2.47)$$

Here we have chosen illustrative lower and upper limits on the redshift of galaxy formation ($z_f = 3$ and ∞ respectively). The actual value of this parameter has not yet been determined, although there are now indications that z_f may be as high as six. Arbitrarily large values of z_f have little physical meaning within the context of calculations of the EBL, since these presume that enough time has elapsed since the big bang for stars and galaxies to have formed, and this implies that z_f is bounded above at some reasonable value. In any case, it may be seen from the previous results that the overall EBL intensity is rather insensitive to this parameter. Increasing z_f lengthens the period over which galaxies radiate, and this increases both Q and Q_{stat} . The ratio Q/Q_{stat} , however, is given by

$$\frac{Q}{Q_{\text{stat}}} = \begin{cases} 7/10 & (z_f = 3) \\ 2/3 & (z_f = \infty) \end{cases} \quad (2.48)$$

and this changes but little. We will find this to be true in general.

Consider next the Einstein–de Sitter model, which has a critical density of dustlike matter ($\Omega_{m,0} = 1$) with $\Omega_{r,0} = \Omega_{\Lambda,0} = 0$. Bolometric EBL intensity in the expanding Universe is, from (2.42),

$$\frac{Q}{Q_*} = \int_1^{1+z_f} \frac{dx}{x^{7/2}} = \begin{cases} 31/80 & (z_f = 3) \\ 2/5 & (z_f = \infty). \end{cases} \quad (2.49)$$

The corresponding static result is given by (2.45) as

$$\frac{Q_{\text{stat}}}{Q_*} = \int_1^{1+z_f} \frac{dx}{x^{5/2}} = \begin{cases} 7/12 & (z_f = 3) \\ 2/3 & (z_f = \infty). \end{cases} \quad (2.50)$$

The ratio of EBL intensity in an expanding Einstein–de Sitter model to that in the equivalent static model is thus

$$\frac{Q}{Q_{\text{stat}}} = \begin{cases} 93/140 & (z_f = 3) \\ 3/5 & (z_f = \infty). \end{cases} \quad (2.51)$$

These numbers are only slightly below those of the radiation case.

A third simple case is the **de Sitter model**, which consists entirely of vacuum energy ($\Omega_{\Lambda,0} = 1$), with $\Omega_{r,0} = \Omega_{m,0} = 0$. Bolometric EBL intensity in the expanding case is, from (2.42),

$$\frac{Q}{Q_*} = \int_1^{1+z_f} \frac{dx}{x^2} = \begin{cases} 3/4 & (z_f = 3) \\ 1 & (z_f = \infty). \end{cases} \quad (2.52)$$

Equation (2.45) gives for the equivalent static case

$$\frac{Q_{\text{stat}}}{Q_*} = \int_1^{1+z_f} \frac{dx}{x} = \begin{cases} \ln 4 & (z_f = 3) \\ \infty & (z_f = \infty). \end{cases} \quad (2.53)$$

The ratio of EBL intensity in an expanding de Sitter model to that in the equivalent static model is then

$$\frac{Q}{Q_{\text{stat}}} = \begin{cases} 3/(4 \ln 4) & (z_f = 3) \\ 0 & (z_f = \infty). \end{cases} \quad (2.54)$$

The de Sitter Universe is older than other models, which means it has more time to fill up with light, so intensities are higher. In fact, Q_{stat} (which is proportional to the lifetime of the galaxies) goes to infinity as $z_f \rightarrow \infty$, driving Q/Q_{stat} to zero in this limit. (It is thus possible to ‘recover Olbers’ paradox’ in the de Sitter model, as noted by White and Scott [11].) Such a limit is, however, unphysical as noted earlier. For realistic values of z_f one obtains values of Q/Q_{stat} which are only slightly lower than those in the radiation and matter cases.

Finally, we consider the **Milne model**, which is empty of all forms of matter and energy ($\Omega_{r,0} = \Omega_{m,0} = \Omega_{\Lambda,0} = 0$), making it an idealization but one which has nevertheless often proved useful as a bridge between the special and general theories of relativity. Bolometric EBL intensity in the expanding case is given by (2.42) as

$$\frac{Q}{Q_*} = \int_1^{1+z_f} \frac{dx}{x^3} = \begin{cases} 15/32 & (z_f = 3) \\ 1/2 & (z_f = \infty) \end{cases} \quad (2.55)$$

which is identical to equation (2.47) for the static radiation model. The corresponding static result is given by (2.45) as

$$\frac{Q_{\text{stat}}}{Q_*} = \int_1^{1+z_f} \frac{dx}{x^2} = \begin{cases} 3/4 & (z_f = 3) \\ 1 & (z_f = \infty) \end{cases} \quad (2.56)$$

which is the same as equation (2.52) for the expanding de Sitter model. The ratio of EBL intensity in an expanding Milne model to that in the equivalent static model is then

$$\frac{Q}{Q_{\text{stat}}} = \begin{cases} 5/8 & (z_f = 3) \\ 1/2 & (z_f = \infty). \end{cases} \quad (2.57)$$

This again lies close to previous results. In fact, we have found in every case (except the $z_f \rightarrow \infty$ limit of the de Sitter model) that the ratio of bolometric EBL intensities with and without expansion lies in the range $0.4 \lesssim Q/Q_{\text{stat}} \lesssim 0.7$.

It may, however, be that models whose total density is neither critical nor zero have different properties. To check this we expand our investigation to the wider class of open and closed models. Equations (2.42) and (2.45) can be solved analytically for these cases, if they are dominated by a single component. We collect the solutions for convenience in appendix A, and plot them in figures 2.3–2.5 for radiation-, matter- and vacuum-dominated models respectively. In each

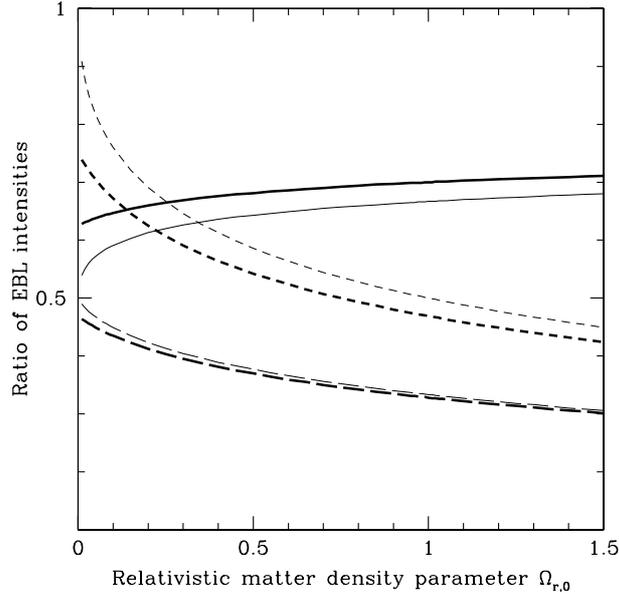


Figure 2.3. Ratios Q/Q_* (long-dashed lines), Q_{stat}/Q_* (short-dashed lines) and Q/Q_{stat} (unbroken lines) as a function of radiation density $\Omega_{r,0}$. Bold lines are calculated for $z_f = 3$ while light ones have $z_f = \infty$.

figure, long-dashed lines correspond to EBL intensity in expanding models (Q/Q_*) while short-dashed ones show the equivalent static quantities (Q_{stat}/Q_*). The ratio of these two quantities (Q/Q_{stat}) is indicated by unbroken lines. The bold curves have $z_f = 3$ while light ones are calculated for $z_f = \infty$.

In the cases of the radiation- and matter-dominated Universes, figures 2.3 and 2.4 show that while the individual intensities Q/Q_* and Q_{stat}/Q_* do vary significantly with $\Omega_{r,0}$ and $\Omega_{m,0}$, their ratio Q/Q_{stat} remains nearly constant across the whole of the phase space, staying inside the range $0.5 \lesssim Q/Q_{\text{stat}} \lesssim 0.7$ for both models.

The case of the vacuum-dominated Universe appears more complicated, but shows the same trend (figure 2.5). Absolute EBL intensities Q/Q_* and Q_{stat}/Q_* differ noticeably from those in the radiation- and matter-dominated models, but their *ratio* (full lines) is again close to flat. The exception occurs as $\Omega_{\Lambda,0} \rightarrow 1$ (the de Sitter model), where Q/Q_{stat} can dip well below 0.5 in the limit of large z_f , as discussed earlier. As $\Omega_{\Lambda,0}$ rises *above* one, the big bang disappears altogether (in models with $\Omega_{r,0} = \Omega_{m,0} = 0$), and one finds instead a **big bounce** (i.e. a positive minimum scale factor at the beginning of the expansionary phase). Hence there

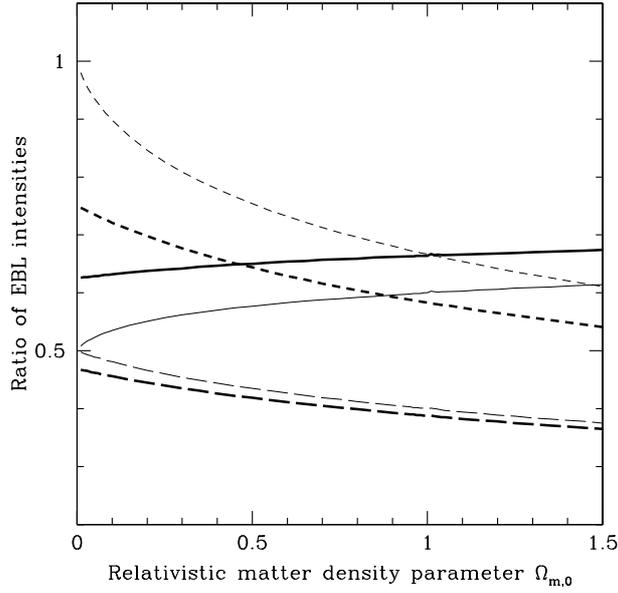


Figure 2.4. Ratios Q/Q_* (long-dashed lines), Q_{stat}/Q_* (short-dashed lines) and Q/Q_{stat} (unbroken lines) as a function of matter density $\Omega_{m,0}$. Bold lines are calculated for $z_f = 3$ while light ones have $z_f = \infty$.

is a maximum possible redshift z_{max} given by

$$1 + z_{\text{max}} = \sqrt{\frac{\Omega_{\Lambda,0}}{\Omega_{\Lambda,0} - 1}}. \quad (2.58)$$

While such models are rarely considered, it is interesting to note that the same pattern persists here. In light of (2.58), one can no longer integrate out to arbitrarily high formation redshift z_f . If one wants to integrate to *at least* z_f , then one is limited to vacuum densities less than $\Omega_{\Lambda,0} < (1 + z_f)^2 / [(1 + z_f)^2 - 1]$. In the case $z_f = 3$ (shown with heavy lines), this corresponds to an upper limit of $\Omega_{\Lambda,0} < 16/15$ (faint dotted line). More generally, for $\Omega_{\Lambda,0} > 1$ the limiting value of EBL intensity (shown with light lines) is reached as $z_f \rightarrow z_{\text{max}}$ rather than $z_f \rightarrow \infty$ for both expanding and static models. Over the entire parameter space $-0.5 \leq \Omega_{\Lambda,0} \leq 1.5$ (except in the immediate vicinity of $\Omega_{\Lambda,0} = 1$), figure 2.5 shows that $0.4 \lesssim Q/Q_{\text{stat}} \lesssim 0.7$ as before.

When more than one component of matter is present, analytic expressions for the bolometric intensity can be found in only a few special cases, and the ratios Q/Q_* and Q_{stat}/Q_* must, in general, be computed numerically. We show the results in figure 2.6 for the case which is of most physical interest: a Universe containing both dustlike matter ($\Omega_{m,0}$, horizontal axis) and vacuum energy ($\Omega_{\Lambda,0}$,

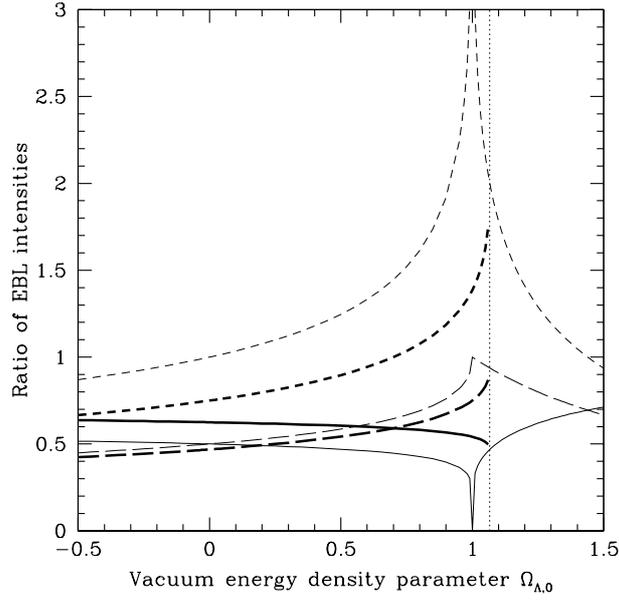


Figure 2.5. Ratios Q/Q_* (long-dashed lines), Q_{stat}/Q_* (short-dashed lines) and Q/Q_{stat} (unbroken lines) as a function of vacuum energy density $\Omega_{\Lambda,0}$. Bold lines are calculated for $z_f = 3$ while light ones have $z_f = \infty$ for $\Omega_{\Lambda,0} \leq 1$ and $z_f = z_{\text{max}}$ for $\Omega_{\Lambda,0} > 1$. The dotted vertical line marks the maximum value of $\Omega_{\Lambda,0}$ for which one can integrate to $z_f = 3$.

vertical axis), with $\Omega_{r,0} = 0$. This is a contour plot, with five bundles of equal-EBL intensity contours for the expanding Universe (labelled $Q/Q_* = 0.37, 0.45, 0.53, 0.61$ and 0.69). The bold (unbroken) lines are calculated for $z_f = 5$, while medium (long-dashed) lines assume $z_f = 10$ and light (short-dashed) lines have $z_f = 50$. Also shown is the boundary between big bang and bounce models (bold line in top left corner), and the boundary between open and closed models (diagonal dashed line). Similar plots could be produced for the $\Omega_{r,0}$ – $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$ – $\Omega_{r,0}$ planes.

Figure 2.6 shows that the bolometric intensity of the EBL is only modestly sensitive to the cosmological parameters $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$. Moving from the lower right-hand corner of the phase space ($Q/Q_* = 0.37$) to the upper left-hand one ($Q/Q_* = 0.69$) changes the value of this quantity by less than a factor of two. Increasing the redshift of galaxy formation from $z_f = 5$ to 10 has little effect, and increasing it again to $z_f = 50$ even less. This means that *essentially all of the light reaching us from outside the Milky Way comes from galaxies at $z < 5$* , regardless of the redshift at which these objects actually formed.

While figure 2.6 confirms that the night sky is dark in any reasonable

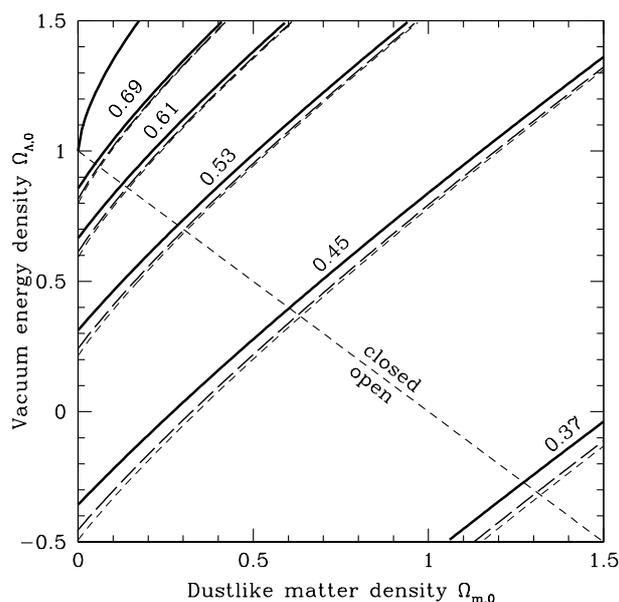


Figure 2.6. The ratio Q/Q_* in an expanding Universe, plotted as a function of matter-density parameter $\Omega_{m,0}$ and vacuum-density parameter $\Omega_{\Lambda,0}$ (with the radiation-density parameter $\Omega_{r,0}$ set to zero). Unbroken lines correspond to $z_f = 5$, while long-dashed lines assume $z_f = 10$ and short-dashed ones have $z_f = 50$.

cosmological model, figure 2.7 shows *why*. It is a contour plot of Q/Q_{stat} , the value of which varies so little across the phase space that we have had to restrict the range of z_f -values in order to keep the diagram from being too cluttered. The bold (unbroken) lines are calculated for $z_f = 4.5$, the medium (long-dashed) lines for $z_f = 5$, and light (short-dashed) lines for $z_f = 5.5$. The spread in contour values is extremely narrow, from $Q/Q_{\text{stat}} = 0.56$ in the upper left-hand corner to 0.64 in the lower right-hand one—a difference of less than 15%. Figure 2.6 confirms our previous analytical results and leaves no doubt about the resolution of Olbers' paradox: *the brightness of the night sky is determined to order of magnitude by the lifetime of the galaxies, and is reduced by no more than a factor of 0.6 ± 0.1 due to the expansion of the Universe.*

2.8 Light at the end of the Universe?

We have obtained a general integral (2.42) and a number of exact expressions for the bolometric intensity Q of the EBL in an expanding Universe. These results are phrased in terms of the densities $\Omega_{r,0}$, $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$ of radiation, matter and vacuum energy, as well as the redshift z_f at which the galaxies formed. All four of

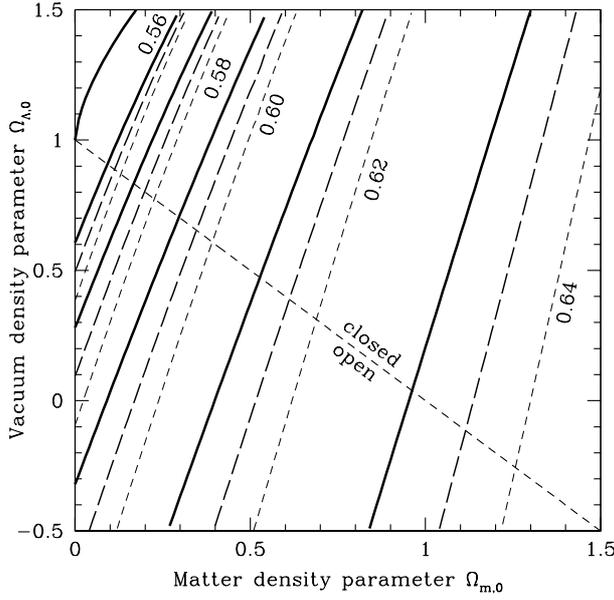


Figure 2.7. The ratio Q/Q_{stat} of EBL intensity in an expanding Universe to that in a static Universe with the same values of the matter-density parameter $\Omega_{m,0}$ and vacuum-density parameter $\Omega_{\Lambda,0}$ (with the radiation-density parameter $\Omega_{r,0}$ set to zero). Unbroken lines correspond to $z_f = 4.5$, while long-dashed lines assume $z = 5$ and short-dashed ones have $z = 5.5$.

these parameters are ones that we can, in principle, measure with our instruments. However, it is occasionally useful to shift attention from redshift z back to time t in order to ask questions which, while not directly connected with experiment, are of conceptual interest. In this section we ask: do the foregoing results mean that the brightness of the night sky is changing with *time*? If so, is it getting brighter or darker? How quickly?

To investigate these questions one would like the EBL intensity expressed as a function of time, as in (2.14). Evaluation of this integral requires a knowledge of the scale factor $R(t)$, which is not well constrained by observation. An approximate expression can, however, be obtained if the Universe has a density very close to the critical one, as suggested by observations of the power spectrum of the CMB (see chapter 4). In this case $k = 0$ and (2.33) simplifies to

$$\left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi G}{3}(\rho_r + \rho_m + \rho_\Lambda) \quad (2.59)$$

where we have set $\rho = \rho_r + \rho_m$ and used (2.34) for ρ_Λ . If we now assume that only one of these three components is dominant at a given time, then we can make

use of (2.32) to obtain

$$\left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi G}{3} \times \begin{cases} \rho_{r,0}(R/R_0)^{-4} & \text{(radiation)} \\ \rho_{m,0}(R/R_0)^{-3} & \text{(matter)} \\ \rho_\Lambda & \text{(vacuum)}. \end{cases} \quad (2.60)$$

These differential equations are separable and have solutions

$$\frac{R(t)}{R_0} = \begin{cases} (t/t_0)^{1/2} & \text{(radiation)} \\ (t/t_0)^{2/3} & \text{(matter)} \\ \exp[H_0(t - t_0)] & \text{(vacuum)}. \end{cases} \quad (2.61)$$

We emphasize that these results assume (i) spatial flatness and (ii) a single-component cosmic fluid which must have the critical density.

Putting (2.61) into (2.14), we can solve for the bolometric intensity under the assumption that the luminosity of the galaxies is constant over their lifetimes, $L(t) = L_0$:

$$\frac{Q(t)}{Q_*} = \begin{cases} (1/3)(t/t_0)^{3/2} & \text{(radiation)} \\ (2/5)(t/t_0)^{5/3} & \text{(matter)} \\ \exp[H_0 t_0(t/t_0 - 1)] & \text{(vacuum)} \end{cases} \quad (2.62)$$

where we have used (2.25) and assumed that $t_f \ll t_0$ and $t_f \ll t$.

The intensity of the light reaching us from intergalactic space climbs as $t^{3/2}$ in a radiation-filled Universe, $t^{5/3}$ in a matter-dominated one, and $\exp(H_0 t)$ in one which contains only vacuum energy. This happens because the horizon of the Universe expands to encompass more and more galaxies, and hence more photons. Clearly it does so at a rate which more than compensates for the dilution and redshifting of existing photons due to expansion. Suppose, for argument's sake, that this state of affairs could continue indefinitely. How long would it take for the night sky to become as bright as a well-lit room ($Q \sim 1000 \text{ erg cm}^{-2} \text{ s}^{-1}$)?

The required increase of $Q(t)$ over Q_* ($= 2.5 \times 10^{-4} \text{ erg cm}^{-2} \text{ s}^{-1}$) is 4 million times. Equation (2.62) then implies that

$$t \approx \begin{cases} 780\,000 \text{ Gyr} & \text{(radiation)} \\ 240\,000 \text{ Gyr} & \text{(matter)} \\ 250 \text{ Gyr} & \text{(vacuum)} \end{cases} \quad (2.63)$$

where we have taken $H_0 t_0 \approx 1$ and $t_0 \approx 16 \text{ Gyr}$ as suggested by observational data (chapter 4).

The last of the numbers in (2.63) is particularly intriguing. In a vacuum-dominated model in which the luminosity of the galaxies could be kept constant indefinitely, the sky would fill up with light over timescales of the same order as the theoretical hydrogen-burning lifetimes of the longest-lived stars, those with $M \lesssim 0.3 M_\odot$ [12]. Of course, the luminosity of galaxies *cannot* stay constant over these timescales, because most of their light comes from much more massive stars which burn themselves out after tens of Gyr or less. Still, the closeness of these numbers prompts us to ask: is this steep increase in theoretical EBL intensity with

time a feature only of the pure de Sitter model; or does it also arise in models containing matter along with vacuum energy?

To answer this, we would like to find a simple expression for $R(t)$ in models with both dustlike matter and vacuum energy. An analytic solution does exist for such models, if they are flat ($\Omega_{\Lambda,0} = 1 - \Omega_{m,0}$). Its usefulness goes well beyond the particular problem at hand and, since we have found it derived only in German [13], it is worth reproducing here. Taking $\rho_m(R)$ and ρ_Λ from (2.32) and using the definitions (2.38), we find that (2.59) leads to the following differential equation in place of (2.60):

$$\left(\frac{\dot{R}}{R}\right)^2 = H_0^2 \left[\Omega_{m,0} \left(\frac{R}{R_0}\right)^{-3} + \Omega_{\Lambda,0} \right]. \quad (2.64)$$

This is solved by making a change of variables from R to $u \equiv R/R_*$, where R_* is the value of the scale factor at the inflection point, $R_*/R_0 = (\Omega_{m,0}/2\Omega_{\Lambda,0})^{1/3}$. It is then found that

$$\dot{u}^2 = H_0^2 \Omega_{\Lambda,0} (u^2 + 2/u). \quad (2.65)$$

This may be inverted for the age of the Universe:

$$t = \int \frac{du}{\dot{u}} = \frac{1}{H_0 \sqrt{\Omega_{\Lambda,0}}} \int \frac{du}{\sqrt{u^2 + 2/u}}. \quad (2.66)$$

One further change of variable to $v \equiv u^3 + 1$ puts this integral into elementary form:

$$t = \frac{1}{3H_0 \sqrt{\Omega_{\Lambda,0}}} \int \frac{dv}{\sqrt{v^2 - 1}}. \quad (2.67)$$

Solving for $t = t(R)$ and reinverting to get $\tilde{R}(t) = R(t)/R_0$, we find

$$\tilde{R}(t) = \left[\sqrt{\frac{\Omega_{m,0}}{1 - \Omega_{m,0}}} \sinh \left(\frac{3}{2} \sqrt{1 - \Omega_{m,0}} H_0 t \right) \right]^{2/3} \quad (2.68)$$

where we have replaced $\Omega_{\Lambda,0}$ with $1 - \Omega_{m,0}$. This elegant formula has many uses and deserves to be known more widely, given the importance of vacuum-dominated models in modern cosmology. Differentiation with respect to time gives the Hubble expansion rate:

$$\tilde{H}(t) = \sqrt{1 - \Omega_{m,0}} \coth \left(\frac{3}{2} \sqrt{1 - \Omega_{m,0}} H_0 t \right). \quad (2.69)$$

This goes over to $\sqrt{\Omega_{\Lambda,0}}$ as $t \rightarrow \infty$, a result which (as noted in section 2.5) holds quite generally for models with $\Lambda > 0$. Alternatively, setting $\tilde{R} = (1 + z)^{-1}$ in (2.68) gives the age of the Universe at a redshift z :

$$t = \frac{2}{3H_0 \sqrt{1 - \Omega_{m,0}}} \sinh^{-1} \sqrt{\frac{1 - \Omega_{m,0}}{\Omega_{m,0}(1 + z)^3}}. \quad (2.70)$$

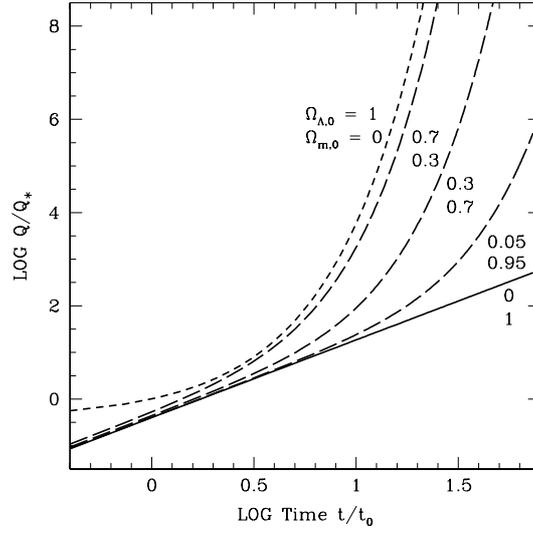


Figure 2.8. Plots of $Q(t)/Q_*$ in flat models containing both dustlike matter ($\Omega_{m,0}$) and vacuum energy ($\Omega_{\Lambda,0}$). The unbroken line is the Einstein–de Sitter model, while the short-dashed line is pure de Sitter, and long-dashed lines represent intermediate models. The curves do not meet at $t = t_0$ because $Q(t_0)$ differs from model to model, and $Q(t_0) = Q_*$ only for the pure de Sitter case.

In particular, this equation with $z = 0$ gives the present **age of the Universe** (t_0). Thus a model with (say) $\Omega_{m,0} = 0.3$ and $\Omega_{\Lambda,0} = 0.7$ has an age of $t_0 = 0.96H_0^{-1}$ or, using (2.19), $t_0 = 9.5h_0^{-1}$ Gyr. Alternatively, in the limit $\Omega_{m,0} \rightarrow 1$, equation (2.70) gives back the standard result for the age of an EdS Universe, $t_0 = 2/(3H_0) = 6.5h_0^{-1}$ Gyr.

Putting (2.68) into the bolometric intensity integral (2.14), holding $L(t) = L_0 = \text{constant}$ as usual, and integrating over time, we obtain the plots of EBL intensity $Q(t)/Q_*$ shown in figure 2.8. This diagram shows that the rapidly-brightening Universe is not a figment only of the pure de Sitter model (short-dashed line). In a less vacuum-dominated model with $\Omega_{m,0} = 0.3$ and $\Omega_{\Lambda,0} = 0.7$, for instance, the ‘reading-room’ intensity of $Q \sim 1000 \text{ erg cm}^{-2} \text{ s}^{-1}$ is still attained after $t \sim 290$ Gyr.

In theory then, it might be thought that our remote descendants could live under skies perpetually ablaze with the light of distant galaxies, in which the rising and setting of their home suns would barely be noticed. That this cannot happen in practice, as we have said, is due to the fact that galaxy luminosities necessarily change with time as their brightest stars burn out. The finite lifetime of the galaxies, in other words, is critical, *not only in the sense of a finite past, but a finite future*. A proper assessment of this requires that we move

from considerations of background cosmology to the astrophysics of the sources themselves. Real galaxies emit light with characteristic spectra which depend strongly on wavelength, and evolve in time as well. In the next chapter, we will adapt our bolometric formalism to take such features into account.

References

- [1] Wesson P S 1999 *Space–Time–Matter* (Singapore: World Scientific)
- [2] McVittie G C 1965 *General Relativity and Cosmology* (Urbana, IL: University of Illinois Press) p 164
- [3] Weinberg S 1972 *Gravitation and Cosmology* (New York: Wiley)
- [4] Wesson P S, Valle K and Stabell R 1987 *Astrophys. J.* **317** 601
- [5] Fukugita M, Hogan C J and Peebles P J E 1998 *Astrophys. J.* **503** 518
- [6] Yasuda N *et al* 2001 *Astron. J.* **122** 1104
- [7] Norberg P *et al* 2002 *Preprint astro-ph/0111011*
- [8] Felten J E 1966 *Astrophys. J.* **144** 241
- [9] Dicke R H 1970 *Gravitation and the Universe* (Philadelphia, PA: American Philosophical Society) p 62
- [10] Ellis G F R 1988 *Class. Quantum Grav.* **5** 891
- [11] White M and Scott D 1996 *Astrophys. J.* **459** 415
- [12] Alexander D R *et al* 1997 *Astron. Astrophys.* **317** 90
- [13] Blome H-J, Hoell J and Priester W 1997 *Bergmann-Schaefer: Lehrbuch der Experimentalphysik* vol 8 (Berlin: de Gruyter) pp 311–427

Chapter 3

The modern resolution and spectra

3.1 The spectral background

The previous chapter has convinced us that the bolometric intensity (or luminosity per unit area at all wavelengths) of the light reaching us is low mostly because the galaxies have not had enough time to fill up the Universe with light. Cosmic expansion does reduce this further, but only by a factor of about one-half. However, galaxies do not emit their light equally in all parts of the spectrum, and it could be that expansion plays a relatively larger role at some wavelengths than others.

This is perhaps best appreciated in the microwave region (at wavelengths from about 1 mm to 10 cm) where we know from decades of radio astronomy that the ‘night sky’ is brighter than its optical counterpart. The majority of this microwave background radiation is thought to come, not from the redshifted light of distant galaxies, but from the fading glow of the big bang itself—the ‘ashes and smoke’ of creation in Lemaître’s words. Since its nature and suspected origin are different from those of the EBL, this part of the extragalactic background has its own name, the **cosmic microwave background (CMB)**. Here expansion is of paramount importance, as emphasized by Peacock [1], since the source radiation in this case was emitted at more or less a single instant in cosmological history (so that the ‘lifetime of the sources’ is negligible). Another way to see this is to take expansion out of the picture, as we did in chapter 2: the CMB intensity we would observe in this ‘equivalent static model’ would be that of the primordial fireball and would roast us alive.

While Olbers’ paradox involves the EBL, not the CMB, this example is still instructive because it prompts us to consider whether similar (though less pronounced) effects could have been operative in the EBL as well. If, for instance, galaxies emitted most of their light in a relatively brief burst of star formation at very early times, this would be a galactic approximation to the picture just described, and could conceivably boost the importance of expansion relative to lifetime, at least in some wavebands. To check on this, we need a way to

calculate EBL intensity as a function of wavelength. This is motivated by other considerations as well. Olbers' paradox has historically been concerned primarily with the *optical* waveband (from approximately 4000 to 8000 Å), and this is still what most people mean when they refer to the 'brightness of the night sky'. And from a practical standpoint, we would like to compare our theoretical predictions with observational data, and these are necessarily taken using detectors which are optimized for finite portions of the electromagnetic spectrum. Our task in this chapter, then, is to convert the bolometric formalism of the previous chapter into a *spectral* one.

3.2 From bolometric to spectral intensity

Rather than considering the entire galaxy luminosity $L(t)$, we focus on the energy emitted per unit time between wavelengths λ and $\lambda + d\lambda$. Let us write this in the form $dL_\lambda \equiv F(\lambda, t) d\lambda$, where $F(\lambda, t)$ is the **spectral energy distribution** (SED) of the galaxy, with dimensions of energy per unit time per unit wavelength. Luminosity is recovered by integrating the SED over all wavelengths:

$$L(t) = \int_0^\infty dL_\lambda = \int_0^\infty F(\lambda, t) d\lambda. \quad (3.1)$$

We then return to (2.12), the bolometric intensity of the spherical shell of galaxies depicted in figure 2.1. Replacing $L(t)$ with dL_λ in this equation gives the intensity of light emitted between λ and $\lambda + d\lambda$:

$$dQ_{\lambda,e} = cn(t)\tilde{R}(t)[F(\lambda, t) d\lambda] dt. \quad (3.2)$$

This light reaches us at the redshifted wavelength $\lambda_0 = \lambda/\tilde{R}(t)$. Redshift also stretches the wavelength interval by the same factor, $d\lambda_0 = d\lambda/\tilde{R}(t)$. So the intensity of light *observed* by us between λ_0 and $\lambda_0 + d\lambda_0$ is

$$dQ_{\lambda_0} = cn(t)\tilde{R}^2(t)F[\tilde{R}(t)\lambda_0, t] d\lambda_0 dt. \quad (3.3)$$

The intensity of the shell *per unit wavelength*, as observed at wavelength λ_0 , is then given simply by

$$4\pi dI_\lambda(\lambda_0) \equiv \frac{dQ_{\lambda_0}}{d\lambda_0} = cn(t)\tilde{R}^2(t)F[\tilde{R}(t)\lambda_0, t] d\lambda_0 dt \quad (3.4)$$

where the factor 4π converts from an all-sky intensity to one measured per steradian. (This is merely a convention, but has become standard.) Integrating over all the spherical shells corresponding to cosmic times t_f and t_0 (as before) we obtain the spectral analogue of our earlier bolometric result, equation (2.14). It reads:

$$I_\lambda(\lambda_0) = \frac{c}{4\pi} \int_{t_f}^{t_0} n(t)F[\tilde{R}(t)\lambda_0, t]\tilde{R}^2(t) dt. \quad (3.5)$$

This is the integrated light from many galaxies, which has been emitted at various wavelengths and redshifted by various amounts, but which is all in the waveband centred on λ_0 when it arrives at us. We refer to this as the **spectral intensity of the EBL** at λ_0 . Equation (3.5), or ones like it, have been considered from the theoretical side principally by McVittie and Wyatt [2], Whitrow and Yallop [3,4] and Wesson *et al* [5,6].

We may convert this integral over time t to one over redshift z using (2.17) as before. This gives

$$I_\lambda(\lambda_0) = \frac{c}{4\pi H_0} \int_0^{z_f} \frac{n(z)F[\lambda_0/(1+z), z] dz}{(1+z)^3 \tilde{H}(z)} \quad (3.6)$$

which is the spectral analogue of (2.20). It may be checked using (3.1) that bolometric intensity is just the integral of spectral intensity over all observed wavelengths, $Q = \int_0^\infty I(\lambda_0) d\lambda_0$. In subsequent chapters we will apply equations (3.5) and (3.6) to calculate the intensity of the EBL, not only from galaxies but from many other sources of radiation as well.

The static analogue (i.e. the equivalent spectral EBL intensity in a Universe without expansion, but with the properties of galaxies unchanged) is also readily obtained in exactly the same way as before (section 2.6). Setting $\tilde{R}(t) = 1$ in (3.5), we obtain

$$I_{\lambda, \text{stat}}(\lambda_0) = \frac{c}{4\pi} \int_{t_f}^{t_0} n(t)F(\lambda_0, t) dt. \quad (3.7)$$

Just as in the bolometric case, we may convert this to an integral over z if we choose. The latter parameter no longer represents physical redshift (since this has been eliminated by hypothesis), but is now merely an algebraic way of expressing the age of the galaxies. This is convenient because it puts (3.7) into a form which may be directly compared with its counterpart (3.6) in the expanding Universe:

$$I_{\lambda, \text{stat}}(\lambda_0) = \frac{c}{4\pi H_0} \int_0^{z_f} \frac{n(z)F(\lambda_0, z) dz}{(1+z)\tilde{H}(z)}. \quad (3.8)$$

If the same values are adopted for H_0 and z_f , and the same functional forms are used for $n(z)$, $F(\lambda, z)$ and $\tilde{H}(z)$, then equations (3.6) and (3.8) allow us to compare model universes which are alike in every way, except that one is expanding while the other stands still.

Some simplification of these expressions is obtained as before in situations where the comoving source number density can be taken as constant, $n(z) = n_0$. However, it is not possible to go farther and pull all the dimensional content out of these integrals, as was done in the bolometric case, until a specific form is postulated for the source SED $F(\lambda, z)$.

In the remainder of chapter 3 we take up this problem, modelling galaxy spectra with several different SEDs and using (3.6) to calculate the resulting optical EBL intensity. Our goals in this exercise are threefold. First, we wish to build up experience with the simpler kinds of SEDs. These will provide a

check of our main results in this chapter, and allow us to model a variety of less conventional radiating sources in later chapters. Second, we would like to get some idea of upper and lower limits on the spectral intensity of the EBL itself, and compare them to the observational data. Third, we return to our original question and divide $I_\lambda(\lambda_0)$ by its static analogue $I_{\lambda,\text{stat}}(\lambda_0)$, as given by (3.8), in order to obtain a quantitative estimate of the importance of *expansion* in the spectral EBL and the resolution of Olbers' paradox.

3.3 The delta-function spectrum

The simplest possible source spectrum is one in which all the energy is emitted at a single peak wavelength λ_p at each redshift z , thus

$$F(\lambda, z) = F_p(z) \delta\left(\frac{\lambda}{\lambda_p} - 1\right). \quad (3.9)$$

The function $F_p(z)$ is obtained in terms of the total source luminosity $L(z)$ by normalizing over all observed wavelengths

$$L(z) \equiv \int_0^\infty F(\lambda, z) d\lambda = F_p(z) \lambda_p \quad (3.10)$$

so that $F_p(z) = L(z)/\lambda_p$. This SED is well suited to sources of electromagnetic radiation such as **elementary particle decays**, which are characterized by specific decay energies and may occur in the dark-matter halos surrounding galaxies. The use of a δ -function is not a very good approximation for the spectra of galaxies, but we will apply it here in this context to lay the foundation for later sections.

Since galaxies shine by starlight, a logical choice for the characteristic wavelength λ_p would be the peak wavelength of a blackbody of 'typical' stellar temperature. Taking the Sun as typical ($T = T_\odot = 5770$ K), this would be $\lambda_p = (0.290 \text{ cm K})/T = 5020 \text{ \AA}$ from **Wiens' law**. Distant galaxies are seen chiefly during periods of intense starburst activity when many stars are much hotter than the Sun, and this would suggest a shift toward shorter wavelengths. But any such effect must be largely offset by the fact that most of the short-wavelength light produced in large starbursting galaxies (as much as 99% in the most massive cases) is absorbed within these galaxies by dust and reradiated in the infrared and microwave regions ($\lambda \gtrsim 10\,000 \text{ \AA}$). It is also important to keep in mind that while distant starburst galaxies are hotter and more luminous than local spirals and ellipticals, the latter are likely to dominate the observed spectrum of the EBL by virtue of their numbers, especially at low redshift. The best that one can do with a single characteristic wavelength is to locate it somewhere within the B-band (3600–5500 \AA). For the purposes of this exercise we associate λ_p with the nominal centre of this band, $\lambda_p = 4400 \text{ \AA}$, corresponding to a blackbody temperature of 6590 K.

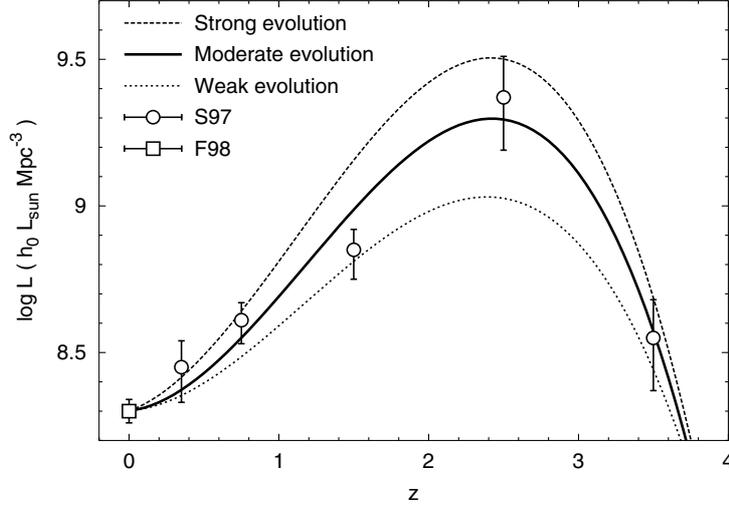


Figure 3.1. The comoving luminosity density of the Universe, as measured at $z = 0$ (F98 [8]) and extrapolated to higher redshifts using galaxies in the Hubble Deep Field (S97 [9]). The unbroken curve is a least-squares fit to the data, while the dashed lines represent upper and lower limits.

Substituting the galaxy SED (3.9) into the spectral intensity integral (3.6) leads to

$$I_\lambda(\lambda_0) = \frac{c}{4\pi H_0 \lambda_p} \int_0^{z_f} \frac{\mathcal{L}(z)}{(1+z)^3 \dot{H}(z)} \delta \left[\frac{\lambda_0}{\lambda_p(1+z)} - 1 \right] dz \quad (3.11)$$

where we have introduced a new shorthand for the **comoving luminosity density** of sources:

$$\mathcal{L}(z) \equiv n(z)L(z). \quad (3.12)$$

For galaxies at redshift $z = 0$ this takes the value \mathcal{L}_0 as given by (2.24). Numerous studies have shown that the product of $n(z)$ and $L(z)$ is approximately conserved with redshift, even when the two quantities themselves appear to be evolving markedly. So it would be reasonable to take $\mathcal{L}(z) = \mathcal{L}_0 = \text{constant}$. However, the latest analyses have been able to benefit from new observational work at deeper redshifts, and a new consensus is emerging that $\mathcal{L}(z)$ does rise slowly but steadily with z , peaking in the range $2 \lesssim z \lesssim 3$, and falling away sharply thereafter [7]. This would be consistent with a picture in which the first generation of massive galaxy formation occurred near $z \sim 3$, being followed at lower redshifts by galaxies whose evolution proceeded more passively.

Figure 3.1 shows the value of \mathcal{L}_0 from (2.24) at $z = 0$ (as taken from Fukugita *et al* [8]), together with the extrapolation of $\mathcal{L}(z)$ to five higher redshifts

($z = 0.35, 0.75, 1.5, 2.5$ and 3.5) as inferred from analysis of photometric galaxy redshifts in the Hubble Deep Field (HDF) by Sawicki *et al* [9]. (Most of the galaxies in this extremely deep sample are too faint for their Doppler or spectroscopic redshifts to be identified.) Let us define a **relative comoving luminosity density** $\tilde{\mathcal{L}}(z)$ by

$$\tilde{\mathcal{L}}(z) \equiv \mathcal{L}(z)/\mathcal{L}_0 \quad (3.13)$$

and fit this to the data with a cubic [$\log \tilde{\mathcal{L}}(z) = \alpha z + \beta z^2 + \gamma z^3$]. The resulting least-squares best fit is shown in figure 3.1 together with plausible upper and lower limits. We will refer to these cases as the ‘moderate’, ‘strong’ and ‘weak’ galaxy evolution scenarios respectively.

Inserting (3.13) into (3.11) makes the latter read:

$$I_\lambda(\lambda_0) = I_\delta \int_0^{z_f} \frac{\tilde{\mathcal{L}}(z)}{(1+z)^3 \tilde{H}(z)} \delta \left[\frac{\lambda_0}{\lambda_p(1+z)} - 1 \right] dz. \quad (3.14)$$

The dimensional content of this integral has been concentrated into a prefactor I_δ , defined by

$$I_\delta = \frac{c\mathcal{L}_0}{4\pi H_0 \lambda_p} = 4.4 \times 10^{-9} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1} \text{ ster}^{-1} \left(\frac{\lambda_p}{4400 \text{ \AA}} \right)^{-1}. \quad (3.15)$$

This constant shares two important properties of its bolometric counterpart Q_* (section 2.3). First, it is explicitly independent of the uncertainty h_0 in Hubble’s constant. Secondly, it is *low* by terrestrial standards. For example, it is well below the intensity of the faint glow known as **zodiacal light**, which is caused by the scattering of sunlight by dust in the plane of the Solar System. This is important, since the value of I_δ sets the scale of the integral (3.14). Indeed, existing observational *bounds* on $I_\lambda(\lambda_0)$ at $\lambda_0 \approx 4400 \text{ \AA}$ are of the same order as I_δ . Toller, for example, set an upper limit of $I_\lambda(4400 \text{ \AA}) < 4.5 \times 10^{-9} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1} \text{ ster}^{-1}$ using data from the Pioneer 10 photopolarimeter [10].

Dividing I_δ of (3.15) by the photon energy $E_0 = hc/\lambda_0$ (where $hc = 1.986 \times 10^{-8} \text{ erg \AA}$) puts the EBL intensity integral (3.14) into new units, sometimes referred to as **continuum units** (CUs):

$$I_\delta = I_\delta(\lambda_0) = \frac{\mathcal{L}_0}{4\pi h H_0} \left(\frac{\lambda_0}{\lambda_p} \right) = 970 \text{ CUs} \left(\frac{\lambda_0}{\lambda_p} \right) \quad (3.16)$$

where $1 \text{ CU} \equiv 1 \text{ photon s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1} \text{ ster}^{-1}$. While both kinds of units (CUs and $\text{erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1} \text{ ster}^{-1}$) are in common use for reporting spectral intensity at near-optical wavelengths, CUs appear most frequently. They are also preferable from a theoretical point of view, because they most faithfully reflect the *energy content* of a spectrum (as emphasized by Henry [11]). A third type

Table 3.1. Cosmological test models

	EdS/SCDM	OCDM	Λ CDM	Λ BDM
$\Omega_{m,0}$	1	0.3	0.3	0.03
$\Omega_{\Lambda,0}$	0	0	0.7	1
k	0	-1	0	+1

of intensity unit, the S_{10} (loosely, the equivalent of one tenth-magnitude star per square degree) is also occasionally encountered, but will be avoided in this book as it is wavelength-dependent and involves other subtleties which differ between workers.

If we let the redshift of formation $z_f \rightarrow \infty$ for simplicity, then equation (3.14) reduces to

$$I_\lambda(\lambda_0) = \begin{cases} I_\delta \left(\frac{\lambda_0}{\lambda_p}\right)^{-2} \frac{\tilde{\mathcal{L}}(\lambda_0/\lambda_p - 1)}{\tilde{H}(\lambda_0/\lambda_p - 1)} & (\text{if } \lambda_0 \geq \lambda_p) \\ 0 & (\text{if } \lambda_0 < \lambda_p). \end{cases} \quad (3.17)$$

The comoving luminosity density $\tilde{\mathcal{L}}(\lambda_0/\lambda_p - 1)$ which appears here is given by the fit (3.13) to the HDF data. The Hubble parameter is given by (2.40) as $\tilde{H}(\lambda_0/\lambda_p - 1) = [\Omega_{m,0}(\lambda_0/\lambda_p)^3 + \Omega_{\Lambda,0} - (\Omega_{m,0} + \Omega_{\Lambda,0} - 1)(\lambda_0/\lambda_p)^2]^{1/2}$ for a Universe containing dustlike matter and vacuum energy with density parameters $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$ respectively.

‘Turning off’ the luminosity density evolution (so that $\tilde{\mathcal{L}} = 1 = \text{constant}$), one can obtain three trivial special cases with

$$I_\lambda(\lambda_0) = I_\delta \times \begin{cases} (\lambda_0/\lambda_p)^{-7/2} & (\text{Einstein–de Sitter}) \\ (\lambda_0/\lambda_p)^{-2} & (\text{de Sitter}) \\ (\lambda_0/\lambda_p)^{-3} & (\text{Milne}). \end{cases} \quad (3.18)$$

This is taken at $\lambda_0 \geq \lambda_p$, where $(\Omega_{m,0}, \Omega_{\Lambda,0}) = (1, 0)$, $(0, 1)$ and $(0, 0)$ respectively for the three models cited (section 2.7). The first of these is the ‘ $\frac{7}{2}$ -law’ which appears frequently in the particle-physics literature as an approximation to the spectrum of EBL contributions from decaying particles. But the second (de Sitter) probably provides a better approximation, given current thinking regarding the values of $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$.

To evaluate the spectral EBL intensity (3.14) and other quantities in a general situation, it will be helpful to define a suite of cosmological test models which span the widest range possible in the parameter space defined by $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$ (table 3.1). Detailed discussion of these is left to chapter 4, but we can summarize the main rationale for each briefly as follows. The **Einstein–de Sitter** (EdS) model has long been favoured on grounds of simplicity and was alternatively

known for some time as the ‘standard cold dark matter’ or SCDM model. It has come under increasing pressure, however, as evidence mounts for levels of $\Omega_{m,0} \lesssim 0.5$ and, most recently, from observations of Type Ia supernovae (SNIa) which indicate that $\Omega_{\Lambda,0} > \Omega_{m,0}$. The **open cold dark matter** (OCDM) model is more consistent with data on $\Omega_{m,0}$ and holds appeal for those who have been reluctant to accept the possibility of a non-zero vacuum energy. It faces the considerable challenge, however, of explaining new data on the spectrum of CMB fluctuations, which imply that $\Omega_{m,0} + \Omega_{\Lambda,0} \approx 1$. The **Λ + cold dark matter** (Λ CDM) model has rapidly become the new standard in cosmology because it agrees best with both the SNIa and CMB observations. However, this model suffers from a ‘coincidence problem’, in that $\Omega_m(t)$ and $\Omega_\Lambda(t)$ evolve so differently with time that the probability of finding ourselves at a moment in cosmic history when they are even of the same order of magnitude appears unrealistically small. This is addressed to some extent in the last model, where we push $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$ to their lowest and highest limits, respectively. In the case of $\Omega_{m,0}$ these limits are set by big-bang nucleosynthesis, which requires a density of at least $\Omega_{m,0} \approx 0.03$ in baryons (hence the **Λ + baryonic dark matter** or Λ BDM model). Upper limits on $\Omega_{\Lambda,0}$ come from various arguments, such as the observed frequency of gravitational lenses and the requirement that the Universe began in a big-bang singularity. Within the context of isotropic and homogeneous cosmology, these four models cover the full range of what would be considered plausible by most workers.

Figure 3.2 shows the solution of the full integral (3.14) for all four test models. The short-wavelength cut-off in these plots is an artefact of the δ -function SED, but the behaviour of $I_\lambda(\lambda_0)$ at wavelengths above $\lambda_p = 4400 \text{ \AA}$ is quite revealing, even in a model as simple as this one. In the EdS case (a), the rapid fall-off in intensity with λ_0 indicates that *nearby* (low-redshift) galaxies dominate. There is a secondary hump at $\lambda_0 \approx 10\,000 \text{ \AA}$, which is an ‘echo’ of the peak in galaxy formation, redshifted into the near infrared. This hump becomes successively enlarged relative to the optical peak at 4400 \AA as the value of $\Omega_{m,0}$ drops relative to $\Omega_{\Lambda,0}$. Eventually one has the situation in the de Sitter-like model (d), where the galaxy-formation peak entirely dominates the observed EBL signal, despite the fact that it comes from distant galaxies at $z \approx 3$. This is because a large $\Omega_{\Lambda,0}$ -term (especially one which is large relative to $\Omega_{m,0}$) inflates comoving volume at high redshifts. Since the comoving *number density* of galaxies is fixed by the fit to $\tilde{\mathcal{L}}(z)$ (figure 3.1), the number of galaxies at these redshifts has no choice but to go up, pushing up the infrared part of the spectrum. Although the δ -function spectrum is an unrealistic one, we will see in subsequent sections that this trend persists in more sophisticated models, providing a clear link between observations of the EBL and the cosmological parameters $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$.

Figure 3.2 is plotted over a broad range of wavelengths from the near ultraviolet (NUV; $2000\text{--}4000 \text{ \AA}$) to the near infrared (NIR; $8000\text{--}40\,000 \text{ \AA}$). We have included a number of reported observational constraints on EBL intensity. These require a bit of comment as regards their origin. Upper limits (full symbols

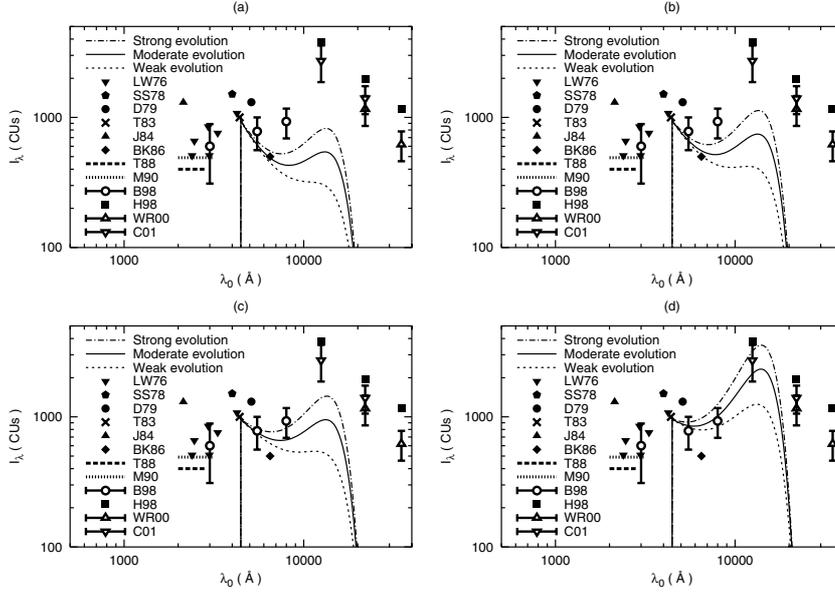


Figure 3.2. The spectral EBL intensity of galaxies whose radiation is modelled by δ -functions at a rest frame wavelength of 4400 \AA , calculated for four different cosmological models: (a) EdS, (b) OCDM, (c) Λ CDM and (d) Λ BDM (table 3.1). Also shown are observational upper limits (full symbols and bold lines) and reported detections (empty symbols) over the waveband $2000\text{--}40\,000 \text{ \AA}$.

and bold lines) have come from analyses of OAO-2 satellite data (LW76 [12]), ground-based telescopes (SS78 [13], D79 [14], BK86 [15]), the aforementioned Pioneer 10 instrument (T83 [10]), sounding rockets (J84 [16], T88 [17]), the Hopkins UVX experiment aboard the Space Shuttle (M90 [18]) and—most recently, in the near infrared—the DIRBE instrument on the COBE satellite (H98 [19]). The past two or three years have also seen the first widely-accepted *detections* of the EBL (figure 3.2, open symbols). In the NIR these have come from continued analysis of DIRBE data in the K-band ($22\,000 \text{ \AA}$) and L-band ($35\,000 \text{ \AA}$; WR00 [20]), as well as the J-band ($12\,500 \text{ \AA}$; C01 [21]). Reported detections in the optical using a combination of Hubble Space Telescope (HST) and Las Campanas telescope observations (B98 [22, 23]) are still preliminary but very important, and we have included them as well.

Figure 3.2 shows that EBL intensities based on the simple δ -function spectrum are in rough agreement with the data. Predicted intensities come in at or just below the optical limits in the low- $\Omega_{\Lambda,0}$ cases (a, b) and remain consistent with the infrared limits even in the high- $\Omega_{\Lambda,0}$ cases (c, d). Vacuum-dominated

models with even higher ratios of $\Omega_{\Lambda,0}$ to $\Omega_{m,0}$ would, however, run afoul of DIRBE limits in the J-band. We will find a similar trend in models with more realistic galaxy spectra.

3.4 Gaussian spectra

The Gaussian distribution provides a useful generalization of the δ -function for modelling sources whose spectra, while essentially monochromatic, are broadened by some physical process. For example, photons emitted by the decay of elementary particles inside dark-matter halos would have their energies Doppler-broadened by the circular velocity $v_c \approx 220 \text{ km s}^{-1}$ of their host galaxies, giving rise to a spread σ_λ of order $\sigma_\lambda(\lambda) = (2v_c/c)\lambda \approx 0.0015\lambda$ in the SED. In the context of galaxies, this extra degree of freedom provides a simple way to model the width of the bright part of the spectrum. If we take this to cover the B-band (3600–5500 Å) then $\sigma_\lambda \sim 1000 \text{ Å}$. The Gaussian SED reads:

$$F(\lambda, z) = \frac{L(z)}{\sqrt{2\pi}\sigma_\lambda} \exp\left[-\frac{1}{2}\left(\frac{\lambda - \lambda_p}{\sigma_\lambda}\right)^2\right] \quad (3.19)$$

where λ_p is the wavelength at which the galaxy emits most of its light. We take $\lambda_p = 4400 \text{ Å}$ as before, and note that integration over λ_0 confirms that $L(z) = \int_0^\infty F(\lambda, z) d\lambda$ as required. Once again we can make the simplifying assumption that $L(z) = L_0 = \text{constant}$; or we can use the empirical fit $\tilde{\mathcal{L}}(z) \equiv n(z)L(z)/\mathcal{L}_0$ to the HDF data of Sawicki *et al* [9]. Taking the latter course and substituting (3.19) into (3.6), we obtain

$$I_\lambda(\lambda_0) = I_g \int_0^{z_f} \frac{\tilde{\mathcal{L}}(z)}{(1+z)^3 \tilde{H}(z)} \exp\left\{-\frac{1}{2}\left[\frac{\lambda_0/(1+z) - \lambda_p}{\sigma_\lambda}\right]^2\right\} dz. \quad (3.20)$$

The dimensional content of this integral has been pulled into a prefactor $I_g = I_g(\lambda_0)$, defined by

$$I_g = \frac{\mathcal{L}_0}{\sqrt{32\pi^3} h H_0} \left(\frac{\lambda_0}{\sigma_\lambda}\right) = 390 \text{ CUs} \left(\frac{\lambda_0}{\sigma_\lambda}\right). \quad (3.21)$$

Here we have divided (3.20) by the photon energy $E_0 = hc/\lambda_0$ to put the result into CUs, as before.

Results are shown in figure 3.3, where we have taken $\lambda_p = 4400 \text{ Å}$, $\sigma_\lambda = 1000 \text{ Å}$ and $z_f = 6$. Aside from the fact that the short-wavelength cut-off has disappeared, the situation is qualitatively similar to that obtained using a δ -function approximation. (This similarity becomes formally exact as σ_λ approaches zero.) One sees, as before, that the expected EBL signal is brightest at optical wavelengths in an EdS Universe (*a*), but that the infrared hump due to the redshifted peak of galaxy formation begins to dominate for higher- $\Omega_{\Lambda,0}$ models

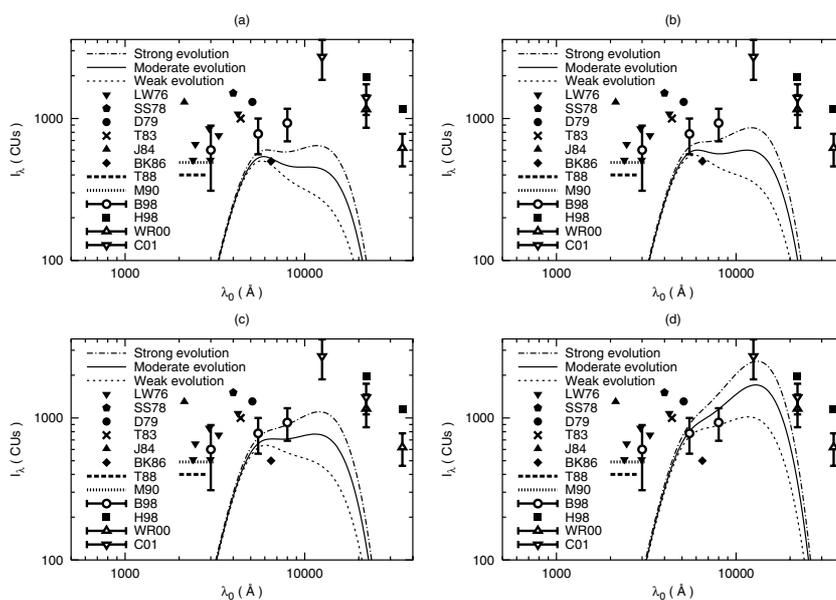


Figure 3.3. The spectral EBL intensity of galaxies whose spectra has been represented by Gaussian distributions with rest-frame peak wavelength 4400 \AA and standard deviation 1000 \AA , calculated for the (a) EdS, (b) OCDM, (c) Λ CDM and (d) Δ BDM cosmologies and compared with observational upper limits (full symbols and bold lines) and reported detections (empty symbols).

(b) and (c), becoming overwhelming in the de Sitter-like model (d). Overall, the best agreement between calculated and observed EBL levels occurs in the Λ CDM model (c). The matter-dominated EdS (a) and OCDM (b) models appear to contain too little light (requiring one to postulate an additional source of optical or near-optical background radiation besides that from galaxies), while the Δ BDM model (d) comes uncomfortably close to containing *too much* light. This is an interesting situation and one which motivates us to reconsider the problem with more realistic models for the galaxy SED.

3.5 Blackbody spectra

The simplest non-trivial approach to a galaxy spectrum is to model it as a blackbody, and this was done by previous workers such as McVittie and Wyatt [2], Whitrow and Yallop [3, 4] and Wesson [6]. Let us suppose that the galaxy SED is a product of the **Planck function** and some wavelength-independent parameter

$C(z)$:

$$F(\lambda, z) = \frac{2\pi hc^2}{\sigma_{\text{SB}}} \frac{C(z)/\lambda^5}{\exp[hc/kT(z)\lambda] - 1}. \quad (3.22)$$

Here $\sigma_{\text{SB}} \equiv 2\pi^5 k^4 / 15c^2 h^3 = 5.67 \times 10^{-5} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ K}^{-1}$ is the Stefan–Boltzmann constant. The function F is normally regarded as an increasing function of redshift (at least out to the redshift of galaxy formation). This can, in principle, be accommodated by allowing $C(z)$ or $T(z)$ to increase with z in (3.22). The former choice would correspond to a situation in which galaxy luminosity decreases with time while its spectrum remains unchanged, as might happen if stars were simply to die. The second choice corresponds to a situation in which galaxy luminosity decreases with time as its spectrum becomes redder, as may happen when its stellar population ages. The latter scenario is more realistic and will be adopted here. The luminosity $L(z)$ is found by integrating $F(\lambda, z)$ over all wavelengths:

$$L(z) = \frac{2\pi hc^2}{\sigma_{\text{SB}}} C(z) \int_0^\infty \frac{\lambda^{-5} d\lambda}{\exp[hc/kT(z)\lambda] - 1} = C(z)[T(z)]^4 \quad (3.23)$$

so that the unknown function $C(z)$ must satisfy $C(z) = L(z)/[T(z)]^4$. If we require that **Stefan’s law** ($L \propto T^4$) hold at each z , then

$$C(z) = \text{constant} = L_0/T_0^4 \quad (3.24)$$

where T_0 is the present ‘galaxy temperature’ (i.e. the blackbody temperature corresponding to a peak wavelength in the B-band). Thus the evolution of galaxy luminosity in this model is just that which is required by Stefan’s law for blackbodies whose *temperatures* evolve as $T(z)$. This is reasonable, since galaxies are made up of stellar populations which cool and redden with time as hot massive stars die out.

Let us supplement this with the minimal assumption that galaxy comoving number density is conserved with redshift, $n(z) = n_0 = \text{constant}$. This is sometimes referred to as the **pure luminosity evolution** or PLE scenario; and while there is some controversy on this point, PLE has been found by many workers to be roughly consistent with observed numbers of galaxies at faint magnitudes, especially if there is a significant vacuum energy density $\Omega_{\Lambda,0} > 0$. Proceeding on this assumption, the comoving galaxy luminosity density can be written

$$\tilde{\mathcal{L}}(z) \equiv \frac{n(z)L(z)}{\mathcal{L}_0} = \frac{L(z)}{L_0} = \left[\frac{T(z)}{T_0} \right]^4. \quad (3.25)$$

This expression can then be inverted to give the blackbody temperature $T(z)$ as a function of redshift, since the form of $\tilde{\mathcal{L}}(z)$ is fixed by our fit to the photometric HDF data (figure 3.1):

$$T(z) = T_0[\tilde{\mathcal{L}}(z)]^{1/4}. \quad (3.26)$$

We can check this by choosing $T_0 = 6600$ K (i.e. a present peak wavelength of 4400 \AA) and reading off values of $\tilde{\mathcal{L}}(z) = \mathcal{L}(z)/\mathcal{L}_0$ at the peaks of the curves marked ‘weak’, ‘moderate’ and ‘strong’ evolution in figure 3.1. Putting these numbers into (3.26) yields blackbody temperatures (and corresponding peak wavelengths) of $10\,000$ K (2900 \AA), $11\,900$ K (2440 \AA) and $13\,100$ K (2210 \AA) respectively at the galaxy-formation peak. These results are consistent with the idea that galaxies would have been dominated by hot UV-emitting stars at this early time.

Inserting the expressions (3.24) for $C(z)$ and (3.26) for $T(z)$ into the SED (3.22), and substituting the latter into the EBL integral (3.6), we obtain

$$I_\lambda(\lambda_0) = I_b \int_0^{z_f} \frac{(1+z)^2 dz}{\{\exp[hc(1+z)/kT(z)\lambda_0] - 1\} \tilde{H}(z)}. \quad (3.27)$$

The dimensional prefactor $I_b = I_b(T_0, \lambda_0)$ reads in this case

$$I_b = \frac{c^2 \mathcal{L}_0}{2H_0 \sigma_{\text{SB}} T_0^4 \lambda_0^4} = 90\,100 \text{ CUs} \left(\frac{T_0}{6600 \text{ K}} \right)^{-4} \left(\frac{\lambda_0}{4400 \text{ \AA}} \right)^{-4}. \quad (3.28)$$

Numerically, the argument of the exponential term may be expressed in the form $hc(1+z)/kT(z)\lambda_0 = 4.95(1+z)/[T(z)/6600 \text{ K}]/(\lambda_0/4400 \text{ \AA})$.

Results are shown in figure 3.4, where we have set $z_f = 6$ following recent observational hints of an epoch of ‘first light’ at this redshift [24]. Overall EBL intensity is, however, quite insensitive to this choice, provided that $z_f \gtrsim 3$. Between $z_f = 3$ and $z_f = 6$, $I_\lambda(\lambda_0)$ rises by less than 1% below $\lambda_0 = 10\,000 \text{ \AA}$ and less than $\sim 5\%$ at $\lambda_0 = 20\,000 \text{ \AA}$ (where most of the signal originates at high redshifts). There is no further increase beyond $z_f > 6$ at the three-figure level of precision.

Figure 3.4 shows some qualitative differences from our earlier results obtained using the δ -function and Gaussian SEDs. Most noticeably, the prominent ‘double-hump’ structure is no longer apparent. The key evolutionary parameter is now blackbody temperature $T(z)$ and this goes as $[\mathcal{L}(z)]^{1/4}$ so that individual features in the comoving luminosity density profile are suppressed. (A similar effect can be achieved with the Gaussian SED by choosing larger values of σ_λ .) One can make out the same general trend as before by focusing on the long-wavelength tail in each panel. This climbs steadily up the right-hand side of the figure as one moves from the $\Omega_{\Lambda,0} = 0$ models (a) and (b) to the $\Omega_{\Lambda,0}$ -dominated models (c) and (d), showing that more and more light is arriving from distant, highly redshifted galaxies and boosting the NIR part of the spectrum.

Absolute EBL intensities in each of these four models are consistent with what we have seen already. This is not surprising because changing the shape of the SED merely shifts light from one part of the spectrum to another. It cannot alter the *total amount* of light in the EBL, which is set by the comoving luminosity density $\tilde{\mathcal{L}}(z)$ of sources, once the background cosmology (and hence the source lifetime) has been chosen. As before, the best match between calculated EBL

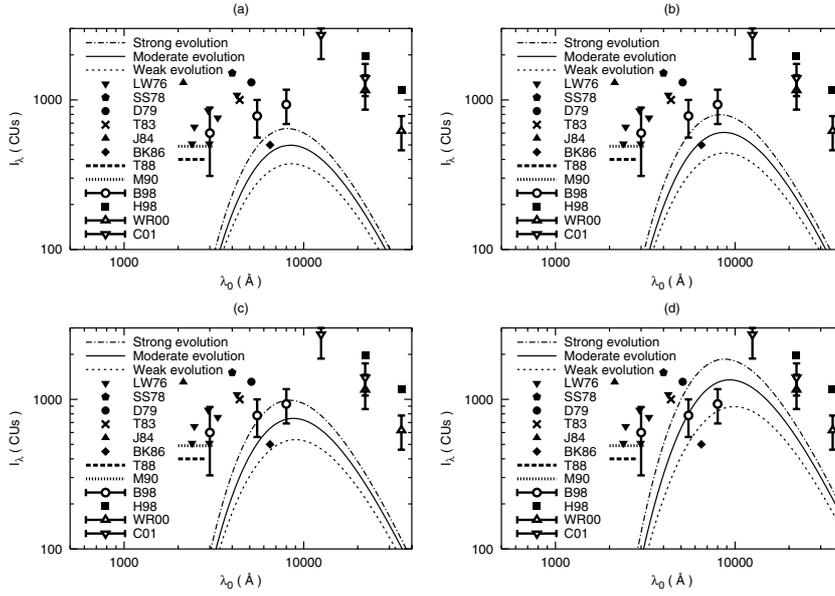


Figure 3.4. The spectral EBL intensity of galaxies, modelled as blackbodies whose characteristic temperatures are such that their luminosities $L \propto T^4$ combine to produce the observed comoving luminosity density $\mathcal{L}(z)$ of the Universe. Results are shown for the (a) EdS, (b) OCDM, (c) Λ CDM and (d) Λ BDM cosmologies. Also shown are observational upper limits (full symbols and bold lines) and reported detections (open symbols).

intensities and the observational detections is found for the $\Omega_{\Lambda,0}$ -dominated models (c) and (d). The fact that the EBL is now spread across a broader spectrum has pulled down its peak intensity slightly, so that the Λ BDM model (d) no longer threatens to violate observational limits and, in fact, fits them rather nicely. The zero- $\Omega_{\Lambda,0}$ models (a) and (b) would again appear to require some additional source of background radiation (beyond that produced by galaxies) if they are to contain enough light to make up the levels of EBL intensity that have been reported.

3.6 Normal and starburst galaxies

The previous sections have shown that simple models of galaxy spectra, combined with data on the evolution of comoving luminosity density in the Universe, can produce levels of spectral EBL intensity in rough agreement with observational limits and reported detections, and even discriminate to a degree between different cosmological models. However, the results obtained up to this point are somewhat

unsatisfactory in that they are sensitive to theoretical input parameters, such as λ_p and T_0 , which are hard to connect with the properties of the actual galaxy population.

A more comprehensive approach would use observational data in conjunction with theoretical models of galaxy evolution to build up an ensemble of evolving galaxy SEDs $F(\lambda, z)$ and comoving number densities $n(z)$ which would depend not only on redshift but on *galaxy type* as well. Increasingly sophisticated work has been carried out along these lines over the years by Partridge and Peebles [25], Tinsley [26], Bruzual [27], Code and Welch [28], Yoshii and Takahara [29] and others. The last-named authors, for instance, divided galaxies into five morphological types (E/SO, Sab, Sbc, Scd and Sdm), with a different evolving SED for each type, and found that their collective EBL intensity at NIR wavelengths was about an order of magnitude below the levels suggested by observation.

Models of this kind, however, are complicated while at the same time containing uncertainties. This makes their use somewhat incompatible with our purpose here, which is primarily to obtain a first-order estimate of EBL intensity so that the importance of *expansion* can be properly ascertained. Also, the increasingly deeper observations of recent years have begun to make it clear that these morphological classifications are of limited value at redshifts $z \gtrsim 1$, where spirals and ellipticals are still in the process of forming [30]. As we have already seen, this is precisely where much of the EBL may originate, especially if luminosity density evolution is strong or if there is a significant $\Omega_{\Lambda,0}$ -term.

What is needed, then, is a simple model which does not distinguish too finely between the spectra of galaxy types as they have traditionally been classified, but which can capture the essence of broad trends in luminosity density evolution over the full range of redshifts $0 \leq z \leq z_f$. For this purpose we will group together the traditional classes (spiral, elliptical, etc) under the single heading of quiescent or **normal galaxies**. At higher redshifts ($z \gtrsim 1$), we will allow a second class of objects to play a role: the active or **starburst galaxies**. Whereas normal galaxies tend to be comprised of older, redder stellar populations, starburst galaxies are dominated by newly-forming stars whose energy output peaks in the ultraviolet (although much of this is absorbed by dust grains and subsequently reradiated in the infrared). One signature of the starburst type is thus a decrease in $F(\lambda)$ as a function of λ over NUV and optical wavelengths, while normal types show an increase [31]. Starburst galaxies also tend to be brighter, reaching bolometric luminosities as high as 10^{12} – $10^{13} L_{\odot}$, versus 10^{10} – $10^{11} L_{\odot}$ for normal types.

There are two ways to obtain SEDs for these objects: by reconstruction from observational data; or as output from theoretical models of galaxy evolution. The former approach has had some success but becomes increasingly difficult at short wavelengths, so that results have typically been restricted to $\lambda \gtrsim 1000 \text{ \AA}$ [31]. This represents a serious limitation if we want to integrate out to redshifts $z_f \sim 6$ (say), since it means that our results are only strictly reliable down to $\lambda_0 = \lambda(1 + z_f) \sim 7000 \text{ \AA}$. In order to integrate out to $z_f \sim 6$ and still go

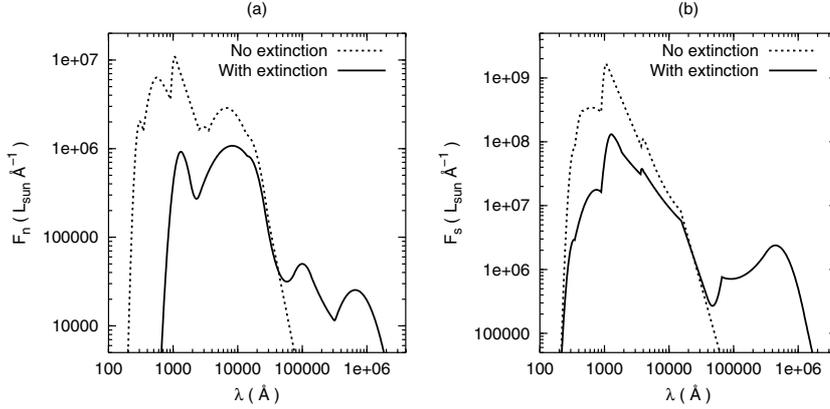


Figure 3.5. Typical galaxy SEDs for (a) normal and (b) starburst type galaxies with and without extinction by dust. These figures are adapted from figures 9 and 10 of Devriendt *et al* [32], where, however, the plotted quantity is $\log[\lambda F(\lambda)]$, not $F(\lambda)$ as shown here. For definiteness we have normalized (over $100 - 3 \times 10^4 \text{ \AA}$) such that $L_n = 1 \times 10^{10} h_0^{-2} L_\odot$ and $L_s = 2 \times 10^{11} h_0^{-2} L_\odot$ with $h_0 = 0.75$. (These values are consistent with what we will later call ‘model 0’ for a comoving galaxy number density of $n_0 = 0.010 h_0^3 \text{ Mpc}^{-3}$.)

down as far as the NUV ($\lambda_0 \sim 2000 \text{ \AA}$), we require SEDs which are good to $\lambda \sim 300 \text{ \AA}$ in the galaxy rest-frame. For this purpose we will make use of theoretical galaxy-evolution models, which have advanced to the point where they cover the entire spectrum from the far ultraviolet to radio wavelengths. This broad range of wavelengths involves diverse physical processes such as star formation, chemical evolution and (of special importance here) dust absorption of ultraviolet light and re-emission in the infrared. Typical normal and starburst galaxy SEDs based on such models are now available down to $\sim 100 \text{ \AA}$ [32]. These functions, displayed in figure 3.5, will constitute our normal and starburst **galaxy SEDs**, $F_n(\lambda)$ and $F_s(\lambda)$.

Figure 3.5 shows the expected increase in $F_n(\lambda)$ with λ at NUV wavelengths ($\sim 2000 \text{ \AA}$) for normal galaxies, as well as the corresponding decrease for starbursts. What is most striking about both templates, however, is their overall multi-peaked structure. These objects are far from pure blackbodies, and the primary reason for this is *dust*. This effectively removes light from the shortest-wavelength peaks (which are due mostly to star formation), and transfers it to the longer-wavelength ones. The dotted lines in figure 3.5 show what the SEDs would look like if this dust reprocessing were ignored. The main difference between normal and starburst types lies in the relative importance of this process. Normal galaxies emit as little as 30% of their bolometric intensity in the infrared, while the equivalent fraction for the largest starburst galaxies can reach 99%. (The

latter are, in fact, usually detected because of their infrared emission, despite the fact that they contain mostly hot, blue stars.) In the full model of Devriendt *et al* [32], these variations are incorporated by modifying input parameters such as star formation timescale and gas density, leading to spectra which are broadly similar in shape to those in figure 3.5 (though they differ in normalization and ‘tilt’ toward longer wavelengths). The results, it should be emphasized, have been successfully matched to a wide range of real galaxy spectra. Here we will be content with the single pair of prototypical SEDs shown in figure 3.5.

We now proceed to calculate the EBL intensity using $F_n(\lambda)$ and $F_s(\lambda)$, with the characteristic luminosities of these two types found as usual by normalization, $\int F_n(\lambda) d\lambda = L_n$ and $\int F_s(\lambda) d\lambda = L_s$. Let us assume that the comoving luminosity density of the Universe at any redshift z is a combination of normal and starburst components

$$\mathcal{L}(z) = n_n(z)L_n + n_s(z)L_s \quad (3.29)$$

where the comoving number densities are

$$n_n(z) \equiv [1 - f(z)]n(z) \quad \text{and} \quad n_s(z) \equiv f(z)n(z). \quad (3.30)$$

In other words, we will account for evolution in $\mathcal{L}(z)$ solely in terms of the changing **starburst fraction** $f(z)$, and a single comoving number density $n(z)$ as before. L_n and L_s are awkward to work with for dimensional reasons, and we will find it more convenient to specify the SED instead by two dimensionless adjustable parameters, the **local starburst fraction** f_0 and **luminosity ratio** ℓ_0 :

$$f_0 \equiv f(0) \quad \ell_0 \equiv L_s/L_n. \quad (3.31)$$

Observations indicate that $f_0 \approx 0.05$ in the local population [31] and Devriendt *et al* were able to fit their templates to a range of normal and starburst galaxies with $40 \lesssim \ell_0 \lesssim 890$ [32]. We will allow these two parameters to vary in the ranges $0.01 \leq f_0 \leq 0.1$ and $10 \leq \ell_0 \leq 1000$. This, in combination with our ‘strong’ and ‘weak’ limits on comoving luminosity-density evolution, gives us the flexibility to obtain upper and lower bounds on EBL intensity.

The functions $n(z)$ and $f(z)$ can now be fixed by equating $\mathcal{L}(z)$ as defined by (3.29) to the comoving luminosity-density curves inferred from HDF data (figure 3.1), and requiring that $f \rightarrow 1$ at peak luminosity (i.e. assuming that the galaxy population is entirely starburst-dominated at the redshift z_p of peak luminosity). These conditions are not difficult to set up. One finds that modest number-density evolution is required in general, if $f(z)$ is not to over or undershoot unity at z_p . We follow [33] and parametrize this with the function $n(z) = n_0(1+z)^\eta$ for $z \leq z_p$. Here η can be termed the **merger parameter** since a value of $\eta > 0$ would imply that the comoving number density of galaxies decreases with time.

Pulling these requirements together, one obtains a model with

$$f(z) = \begin{cases} \left(\frac{1}{\ell_0 - 1}\right) [\ell_0(1+z)^{-\eta} \mathcal{N}(z) - 1] & (z \leq z_p) \\ 1 & (z > z_p) \end{cases}$$

$$n(z) = n_0 \times \begin{cases} (1+z)^\eta & (z \leq z_p) \\ \mathcal{N}(z) & (z > z_p). \end{cases} \quad (3.32)$$

Here $\mathcal{N}(z) \equiv [1/\ell_0 + (1 - 1/\ell_0)f_0]\tilde{\mathcal{L}}(z)$ and $\eta = \ln[\mathcal{N}(z_p)]/\ln(1+z_p)$. The evolution of $f(z)$, $n_n(z)$ and $n_s(z)$ is plotted in figure 3.6 for five models: a best-fit **model 0**, corresponding to the moderate evolution curve in figure 3.1 with $f_0 = 0.05$ and $\ell_0 = 20$, and four other models chosen to produce the widest possible spread in EBL intensities across the optical band. **Models 1** and **2** are the *most* starburst-dominated, with initial starburst fraction and luminosity ratio at their upper limits ($f_0 = 0.1$ and $\ell_0 = 1000$). **Models 3** and **4** are the *least* starburst-dominated, with the same quantities at their lower limits ($f_0 = 0.01$ and $\ell_0 = 10$). Luminosity density evolution is set to ‘weak’ in the odd-numbered models 1 and 3, and ‘strong’ in the even-numbered models 2 and 4. (In principle one could identify four other ‘extreme’ combinations, such as maximum f_0 with minimum ℓ_0 , but these will be intermediate to models 1–4.) We find merger parameters η between +0.4, 0.5 in the strong-evolution models 2 and 4, and –0.5, –0.4 in the weak-evolution models 1 and 3, while $\eta = 0$ for model 0. These are well within the usual range [34].

The information contained in figure 3.6 can be summarized in words as follows: starburst galaxies formed near $z_f \sim 4$ and increased in comoving number density until $z_p \sim 2.5$ (the redshift of peak comoving luminosity density in figure 3.1). They then gave way to a steadily growing population of fainter normal galaxies which began to dominate between $1 \lesssim z \lesssim 2$ (depending on the model) and now make up 90–99% of the total galaxy population at $z = 0$. This is a reasonable scenario and agrees with others that have been constructed to explain the observed **faint blue excess** in galaxy number counts [35].

We are now in a position to compute the total spectral EBL intensity, which is the sum of normal and starburst components:

$$I_\lambda(\lambda_0) = I_\lambda^n(\lambda_0) + I_\lambda^s(\lambda_0). \quad (3.33)$$

These components are found as usual by substituting the SEDs (F_n , F_s) and comoving number densities (3.30) into (3.6). This leads to:

$$I_\lambda^n(\lambda_0) = I_{ns} \int_0^{z_f} \tilde{n}(z)[1 - f(z)]F_n\left(\frac{\lambda_0}{1+z}\right) \frac{dz}{(1+z)^3 \tilde{H}(z)}$$

$$I_\lambda^s(\lambda_0) = I_{ns} \int_0^{z_f} \tilde{n}(z)f(z)F_s\left(\frac{\lambda_0}{1+z}\right) \frac{dz}{(1+z)^3 \tilde{H}(z)}. \quad (3.34)$$

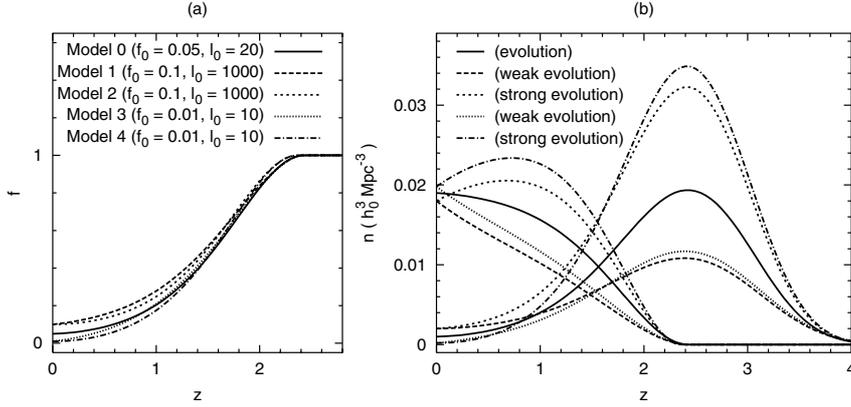


Figure 3.6. Evolution of (a) starburst fraction $f(z)$ and (b) comoving normal and starburst galaxy number densities $n_n(z)$ and $n_s(z)$, where total comoving luminosity density $\mathcal{L}(z) = n_n(z)L_n + n_s(z)L_s$ is matched to the ‘moderate’, ‘weak’ and ‘strong’ evolution curves in figure 3.1. Model 0 lies midway between models 1–4, which have maximum and minimum possible values of the two adjustable parameters $f_0 \equiv f(0)$ and $\ell_0 \equiv L_s/L_n$.

Here $\tilde{n}(z) \equiv n(z)/n_0$ is the **relative comoving number density** and the dimensional content of both integrals has been pulled into the prefactor

$$I_{\text{ns}} = I_{\text{ns}}(\lambda_0) = \frac{\mathcal{L}_0}{4\pi h H_0} \left(\frac{\lambda_0}{\text{\AA}} \right) = 970 \text{ CUs} \left(\frac{\lambda_0}{\text{\AA}} \right). \quad (3.35)$$

This is explicitly independent of h_0 , as before, and it is important to realize how this comes about. There is an implicit factor of L_0 contained in the galaxy SEDs $F_n(\lambda)$ and $F_s(\lambda)$ via their *normalizations* (i.e. via the fact that we require the galaxy population to make up the observed comoving luminosity density of the Universe). To see this, note that equation (3.29) reads $\mathcal{L}_0 = n_0 L_n [1 + (\ell_0 - 1)f_0]$ at $z = 0$. Since $\mathcal{L}_0 \equiv n_0 L_0$, it follows that $L_n = L_0 / [1 + (\ell_0 - 1)f_0]$ and $L_s = L_0 \ell_0 / [1 + (\ell_0 - 1)f_0]$. If the factor of L_0 is divided out of these expressions when normalizing the functions F_n and F_s , it can be put directly into the integrals (3.34), forming the quantity \mathcal{L}_0 in (3.35) and cancelling out the factor of h_0 in H_0 as required.

The spectral intensity (3.33) is plotted in figure 3.7, where we have set $z_f = 6$ as usual. (Results are insensitive to this choice, increasing by less than 5% as one moves from $z_f = 3$ to $z_f = 6$, with no further increase for $z_f \geq 6$ at three-figure precision.) These plots show that the most starburst-dominated models (1 and 2) produce the bluest EBL spectra, as might be expected. For these two models, EBL contributions from normal galaxies remain well below those from starbursts at all wavelengths, so that the bump in the observed spectrum at

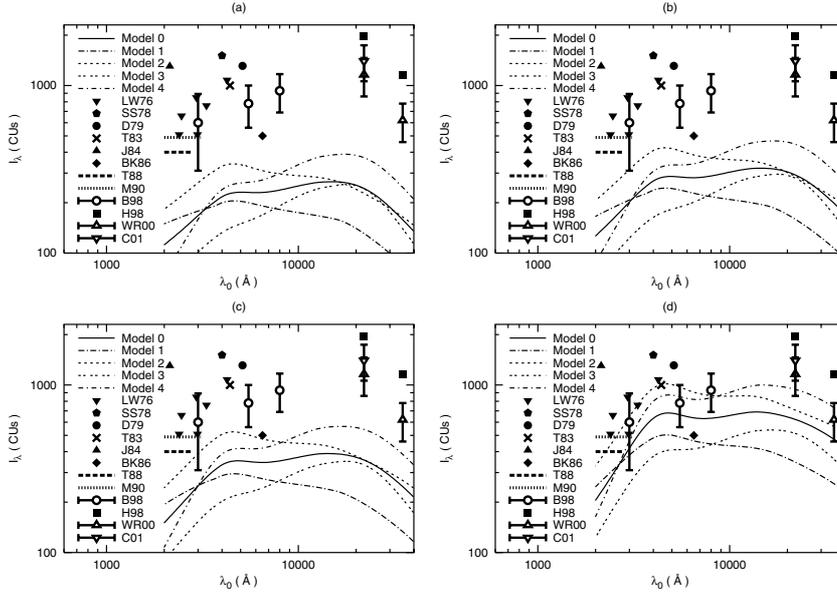


Figure 3.7. The spectral EBL intensity of a combined population of normal and starburst galaxies, with SEDs as shown in figure 3.5. The evolving number densities are such as to reproduce the total comoving luminosity density seen in the HDF (figure 3.1). Results are shown for the (a) EdS, (b) OCDM, (c) Λ CDM and (d) Λ BDM cosmologies. Also shown are observational upper limits (full symbols and bold lines) and reported detections (open symbols).

$\lambda_0 \sim 4000 \text{ \AA}$ is essentially an echo of the peak at $\sim 1100 \text{ \AA}$ in the starburst SED (figure 3.5), redshifted by a factor $(1 + z_p)$ from the epoch $z_p \approx 2.5$ of maximum comoving luminosity density. In the least starburst-dominated models (3 and 4), by contrast, EBL contributions from normal galaxies catch up to and exceed those from starbursts at $\lambda_0 \gtrsim 10000 \text{ \AA}$, giving rise to the bump seen at $\lambda_0 \sim 20000 \text{ \AA}$ in these models. Absolute EBL intensities are highest in the strong-evolution models (2 and 4) and lowest in the weak-evolution models (1 and 3). We emphasize that, for a given cosmology, the total amount of light in the EBL is set by a choice of luminosity density profile. A choice of SED merely shifts this from one part of the spectrum to another. Within the context of the simple two-component model adopted here, and the constraints imposed on luminosity density by the HDF data (section 3.3), the curves in figure 3.7 represent upper and lower limits on the spectral intensity of the EBL at near-optical wavelengths.

These curves are spread over a broader range of wavelengths than those obtained earlier using single-component Gaussian and blackbody spectra. This leads to a drop in peak intensity, as we can appreciate by noting that there now

appears to be a significant gap between theory and observation in all but the most vacuum-dominated cosmology, Λ BDM (d). This is so even for the models with the strongest luminosity density evolution (models 2 and 4). In the case of the EdS cosmology (a), this gap is nearly an order of magnitude, as found by Yoshii and Takahara [29]. Similar conclusions have been reached more recently from an analysis of Subaru Deep Field data by Totani *et al* [36], who suggest that the shortfall could be made up by a very diffuse, previously undetected component of background radiation not associated with galaxies. However, other workers have argued that existing galaxy populations are enough to explain the data, if different assumptions are made about their SEDs [37] or if allowance is made for faint **low-surface-brightness galaxies** below the detection limit of existing surveys [23]. In this connection it may be worthwhile to remember that there are preliminary indications from the Sloan Digital Sky Survey that the traditional value of \mathcal{L}_0 (as we have used it here) may go up by as much as 40% when corrected for a downward bias which is inherent in the way that galaxy magnitudes have been measured at high redshift [38].

3.7 Back to Olbers

Having obtained quantitative estimates of the spectral intensity of the near-optical EBL which are in reasonable agreement with observation, let us return to the question with which we began: why precisely is the sky dark at night? By ‘dark’ we now mean specifically dark at *near-optical wavelengths*. We can provide a quantitative answer to this question by using a spectral version of the bolometric argument in chapter 2. That is, we compute the EBL intensity $I_{\lambda, \text{stat}}$ in model universes which are equivalent to expanding ones in every way *except* expansion, and then take the ratio $I_{\lambda}/I_{\lambda, \text{stat}}$. If this is of order unity, then expansion plays a minor role and the darkness of the optical sky (like the bolometric one) must be attributed mainly to the fact that the Universe is too young to have filled up with light. If $I_{\lambda}/I_{\lambda, \text{stat}} \ll 1$, however, then we would have a situation qualitatively different from the bolometric one, and expansion would play a crucial role in the resolution to Olbers’ paradox.

The spectral EBL intensity for the equivalent static model is obtained by putting the functions $\tilde{n}(z)$, $f(z)$, $F_n(\lambda)$, $F_s(\lambda)$ and $\tilde{H}(z)$ into (3.8) rather than (3.6). This results in

$$I_{\lambda, \text{stat}}(\lambda_0) = I_{\lambda, \text{stat}}^n(\lambda_0) + I_{\lambda, \text{stat}}^s(\lambda_0) \quad (3.36)$$

where the normal and starburst contributions are given by

$$\begin{aligned} I_{\lambda, \text{stat}}^n(\lambda_0) &= I_{\text{ns}} F_n(\lambda_0) \int_0^{z_f} \frac{\tilde{n}(z)[1 - f(z)] dz}{(1+z)\tilde{H}(z)} \\ I_{\lambda, \text{stat}}^s(\lambda_0) &= I_{\text{ns}} F_s(\lambda_0) \int_0^{z_f} \frac{\tilde{n}(z)f(z) dz}{(1+z)\tilde{H}(z)}. \end{aligned} \quad (3.37)$$

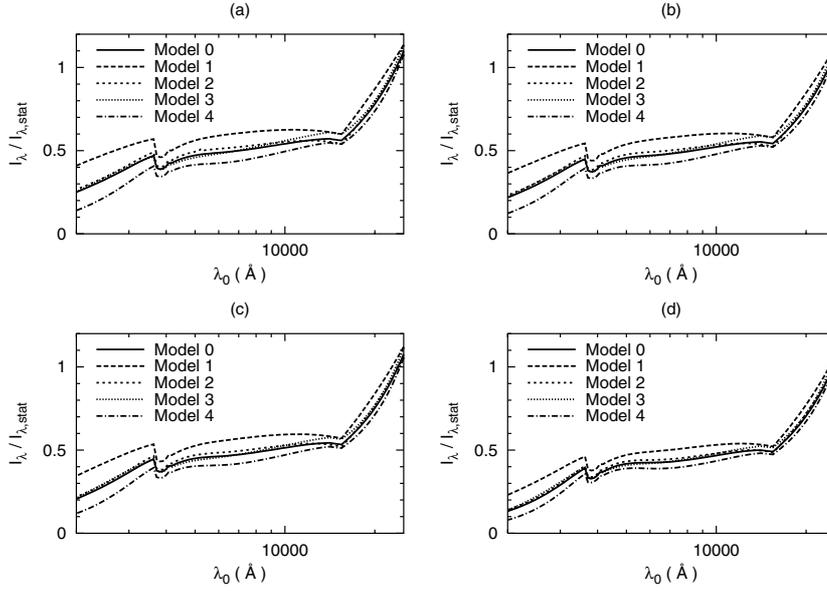


Figure 3.8. The ratio $I_\lambda/I_{\lambda,\text{stat}}$ of spectral EBL intensity in expanding models to that in equivalent static models, for the (a) EdS, (b) OCDM, (c) Λ CDM and (d) Λ BDM models. The fact that this ratio lies between 0.3 and 0.6 in the B-band (4000–5000 Å) tells us that expansion reduces the intensity of the night sky at optical wavelengths by a factor of between two and three.

Despite a superficial resemblance to their counterparts (3.34) in the expanding Universe, these are vastly different expressions. Most importantly, the SEDs $F_n(\lambda_0)$ and $F_s(\lambda_0)$ no longer depend on z and have been pulled out of the integrals. The quantity $I_{\lambda,\text{stat}}(\lambda_0)$ is effectively a *weighted mean* of the SEDs $F_n(\lambda_0)$ and $F_s(\lambda_0)$. The weighting factors (i.e. the integrals over z) are related to the age of the galaxies, $\int_0^{z_f} dz/(1+z)\tilde{H}(z)$, but modified by factors of $n_n(z)$ and $n_s(z)$ under the integral. This latter modification is important because it prevents the integrals from increasing without limit as z_f becomes arbitrarily large, a problem that would otherwise introduce considerable uncertainty into any attempt to put bounds on the ratio $I_{\lambda,\text{stat}}/I_\lambda$ [6]. A numerical check confirms that $I_{\lambda,\text{stat}}$ is nearly as insensitive to the value of z_f as I_λ , increasing by up to 8% as one moves from $z_f = 3$ to $z_f = 6$, but with no further increase as $z_f \geq 6$.

The ratio of $I_\lambda/I_{\lambda,\text{stat}}$ is plotted over 2000–25 000 Å in figure 3.8, where we have set $z_f = 6$. (Results are insensitive to this choice, as we have mentioned earlier and it may be noted that they are also independent of uncertainty in constants such as \mathcal{L}_0 since these are common to both I_λ and $I_{\lambda,\text{stat}}$.) Several features in this figure deserve notice. First, the average value of $I_\lambda/I_{\lambda,\text{stat}}$ across

the spectrum is about 0.6, consistent with bolometric expectations (chapter 2). Second, the diagonal, bottom-left to top-right orientation arises largely because $I_\lambda(\lambda_0)$ drops off at short wavelengths, while $I_{\lambda,\text{stat}}(\lambda_0)$ does so at long ones. The reason why $I_\lambda(\lambda_0)$ drops off at short wavelengths is that ultraviolet light reaches us only from the nearest galaxies; anything from more distant ones is redshifted into the optical. The reason why $I_{\lambda,\text{stat}}(\lambda_0)$ drops off at long wavelengths is because it is a weighted mixture of the galaxy SEDs, and drops off at exactly the same place that they do: $\lambda_0 \sim 3 \times 10^4 \text{ \AA}$. In fact, the weighting is heavily tilted toward the dominant starburst component, so that the two sharp bends apparent in figure 3.8 are essentially (inverted) reflections of features in $F_s(\lambda_0)$; namely, the small bump at $\lambda_0 \sim 4000 \text{ \AA}$ and the shoulder at $\lambda_0 \sim 11\,000 \text{ \AA}$ (figure 3.5).

Finally, the numbers: figure 3.8 shows that the ratio of $I_\lambda/I_{\lambda,\text{stat}}$ is remarkably consistent across the B-band (4000–5000 \AA) in all four cosmological models, varying from a high of 0.46 ± 0.10 in the EdS model to a low of 0.39 ± 0.08 in the Λ BDM model. These numbers should be compared with the bolometric result of $Q/Q_{\text{stat}} \approx 0.6 \pm 0.1$ from chapter 2. They tell us that expansion *does* play a greater role in determining B-band EBL intensity than it does across the spectrum as a whole—but not by much. If its effects were removed, the night sky at optical wavelengths would be anywhere from twice as bright (in the EdS model) to three times brighter (in the Λ BDM model). These results depend modestly on the makeup of the evolving galaxy population and figure 3.8 shows that $I_\lambda/I_{\lambda,\text{stat}}$ in every case is highest for the weak-evolution model 1, and lowest for the strong-evolution model 4. This is as we would expect, based on our discussion at the beginning of this chapter: models with the strongest evolution effectively ‘concentrate’ their light production over the shortest possible interval in time, so that the importance of the lifetime factor drops relative to that of expansion. Our numerical results, however, prove that this effect cannot qualitatively alter the resolution of Olbers’ paradox. Whether intensity is reduced by two times or three due to expansion, its *order of magnitude* is still set by the lifetime of the galaxies.

There is one factor which we have not considered in this chapter, and that is the possibility of **absorption** by intergalactic dust and neutral hydrogen, both of which are strongly absorbing at ultraviolet wavelengths. The effect of this would primarily be to remove ultraviolet light from high-redshift galaxies and transfer it into the infrared—light that would otherwise be redshifted into the optical and contribute to the EBL. (Dust extinction thus acts in the opposite sense to enhanced galaxy evolution, which effectively *injects* ultraviolet light at high redshifts.) The EBL intensity $I_\lambda(\lambda_0)$ would therefore drop, and one could expect reductions over the B-band in particular. The size of this effect is difficult to assess because we have limited data on the character and distribution of dust beyond our own galaxy. We will find indications in chapter 7, however, that the reduction could be significant at the shortest wavelengths considered here ($\lambda_0 \approx 2000 \text{ \AA}$) for the most extreme dust models. This would further widen the gap between observed and predicted EBL intensities noted at the end of section 3.6.

Absorption would play far less of a role in the equivalent static models,

where there is no redshift. (Ultraviolet light would still be absorbed, but the effect would not carry over into the optical.) Therefore, the ratio $I_\lambda/I_{\lambda,\text{stat}}$ would be expected to drop in nearly direct proportion to the drop in I_λ . In this sense Olbers may actually have had part of the solution after all—not (as he thought) because intervening matter ‘blocks’ the light from distant sources but because it transfers it out of the optical. The importance of this effect, which would be somewhere below that of expansion, is, however, a separate issue from the one we have concerned ourselves with in this chapter. What we have shown is that the optical sky, like the bolometric one, is dark at night *mainly because it has not had time to fill up with light from distant galaxies*. Cosmic expansion darkens it further by a factor which depends on background cosmology and galaxy evolution, but which lies between two and three in any case.

References

- [1] Peacock J A 1999 *Cosmological Physics* (Cambridge: Cambridge University Press) p 354
- [2] McVittie G C and Wyatt S P 1959 *Astrophys. J.* **130** 1
- [3] Whitrow G J and Yallop B D 1964 *Mon. Not. R. Astron. Soc.* **127** 130
- [4] Whitrow G J and Yallop B D 1965 *Mon. Not. R. Astron. Soc.* **130** 31
- [5] Wesson P S, Valle K and Stabell R 1987 *Astrophys. J.* **317** 601
- [6] Wesson P S 1991 *Astrophys. J.* **367** 399
- [7] Fukugita M, Hogan C J and Peebles P J E 1996 *Nature* **381** 489
- [8] Fukugita M, Hogan C J and Peebles P J E 1998 *Astrophys. J.* **503** 518
- [9] Sawicki M J, Lin H and Yee H K C 1997 *Astron. J.* **113** 1
- [10] Toller G N 1983 *Astrophys. J.* **266** L79
- [11] Henry R C 1999 *Astrophys. J.* **516** L49
- [12] Lillie C F and Witt A N 1976 *Astrophys. J.* **208** 64
- [13] Spinrad H and Stone R P S 1978 *Astrophys. J.* **226** 609
- [14] Dube R R, Wickes W C and Wilkinson D T 1979 *Astrophys. J.* **232** 333
- [15] Boughn S P and Kuhn J R 1986 *Astrophys. J.* **309** 33
- [16] Jakobsen P *et al* 1984 *Astron. Astrophys.* **139** 481
- [17] Tennyson P D *et al* 1988 *Astrophys. J.* **330** 435
- [18] Murthy J *et al* 1990 *Astron. Astrophys.* **231** 187
- [19] Hauser M G *et al* 1998 *Astrophys. J.* **508** 25
- [20] Wright E L and Reese E D 2000 *Astrophys. J.* **545** 43
- [21] Cambr esy L *et al* 2001 *Astrophys. J.* **555** 563
- [22] Bernstein R A 1998 *PhD Thesis* California Institute of Technology
- [23] Bernstein R A 1999 *The Low Surface Brightness Universe* (Astronomical Society of the Pacific Conference Series, Volume 170) ed J I Davies, C Impey and S Phillips (San Francisco, CA: ASP) p 341
- [24] Ellis R *et al* 2001 *Astrophys. J.* **560** L119
- [25] Partridge R B and Peebles P J E 1967 *Astrophys. J.* **148** 377
- [26] Tinsley N M 1973 *Astron. Astrophys.* **24** 89
- [27] Bruzual A G 1981 *PhD Thesis* University of California, Berkeley

- [28] Code A D and Welch G A 1982 *Astrophys. J.* **256** 1
- [29] Yoshii Y and Takahara F 1988 *Astrophys. J.* **326**
- [30] Abraham R G and van den Bergh S 2001 *Science* **293** 1273
- [31] Kinney A L *et al* 1996 *Astrophys. J.* **467** 38
- [32] Devriendt J E G, Guiderdoni B and Sadat R 1999 *Astron. Astrophys.* **350** 381
- [33] Totani T and Yoshii Y 2000 *Astrophys. J.* **540** 81
- [34] Totani T and Yoshii Y 1998 *Astrophys. J.* **501** L177
- [35] Pearson C and Rowan-Robinson M 1996 *Mon. Not. R. Astron. Soc.* **283** 174
- [36] Totani T *et al* 2001 *Astrophys. J.* **550** L137
- [37] Jimenez R and Kashlinsky A 1999 *Astrophys. J.* **511** 16
- [38] Yasuda N *et al* 2001 *Astron. J.* **122** 1104

Chapter 4

The dark matter

4.1 From light to dark matter

In calculating the intensity of the summed light from distant galaxies, we have effectively taken a census of one component of the Universe: its luminous matter. In astronomy, where nearly everything we know comes to us in the form of light signals from vast distances, one might be forgiven for thinking that luminous matter was the only kind that counted. This supposition, however, turns out to be spectacularly wrong. The density Ω_{lum} of luminous matter is now thought to comprise less than 1% of the total density Ω_{tot} of all forms of matter and energy put together. (Here as in chapters 2 and 3, we express densities in units of the critical density, equation (2.36), and denote them with the symbol Ω .) The remaining 99% or more consists of **dark matter** which, while not seen directly, is inferred to exist from its gravitational influence on luminous matter as well as the geometry of the Universe.

The identity of this unseen material, whose existence was first suspected in the 1920s and 1930s by astronomers such as Kapteyn [1], Oort [2] and Zwicky [3], has now become the central mystery of modern cosmology. In subsequent chapters, we will examine some of the theoretical dark-matter candidates which have been proposed, concentrating on the fact that many are *not perfectly black*. For those which can decay either directly or indirectly into photons (and this turns out to be most of them), we can use the same formalism we have applied to galaxies in chapters 2 and 3. That is, we can calculate the intensity of their contributions to the spectrum of background radiation reaching our Galaxy, and compare with observational bounds over wavebands stretching from long-wavelength radio waves to high-energy γ -rays. We will find that this procedure allows us to narrow down the list of plausible contenders quite substantially.

In the present chapter, we will outline what is known about dark matter in general, and how it is that we know it exists when it does not shine. Observation and experiment over the past few years have increasingly suggested that there are, in fact, *four distinct categories* of dark matter, three of which imply new

physics beyond the existing standard model of particle interactions. This is an unexpected development, and one whose supporting evidence deserves to be carefully scrutinized. We proceed to give a brief overview of the current situation, followed by a closer look at the arguments for all four parts of nature's 'dark side'.

4.2 The four elements of modern cosmology

At least some of the dark matter, such as that contained in planets and 'failed stars' too dim to see, must be composed of ordinary atoms and molecules. The same applies to dark gas and dust (although these can sometimes be seen in absorption, if not emission). Such contributions are termed **baryonic dark matter** (BDM), which combined together with luminous matter gives the *total baryonic* matter density $\Omega_{\text{bar}} \equiv \Omega_{\text{lum}} + \Omega_{\text{bdm}}$. If our understanding of big-bang theory and the formation of the light elements is correct, then we will see that Ω_{bar} cannot represent more than 5% of the critical density.

What, then, makes up the bulk of the Universe? Besides the dark baryons, it now appears that three other varieties of dark matter play a role. The first of these is **cold dark matter** (CDM), the existence of which has been inferred from the behaviour of visible matter on scales larger than the Solar System (e.g. galaxies and clusters of galaxies). CDM is thought to consist of particles (sometimes referred to as 'exotic' dark-matter particles) whose interactions with ordinary matter are so weak that they are seen primarily via their gravitational influence. While they have not been detected (and are indeed hard to detect by definition), such particles are predicted by plausible extensions of the standard model. The overall CDM density Ω_{cdm} is believed by many cosmologists to exceed that of the baryons (Ω_{bar}) by at least an order of magnitude.

Another piece of the puzzle is provided by **neutrinos**, particles whose existence is unquestioned but whose collective density (Ω_{ν}) depends on their rest mass, which is not yet known. If neutrinos are massless, or nearly so, then they remain relativistic throughout the history of the Universe and behave for dynamical purposes like photons. In this case neutrino contributions combine with those of photons (Ω_{γ}) to give the present radiation density as $\Omega_{\text{r},0} = \Omega_{\nu} + \Omega_{\gamma}$. This is known to be very small. If, however, neutrinos are sufficiently massive, then they are no longer relativistic on average, and belong together with baryonic and cold dark matter under the category of pressureless matter, with present density $\Omega_{\text{m},0} = \Omega_{\text{bar}} + \Omega_{\text{cdm}} + \Omega_{\nu}$. These neutrinos could play a significant dynamical role, especially in the formation of large-scale structure in the early Universe, where they are sometimes known as **hot dark matter** (HDM). Recent experimental evidence suggests that neutrinos do contribute to $\Omega_{\text{m},0}$ but at levels below those of the baryons.

Influential only over the largest scales—those comparable to the cosmological horizon itself—is the last component of the unseen Universe: **vacuum energy**. Its many alternative names (the zero-point field, quintessence,

dark energy and the cosmological constant Λ) betray the fact that there is, at present, no consensus as to where vacuum energy originates; or how to calculate its energy density (Ω_Λ) from first principles. Existing theoretical estimates of this latter quantity range over some 120 orders of magnitude, prompting many cosmologists until very recently to disregard it altogether. Observations of distant supernovae and CMB fluctuations, however, increasingly imply that vacuum energy is not only real but that its present energy density ($\Omega_{\Lambda,0}$) exceeds that of all other forms of matter ($\Omega_{m,0}$) and radiation ($\Omega_{r,0}$) put together.

If this account is correct, then the real Universe hardly resembles the one we see at night. It is composed, to a first approximation, of invisible vacuum energy whose physical origin remains obscure. A significant minority of its total energy density is found in CDM particles, whose ‘exotic’ nature is also not yet understood. Close inspection is needed to make out the further contribution of neutrinos, although this too is non-zero. And baryons, the stuff of which we are made, are little more than a cosmic afterthought. This picture, if confirmed, constitutes a revolution of Copernican proportions, for it is not only our location in space which turns out to be undistinguished, but our very makeup. The ‘four elements’ of modern cosmology are shown schematically in figure 4.1.

4.3 Baryons

Let us now go over the evidence for these four species of dark matter more carefully, beginning with the baryons. The total present density of *luminous* baryonic matter can be inferred from the observed luminosity density of the Universe, if various reasonable assumptions are made about the fraction of galaxies of different morphological type, their ratios of disc-type to bulge-type stars, and so on. A recent and thorough such estimate is that of Fukugita *et al* [4]:

$$\Omega_{\text{lum}} = (0.0027 \pm 0.0014)h_0^{-1}. \quad (4.1)$$

Here h_0 is, as usual, the value of **Hubble’s constant** expressed in units of $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$. While this parameter (and hence the experimental uncertainty in H_0) factored out of the EBL intensities in chapters 2 and 3, it must be squarely faced where densities are concerned. We therefore digress briefly to discuss the observational status of h_0 .

Using various relative-distance methods, all calibrated against the distance to Cepheid variables in the Large Magellanic Cloud (LMC), the Hubble Key Project (HKP) team has determined that $h_0 = 0.72 \pm 0.08$ [5]. Independent ‘absolute’ methods (e.g. time delays in gravitational lenses, the Sunyaev–Zeldovich effect in the CMB, and the Baade–Wesselink method applied to supernovae) have higher uncertainties but are roughly consistent with this, giving $h_0 \approx 0.55\text{--}0.74$ [6]. This level of agreement is a great improvement over the factor-two discrepancies of previous decades.

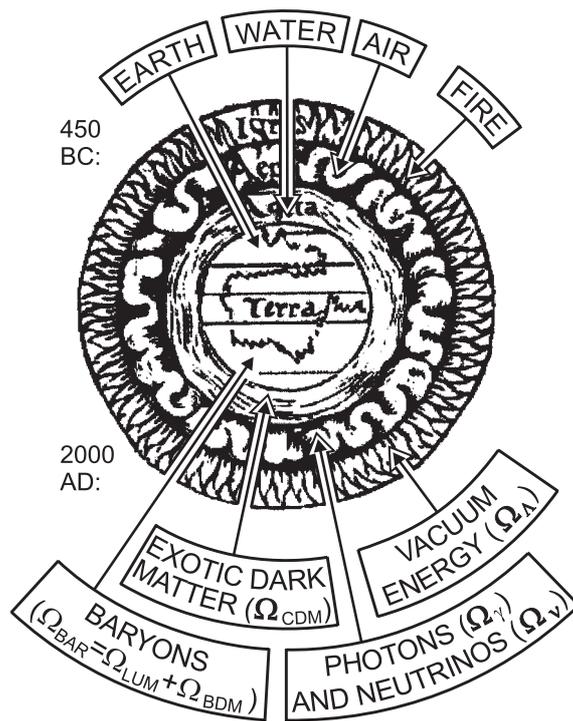


Figure 4.1. Top: the ‘four elements’ of ancient cosmology. These are widely attributed to the Greek philosopher Empedocles, who introduced them as follows, ‘Hear first the roots of all things: bright Zeus (fire), life-giving air (Hera), Aidoneus (earth) and Nestis (water), who moistens the springs of mortals with her tears’ (*Fragments*, c. 450 BC). Bottom: the modern counterparts. (Figure adapted from a 1519 edition of Aristotle’s *De caelo* by Johann Eck.)

There are signs, however, that we are still some way from ‘precision’ values with uncertainties of less than 10%. A recalibrated LMC Cepheid period–luminosity relation based on a much larger sample (from the OGLE microlensing survey) leads to considerably higher values, namely $h_0 = 0.85 \pm 0.05$ [7]. A new, purely geometric technique, based on the use of long-baseline radio interferometry to measure the transverse velocity of water masers [8], also implies that the traditional calibration is off, raising all Cepheid-based estimates by $12 \pm 9\%$ [9]. This would boost the HKP value to $h_0 = 0.81 \pm 0.09$. There is some independent support for such a recalibration in recent observations of eclipsing binaries [10] and ‘red-clump stars’ in the LMC [11], although this—like most conclusions involving the value of h_0 —has been disputed [12]. If the LMC-based numbers do go up, then one would be faced with explaining the lower h_0 values from the absolute methods. Here the choice of cosmological model

(e.g. as defined in table 3.1) can be important. Gravitational lensing values of h_0 (routinely reported assuming an EdS model) rise on average by 9% in OCDM, 7% in Λ CDM and 3% in Λ BDM for the six lens systems whose time delays have been measured so far.

On this subject, history encourages caution. Where it is necessary to specify the value of h_0 in this book, we will adopt the range

$$h_0 = 0.75 \pm 0.15. \quad (4.2)$$

Values at the edges of this range can discriminate powerfully between different cosmological models. To a large extent this is a function of their *ages*, which can be computed by integrating (2.44) or (in the case of flat models) directly from (2.70). Alternatively, one can integrate the Friedmann–Lemaître equation (2.40) numerically backward in time. Since this equation defines the expansion rate $H \equiv \dot{R}/R$, its integral gives the scale factor $R(t)$. We show the results in figure 4.2 for our four standard cosmological test models (EdS, OCDM, Λ CDM and Λ BDM). These are seen to have ages of $7h_0^{-1}$, $8h_0^{-1}$, $10h_0^{-1}$ and $17h_0^{-1}$ Gyr respectively. Abundances of radioactive thorium and uranium in metal-poor halo stars imply a Milky Way age 16 ± 5 Gyr [13], setting a firm lower limit of 11 Gyr on the age of the Universe. If h_0 lies at the upper end of its allowed range ($h_0 = 0.9$), then the EdS and OCDM models would be ruled out on the basis that they are not old enough to contain these stars (this is known as the **age crisis** in low- $\Omega_{\Lambda,0}$ models). With h_0 at the bottom of the range ($h_0 = 0.6$), however, only EdS comes close to being excluded. The EdS model thus defines one edge of the spectrum of observationally viable models.

The Λ BDM model faces the opposite problem: figure 4.2 gives its age as $17h_0^{-1}$ Gyr, or as high as 28 Gyr (if $h_0 = 0.6$). The latter number, in particular, is well beyond the age of anything seen in our Galaxy. Of course, upper limits on age are not as easy to set as lower ones. But following Copernican reasoning, we do not expect to live in a galaxy which is unusually young. The age of a ‘typical’ galaxy can be estimated by recalling from chapter 3 that most galaxies appear to have formed in the redshift range $2 \lesssim z_f \lesssim 4$. Since $R/R_0 = (1+z)^{-1}$ from (2.15), this corresponds to a range of scale factors $0.33 \gtrsim R/R_0 \gtrsim 0.2$. For the Λ BDM model, figure 4.2 shows that $R(t)/R_0$ does not reach these values until $(5 \pm 2)h_0^{-1}$ Gyr after the big bang. Thus galaxies would have an age of about $(12 \pm 2)h_0^{-1}$ Gyr in the Λ BDM model, and not more than 21 Gyr in any case. This is close to upper limits which have been set on the age of the Universe in models of this type, $t_0 < 24 \pm 2$ Gyr [14]. The Λ BDM model, or something close to it, probably defines a position opposite that of EdS on the spectrum of observationally viable models.

For the other three models, figure 4.2 shows that galaxy formation must be accomplished within less than 2 Gyr after the big bang. The reason this is able to occur so quickly is that these models all contain significant amounts of CDM, which decouples from the primordial plasma before the baryons and prepares

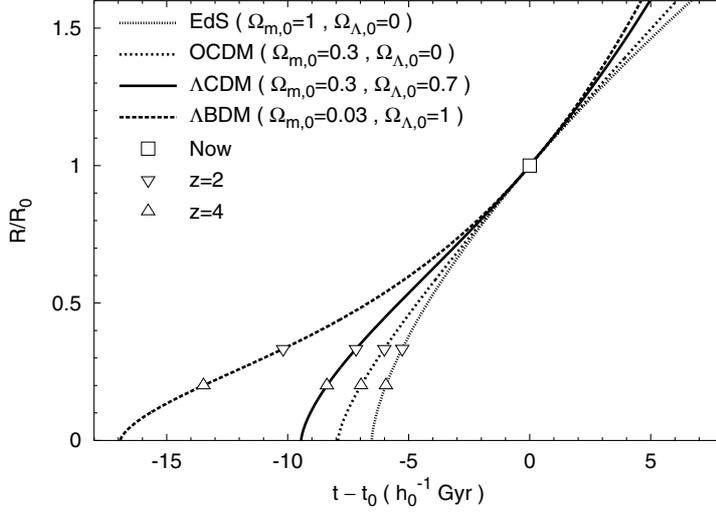


Figure 4.2. Evolution of the cosmological scale factor $\tilde{R}(t) \equiv R(t)/R_0$ as a function of time for the cosmological test models introduced in table 3.1. Triangles indicate the range $2 \leq z_f \leq 4$ where the bulk of galaxy formation may have taken place.

potential wells for the baryons to fall into. This, as we will see, is one of the main motivations for CDM.

Returning now to the density of luminous matter, we find with our values for h_0 that equation (4.1) gives

$$\Omega_{\text{lum}} = 0.0036 \pm 0.0020. \quad (4.3)$$

This is the basis for our statement (section 4.1) that the visible Universe makes up less than 1% of the critical density.

It is, however, conceivable that most of the baryons are not visible. How significant could such dark baryons be? The theory of **primordial big-bang nucleosynthesis** provides us with an independent method for determining the density of *total* baryonic matter in the Universe, based on the assumption that the light elements we see today were forged in the furnace of the hot big bang. Results using different light elements are roughly consistent, which is impressive in itself. The primordial abundances of ^4He (by mass) and ^7Li (relative to H) imply a baryon density of $\Omega_{\text{bar}} = (0.011 \pm 0.005)h_0^{-2}$ [15]. By contrast, new measurements based exclusively on the primordial D/H abundance give a higher value with lower uncertainty: $\Omega_{\text{bar}} = (0.019 \pm 0.002)h_0^{-2}$ [16]. Since it appears premature at present to exclude either of these results, we choose an intermediate value of $\Omega_{\text{bar}} = (0.016 \pm 0.005)h_0^{-2}$. Combining this with our range

of values (4.2) for h_0 , we conclude that

$$\Omega_{\text{bar}} = 0.028 \pm 0.012. \quad (4.4)$$

This agrees very well with an entirely independent estimate of Ω_{bar} obtained by adding up individual mass contributions from all known repositories of baryonic matter via their estimated *mass-to-light ratios* [4]. It also provides the rationale for our choice of $\Omega_{\text{m},0} = 0.03$ in the baryonic dark-matter model (Λ BDM). If $\Omega_{\text{tot},0}$ is close to unity, then it follows from equation (4.4) that all the atoms and molecules in existence combine to make up less than 5% of the Universe by mass.

The vast majority of these baryons, moreover, are invisible. The baryonic dark-matter fraction $f_{\text{bdm}} (\equiv \Omega_{\text{bdm}}/\Omega_{\text{bar}}) = 1 - \Omega_{\text{lum}}/\Omega_{\text{bar}}$ is

$$f_{\text{bdm}} = (87 \pm 8)\%. \quad (4.5)$$

Here we have used equations (4.1) and (4.2), together with the range of values for $\Omega_{\text{bar}}h_0^2$ from nucleosynthesis. Where could these dark baryons be? One possibility is that they are smoothly distributed in a gaseous intergalactic medium, which would have to be strongly ionized in order to explain why it has not left a more obvious absorption signature in the light from distant quasars. Observations using OVI absorption lines as a tracer of ionization suggest that the contribution of such material to Ω_{bar} is at least $0.003h_0^{-1}$ [17], comparable to Ω_{lum} . Simulations are able to reproduce many observed features of the ‘forest’ of **Lyman- α** (**Ly α**) **absorbers** with as much as 80–90% of the baryons in this form [18].

Dark baryonic matter could also be bound up in clumps of matter such as substellar objects (jupiters, brown dwarfs) or stellar remnants (white, red and black dwarfs, neutron stars, black holes). Substellar objects are not likely to make a large contribution, given their small masses. Black holes are limited in the opposite sense: they cannot be more massive than about $10^5 M_{\odot}$ since this would lead to dramatic tidal disruptions and lensing effects which are not seen [19]. The baryonic dark-matter clumps of most interest are therefore ones whose mass is within a few orders of magnitude of M_{\odot} . **Gravitational microlensing** constraints based on quasar variability do not seriously limit the cosmological density of such objects at present, setting an upper bound of 0.1 (well above Ω_{bar}) on their combined contributions to $\Omega_{\text{m},0}$ in an EdS Universe [20].

The existence of dark **massive compact halo objects** (MACHOs) within our own galactic halo has been confirmed by the MACHO microlensing survey of LMC stars [21]. The inferred lensing masses lie in the range $(0.15\text{--}0.9)M_{\odot}$ and would account for between 8% and 50% of the high rotation velocities seen in the outer parts of the Milky Way. (This depends on the choice of halo model and is an extrapolation from the 15 or so events actually seen.) The identity of these objects has been hotly debated, with some authors linking them to faint, fast-moving objects apparently detected in the HDF [22]. It is unlikely that they could be traditional white dwarfs, since these are formed from massive progenitors whose metal-rich ejecta we do not see [23]. Ancient, low-mass ($\lesssim 0.6M_{\odot}$) red dwarfs

which have cooled into invisibility are one possibility. Existing bounds on such objects [24] depend on many extrapolations from known populations. Degenerate ‘beige dwarfs’, which might be able to form above the hydrogen-burning mass limit of $0.08M_{\odot}$ [25], have also been suggested.

4.4 Cold dark matter

The introduction of a second species of unseen dark matter into the Universe has been primarily motivated in three ways:

- (1) a range of *observational arguments* imply that the density parameter of total gravitating non-relativistic matter ($\Omega_{m,0} = \Omega_{\text{bar}} + \Omega_{\text{cdm}} + \Omega_{\nu}$) is higher than that provided by baryons and neutrinos alone;
- (2) our current understanding of *large-scale structure formation* requires the process to be helped along by quantities of non-relativistic, weakly interacting matter in the early Universe, creating the potential wells for infalling baryons; and
- (3) theoretical physics supplies several plausible (albeit still undetected) *candidate CDM particles* with the right properties.

Since our ideas on structure formation may change, and the candidate particles may not materialize, the case for CDM turns at present on the observational arguments. At one time, these were compatible with $\Omega_{\text{cdm}} \approx 1$, raising hopes that CDM would resolve two of the biggest challenges in cosmology at a single stroke: accounting for observations of large-scale structure *and* providing all the dark matter necessary to make $\Omega_{m,0} = 1$, vindicating the EdS model (and with it, the simplest models of inflation). Observations, however, no longer support values of $\Omega_{m,0}$ this high, and independent evidence now points to at least two other kinds of non-baryonic dark matter (neutrinos and vacuum energy). The CDM hypothesis is, therefore, no longer as compelling as it once was. With this in mind we will pay special attention to the observational arguments in this section. The *lower limits* on $\Omega_{m,0}$, in particular, are crucial: only if $\Omega_{m,0} > \Omega_{\text{bar}} + \Omega_{\nu}$ do we require $\Omega_{\text{cdm}} > 0$.

The arguments can be broken into two classes: those which are purely empirical; and those which assume, in addition, the validity of the gravitational instability theory of structure formation. Let us begin with the empirical arguments. The first has already been encountered in section 4.3: the spiral **galaxy rotation curve**. If the MACHO results are taken at face value, and if the Milky Way is typical, then compact objects make up less than 50% of the mass of the halos of spiral galaxies. If, as has been argued [26], the remaining halo mass cannot be attributed to baryonic matter in known forms such as dust, rocks, planets, gas or hydrogen snowballs, then a more exotic form of dark matter is required.

The total mass of dark matter in galaxies, however, is limited. The easiest way to see this is to compare the **mass-to-light ratio** (M/L) of our own Galaxy

to that of the Universe as a whole. If the latter is flat, then its M/L ratio is just the ratio of the critical density to its luminosity density. This is $(M/L)_{\text{crit},0} = \rho_{\text{crit},0}/\mathcal{L}_0 = (1040 \pm 230)M_\odot/L_\odot$, where we have used (2.24) for \mathcal{L}_0 , (2.36) for $\rho_{\text{crit},0}$ and (4.2) for h_0 . The corresponding value for the Milky Way is $(M/L)_{\text{mw}} = (21 \pm 7)M_\odot/L_\odot$, since the latter's luminosity is $L_{\text{mw}} = (2.3 \pm 0.6) \times 10^{10}L_\odot$ (in the B-band) and its total dynamical mass (including that of any unseen halo component) is $M_{\text{mw}} = (4.9 \pm 1.1) \times 10^{11}M_\odot$ inside 50 kpc from the motions of galactic satellites [27]. The ratio of $(M/L)_{\text{mw}}$ to $(M/L)_{\text{crit},0}$ is thus less than 3%, and even if we multiply this by a factor of a few (to account for possible halo mass outside 50 kpc), it is clear that galaxies like our own cannot make up more than 10% of the critical density.

Most of the mass of the Universe, in other words, is spread over scales larger than galaxies, and it is here that the arguments for CDM take on the most force. The **M/L -ratio method** is, in fact, the most straightforward: one measures M/L for a chosen region, corrects for the corresponding value in the 'field' and divides by $(M/L)_{\text{crit},0}$ to obtain $\Omega_{\text{m},0}$. Much, however, depends on the choice of region. A widely respected application of this approach is that of the CNOC team [28], which uses rich clusters of galaxies. These systems sample large volumes of the early Universe, have dynamical masses which can be measured by three independent methods (the virial theorem, x-ray gas temperatures and gravitational lensing) and are subject to reasonably well-understood evolutionary effects. They are found to have $M/L \sim 200M_\odot/L_\odot$ on average, giving $\Omega_{\text{m},0} = 0.19 \pm 0.06$ when $\Omega_{\Lambda,0} = 0$ [28]. This result scales as $(1 - 0.4\Omega_{\Lambda,0})$ [29], so that $\Omega_{\text{m},0}$ drops to 0.11 ± 0.04 in a model with $\Omega_{\Lambda,0} = 1$.

The weak link in this chain of inference is that rich clusters may not be characteristic of the Universe as a whole. Only about 10% of galaxies are found in such systems. If *individual* galaxies (like the Milky Way, with $M/L \approx 21M_\odot/L_\odot$) are substituted for clusters, then the inferred value of $\Omega_{\text{m},0}$ drops by a factor of ten, approaching Ω_{bar} and removing the need for CDM. A recent effort to address the impact of scale on M/L arguments concludes that $\Omega_{\text{m},0} = 0.16 \pm 0.05$ (in flat models), when regions of all scales are considered from individual galaxies to superclusters [30].

Another line of argument uses the ratio of baryonic to total gravitating mass in clusters, or **cluster baryon fraction** ($M_{\text{bar}}/M_{\text{tot}}$). Baryonic matter is defined as the sum of visible galaxies and hot cluster gas (the mass of which can be inferred from the x-ray temperature). Total cluster mass is measured by one or all of the three methods listed earlier (virial theorem, x-ray temperature or gravitational lensing). At sufficiently large radii, the cluster may be taken as representative of the Universe as a whole, so that $\Omega_{\text{m},0} = \Omega_{\text{bar}}/(M_{\text{bar}}/M_{\text{tot}})$, where Ω_{bar} is fixed by big-bang nucleosynthesis (section 4.3). Applied to various clusters, this procedure leads to $\Omega_{\text{m},0} = 0.3 \pm 0.1$ [31]. This result is probably an upper limit, partly because baryon enrichment is more likely to take place inside the cluster than out, and partly because *dark* baryonic matter (such as MACHOs) is not taken into account; this would raise M_{bar} and lower $\Omega_{\text{m},0}$.

A final, recent entry into the list of purely empirical methods uses the separation of **radio galaxy lobes** as standard rulers in a variation on the classical angular size–distance test in cosmology. The widths, propagation velocities and magnetic field strengths of these lobes have been calibrated for 14 radio galaxies with the aid of long-baseline radio interferometry. This leads to the constraints $\Omega_{m,0} < 0.10$ (for $\Omega_{\Lambda,0} = 0$) and $\Omega_{m,0} = 0.10^{+0.25}_{-0.10}$ for flat models with $\Omega_{\Lambda,0} = 1 - \Omega_{m,0}$ [32].

Let us consider, next, the measurements of $\Omega_{m,0}$ based on the idea that large-scale structure formed via **gravitational instability** from a Gaussian spectrum of primordial density fluctuations (GI theory for short). These are somewhat circular, in the sense that such a process could not have taken place as it did *unless* $\Omega_{m,0}$ is considerably larger than Ω_{bar} . But inasmuch as GI theory is the only structure-formation theory we have which is both fully worked out and in good agreement with observation (with some potential difficulties on small scales [33, 34]), this way of measuring $\Omega_{m,0}$ should be taken seriously.

According to GI theory, formation of large-scale structure is more or less complete by $z \approx \Omega_{m,0}^{-1} - 2$ [35]. Therefore, one way to constrain $\Omega_{m,0}$ is to look for **number density evolution** in large-scale structures such as galaxy clusters. In a low-matter-density Universe, this would be relatively constant out to at least $z \sim 1$, whereas in a high-matter-density Universe one would expect the abundance of clusters to drop rapidly with z because they are still in the process of forming. The fact that massive clusters are seen at redshifts as high as $z = 0.83$ leads to limits $\Omega_{m,0} = 0.17^{+0.14}_{-0.09}$ for $\Omega_{\Lambda,0} = 0$ models and $\Omega_{m,0} = 0.22^{+0.13}_{-0.07}$ for flat ones [36].

Studies of the **Fourier power spectra** $P(k)$ of the distributions of galaxies or other structures can be used in a similar way. In GI theory, structures of a given mass form by the collapse of large volumes in a low-matter-density Universe or smaller volumes in a high-matter-density Universe. Thus $\Omega_{m,0}$ can be constrained by changes in $P(k)$ between one redshift and another. Comparison of the mass power spectrum of Ly α absorbers at $z \approx 2.5$ with that of local galaxy clusters at $z = 0$ has led to an estimate of $\Omega_{m,0} = 0.46^{+0.12}_{-0.10}$ for $\Omega_{\Lambda,0} = 0$ models [37]. This result goes as approximately $(1 - 0.4\Omega_{\Lambda,0})$, so that the central value of $\Omega_{m,0}$ drops to 0.34 in a flat model and 0.28 if $\Omega_{\Lambda,0} = 1$. One can also constrain $\Omega_{m,0}$ from the local galaxy power spectrum alone, although this involves some assumptions about the extent to which ‘light traces mass’ (i.e. to which visible galaxies trace the underlying density field). Early results from the 2dF survey give $\Omega_{m,0} = (0.20 \pm 0.03)h_0^{-1}$ for flat models [38], or $\Omega_{m,0} = 0.27 \pm 0.07$ with our range of h_0 values as given by (4.2).

A final group of measurements, and one which has traditionally yielded the highest estimates of $\Omega_{m,0}$, comes from the analysis of **galaxy peculiar velocities**. These are produced by the gravitational potential of locally over- or under-dense regions relative to the mean matter density. The power spectra of the velocity and density distributions can be related to each other in the context of GI theory in a way which depends explicitly on $\Omega_{m,0}$. Tests of this kind probe

relatively small volumes and are hence insensitive to $\Omega_{\Lambda,0}$, but they can depend significantly on h_0 as well as the spectral index n of the density distribution. In [39], where the latter is normalized to CMB fluctuations, results take the form $\Omega_{m,0}h_0^{1.3}n^2 \approx 0.33 \pm 0.07$ or (taking $n = 1$ and using our values of h_0) $\Omega_{m,0} \approx 0.48 \pm 0.15$.

In summarizing these results, one is struck by the fact that arguments based on gravitational instability (GI) theory favour values of $\Omega_{m,0} \gtrsim 0.2$ and *higher*, whereas purely empirical arguments require $\Omega_{m,0} \lesssim 0.4$ and *lower*. The latter are, in fact, compatible in some cases with values of $\Omega_{m,0}$ as low as Ω_{bar} , raising the possibility that CDM might not, in fact, be necessary. The results from GI-based arguments, however, cannot be stretched this far. What is sometimes done is to ‘go down the middle’ and blend the results of both kinds of arguments into a single bound of the form $\Omega_{m,0} \approx 0.3 \pm 0.1$. Any such bound with $\Omega_{m,0} > 0.05$ constitutes a proof of the existence of CDM, since $\Omega_{\text{bar}} \leq 0.04$ from (4.4). (Massive neutrinos do not affect this argument, as we will indicate in section 4.5.) A more conservative interpretation of the data, bearing in mind the full range of $\Omega_{m,0}$ values implied above ($\Omega_{\text{bar}} \lesssim \Omega_{m,0} \lesssim 0.6$), is

$$\Omega_{\text{cdm}} = 0.3 \pm 0.3. \quad (4.6)$$

Values of Ω_{cdm} at the bottom of this range, however, carry with them the (uncomfortable) implication that the conventional picture of structure formation via gravitational instability is incomplete. Conversely, *if our current understanding of structure formation is correct, then CDM must exist and $\Omega_{\text{cdm}} > 0$.*

The question, of course, becomes moot if CDM is discovered in the laboratory. From a large field of theoretical particle candidates, two have emerged as frontrunners: the **axion** and the supersymmetric **weakly-interacting massive particle** (WIMP). The plausibility of both candidates rests on three properties. Either one, if it exists, is:

- (1) *weakly interacting* (i.e. ‘noticed’ by ordinary matter primarily via its gravitational influence);
- (2) *cold* (i.e. non-relativistic in the early Universe, when structures begin to form); and
- (3) *expected* on theoretical grounds to have a collective density within a few orders of magnitude of the critical one.

We will return to these particles in chapters 6 and 8 respectively.

4.5 Neutrinos

Since neutrinos indisputably exist in great numbers (their number density n_ν is 3/11 times that of the CMB photons, or 112 cm^{-3} per species [40]), they have been leading dark-matter candidates for longer than either the axion or the WIMP.

They gained prominence in 1980 when teams in the USA and the Soviet Union both reported evidence of non-zero neutrino rest masses. While these claims did not stand up, a new round of experiments now indicates once again that m_ν (and hence Ω_ν) > 0 .

Dividing $n_\nu m_\nu$ by the critical density (2.36), one obtains

$$\Omega_\nu = \left(\sum m_\nu c^2 / 93.8 \text{ eV} \right) h_0^{-2} \quad (4.7)$$

where the sum is over the three neutrino species. Here, we follow convention and specify particle masses in units of eV/c^2 , where $1 \text{ eV}/c^2 = 1.602 \times 10^{-12} \text{ erg}/c^2 = 1.783 \times 10^{-33} \text{ g}$. The calculations in this section are strictly valid only for $m_\nu c^2 \lesssim 1 \text{ MeV}$. More massive neutrinos with $m_\nu c^2 \sim 1 \text{ GeV}$ were once considered as CDM candidates but are no longer viable since experiments at the LEP collider rule out additional neutrino species with masses up to at least half of that of the Z_0 ($m_{Z_0} c^2 = 91 \text{ GeV}$).

Current laboratory upper bounds on neutrino rest masses are $m_{\nu_e} c^2 < 3 \text{ eV}$, $m_{\nu_\mu} c^2 < 0.19 \text{ MeV}$ and $m_{\nu_\tau} c^2 < 18 \text{ MeV}$, so it would appear feasible in principle for these particles to close the Universe. In fact m_{ν_μ} and m_{ν_τ} are limited far more stringently by (4.7) than by laboratory bounds. Perhaps the best-known theory along these lines is that of Sciama [41], who postulated a population of 29 eV τ neutrinos which, if they decayed into lighter neutrinos on suitable timescales, could be shown to solve a number of longstanding astrophysical puzzles related to interstellar and intergalactic ionization. Equation (4.7) shows that such neutrinos would also account for much of the dark matter, contributing a *minimum* collective density of $\Omega_\nu \geq 0.38$ (assuming as usual that $h_0 \leq 0.9$). We will consider the decaying-neutrino hypothesis in more detail in chapter 7.

Strong upper limits can be set on Ω_ν within the context of the gravitational instability picture. Neutrinos constitute **hot dark matter** (i.e. they are relativistic when they decouple from the primordial fireball) and are therefore able to stream freely out of density perturbations in the early Universe, erasing them before they have a chance to grow. Good agreement with observations of large-scale structure can be achieved in models with Ω_ν as high as 0.2, but only if $\Omega_{\text{bar}} + \Omega_{\text{cdm}} = 0.8$ and $h_0 = 0.5$ [42]. Statistical exploration of a larger set of (flat) models leads to an upper limit on m_ν which is roughly proportional to the CDM density Ω_{cdm} and peaks (when $\Omega_{\text{cdm}} = 0.6$) at $m_\nu c^2 \leq 5.5 \text{ eV}$ [43]. Since $0 \leq \Omega_{\text{cdm}} \leq 0.6$ from (4.6), this implies

$$m_\nu c^2 \leq (9.2 \text{ eV}) \Omega_{\text{cdm}}. \quad (4.8)$$

If $h_0 \geq 0.6$ and $\Omega_{\text{cdm}} \leq 0.6$, and if structure grows via gravitational instability as is generally assumed, then equations (4.7) and (4.8) together allow a collective neutrino density as high as $\Omega_\nu = 0.16$. Thus neutrino contributions could, in principle, be as much as four to ten times as high as those from baryons, equation (4.4). This is perfectly consistent with the neglect of relativistic matter during the matter-dominated era (section 2.5). Neutrinos, while relativistic at decoupling, lose energy and become *non-relativistic* on timescales $t_{\text{nr}} \approx$

190 000 yr $(m_\nu c^2/\text{eV})^{-2}$ [44]. This is well before the present epoch for neutrinos which are massive enough to be of interest.

New *lower* limits on Ω_ν have been reported in recent years from atmospheric (Super-Kamiokande [45]), Solar (SAGE [46], Homestake [47], GALLEX [48]), and accelerator-based (LSND [49]) neutrino experiments. In each case it appears that two species are converting into each other in a process known as **neutrino oscillation**, which can only take place if both have non-zero rest masses. The strongest evidence comes from Super-Kamiokande, which has reported oscillations between τ and μ neutrinos with $5 \times 10^{-4} \text{ eV}^2 < \Delta m_{\tau\mu}^2 c^2 < 6 \times 10^{-3} \text{ eV}^2$, where $\Delta m_{\tau\mu}^2 \equiv |m_{\nu_\tau}^2 - m_{\nu_\mu}^2|$ [45]. Oscillation experiments measure the square of the mass difference between two species, and cannot fix the mass of any one species unless they are combined with other evidence such as that from **neutrinoless double-beta decay** [50]. Nevertheless, if neutrino masses are hierarchical, like those of other fermions, then $m_{\nu_\tau} \gg m_{\nu_\mu}$ and the Super-Kamiokande results imply that $m_{\nu_\tau} c^2 > 0.02 \text{ eV}$. In this case it follows from (4.7) that $\Omega_\nu \geq 0.0003$ (with $h_0 \leq 0.9$ as usual). If, instead, neutrino masses are nearly degenerate, then Ω_ν could, in principle, be much larger than this but will still lie below the upper bound imposed by structure formation. Putting (4.8) into (4.7), we conclude that

$$0.0003 \leq \Omega_\nu < 0.3\Omega_{\text{cdm}}. \quad (4.9)$$

The neutrino contribution to $\Omega_{\text{tot},0}$ is anywhere from an order of magnitude below that of the visible stars and galaxies (section 4.3) up to nearly one-third the amount conventionally attributed to CDM (section 4.4). We emphasize that low values of Ω_{cdm} imply low values of Ω_ν also. Physically, this reflects the fact that neutrinos interfere with structure formation in the early Universe unless something like CDM is present to help hold primordial density perturbations together. In theories (like that to be discussed in chapter 6) where the density of neutrinos exceeds that in (4.9), one would need to modify the standard gravitational instability picture by encouraging the growth of early structures in some other way, as for instance by ‘seeding’ them with loops of cosmic string.

4.6 Vacuum energy

There are at least four reasons to include a **cosmological constant** (Λ) in Einstein’s field equations (2.1). The first is *mathematical*: Λ plays a role in these equations similar to that of the additive constant in an indefinite integral [51]. The second is *dimensional*: Λ specifies the radius of curvature $R_\Lambda \equiv \Lambda^{-1/2}$ in closed models at the moment when the matter density parameter Ω_m goes through its maximum, thereby providing a fundamental length scale for cosmology [52]. The third is *dynamical*: Λ determines the asymptotic expansion rate of the Universe according to equation (2.41). And the fourth is *material*: Λ is related to the energy density of the vacuum via equation (2.38), $\Omega_{\Lambda,0} = \Lambda c^2/3H_0^2$.

Table 4.1. Theoretical estimates of $\Omega_{\Lambda,0}$.

Theory	Estimated value of ρ_{Λ}	$\Omega_{\Lambda,0}$
QCD	$(0.3 \text{ GeV})^4 \hbar^{-3} c^{-5} = 10^{16} \text{ g cm}^{-3}$	$10^{44} h_0^{-2}$
EW	$(200 \text{ GeV})^4 \hbar^{-3} c^{-5} = 10^{26} \text{ g cm}^{-3}$	$10^{55} h_0^{-2}$
GUTs	$(10^{19} \text{ GeV})^4 \hbar^{-3} c^{-5} = 10^{93} \text{ g cm}^{-3}$	$10^{122} h_0^{-2}$

With all these reasons to take this term seriously, why have many cosmologists since Einstein set $\Lambda = 0$? Mathematical simplicity is one explanation. Another is the smallness of most effects associated with the Λ -term. Einstein himself set $\Lambda = 0$ in 1931 ‘for reasons of logical economy’, because he saw no hope of measuring this quantity experimentally at the time. He is often quoted as adding that its introduction in 1915 was the biggest blunder of his life. This comment (which does not appear in Einstein’s writings but was rather attributed to him by Gamow [53]) is sometimes interpreted as a rejection of the very idea of a cosmological constant. It more likely represents Einstein’s rueful recognition that, by invoking the Λ -term solely to obtain a static solution of the field equations, he had narrowly missed what would surely have been one of the greatest *triumphs* of his life: the prediction of cosmic expansion.

The relation between Λ and the energy density of the vacuum has led to a quandary in more recent times: the fact that ρ_{Λ} as estimated in the context of quantum field theories such as quantum chromodynamics (QCD), electroweak (EW) and grand unified theories (GUTs) implies impossibly large values of $\Omega_{\Lambda,0}$ (table 4.1). These theories have been very successful in the microscopic realm. Here, however, they are in gross disagreement with the observed facts of the macroscopic world, which tell us that $\Omega_{\Lambda,0}$ cannot be much larger than order *unity*. This **cosmological-constant problem** has been reviewed by several workers but there is no consensus on how to solve it [54]. It is undoubtedly another reason why many cosmologists have preferred to set $\Lambda = 0$, rather than deal with a parameter whose microphysical origins are still unclear.

Setting Λ to zero, however, is not really an appropriate response because observations indicate that $\Omega_{\Lambda,0}$, while nowhere near the size suggested by table 4.1, is nevertheless greater than zero. The cosmological constant problem has therefore become *more* baffling, in that an explanation of this parameter must apparently contain a cancellation mechanism which is not only good to some 44 (or 122) decimal places, but which begins to fail at precisely the 45th (or 123rd).

One suggestion for understanding the possible nature of such a cancellation has been to treat the vacuum energy field literally as an Olbers-type summation of contributions from different places in the Universe [55]. It can then be handled with the same formalism that we have developed in chapters 2 and 3 for background radiation. This has the virtue of framing the problem in concrete

terms, and raises some interesting possibilities, but does not in itself explain why the energy density inherent in such a field does not gravitate in the conventional way [56]. Another idea is that theoretical expectations for the value of Λ might refer only to the latter's primordial or 'bare' value, which could have been progressively 'screened' over time. The cosmological constant would thus become a **variable cosmological term** (see [57] for a review). In such a scenario the present 'low' value of $\Omega_{\Lambda,0}$ would simply reflect the fact that the Universe is *old*. In general, however, this means modifying Einstein's field equations (2.1) and/or introducing new forms of matter such as scalar fields. We will look at this suggestion in more detail in chapter 5.

A third possibility occurs in **higher-dimensional gravity**, where the cosmological constant can arise as an artefact of dimensional reduction (i.e. in extracting the appropriate four-dimensional limit from the theory). In such theories it is possible that the 'effective' Λ_4 could be small while its N -dimensional analogue Λ_N is large [58]. We will consider some aspects of higher-dimensional gravity in chapter 9. Some workers, finally, have argued that a Universe in which Λ was too large might be incapable of giving rise to intelligent observers. That is, the fact of our own existence might already 'require' $\Omega_{\Lambda,0} \sim 1$ [59]. This is an application of the **anthropic principle** whose status, however, remains unclear.

Let us pass now to what is known about the value of $\Omega_{\Lambda,0}$ from cosmology. It is widely believed that the Universe originated in a big-bang singularity rather than passing through a 'big bounce' at the beginning of the current expansionary phase. By differentiating the Friedmann–Lemaître equation (2.40) and setting both the expansion rate and its time derivative to zero, one finds that this implies an upper limit (sometimes called the Einstein limit $\Omega_{\Lambda,E}$) on $\Omega_{\Lambda,0}$ as a function of $\Omega_{m,0}$. Models with $\Omega_{\Lambda} = \Omega_{\Lambda,E}$, known as **Eddington–Lemaître models**, are asymptotic to Einstein's original static model in the infinite past. The quantity $\Omega_{\Lambda,E}$ can be expressed as follows for cases with $\Omega_{m,0} < 0.5$ [60]:

$$\Omega_{\Lambda,E} = 1 - \Omega_{m,0} + \frac{3}{2}\Omega_{m,0}^{2/3} \left[\left(1 - \Omega_{m,0} + \sqrt{1 - 2\Omega_{m,0}} \right)^{1/3} + \left(1 - \Omega_{m,0} - \sqrt{1 - 2\Omega_{m,0}} \right)^{1/3} \right]. \quad (4.10)$$

For $\Omega_{m,0} = 0.3$ the requirement that $\Omega_{\Lambda,0} < \Omega_{\Lambda,E}$ implies $\Omega_{\Lambda,0} < 1.71$, a limit that tightens to $\Omega_{\Lambda,0} < 1.16$ for $\Omega_{m,0} = 0.03$.

A slightly stronger constraint can be formulated (for closed models) in terms of the **antipodal redshift** (z_a). The antipodes are defined as the set of points located on the 'other side of the Universe' at $\chi = \pi$, where $d\chi$ (the radial coordinate distance element) is given by $d\chi = (1 - kr^2)^{-1/2} dr$ from the metric (2.2). Using (2.10) and (2.18) this can be put into the form $d\chi = -(c/H_0 R_0) dz/\tilde{H}(z)$, which can be integrated from $z = 0$ to z_a with the help of (2.39) and (2.40). Gravitational lensing of sources beyond the antipodes cannot

give rise to normal (multiple) images [61], so the redshift z_a of the antipodes must exceed that of the most distant normally-lensed object, currently a galaxy at $z = 4.92$ [62]. Requiring that $z_a > 4.92$ leads to the upper bound $\Omega_{\Lambda,0} < 1.57$ if $\Omega_{m,0} = 0.3$. This tightens to $\Omega_{\Lambda,0} < 1.15$ for Λ BDM-type models with $\Omega_{m,0} = 0.03$.

Gravitational lensing also leads to a different and considerably more stringent upper limit (for all models) based on **lensing statistics**. The increase in path length to a given redshift in vacuum-dominated models (relative to, say, the EdS model) means that there are more sources to be lensed, and presumably more lensed objects to be seen. The observed frequency of lensed quasars, however, is rather modest, leading to the bound $\Omega_{\Lambda,0} < 0.66$ for flat models [63]. Dust could hide distant sources [64]. However, radio lenses should be far less affected and these give only slightly weaker constraints: $\Omega_{\Lambda,0} < 0.73$ (for flat models) or $\Omega_{\Lambda,0} \lesssim 0.4 + 1.5\Omega_{m,0}$ (for curved ones) [65]. Uncertainties arise in lens modelling and evolution as well as source redshifts and survey completeness. A recent conservative limit from radio lenses is $\Omega_{\Lambda,0} < 0.95$ for flat models [66].

Tentative *lower* limits have been set on $\Omega_{\Lambda,0}$ using faint **galaxy number counts**. This method is based on a similar premise to that of lensing statistics: the enhanced comoving volume at large redshifts in vacuum-dominated models should lead to greater (projected) galaxy number densities at faint magnitudes. In practice, it has proven difficult to disentangle this effect from galaxy luminosity evolution. Early claims of a best fit at $\Omega_{\Lambda,0} \approx 0.9$ [67] have been disputed on the basis that the steep increase seen in numbers of blue galaxies is not matched in the K-band, where luminosity evolution should be less important [68]. Attempts to account for evolution in a comprehensive way have subsequently led to a lower limit of $\Omega_{\Lambda,0} > 0.53$ [69] and, most recently, a reasonable fit (for flat models) with a vacuum density parameter of $\Omega_{\Lambda,0} = 0.8$ [70].

Other evidence for a significant $\Omega_{\Lambda,0}$ term has come from numerical simulations of **large-scale structure formation**. Figure 4.3 shows the evolution of massive structures between $z = 3$ and $z = 0$ in simulations by the VIRGO Consortium [71]. Of the two models shown, Λ CDM (top row) is qualitatively closer to the observed distribution of galaxies in the real Universe than EdS ('SCDM', bottom row). The improvement is especially marked at higher redshifts (left-hand panels). Power spectrum analysis, however, reveals that the match is not particularly good in *either* case [71]. This may be a reflection of bias (i.e. of a systematic discrepancy between the distributions of mass and light). Different combinations of $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$ might also provide better fits. Simulations of closed Λ BDM-type models would be of particular interest [72, 73].

The first measurements to put both lower *and* upper bounds on $\Omega_{\Lambda,0}$ have come from **Type Ia supernovae** (SNIa). These objects have luminosities which are both high and consistent, and they are thought not to evolve significantly with redshift. All of these properties make them ideal for use in the classical magnitude–redshift relation. Two independent groups (HZT [74] and SCP [75]) have reported a systematic dimming of SNIa at $z \approx 0.5$ by about 0.25 magnitudes

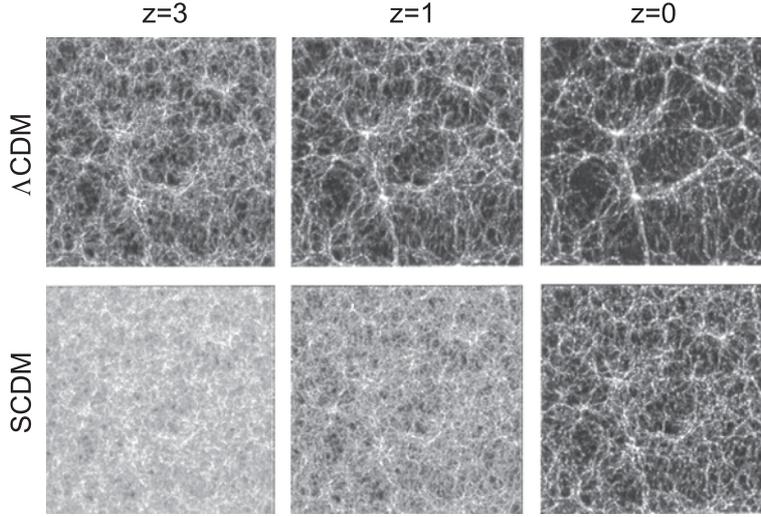


Figure 4.3. Numerical simulations of structure formation. In the top row is the Λ CDM model with $\Omega_{m,0} = 0.3$, $\Omega_{\Lambda,0} = 0.7$ and $h_0 = 0.7$. The bottom row shows the EdS (‘SCDM’) model with $\Omega_{m,0} = 1$, $\Omega_{\Lambda,0} = 0$ and $h_0 = 0.5$. The panel size is comoving with the Hubble expansion, and time runs from left ($z = 3$) to right ($z = 0$). (Images courtesy of J Colberg and the VIRGO Consortium.)

relative to that expected in an EdS model, suggesting that space at these redshifts is ‘stretched’ by a significant vacuum term. The 2σ confidence intervals from both studies may be summarized as $0.8\Omega_{m,0} - 0.6\Omega_{\Lambda,0} = -0.2 \pm 0.3$, or

$$\Omega_{\Lambda,0} = \frac{4}{3}\Omega_{m,0} + \frac{1}{3} \pm \frac{1}{2}. \quad (4.11)$$

From this result it follows that $\Omega_{\Lambda,0} > 0$ for any $\Omega_{m,0} > 1/8$, and that $\Omega_{\Lambda,0} > 0.2$ if $\Omega_{m,0} = 0.3$. These numbers are in disagreement with both the EdS and OCDM models. To extract limits on $\Omega_{\Lambda,0}$ alone, we recall that $\Omega_{m,0} \geq \Omega_{\text{bar}} \geq 0.02$ (section 4.3) and $\Omega_{m,0} \leq 0.6$ (section 4.4). If we combine these conservative bounds with a single limit, $\Omega_{m,0} = 0.31 \pm 0.29$, then (4.11) gives

$$\Omega_{\Lambda,0} = 0.75 \pm 0.63. \quad (4.12)$$

This is not a high-precision measurement but it is enough to establish that $\Omega_{\Lambda,0} > 0.1$ and hence that *the vacuum energy is real*. Several words of caution are in order, however. Intergalactic dust could also cause a systematic dimming (without reddening) if it were ‘sifted’ during the process of ejection from galaxies [76]. The neglect of evolution may be more serious than claimed [77]. And much remains to be understood about the physics of SNIa explosions [78]. It is likely that observations will have to reach $z \sim 2$ before the SNIa magnitude–

redshift relation is able to discriminate statistically between models (like Λ CDM and Λ BDM) with different ratios of $\Omega_{m,0}$ to $\Omega_{\Lambda,0}$.

Further support for the existence of ‘dark energy’ has arisen from a completely independent source: the **angular power spectrum of CMB fluctuations**. These are produced by density waves in the primordial plasma, termed by Lineweaver ‘the oldest music in the Universe’ [79]. The first peak in their power spectrum picks out the angular size of the largest fluctuations in this plasma at the moment when the Universe became transparent to light. Because it is seen through the ‘lens’ of a curved Universe, the location of this peak is sensitive to the latter’s total density $\Omega_{\Lambda,0} + \Omega_{m,0}$. Results have been reported from three different experiments (BOOMERANG [80], MAXIMA [81] and DASI [82]) and combined with earlier data from the COBE satellite [83]. BOOMERANG and MAXIMA conclusions have been analysed together to yield $\Omega_{m,0} + \Omega_{\Lambda,0} = 1.11^{+0.13}_{-0.12}$ [84], while those from the more recent DASI team give $\Omega_{m,0} + \Omega_{\Lambda,0} = 1.047^{+0.103}_{-0.120}$ [82]. (These are reported 95% uncertainties using the most conservative statistical priors.) Let us summarize these results as follows:

$$\Omega_{m,0} + \Omega_{\Lambda,0} = 1.08 \pm 0.16. \quad (4.13)$$

The Universe is therefore spatially flat or very close to it. To extract a value for $\Omega_{\Lambda,0}$ alone, we can do as in the SNIa case and substitute our matter-density bounds ($\Omega_{m,0} = 0.31 \pm 0.29$) into (4.13) to obtain

$$\Omega_{\Lambda,0} = 0.77 \pm 0.33. \quad (4.14)$$

This is consistent with (4.12), but has error bars which have been reduced by 50% and are now due almost entirely to the uncertainty in $\Omega_{m,0}$. This measurement is impervious to most of the uncertainties of the earlier ones, because it leapfrogs ‘local’ systems whose interpretation is complex (supernovae, galaxies, and quasars), going directly back to the radiation-dominated era when physics was simpler. Equation (4.14) is sufficient to establish that $\Omega_{\Lambda,0} > 0.4$, and hence that *vacuum energy not only exists, but may very well dominate the energy density of the Universe.*

The CMB power spectrum favours vacuum-dominated models but is not yet resolved with sufficient precision to discriminate (on its own) between models which have exactly the critical density (like Λ CDM) and those which are close to the critical density (like Λ BDM). As it stands, the location of the first peak in these data actually hints at ‘marginally closed’ models, although the implied departure from flatness is not statistically significant and could also be explained in other ways [85].

Much attention is focused on the second- and higher-order peaks of the spectrum, which contain valuable clues about the matter component. Odd-numbered peaks are produced by regions of the primordial plasma which have been maximally compressed by infalling material, and even ones correspond to maximally rarefied regions which have rebounded due to photon pressure. A high

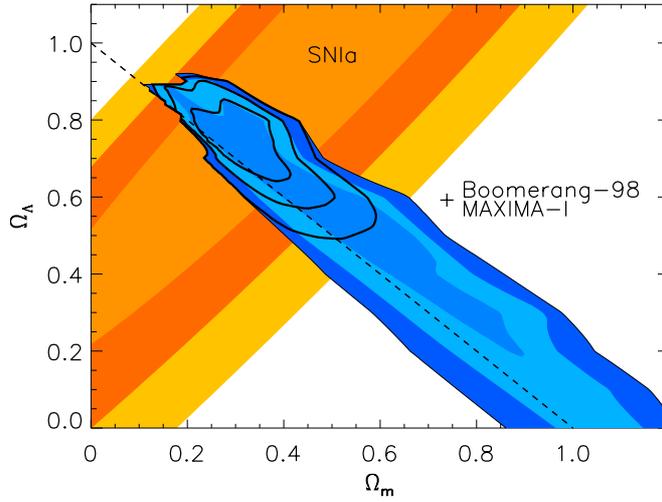


Figure 4.4. Observational constraints on the values of $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$ from supernovae data (SNIa) and the power spectrum of CMB fluctuations (BOOMERANG, MAXIMA). Shown are 68%, 95% and 99.7% confidence intervals inferred both separately and jointly from the data. The dashed line indicates spatially flat models with $k = 0$; models on the lower left are closed while those on the upper right are open. (Reprinted from [84] by permission of A Jaffe and P L Richards.)

baryon-to-photon ratio enhances the compressions and retards the rarefactions, thus suppressing the size of the second peak relative to the first. The strength of this effect depends on the fraction of baryons (relative to the more weakly-bound neutrinos and CDM particles) in the overdense regions. The BOOMERANG and MAXIMA data show an unexpectedly weak second peak. While there are a number of ways to account for this in Λ CDM models (e.g. by ‘tilting’ the primordial spectrum), the data are fit most naturally by the Λ BDM model with $\Omega_{m,0} = \Omega_{\text{bar}}$, $\Omega_{\text{cdm}} = 0$ and $\Omega_{\Lambda,0} \approx 1$ [86]. Models of this kind have been advocated on other grounds over the years, notably in connection with an analysis of Ly α absorption spectra [87] (see [88] for a review). Subsequent DASI data, however, show a stronger second peak and are better fit by Λ CDM [82]. These issues will be resolved as data begin to come in from other experiments such as the MAP and Planck satellites.

The best constraints on $\Omega_{\Lambda,0}$ come from taking *both* the supernovae and microwave background results at face value and substituting one into the other. This provides a valuable cross-check on the matter density, because the SNIa and CMB constraints are very nearly orthogonal in the $\Omega_{m,0}$ – $\Omega_{\Lambda,0}$ plane (figure 4.4). Thus, forgetting all about our conservative bounds on $\Omega_{m,0}$ and

merely substituting (4.13) into (4.11), we find

$$\Omega_{\Lambda,0} = 0.76 \pm 0.23. \quad (4.15)$$

Alternatively, extracting the matter density parameter, we obtain

$$\Omega_{m,0} = 0.32 \pm 0.22. \quad (4.16)$$

These results further tighten the case for a vacuum-dominated Universe. Equation (4.15) also implies that $\Omega_{\Lambda,0} < 1$, which begins to put pressure on the Λ BDM model. Perhaps most importantly, equation (4.16) establishes that $\Omega_{m,0} \geq 0.1$, which is inconsistent with Λ BDM and *requires the existence of CDM*. Moreover, the fact that the range of values picked out by (4.16) agrees so well with that derived in section 4.4 constitutes solid evidence for the Λ CDM model in particular, and for the gravitational instability picture of large-scale structure formation in general.

The depth of the change in thinking that has been triggered by these developments on the observational side can hardly be exaggerated. Only a few years ago, it was still routine to set $\Lambda = 0$ and cosmologists had two main choices: the ‘one true faith’ (flat, with $\Omega_{m,0} \equiv 1$) or the ‘reformed’ (open, with individual believers being free to choose their own values near $\Omega_{m,0} \approx 0.3$). All this has been irrevocably altered by the CMB experiments. If there is a guiding principle now, it is no longer $\Omega_{m,0} \approx 0.3$, and certainly not $\Omega_{\Lambda,0} = 0$; it is $\Omega_{\text{tot},0} \approx 1$ from the power spectrum of the CMB. Cosmologists have been obliged to accept a Λ -term, and it is not so much a question of whether or not it dominates the energy budget of the Universe but by *how much*.

4.7 The coincidental Universe

The observational evidence reviewed in the foregoing sections has led us progressively toward the corner of parameter space occupied by vacuum-dominated models with close to (or exactly) the critical density. The resulting picture is self-consistent and agrees with nearly all the data. Major questions, however, remain on the theoretical side. Prime among these is the problem of the cosmological constant, which (as previously described) is particularly acute in models with non-zero values of Λ , because one can no longer hope that a simple symmetry of nature will eventually be found which requires $\Lambda = 0$.

A related concern has to do with the *evolution* of the matter and vacuum energy density parameters Ω_m and Ω_Λ over time. Equations (2.32) and (2.35) combine to give

$$\Omega_m(t) \equiv \frac{\rho_m(t)}{\rho_{\text{crit}}(t)} = \frac{\Omega_{m,0}}{\tilde{R}^3(t)\tilde{H}^2(t)} \quad \Omega_\Lambda(t) \equiv \frac{\rho_\Lambda(t)}{\rho_{\text{crit}}(t)} = \frac{\Omega_{\Lambda,0}}{\tilde{H}^2(t)}. \quad (4.17)$$

Here $\tilde{H}[z(t)]$ is given by (2.40) as usual and $z(t) = 1/\tilde{R}(t) - 1$ from (2.15). Equations (4.17) can be solved exactly for flat models using (2.68) and (2.69) for

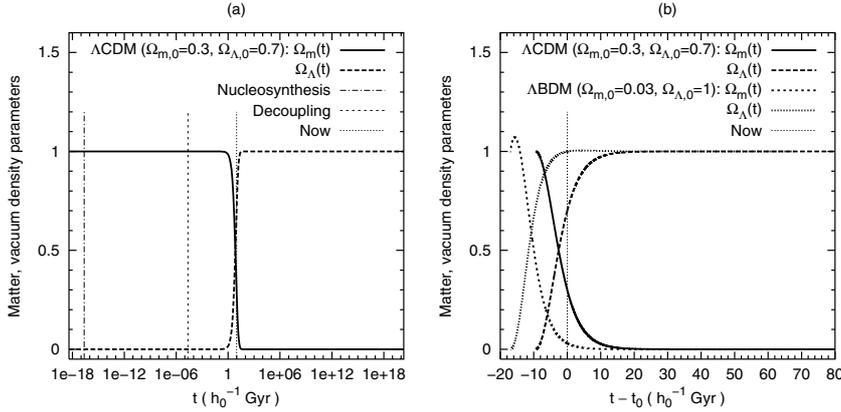


Figure 4.5. The evolution of $\Omega_m(t)$ and $\Omega_\Lambda(t)$ in vacuum-dominated models. The left-hand panel (a) shows a single model (Λ CDM) over 20 powers of time in either direction (after a similar plot against scale factor R in [54]). Plotted this way, we are seen to live at a very special time (marked ‘Now’). Standard nucleosynthesis ($t_{\text{nuc}} \sim 1$ s) and matter–radiation decoupling times ($t_{\text{dec}} \sim 10^{11}$ s) are included for comparison. The right-hand panel (b) shows both the Λ CDM and Λ BDM models on a linear rather than logarithmic scale, for the first $100h_0^{-1}$ Gyr after the big bang (i.e. the lifetime of the stars and galaxies).

$\tilde{R}(t)$ and $\tilde{H}(t)$. Results for the Λ CDM model are illustrated in figure 4.5(a). At early times, the cosmological term is insignificant ($\Omega_\Lambda \rightarrow 0$ and $\Omega_m \rightarrow 1$), while at late ones it entirely dominates and the Universe is effectively empty ($\Omega_\Lambda \rightarrow 1$ and $\Omega_m \rightarrow 0$).

What is remarkable in this figure is the location of the present (marked ‘Now’) in relation to the values of Ω_m and Ω_Λ . *We have apparently arrived on the scene at the precise moment when these two parameters are in the midst of switching places.* (We have not considered radiation density Ω_r here, but similar considerations apply to it.) This has come to be known as the **coincidence problem** and Carroll [54] has aptly described such a Universe as ‘preposterous’, writing: ‘This scenario staggers under the burden of its unnaturalness, but nevertheless crosses the finish line well ahead of any of its competitors by agreeing so well with the data’. Cosmology may indeed be moving toward a position like that of particle physics, in which there is a standard model which accurately accounts for all observed phenomena but which appears to be founded on a series of finely-tuned parameter values which leave one with the distinct impression that the underlying reality has not yet been grasped.

Figure 4.5(b) is a close-up view of figure 4.5(a), with one difference: it is plotted on a *linear* scale in time for the first $100h_0^{-1}$ Gyr after the big bang, rather than a logarithmic scale over $10^{\pm 20}h_0^{-1}$ Gyr. The rationale for this is simple:

100 Gyr is approximately the lifespan of the galaxies (as determined by their main-sequence stellar populations). One would not, after all, expect observers to appear on the scene long after all the galaxies had disappeared or, for that matter, in the early stages of the expanding fireball. Seen from the perspective of figure 4.5(b), the coincidence, while still striking, is perhaps no longer so preposterous. However, the Λ CDM model still appears fine-tuned, in that ‘Now’ follows rather quickly on the heels of the epoch of matter–vacuum equality. In the Λ BDM model, $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$ are closer to the cosmological time-averages of $\Omega_m(t)$ and $\Omega_\Lambda(t)$, namely zero and one respectively. In such a picture it might be easier to believe that we have not arrived on the scene at a special time but merely a *late* one. Whether or not this is a helpful way to approach the coincidence problem is, to a certain extent, a matter of taste.

To summarize the contents of this chapter: we have learned that what can be seen with our telescopes constitutes no more than 1% of the density of the Universe. The rest is dark. A small portion (no more than 5%) of this dark matter is made up of ordinary baryons. Many observational arguments hint at the existence of a second, more exotic species known as cold dark matter (though they do not quite establish its existence unless they are combined with ideas about the formation of large-scale structure). Experiments also imply the existence of a third dark-matter species, the massive neutrino, although its role appears to be more limited. Finally, all these components are dwarfed in importance by a newcomer whose physical origin remains shrouded in obscurity: the energy of the vacuum.

In the chapters that follow, we will explore the leading contenders for the dark matter in more depth: vacuum energy, axions, neutrinos, supersymmetric weakly-interacting massive particles (WIMPs) and objects like black holes. Our focus will be on what can be learned about each one from its possible contributions to the extragalactic background light at all wavelengths, from the radio region to the γ -ray. In the spirit of Olbers’ paradox and what we have done so far, our main question for each candidate will thus be: Just *how* dark is it?

References

- [1] Kapteyn J C 1922 *Astrophys. J.* **55** 302
- [2] Oort J H 1932 *Bull. Astron. Inst. Neth.* **6** 249
- [3] Zwicky F 1937 *Astrophys. J.* **86** 217
- [4] Fukugita M, Hogan C J and Peebles P J E 1998 *Astrophys. J.* **503** 518
- [5] Freedman W L *et al* 2001 *Astrophys. J.* **553** 47
- [6] Peacock J A 1999 *Cosmological Physics* (Cambridge: Cambridge University Press) pp 141–5
- [7] Willick J A and Batra P 2001 *Astrophys. J.* **548** 564
- [8] Herrnstein J R *et al* 1999 *Nature* **400** 539
- [9] Maoz E *et al* 1999 *Nature* **401** 351
- [10] Guinan E F *et al* 1998 *Astrophys. J.* **509** L21

- [11] Udalski A 2000 *Astrophys. J.* **531** L25
- [12] Romaniello M *et al* 2000 *Astrophys. J.* **530** 738
- [13] Cowan J J *et al* 1999 *Astrophys. J.* **521** 194
- [14] Griffiths L M, Melchiorri A and Silk J 2001 *Astrophys. J.* **553** L5
- [15] Olive K A 2000 *Nucl. Phys. Proc. Suppl.* **80** 79
- [16] Tytler D, O'Meara J M, Suzuki N and Lubin D 2000 *Phys. Scr.* **85** 12
- [17] Tripp T M, Savage B D and Jenkins E B 2000 *Astrophys. J.* **534** L1
- [18] Miralda-Escudé J *et al* 1996 *Astrophys. J.* **471** 582
- [19] Carr B J and Sakellariadou M 1998 *Astrophys. J.* **516** 195
- [20] Schneider P 1993 *Astron. Astrophys.* **279** 1
- [21] Alcock C *et al* 2000 *Astrophys. J.* **542** 281
- [22] Ibata R A *et al* 1999 *Astrophys. J.* **524** L95
- [23] Fields B D, Freese K and Graff D S 2000 *Astrophys. J.* **534** 265
- [24] Graff D S and Freese K 1996 *Astrophys. J.* **456** L49
- [25] Hansen B M S 1999 *Astrophys. J.* **517** L39
- [26] Hegyi D J and Olive K A 1986 *Astrophys. J.* **303** 56
- [27] Kochanek C S 1996 *Astrophys. J.* **457** 228
- [28] Carlberg R G, Yee H K C and Ellingson E 1997 *Astrophys. J.* **478** 462
- [29] Carlberg R G *et al* 1999 *Astrophys. J.* **516** 552
- [30] Bahcall N A *et al* 2000 *Astrophys. J.* **541** 1
- [31] Bahcall N A *et al* P J 1999 *Science* **284** 1481
- [32] Guerra E J, Daly R A and Wan L 2000 *Astrophys. J.* **544** 659
- [33] Moore B 1994 *Nature* **370** 629
- [34] Klypin A A *et al* 1999 *Astrophys. J.* **522** 82
- [35] Padmanabhan T 1993 *Structure Formation in the Universe* (Cambridge: Cambridge University Press) pp 63, 135–6
- [36] Bahcall N A and Fan X 1998 *Astrophys. J.* **504** 1
- [37] Weinberg D H *et al* 1999 *Astrophys. J.* **522** 563
- [38] Peacock J A *et al* 2001 *Proc. 20th Texas Symposium on Relativistic Astrophysics* ed J C Wheeler and H Martel (New York: American Institute of Physics) p 938
- [39] Zehavi I and Dekel A 1999 *Nature* **401** 252
- [40] Peebles P J E 1993 *Principles of Physical Cosmology* (Princeton, NJ: Princeton University Press) pp 158–63
- [41] Sciama D W 1993 *Modern Cosmology and the Dark Matter Problem* (Cambridge: Cambridge University Press)
- [42] Gawiser E and Silk J 1998 *Science* **280** 1405
- [43] Croft R A C, Hu W and Davé R 1999 *Phys. Rev. Lett.* **83** 1092
- [44] Kolb E W and Turner M S 1990 *The Early Universe* (Reading, MA: Addison-Wesley) p 352
- [45] Fukuda Y *et al* 1998 *Phys. Rev. Lett.* **81** 1562
- [46] Abdurashitov N *et al* 1999 *Phys. Rev. Lett.* **83** 4686
- [47] Cleveland B T *et al* 1998 *Astrophys. J.* **496** 505
- [48] Hampel W *et al* 1999 *Phys. Lett. B* **447** 127
- [49] Athanassopoulos C *et al* 1998 *Phys. Rev. Lett.* **81** 1774
- [50] Klapdor-Kleingrothaus H V, Päs H and Smirnov A Y 2001 *Phys. Rev. D* **63** 073005
- [51] Rindler W 1977 *Essential Relativity* (Berlin: Springer) p 226
- [52] Priester W, Hoell J and Blome H-J 1995 *Comments Astrophys.* **17** 327
- [53] Gamow G 1970 *My World Line* (New York: Viking Press) p 44

- [54] Carroll S M 2001 *Living Reviews in Relativity* **4**
<http://www.livingreviews.org/Articles/Volume4/2001-1carroll>
- [55] Puthoff H E 1989 *Phys. Rev. A* **40** 4857
- [56] Wesson P S 1991 *Astrophys. J.* **378** 466
- [57] Overduin J M and Cooperstock F I 1998 *Phys. Rev. D* **58** 043506
- [58] Wesson P S and Liu H 2001 *Int. J. Mod. Phys. D* **10** 905
- [59] Weinberg S 2001 *Sources and Detection of Dark Matter in the Universe* ed D B Cline (Berlin: Springer) p 18
- [60] Blome H-J and Priester W 1985 *Astrophys. Sp. Sci.* **117** 327
- [61] Gott J R III, Park M-G and Lee H M 1989 *Astrophys. J.* **338** 1
- [62] Franx M *et al* 1997 *Astrophys. J.* **486** L75
- [63] Kochanek C S 1996 *Astrophys. J.* **466** 638
- [64] Malhotra S, Rhoads J E and Turner E L 1997 *Mon. Not. R. Astron. Soc.* **288** 138
- [65] Falco E E, Kochanek C S and Muñoz J A 1998 *Astrophys. J.* **494** 47
- [66] Cooray A R 1999 *Astron. Astrophys.* **342** 353
- [67] Fukugita M *et al* 1990 *Astrophys. J.* **361** L1
- [68] Gardner J P, Cowie L L and Wainscoat R J 1993 *Astrophys. J.* **415** L9
- [69] Totani T, Yoshii Y and Sato K 1997 *Astrophys. J.* **483** L75
- [70] Totani T and Yoshii Y 2000 *Astrophys. J.* **540** 81
- [71] Jenkins A *et al* 1998 *Astrophys. J.* **499** 20
- [72] Feldman H A and Evrard A E 1993 *Int. J. Mod. Phys. D* **2** 113
- [73] White M and Scott D 1996 *Astrophys. J.* **459** 415
- [74] Riess A G *et al* 1998 *Astron. J.* **116** 1009
- [75] Perlmutter S *et al* 1999 *Astrophys. J.* **517** 565
- [76] Aguirre A and Haiman Z 2000 *Astrophys. J.* **532** 28
- [77] Drell P S, Loredo T J and Wasserman I 2000 *Astrophys. J.* **530** 593
- [78] Hillebrandt W and Niemeyer J C 2000 *Ann. Rev. Astron. Astrophys.* **38** 191
- [79] Lineweaver C H 2001 *Gravitational Lensing: Recent Progress and Future Goals* (Astronomical Society of the Pacific Conference Series, Volume 237) ed T G Brainerd and C S Kochanek (San Francisco, CA: ASP)
- [80] de Bernardis P *et al* 2000 *Nature* **404** 955
- [81] Hanany S *et al* 2000 *Astrophys. J.* **545** L5
- [82] Pryke C *et al* 2001 *Astrophys. J.* **568** 46
- [83] Bennett C *et al* 1996 *Astrophys. J.* **464** L1
- [84] Jaffe A H *et al* 2000 *Phys. Rev. Lett.* **86** 3475
- [85] White M, Scott D and Pierpaoli E 2000 *Astrophys. J.* **545** 1
- [86] McGaugh S S 2000 *Astrophys. J.* **541** L33
- [87] Liebscher D-E, Priester W and Hoell J 1992 *Astron. Astrophys.* **261** 377
- [88] Overduin J M and Priester W 2001 *Naturwissenschaften* **88** 229

Chapter 5

The vacuum

5.1 Vacuum decay

It may be surprising to find the vacuum, traditionally defined as the ‘absence of anything’, playing a large and possibly dominant role in the dynamics of the Universe. But quantum mechanics has taught us to regard the vacuum as the *ground state* of all possible fields, and ground states in quantum field theory can have very high energies indeed. Of course, it is the astonishing gap between these theoretical energy densities and those which now seem to be observed in cosmology which constitutes the cosmological-constant problem, perhaps the single deepest unsolved puzzle in theoretical physics today.

Many authors have sought to address the crisis by finding a mechanism by which the energy density ρ_v of the vacuum could *decay* with time. This would be equivalent to a **variable cosmological term** since $\Lambda c^2 = 8\pi G\rho_v$ from (2.34). With such a mechanism in hand, the problem would be reduced to explaining why the Universe is of intermediate age: old enough that Λ has relaxed from primordial values like those suggested by quantum field theory to the values which we measure now, but young enough that $\Omega_\Lambda \equiv \rho_v/\rho_{\text{crit}}$ has not yet reached its asymptotic value of unity.

Energy conservation requires that any decrease in the energy density of the vacuum be made up by a corresponding increase somewhere else. In some scenarios, vacuum energy goes into the kinetic energy of new forms of matter such as scalar fields, which have yet to be observed in nature. In others it is channelled instead into baryons, photons or neutrinos. Baryonic decays would produce equal amounts of matter and antimatter, whose subsequent annihilation would flood the Universe with γ -rays. Radiative decays would similarly pump photons into intergalactic space, but are harder to constrain because they could, in principle, involve any part of the electromagnetic spectrum. As we will see, however, robust limits can be set on any such process under conservative assumptions.

We proceed in the remainder of the chapter to describe how cosmology is modified in the presence of a decaying cosmological ‘constant’ and then

specialize to the case of vacuum decay into radiation. Our objective is to assess the impact of such a process on both the bolometric and spectral EBL intensity, using the same formalism that was applied in chapters 2 and 3 to the light from ordinary galaxies.

5.2 The variable cosmological ‘constant’

Einstein originally introduced the Λ -term in 1917 as a constant of nature akin to c or G , and it is worth beginning by asking how such a quantity can be allowed to vary. Suppose that we take the covariant divergence of Einstein’s field equations (2.1) and apply the **Bianchi identities**. The latter are a mathematical statement about Riemannian geometry and read $\nabla^\nu(\mathcal{R}_{\mu\nu} - \frac{1}{2}\mathcal{R}g_{\mu\nu}) = 0$. With these conditions, and the fact that the metric tensor itself has a vanishing covariant derivative ($\nabla^\nu g_{\mu\nu} = 0$), we find

$$\partial_\mu \Lambda = \frac{8\pi G}{c^4} \nabla^\nu \mathcal{T}_{\mu\nu}. \quad (5.1)$$

Assuming that matter and energy (as contained in $\mathcal{T}_{\mu\nu}$) are conserved, it follows that $\partial_\mu \Lambda = 0$ and, hence, that $\Lambda = \text{constant}$.

In variable- Λ theories, one must therefore do one of three things: abandon matter–energy conservation, modify general relativity or stretch the definition of what is conserved. The first of these routes was explored in 1933 by Bronstein [1], who sought to connect energy non-conservation with the observed fact that the Universe is expanding rather than contracting. (Einstein’s equations actually allow both types of solutions, and a convincing explanation for this **cosmological arrow of time** has yet to be given.) Bronstein was executed in Stalin’s Soviet Union a few years later and his work is not widely known [2].

Today, few physicists would be willing to sacrifice energy conservation outright. Some, however, would be willing to modify general relativity, or to consider new forms of matter and energy. Historically, these two approaches have sometimes been seen as distinct, with one being a change to the ‘geometry of nature’ while the other is concerned with the material content of the Universe. The modern tendency, however, is to regard them as equivalent. This viewpoint is best personified by Einstein, who in 1936 compared the left-hand (geometrical) and right-hand (matter) sides of his field equations to ‘fine marble’ and ‘low-grade wooden’ wings of the same house [3]. In a more complete theory, he argued, matter fields of all kinds would be seen to be just as geometrical as the gravitational one.

Let us see how this works in one of the oldest and simplest variable- Λ theories: a modification of general relativity in which the metric tensor $g_{\mu\nu}$ is supplemented by a **scalar field** φ whose coupling to matter is determined by a parameter ω . The idea for such a theory goes back to Jordan in 1949 [4], Fierz in 1956 [5] and Brans and Dicke in 1961 [6]. In those days, new scalar fields

were not introduced into theoretical physics as routinely as they are today, and all these authors sought to associate φ with a known quantity. Various lines of argument (notably **Mach's principle**) pointed to an identification with Newton's gravitational 'constant' such that $G \sim 1/\varphi$. By 1968 it was appreciated that Λ and ω too would depend on φ in general [7]. The original **Brans–Dicke theory** (with $\Lambda = 0$) has thus been extended to generalized scalar–tensor theories in which $\Lambda = \Lambda(\varphi)$ [8], $\Lambda = \Lambda(\varphi)$, $\omega = \omega(\varphi)$ [9] and $\Lambda = \Lambda(\varphi, \psi)$, $\omega = \omega(\varphi)$ where $\psi \equiv \partial^\mu \varphi \partial_\mu \varphi$ [10]. Let us consider the last and most general of these cases, for which the field equations are found to read:

$$\begin{aligned} \mathcal{R}_{\mu\nu} - \frac{1}{2}\mathcal{R}g_{\mu\nu} + \frac{1}{\varphi}[\nabla_\mu(\partial_\nu\varphi) - \square\varphi g_{\mu\nu}] + \frac{\omega(\varphi)}{\varphi^2}\left(\partial_\mu\varphi\partial_\nu\varphi - \frac{1}{2}\psi g_{\mu\nu}\right) \\ - \Lambda(\varphi, \psi)g_{\mu\nu} + 2\frac{\partial\Lambda(\varphi, \psi)}{\partial\psi}\partial_\mu\varphi\partial_\nu\varphi = -\frac{8\pi}{\varphi c^4}\mathcal{T}_{\mu\nu} \end{aligned} \quad (5.2)$$

where $\square\varphi \equiv \nabla^\mu(\partial_\mu\varphi)$ is the D'Alembertian. These reduce to Einstein's equations (2.1) when $\varphi = \text{constant} = 1/G$.

If we now repeat the exercise on the previous page and take the covariant derivative of the field equations (5.2) with the Bianchi identities, we obtain a generalized version of the equation (5.1) faced by Bronstein:

$$\begin{aligned} \partial_\mu\varphi\left\{\frac{\mathcal{R}}{2} + \frac{\omega(\varphi)}{2\varphi^2}\psi - \frac{\omega(\varphi)}{\varphi}\square\varphi + \Lambda(\varphi, \psi) + \varphi\frac{\partial\Lambda(\varphi, \psi)}{\partial\varphi} - \frac{\psi}{2\varphi}\frac{d\omega(\varphi)}{d\varphi}\right. \\ \left. - 2\varphi\square\varphi\frac{\partial\Lambda(\varphi, \psi)}{\partial\psi} - 2\partial^\kappa\varphi\partial_\kappa\left[\varphi\frac{\partial\Lambda(\varphi, \psi)}{\partial\psi}\right]\right\} = \frac{8\pi}{c^4}\nabla^\nu\mathcal{T}_{\mu\nu}. \end{aligned} \quad (5.3)$$

Now energy conservation ($\nabla^\nu\mathcal{T}_{\mu\nu} = 0$) no longer requires $\Lambda = \text{constant}$. In fact, it is generally *incompatible* with constant Λ , unless an extra condition is imposed on the terms inside the curly brackets in (5.3). (This cannot come from the wave equation for φ , which merely confirms that the terms inside the curly brackets sum to zero, in agreement with energy conservation.) Similar conclusions hold for other scalar–tensor theories in which φ is no longer associated with G . Examples include models with non-minimal couplings between φ and the curvature scalar \mathcal{R} [11], conformal rescalings of the metric tensor by functions of φ [12] and non-zero potentials $V(\varphi)$ for the scalar field [13–15]. (Theories of this last kind are now known as **quintessence scenarios** [16].) For all such cases, the cosmological 'constant' becomes a dynamical variable.

In the modern approach to variable- Λ cosmology, which goes back to Zeldovich in 1968 [17], all extra terms of the kind just described—including Λ —are moved to the right-hand side of the field equations (5.2), leaving only the **Einstein tensor** ($\mathcal{R}_{\mu\nu} - \frac{1}{2}\mathcal{R}g_{\mu\nu}$) to make up the 'geometrical' left-hand side. The cosmological term, along with scalar (or other) additional fields, are thus effectively reinterpreted as *new kinds of matter*. The field equations (5.2) may then be written

$$\mathcal{R}_{\mu\nu} - \frac{1}{2}\mathcal{R}g_{\mu\nu} = -\frac{8\pi}{\varphi c^4}\mathcal{T}_{\mu\nu}^{\text{eff}} + \Lambda(\varphi)g_{\mu\nu}. \quad (5.4)$$

Here $\mathcal{T}_{\mu\nu}^{\text{eff}}$ is an **effective energy–momentum tensor** describing the sum of ordinary matter plus whatever scalar (or other) fields have been added to the theory. For generalized scalar–tensor theories as described above, this could be written as $\mathcal{T}_{\mu\nu}^{\text{eff}} \equiv \mathcal{T}_{\mu\nu} + \mathcal{T}_{\mu\nu}^{\varphi}$ where $\mathcal{T}_{\mu\nu}$ refers to ordinary matter and $\mathcal{T}_{\mu\nu}^{\varphi}$ to the scalar field. For the case with $\Lambda = \Lambda(\varphi)$ and $\omega = \omega(\varphi)$, for instance, the latter would be defined by (5.2) as

$$\mathcal{T}_{\mu\nu}^{\varphi} \equiv \frac{1}{\varphi} [\nabla_{\mu}(\partial_{\nu}\varphi) - \square\varphi g_{\mu\nu}] + \frac{\omega(\varphi)}{\varphi^2} \left(\partial_{\mu}\varphi\partial_{\nu}\varphi - \frac{1}{2}\psi g_{\mu\nu} \right). \quad (5.5)$$

The covariant derivative of the field equations (5.4) with the Bianchi identities now gives

$$0 = \nabla^{\nu} \left[\frac{8\pi}{\varphi c^4} \mathcal{T}_{\mu\nu}^{\text{eff}} - \Lambda(\varphi) g_{\mu\nu} \right]. \quad (5.6)$$

Equation (5.6) carries the same physical content as (5.3), but is more general in form and can readily be extended to other theories. Physically, it says that energy *is* conserved in variable- Λ cosmology—where ‘energy’ is now understood to refer to the energy of ordinary matter along with that in any additional fields which may be present, *and along with that in the vacuum*, as represented by Λ . In general, the latter parameter can vary as it likes, so long as the conservation equation (5.6) is satisfied.

It is natural to wonder whether the evolution of Λ in these theories actually helps with the cosmological ‘constant’ problem, in the sense that Λ drops from large primordial values to ones like those seen today without fine-tuning. Behaviour of this kind was noted at least as early as 1977 [8] in the context of models with $\Lambda = \Lambda(\varphi)$ and $\omega = \text{constant}$, which have solutions for $\varphi(t)$ such that $\Lambda \propto t^{-2}$. In precursors to the modern quintessence scenarios, Barr [13] found models in which $\Lambda \propto t^{-\ell}$ at late times, while Peebles and Ratra [14] discussed a theory in which $\Lambda \propto R^{-m}$ at early ones (here ℓ and m are powers). There is now a rich literature on Λ -decay laws of this kind (see [18] for a review). Their appeal is easy to understand and can be illustrated with a simple dimensional argument for the case in which $\Lambda \propto R^{-2}$ [19]. Since Λ already has dimensions of L^{-2} , the proportionality factor in this case is a pure number (α , say) which is presumably of order unity. Taking $\alpha \sim 1$ and identifying R with a suitable length scale in cosmology (namely the Hubble distance c/H_0), one finds that $\Lambda_0 \sim H_0^2/c^2$. The present vacuum density parameter is then $\Omega_{\Lambda,0} \equiv \Lambda_0 c^2/3H_0^2 \sim 1/3$, close to the values implied by current supernovae data (section 4.6). A natural choice for the primordial value of Λ , moreover, would have $R \sim \ell_{\text{Pl}}$ so that $\Lambda_{\text{Pl}} \sim \alpha \ell_{\text{Pl}}^{-2}$. This leads to a ratio $\Lambda_{\text{Pl}}/\Lambda_0 \sim (c/H_0 \ell_{\text{Pl}})^2 \sim 10^{122}$, which may be compared with the values in table 4.1.

While this would seem to be a promising approach, two cautions must be kept in mind. The first is theoretical. Insofar as the mechanisms discussed so far are entirely classical, they do not address the underlying problem. For this, one would also need to explain why net contributions to Λ from the *quantum*

vacuum do not remain at the primordial level or how they are suppressed with time. Polyakov [20] and Adler [21] in 1982 were the first to speculate explicitly that such a suppression might come about if the ‘bare’ cosmological term implied by quantum field theory were progressively screened by an ‘induced’ counterterm of opposite sign, driving the effective value of $\Lambda(t)$ toward zero at late times. Many theoretical **adjustment mechanisms** have now been identified as potential sources of such a screening effect, beginning with a 1983 suggestion by Dolgov [22] based on non-minimally-coupled scalar fields. Subsequent proposals have involved scalar fields [23–25], fields of higher spin [26–28], quantum effects during inflation [29–31] and other phenomena [32–34]. In most of these cases, no analytic expression is found for Λ in terms of time or other cosmological parameters; the intent is merely to demonstrate that decay (and preferably near-cancellation) of the cosmological term is possible in principle. None of these mechanisms has been widely accepted as successful to date. In fact, there is a general argument due to Weinberg to the effect that a successful mechanism based on scalar fields would necessarily be so finely tuned as to be just as mysterious as the original problem [35]. Similar concerns have been raised in the case of vector and tensor-based proposals [36]. Nevertheless, the idea of the adjustment mechanism remains feasible in principle and continues to attract more attention than any other approach to the cosmological-constant problem.

The second caution is empirical. Observational data place increasingly strong restrictions on the way in which Λ can vary with time. Among the most important are early-time bounds on the vacuum energy density $\rho_\Lambda c^2 = \Lambda c^4/8\pi G$. The success of standard primordial nucleosynthesis theory implies that ρ_Λ was smaller than ρ_r and ρ_m during the radiation-dominated era, and large-scale structure formation could not have proceeded in the conventional way unless $\rho_\Lambda < \rho_m$ during the early matter-dominated era. Since $\rho_r \propto R^{-4}$ and $\rho_m \propto R^{-3}$ from (2.32), these requirements mean in practice that the vacuum energy density must climb *less steeply than* R^{-3} in the past direction, if it is comparable to that of matter or radiation at present [37, 38]. The variable- Λ term must also satisfy late-time bounds like those which have been placed on the cosmological constant (section 4.6). Tests of this kind have been carried out using data on the age of the Universe [39, 40], structure formation [41–43], galaxy number counts [44], the CMB power spectrum [45, 46], gravitational lensing statistics [46–48] and Type Ia supernovae [46, 49]. Some of these tests are less restrictive in the case of a variable Λ -term than they are for $\Lambda = \text{constant}$, and this can open up new regions of parameter space. Observation may even be compatible with some non-singular models whose expansion originates in a hot, dense ‘big bounce’ rather than a big bang [50], a possibility which can be ruled out on quite general grounds if $\Lambda = \text{constant}$.

A third group of limits comes from asking what the vacuum decays *into*. In quintessence theories, vacuum energy is transferred to the kinetic energy of a scalar field as it ‘rolls’ down a gradient toward the minimum of its potential. This may have observable consequences if the scalar field is coupled strongly to

ordinary matter, but is hard to constrain in general. A simpler situation is that in which the vacuum decays into known particles such as baryons, photons or neutrinos. The baryonic decay channel would produce excessive levels of γ -ray background radiation due to matter–antimatter annihilation unless the energy density of the vacuum component is less than 3×10^{-5} times that of matter [37]. This limit can be weakened if the decay process violates baryon number or if it takes place in such a way that matter and antimatter are segregated on large scales, but such conditions are hard to arrange in a natural way. The radiative decay channel is more promising, but also faces a number of tests. The decay process should meet certain criteria of thermodynamic stability [51] and adiabaticity [52]. The *shape* of the spectrum of decay photons offers another possibility. If this differs from that of pre-existing background radiation, then distortions will arise. Freese *et al* have argued on this basis that a vacuum decaying primarily into low-energy photons could have a density no more than 4×10^{-4} times that of radiation [37].

It may be, however, that vacuum-decay photons blend into the spectrum of background radiation without distorting it. Figure 2.1 shows that the best place to ‘hide’ large quantities of excess energy would be the microwave and infrared regions where the energy density of background radiation is highest. Could part of the CMB originate in a decaying vacuum? We know from the COBE satellite that its spectrum is very nearly that of a perfect blackbody [53]. Freese *et al* pointed out that vacuum-decay photons would be thermalized efficiently by brehmsstrahlung and double-Compton scattering in the early Universe, and might continue to assume a blackbody spectrum at later times if pre-existing CMB photons played a role in ‘inducing’ the vacuum to decay [37]. Subsequent work has shown that this would require a special combination of thermodynamical parameters [54]. Such a possibility is important in practice, however, because it leads to the most conservative limits on the theory. If the radiation produced by vacuum decay does not distort the background, it will in any case contribute to the latter’s *absolute intensity*. We can calculate the size of these contributions to the background radiation using the methods that have been laid out in chapters 2 and 3.

5.3 Energy density

The first step in this problem is to solve the field equations and conservation equations for the energy density of the decaying vacuum. We will do this in the context of a general **phenomenological model**. This means that we retain the field equations (5.4) and the conservation law (5.6) without specifying the form of the effective energy–momentum tensor in terms of scalar (or other) fields. These equations may be written

$$\mathcal{R}_{\mu\nu} - \frac{1}{2}\mathcal{R}g_{\mu\nu} = -\frac{8\pi G}{c^4}(\mathcal{T}_{\mu\nu}^{\text{eff}} - \rho_{\Lambda}c^2g_{\mu\nu}) \quad (5.7)$$

$$0 = \nabla^\nu (\mathcal{T}_{\mu\nu}^{\text{eff}} - \rho_\Lambda c^2 g_{\mu\nu}). \quad (5.8)$$

Here $\rho_\Lambda c^2 \equiv \Lambda c^4 / 8\pi G$ from (2.34) and we have put back G in place of $1/\varphi$. Equations (5.7) and (5.8) have the same form as their counterparts (2.1) and (2.29) in standard cosmology, the key difference being that the cosmological term has migrated to the right-hand side and is *no longer necessarily constant*. Its dynamical evolution is now governed by the conservation equations (5.8), which require only that any change in $\rho_\Lambda c^2 g_{\mu\nu}$ be offset by an equal and opposite change in the energy–momentum tensor $\mathcal{T}_{\mu\nu}^{\text{eff}}$.

While the latter is model-dependent in general, it is reasonable to assume in the context of isotropic and homogeneous cosmology that its form is that of a perfect fluid, as given by (2.26):

$$\mathcal{T}_{\mu\nu}^{\text{eff}} = (\rho_{\text{eff}} + p_{\text{eff}}/c^2) U_\mu U_\nu + p_{\text{eff}} g_{\mu\nu}. \quad (5.9)$$

Comparison of equations (5.8) and (5.9) shows that the conserved quantity in (5.8) must then *also* have the form of a perfect-fluid energy–momentum tensor, with density and pressure given by

$$\rho = \rho_{\text{eff}} + \rho_\Lambda \quad p = p_{\text{eff}} - \rho_\Lambda c^2. \quad (5.10)$$

The conservation law (5.8) may then be simplified at once by analogy with equation (2.29):

$$\frac{1}{R^3} \frac{d}{dt} [R^3 (\rho_{\text{eff}} c^2 + p_{\text{eff}})] = \frac{d}{dt} (p_{\text{eff}} - \rho_\Lambda c^2). \quad (5.11)$$

This reduces to the standard result (2.30) for the case of a constant cosmological term, $\rho_\Lambda = \text{constant}$.

In this chapter, we will allow the cosmological term to contain both a constant part *and* a time-varying part so that

$$\rho_\Lambda = \rho_c + \rho_v(t) \quad \rho_c = \text{constant}. \quad (5.12)$$

Let us assume, in addition, that the perfect fluid described by $\mathcal{T}_{\mu\nu}^{\text{eff}}$ consists of a mixture of dustlike matter ($p_m = 0$) and radiation ($p_r = \frac{1}{3}\rho_r c^2$):

$$\rho_{\text{eff}} = \rho_m + \rho_r \quad p_{\text{eff}} = \frac{1}{3}\rho_r c^2. \quad (5.13)$$

The conservation equation (5.11) then reduces to

$$\frac{1}{R^4} \frac{d}{dt} (R^4 \rho_r) + \frac{1}{R^3} \frac{d}{dt} (R^3 \rho_m) + \frac{d\rho_v}{dt} = 0. \quad (5.14)$$

From this equation it is clear that one (or both) of the radiation and matter densities can no longer obey the usual relations $\rho_r \propto R^{-4}$ and $\rho_m \propto R^{-3}$ in

a theory with $\Lambda \neq \text{constant}$. Any change in Λ (or ρ_Λ) must be accompanied by a change in radiation and/or matter densities.

To go further, some simplifying assumptions must be made. Let us take to begin with

$$\frac{d}{dt}(R^3 \rho_m) = 0. \quad (5.15)$$

This is equivalent to imposing **conservation of particle number**, as may be seen by replacing ‘galaxies’ with ‘particles’ in equation (2.7). Such an assumption is well justified during the matter-dominated era by the stringent constraints on matter creation discussed in section 5.2. It is equally well justified during the radiation-dominated era, when the matter density is small so that the ρ_m term is of secondary importance compared to the other terms in (5.14) in any case.

In light of equations (5.14) and (5.15), the vacuum can exchange energy only with radiation. As a model for this process, let us follow Pollock in 1980 [55] and assume that it takes place in such a way that the energy density of the decaying vacuum component remains proportional to that of radiation, $\rho_v \propto \rho_r$. We adopt the notation of Freese *et al* in 1987 and write the proportionality factor as $x/(1-x)$ with x the **coupling parameter** of the theory [37]. If this is allowed to take (possibly different) constant values during the radiation- and matter-dominated eras, then

$$x \equiv \frac{\rho_v}{\rho_r + \rho_v} = \begin{cases} x_r & (t < t_{\text{eq}}) \\ x_m & (t \geq t_{\text{eq}}). \end{cases} \quad (5.16)$$

Here t_{eq} refers to the **epoch of matter–radiation equality** radiation when $\rho_r = \rho_m$. Standard cosmology is recovered in the limits $x_r \rightarrow 0$ and $x_m \rightarrow 0$. The most natural situation is that in which the value of x stays constant so that $x_r = x_m$. However, since observational constraints on x are, in general, different for the radiation- and matter-dominated eras, the most conservative limits on the theory are obtained by letting x_r and x_m take different values. Physically, this would correspond to a **phase transition** or sudden change in the expansion rate \dot{R}/R of the Universe at $t = t_{\text{eq}}$.

With the assumptions (5.15) and (5.16), the conservation equation (5.14) reduces to

$$\frac{\dot{\rho}_v}{\rho_v} + 4(1-x) \frac{\dot{R}}{R} = 0 \quad (5.17)$$

where overdots denote derivatives with respect to time. This may be integrated to give

$$\rho_v(R) = \alpha_v R^{-4(1-x)} \quad (5.18)$$

where $\alpha_v = \text{constant}$. The cosmological term Λ is thus an inverse power-law function of the scale factor R . This is a scenario that has been widely studied, also in cases where vacuum energy density is not proportional to that of radiation [18]. Equation (5.18) shows that the conserved quantity in this theory has a form intermediate between that of ordinary radiation entropy ($R^4 \rho_r$) and particle number ($R^3 \rho_m$) when $0 < x < \frac{1}{4}$.

The fact that $\rho_r \propto \rho_v \propto R^{-4(1-x)}$ places an immediate upper limit of $\frac{1}{4}$ on x (in both eras), since higher values would erase the dynamical distinction between radiation and matter. With $x \leq \frac{1}{4}$ it then follows from (5.16) that $\rho_v \leq \frac{1}{3}\rho_r$. This is consistent with section 5.2 where we noted that a vacuum energy component which climbs more steeply than R^{-3} in the past direction cannot have an energy density greater than that of radiation at present. Freese *et al* [37] set a stronger bound by showing that $x \leq 0.07$ if the **baryon-to-photon ratio** η is to be consistent with both primordial nucleosynthesis during the radiation-dominated era and CMB observations at present. (This argument assumes that $x = x_r = x_m$.) As a guideline in what follows, then, we will allow x_r and x_m to take values between zero and 0.07 and consider, in addition, the theoretical possibility that x_m could increase to 0.25 in the matter-dominated era.

With $\rho_m(R)$ specified by (5.15), ρ_r related to ρ_v by (5.16) and $\rho_v(R)$ given by (5.18), we can solve for all three components as functions of time if the scale factor $R(t)$ is known. This comes as usual from the field equations (5.7). Since these are the same as equations (2.1) for standard cosmology, they lead to the same result, equation (2.33):

$$\left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi G}{3}(\rho_m + \rho_r + \rho_v + \rho_c). \quad (5.19)$$

Here we have used equations (5.12) to replace ρ_Λ with $\rho_v + \rho_c$ and (5.13) to replace ρ_{eff} with $\rho_m + \rho_r$. We have also set $k = 0$ since observations indicate that these components together make up very nearly the critical density (section 4.6).

Equation (5.19) can be solved analytically for the three cases which are of greatest physical interest. We define these as follows:

- (1) the radiation-dominated regime ($t < t_{\text{eq}}$, $\rho_r + \rho_v \gg \rho_m + \rho_c$);
- (2) the matter-dominated regime ($t \geq t_{\text{eq}}$, $\rho_r + \rho_v \ll \rho_m$, $\rho_c = 0$); and
- (3) the vacuum-dominated regime ($t \geq t_{\text{eq}}$, $\rho_r + \rho_v \ll \rho_m + \rho_c$).

The distinction between regimes 2 and 3 is important because we will find in practice that the former describes models which are close to the EdS one, while the latter is needed to model vacuum-dominated cosmologies like Λ CDM or Λ BDM (table 3.1). The definitions of these terms should be amended slightly for this chapter, since we now consider flat models containing not only matter and a cosmological constant, but radiation and a decaying-vacuum component as well. The densities of the latter two components, however, are at least four orders of magnitude below that of matter at present. Thus models with $\rho_c = 0$, for example, have $\Omega_{m,0} = 1$ to four-figure precision or better and are dynamically indistinguishable from EdS during all but the first fraction (of order 10^{-4} or less) of their lifetimes. For definiteness, we will use the terms ‘EdS’, ‘ Λ CDM’ and ‘ Λ BDM’ in this chapter to refer to flat models in which $\Omega_{m,0} = 1$, 0.3 and 0.03 respectively, with the remainder (if any) of the present critical density being effectively made up by the constant-density component of the vacuum energy.

It remains to solve equations (5.15), (5.16), (5.18) and (5.19) for the four dynamical quantities R , ρ_m , ρ_r and ρ_v in terms of the constants $\rho_{m,0}$, $\rho_{r,0}$, x_r and x_m . We relegate the details of this exercise to appendix B and summarize the results here. The scale factor is found as

$$R(t) \propto \begin{cases} \left(\frac{t}{t_0}\right)^{1/2(1-x_r)} & (t < t_{\text{eq}}) \\ \left[\frac{\mathcal{S}_m(t)}{\mathcal{S}_m(t_0)}\right]^{2/3} & (t \geq t_{\text{eq}}). \end{cases} \quad (5.20)$$

The vacuum density is given by

$$\rho_v(t) = \begin{cases} \frac{\alpha x_r}{(1-x_r)^2} t^{-2} & (t < t_{\text{eq}}) \\ \left(\frac{x_m}{1-x_m}\right) \rho_r(t) & (t \geq t_{\text{eq}}) \end{cases} \quad (5.21)$$

where $\alpha = 3/(32\pi G) = 4.47 \times 10^5 \text{ g cm}^{-2} \text{ s}^2$. For the density of radiation we find

$$\rho_r(t) = \begin{cases} \left(\frac{1-x_r}{x_r}\right) \rho_v(t) & (t < t_{\text{eq}}) \\ \rho_{r,0} \left[\frac{\mathcal{S}_m(t)}{\mathcal{S}_m(t_0)}\right]^{-8(1-x_m)/3} & (t \geq t_{\text{eq}}). \end{cases} \quad (5.22)$$

The matter density can be written

$$\rho_m(t) = \rho_{m,0} \times \begin{cases} \left[\frac{\mathcal{S}_m(t_{\text{eq}})}{\mathcal{S}_m(t_0)}\right]^{-2} \left(\frac{t}{t_{\text{eq}}}\right)^{-3/2(1-x_r)} & (t < t_{\text{eq}}) \\ \left[\frac{\mathcal{S}_m(t)}{\mathcal{S}_m(t_0)}\right]^{-2} & (t \geq t_{\text{eq}}). \end{cases} \quad (5.23)$$

Here we have applied $\rho_{m,0} = \Omega_{m,0} \rho_{\text{crit},0}$ and $\rho_{r,0} = \Omega_{r,0} \rho_{\text{crit},0}$ as boundary conditions, with the values of $\Omega_{m,0}$ and $\Omega_{r,0}$ to be specified and $\rho_{\text{crit},0}$ given by (2.36). The function $\mathcal{S}_m(t)$ is defined as

$$\mathcal{S}_m(t) \equiv \begin{cases} t & (\Omega_{m,0} = 1) \\ \sinh(t/\tau_0) & (0 < \Omega_{m,0} < 1) \end{cases} \quad (5.24)$$

where $\tau_0 \equiv 2/(3H_0\sqrt{1-\Omega_{m,0}})$ and H_0 is given by (2.19). The age of the Universe is

$$t_0 = \begin{cases} 2/(3H_0) & (\Omega_{m,0} = 1) \\ \tau_0 \sinh^{-1} \chi_0 & (0 < \Omega_{m,0} < 1) \end{cases} \quad (5.25)$$

where $\chi_0 \equiv \sqrt{(1-\Omega_{m,0})/\Omega_{m,0}}$. The second of these expressions comes from (2.70). Corrections from the radiation-dominated era can be ignored since $t_0 \gg t_{\text{eq}}$ in all cases.

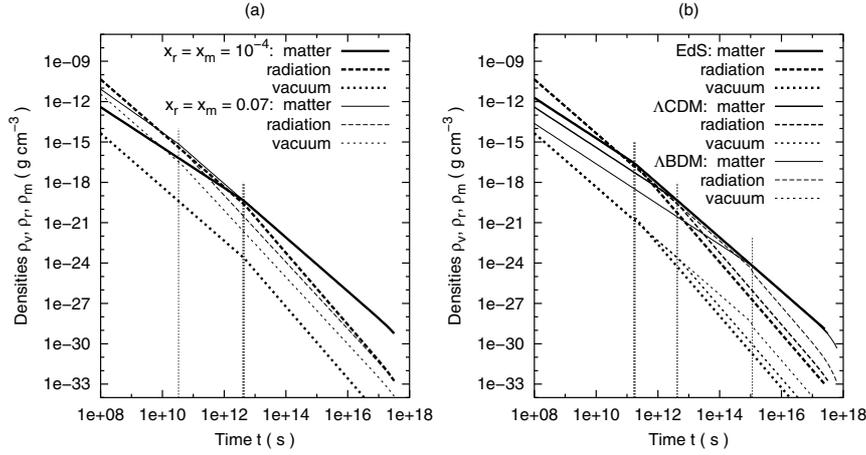


Figure 5.1. The densities of decaying vacuum energy (ρ_v), radiation (ρ_r) and matter (ρ_m) as functions of time. The left-hand panel (a) shows the effects of changing the values of x_r and x_m , assuming a model with $\Omega_{m,0} = 0.3$ (similar to Λ CDM). The right-hand panel (b) shows the effects of changing the cosmological model, assuming $x_r = x_m = 10^{-4}$. The vertical lines indicate the epochs when the densities of matter and radiation are equal (t_{eq}). All curves assume $h_0 = 0.75$.

The parameter t_{eq} is obtained as in standard cosmology by setting $\rho_r(t_{\text{eq}}) = \rho_m(t_{\text{eq}})$ in equations (5.22) and (5.23). This leads to

$$t_{\text{eq}} = \begin{cases} t_0 \Omega_{r,0}^{3/2(1-4x_m)} & (\Omega_{m,0} = 1) \\ \tau_0 \sinh^{-1} \left[\chi_0 \left(\frac{\Omega_{r,0}}{\Omega_{m,0}} \right)^{3/2(1-4x_m)} \right] & (0 < \Omega_{m,0} < 1) . \end{cases} \quad (5.26)$$

The epoch of matter–radiation equality plays a crucial role in this chapter because it is at about this time that the Universe became transparent to radiation (the two events are not simultaneous but the difference between them is minor for our purposes). Decay photons created before t_{eq} would simply have been thermalized by the primordial plasma and eventually re-emitted as part of the CMB. It is the decay photons emitted *after* this time which can contribute to the EBL, and whose contributions we wish to calculate. The quantity t_{eq} is thus analogous to the galaxy formation time t_f in previous chapters.

The densities $\rho_m(t)$, $\rho_r(t)$ and $\rho_v(t)$ are plotted as functions of time in figure 5.1. The left-hand panel (a) shows the effects of varying the parameters x_r and x_m within a given cosmological model (here, Λ CDM). Raising the value of x_m (i.e. moving from bold to light curves) leads to a proportionate increase in ρ_v and a modest drop in ρ_r . It also flattens the *slope* of both components. This change in slope (relative to that of the matter component) pushes the epoch

of equality back toward the big bang (vertical lines). Such an effect could, in principle, allow more time for structure to form during the early matter-dominated era [37], although the ‘compression’ of the radiation-dominated era rapidly becomes unrealistic for values of x_m close to $\frac{1}{4}$. Thus figure 5.1(a) shows that the value of t_{eq} is reduced by a factor of over 100 in going from a model with $x_m = 10^{-4}$ to one with $x_m = 0.07$. In the limit $x_m \rightarrow \frac{1}{4}$, the duration of the radiation-dominated era in fact dwindles to nothing, as remarked earlier and as shown explicitly by equations (5.26).

Figure 5.1(b) shows the effects of changes in cosmological model for fixed values of x_r and x_m (here both set to 10^{-4}). Moving from the matter-filled EdS model toward vacuum-dominated ones such as Λ CDM and Λ BDM does three things. The first is to increase the age (t_0) of the Universe. This increases the density of radiation at any given time, since the latter is fixed at present and climbs at the same rate in the past direction. Based on our experience with the galactic EBL in previous chapters, we may expect that this should lead to significantly higher levels of background radiation when integrated over time. However, there is a second effect in the present theory which acts in the opposite direction: smaller values of $\Omega_{m,0}$ imply higher values of t_{eq} as well as t_0 , thus delaying the onset of the matter-dominated era (vertical lines). As we will see, these two changes all but cancel each other out as far as vacuum decay contributions to the background are concerned. The third important consequence of vacuum-dominated cosmologies is ‘late-time inflation’, the sharp increase in the expansion rate at recent times (figure 4.2). This translates in figure 5.1(b) into a noticeable ‘droop’ in the densities of matter and radiation at the right-hand edge of the figure for the Λ BDM model in particular.

5.4 Source regions and luminosity

In order to proceed with the formalism we have developed in chapters 2 and 3, we need to define discrete ‘sources’ of vacuum-decay radiation, analogous to the galaxies of previous chapters. For this purpose we carve up the Universe into hypothetical regions of arbitrary comoving volume V_0 . The **comoving number density** of these source regions is just

$$n(t) = n_0 = V_0^{-1} = \text{constant}. \quad (5.27)$$

These regions are introduced for convenience, and are not physically significant since the vacuum decays uniformly throughout space. We therefore expect that the parameter V_0 will not appear in our final results.

The next step is to identify the ‘source luminosity’. There are at least two ways to approach this question [56]. One could simply regard the source region as a ball of physical volume $V(t) = \tilde{R}^3(t)V_0$ filled with fluctuating vacuum energy. As the density of this energy drops by $-d\rho_v$ during time dt , the ball loses energy at a rate $-d\rho_v/dt$. If some fraction β of this energy flux goes into photons, then

the luminosity of the ball is

$$L_v(t) = -\beta c^2 \dot{\rho}_v(t) V(t). \quad (5.28)$$

This is the definition of vacuum luminosity which has been assumed implicitly by several workers including Pavón [51], who investigated the thermodynamical stability of the vacuum decay process by requiring that fluctuations in $\dot{\rho}_v$ not grow larger than the mean value of $\dot{\rho}_v$ with time. For convenience we will refer to (5.28) as the **thermodynamical definition** of vacuum luminosity (L_{th}).

A second approach is to treat this as a problem involving spherical symmetry within general relativity. The assumption of spherical symmetry allows the total mass–energy (Mc^2) of a localized region of perfect fluid to be identified unambiguously. Luminosity can then be related to the time rate of change of this mass–energy. Assuming once again that the two are related by a factor β , one has

$$L_v(t) = \beta \dot{M}(t) c^2. \quad (5.29)$$

Application of Einstein’s field equations leads to the following expression [57] for the rate of change of mass–energy in terms of the pressure p_v at the region’s surface:

$$\dot{M}(t) c^2 = -4\pi p_v(t) [r(t)]^2 \dot{r}(t) \quad (5.30)$$

where $r(t) = \tilde{R}(t)r_0$ is the region’s physical radius. Taking $V = \frac{4}{3}\pi r^3$, applying the vacuum equation of state $p_v = -\rho_v c^2$ and substituting (5.30) into (5.29), we find that the latter can be written in the form

$$L_v(t) = \beta c^2 \rho_v(t) \dot{V}(t). \quad (5.31)$$

This is just as appealing dimensionally as equation (5.28), and shifts the emphasis physically from fluctuations in the material content of the source region toward changes in its geometry. We will refer to (5.31) for convenience as the **relativistic definition** of vacuum luminosity (L_{rel}).

It is not obvious which of the two definitions (5.28) and (5.31) more correctly describes vacuum luminosity; this is a conceptual issue. Before choosing between them, let us inquire whether the two expressions might not be equivalent. We can do this by taking the ratio

$$\frac{L_{\text{th}}}{L_{\text{rel}}} = -\frac{\dot{\rho}_v V}{\rho_v \dot{V}} = -\frac{1}{3} \frac{\dot{\rho}_v R}{\rho_v \dot{R}}. \quad (5.32)$$

Differentiating equations (5.20) and (5.21) with respect to time, we find

$$\frac{\dot{R}}{R} = \begin{cases} \frac{2}{3t} & (\Omega_{m,0} = 1) \\ \frac{2}{3\tau_0} \coth\left(\frac{t}{\tau_0}\right) & (0 < \Omega_{m,0} < 1) \end{cases} \quad (5.33)$$

$$\frac{\dot{\rho}_v}{\rho_v} = \begin{cases} \frac{-8}{3t}(1 - x_m) & (\Omega_{m,0} = 1) \\ \frac{-8}{3\tau_0}(1 - x_m) \coth\left(\frac{t}{\tau_0}\right) & (0 < \Omega_{m,0} < 1). \end{cases} \quad (5.34)$$

The ratio of L_{th} to L_{rel} is therefore a constant:

$$\frac{L_{\text{th}}}{L_{\text{rel}}} = \frac{4}{3}(1 - x_{\text{m}}). \quad (5.35)$$

This takes numerical values between $\frac{4}{3}$ (in the limit $x_{\text{m}} \rightarrow 0$ where standard cosmology is recovered) and 1 (in the opposite limit where x_{m} takes its maximum theoretical value of $\frac{1}{4}$). There is thus little difference between the two scenarios in practice, at least where this model of vacuum decay is concerned. We will proceed using the relativistic definition (5.31) which gives lower intensities and hence more conservative limits on the theory. At the end of the chapter it will be a small matter to obtain the corresponding intensity for the thermodynamical case (5.28) by multiplying through by $\frac{4}{3}(1 - x_{\text{m}})$.

We now turn to the question of the **branching ratio** β , or fraction of vacuum decay energy which goes into photons as opposed to other forms of radiation such as massless neutrinos. This is model-dependent in general. If the vacuum-decay radiation reaches equilibrium with that already present, however, then we may reasonably set this equal to the ratio of photon-to-total radiation energy densities in the CMB:

$$\beta = \Omega_{\gamma} / \Omega_{\text{r},0}. \quad (5.36)$$

The density parameter Ω_{γ} of CMB photons is given in terms of their blackbody temperature T by Stefan's law. Using the COBE value $T_{\text{cmb}} = 2.728$ K [53], we get

$$\Omega_{\gamma} = \frac{4\sigma_{\text{SB}}T^4}{c^3\rho_{\text{crit},0}} = 2.48 \times 10^{-5}h_0^{-2}. \quad (5.37)$$

The total radiation density $\Omega_{\text{r},0} = \Omega_{\gamma} + \Omega_{\nu}$ is harder to determine, since there is little prospect of detecting the neutrino component directly. What is done in standard cosmology is to calculate the size of neutrino contributions to $\Omega_{\text{r},0}$ under the assumption of **entropy conservation**. With three massless neutrino species, this leads to

$$\Omega_{\text{r},0} = \Omega_{\gamma} \left[1 + 3 \times \frac{7}{8} \left(\frac{T_{\nu}}{T} \right)^4 \right] \quad (5.38)$$

where T_{ν} is the blackbody temperature of the relic neutrinos and the factor of $7/8$ arises from the fact that these particles obey Fermi rather than Bose–Einstein statistics [58]. During the early stages of the radiation-dominated era, neutrinos were in thermal equilibrium with photons so that $T_{\nu} = T$. They dropped out of equilibrium, however, when the temperature of the expanding fireball dropped below about $kT \sim 1$ MeV (the energy scale of weak interactions). Shortly thereafter, when the temperature dropped to $kT \sim m_e c^2 = 0.5$ MeV, electrons and positrons began to annihilate, transferring their entropy to the remaining photons in the plasma. This raised the photon temperature by a factor of $(1 + 2 \times \frac{7}{8} = \frac{11}{4})^{1/3}$ relative to that of the neutrinos. In standard cosmology, the

ratio of T_v/T has remained at $(4/11)^{1/3}$ down to the present day, so that (5.38) gives

$$\Omega_{r,0} = 1.68\Omega_\gamma = 4.17 \times 10^{-5}h_0^{-2}. \quad (5.39)$$

Using (5.36) for β , this would imply:

$$\beta = 1/1.68 = 0.595. \quad (5.40)$$

We will take these as our ‘standard values’ of $\Omega_{r,0}$ and β in what follows. They are conservative ones, in the sense that most alternative lines of argument would imply higher values of β . Birkel and Sarkar [38], for instance, have argued that vacuum decay into radiation with $x_r = \text{constant}$ would be easier to reconcile with processes such as electron–positron annihilation if the vacuum coupled to photons but not neutrinos. This would complicate the theory, breaking the radiation density ρ_r in (5.14) into a photon part ρ_γ and a neutrino part with different dependencies on R . One need not solve this equation, however, in order to appreciate the main impact of such a modification. Decay into photons alone would pump entropy into the photon component relative to the neutrino component in an effectively *ongoing version* of the electron–positron annihilation argument outlined above. The neutrino temperature T_v (and density ρ_v) would continue to be driven down relative to T (and ρ_γ) throughout the radiation-dominated era and into the matter-dominated one. In the limit $T_v/T \rightarrow 0$ one sees from (5.36) and (5.38) that such a scenario would lead to

$$\Omega_{r,0} = \Omega_\gamma = 2.48 \times 10^{-5}h_0^{-2} \quad \beta = 1. \quad (5.41)$$

In other words, the present energy density of radiation would be lower, but it would effectively *all* be in the form of photons. Insofar as the decrease in $\Omega_{r,0}$ is precisely offset by the increase in β , these changes cancel each other out. The drop in $\Omega_{r,0}$, however, has an added consequence which is not cancelled: it pushes t_{eq} farther into the past, increasing the length of time over which the decaying vacuum has been contributing to the background. This raises the latter’s intensity, particularly at longer wavelengths. The effect can be significant, and we will return to this possibility at the end of the chapter. For the most part, however, we will stay with the values of $\Omega_{r,0}$ and β given by equations (5.39) and (5.40).

Armed with a definition for vacuum luminosity, equation (5.31), and a value for β , equation (5.40), we are in a position to calculate the luminosity of the vacuum. Noting that $\dot{V} = 3(R/R_0)^3(\dot{R}/R)V_0$ and substituting equations (5.20), (5.21) and (5.33) into (5.31), we find that

$$L_v(t) = \mathcal{L}_{v,0}V_0 \times \left\{ \begin{array}{l} \left(\frac{t}{t_0}\right)^{-(5-8x_m)/3} \\ \left[\frac{\cosh(t/\tau_0)}{\cosh(t_0/\tau_0)} \right] \left[\frac{\sinh(t/\tau_0)}{\sinh(t_0/\tau_0)} \right]^{-(5-8x_m)/3} \end{array} \right. \quad (5.42)$$

The first of these solutions corresponds to models with $\Omega_{m,0} = 1$ while the second holds for the general case ($0 < \Omega_{m,0} < 1$). Both results reduce at the present time $t = t_0$ to

$$L_{v,0} = \mathcal{L}_{v,0} V_0 \quad (5.43)$$

where $\mathcal{L}_{v,0}$ is the **comoving luminosity density of the vacuum**

$$\begin{aligned} \mathcal{L}_{v,0} &= \frac{9c^2 H_0^3 \Omega_{r,0} \beta x_m}{8\pi G(1-x_m)} \\ &= 4.1 \times 10^{-30} h_0 \text{ erg s}^{-1} \text{ cm}^{-3} \left(\frac{x_m}{1-x_m} \right). \end{aligned} \quad (5.44)$$

Numerically, we find for example that

$$\mathcal{L}_{v,0} = \begin{cases} 3.1 \times 10^{-31} h_0 \text{ erg s}^{-1} \text{ cm}^{-3} & (x_m = 0.07) \\ 1.4 \times 10^{-30} h_0 \text{ erg s}^{-1} \text{ cm}^{-3} & (x_m = 0.25). \end{cases} \quad (5.45)$$

In principle, then, the comoving luminosity density of the decaying vacuum can reach levels as high as 10 or even 50 times that of galaxies, as given by (2.24). Raising the value of the branching ratio β to 1 instead of 0.595 does not affect these results, since this must be accompanied by a proportionate drop in the value of $\Omega_{r,0}$ as argued earlier. The numbers in (5.45) do go up if one replaces the relativistic definition (5.31) of vacuum luminosity with the thermodynamical one (5.28) but the change is modest, raising $\mathcal{L}_{v,0}$ by no more than a factor of 1.2 (for $x_m = 0.07$). The primary reason for the high luminosity of the decaying vacuum lies in the fact that it converts nearly 60% of its energy density into photons. By comparison, less than 1% of the rest energy of ordinary luminous matter has gone into photons so far in the history of the Universe.

5.5 Bolometric intensity

We showed in chapter 2 that the bolometric intensity of an arbitrary distribution of sources with comoving number density $n(t)$ and luminosity $L(t)$ could be expressed as an integral over time by (2.14). Let us apply this result here to regions of decaying vacuum energy, for which $n_v(t)$ and $L_v(t)$ are given by (5.27) and (5.42) respectively. Putting these equations into (2.14) along with (5.20) for the scale factor, we find that

$$Q = c \mathcal{L}_{v,0} \times \left\{ \int_{t_{\text{eq}}}^{t_0} \left(\frac{t}{t_0} \right)^{-(1-8x_m)/3} dt \right. \\ \left. \int_{t_{\text{eq}}}^{t_0} \left[\frac{\cosh(t/\tau_0)}{\cosh(t_0/\tau_0)} \right] \left[\frac{\sinh(t/\tau_0)}{\sinh(t_0/\tau_0)} \right]^{-(1-8x_m)/3} dt. \right. \quad (5.46)$$

The first of these integrals corresponds to models with $\Omega_{m,0} = 1$ while the second holds for the general case ($0 < \Omega_{m,0} < 1$). The latter may be simplified

with a change of variables to $y \equiv [\sinh(t/\tau_0)]^{8x_m/3}$. Using the facts that $\sinh(t_0/\tau_0) = \sqrt{(1 - \Omega_{m,0})/\Omega_{m,0}}$ and $\cosh(t_0/\tau_0) = 1/\sqrt{\Omega_{m,0}}$ along with the definition (5.26) of t_{eq} , one can show that both of these integrals reduce to the same formula,

$$Q = Q_v \left[1 - \left(\frac{\Omega_{r,0}}{\Omega_{m,0}} \right)^{4x_m/(1-4x_m)} \right]. \quad (5.47)$$

Here Q_v is found with the help of (5.44) as

$$Q_v \equiv \frac{c\mathcal{L}_{v,0}}{4H_0 x_m} = \frac{9c^3 H_0^2 \Omega_{r,0} \beta}{32\pi G(1-x_m)} = \frac{0.0094 \text{ erg cm}^{-2} \text{ s}^{-1}}{(1-x_m)}. \quad (5.48)$$

There are several points to note about this result. First, it does not depend on V_0 , as expected. There is also no dependence on the uncertainty h_0 in Hubble's constant, since the two factors of h_0 in H_0^2 are cancelled out by those in $\Omega_{r,0}$. In the limit $x_m \rightarrow 0$ one sees that $Q \rightarrow 0$ as expected. In the opposite limit where $x_m \rightarrow \frac{1}{4}$, the vacuum reaches a maximum possible bolometric intensity of $Q \rightarrow Q_v = 0.013 \text{ erg cm}^{-2} \text{ s}^{-1}$. This is 50 times the approximate bolometric intensity due to galaxies, as given by (2.25).

The matter density $\Omega_{m,0}$ enters only weakly into this result, and plays no role at all in the limit $x_m \rightarrow \frac{1}{4}$. Based on our experience with the EBL due to galaxies, we might have expected that Q would rise significantly in models with smaller values of $\Omega_{m,0}$ since these have longer ages, giving more time for the Universe to fill up with light. What is happening here, however, is that the larger values of t_0 are offset by larger values of t_{eq} (which follow from the fact that smaller values of $\Omega_{m,0}$ imply smaller ratios of $\Omega_{m,0}/\Omega_{r,0}$). This removes contributions from the early matter-dominated era and thereby *reduces* the value of Q . In the limit $x_m \rightarrow \frac{1}{4}$ these two effects cancel each other out. For smaller values of x_m , the t_{eq} effect proves to be the stronger of the two, and one finds an overall decrease in Q for these cases. With $x_m = 0.07$, for instance, the value of Q drops by 2% when moving from the EdS model to Λ CDM, and by another 6% when moving from Λ CDM to Λ BDM.

5.6 Spectral energy distribution

To obtain limits on the parameter x_m , we would like to calculate the spectral intensity of the background due to vacuum decay, just as we did for galaxies in chapter 3. For this we need to know the spectral energy distribution (SED) of the decay photons. As discussed in section 5.2, theories in which these photons are distributed with a *non*-thermal spectrum can be strongly constrained by means of distortions in the CMB. We therefore restrict ourselves here to the case of a blackbody SED, as given by equation (3.22):

$$F_v(\lambda, t) = \frac{C(t)/\lambda^5}{\exp[hc/kT(t)\lambda] - 1} \quad (5.49)$$

where $T(t)$ is the blackbody temperature. The function $C(t)$ is found as usual by normalization, equation (3.1). Changing integration variables from λ to $\nu = c/\lambda$ for convenience, we find

$$L_\nu(t) = \frac{C(t)}{c^4} \int_0^\infty \frac{\nu^3 d\nu}{\exp[h\nu/kT(t)] - 1} = \frac{C(t)}{c^4} \left[\frac{h}{kT(t)} \right]^{-4} \Gamma(4)\zeta(4). \quad (5.50)$$

Inserting our result (5.42) for $L_\nu(t)$ and using the facts that $\Gamma(4) = 3! = 6$ and $\zeta(4) = \pi^4/90$, we then obtain for $C(t)$:

$$C(t) = \frac{15\mathcal{L}_{\nu,0}V_0}{\pi^4} \left[\frac{hc}{kT(t)} \right]^4 \times \begin{cases} \left(\frac{t}{t_0} \right)^{-(5-8x_m)/3} \\ \left[\frac{\cosh(t/\tau_0)}{\cosh(t_0/\tau_0)} \right] \left[\frac{\sinh(t/\tau_0)}{\sinh(t_0/\tau_0)} \right]^{-(5-8x_m)/3} \end{cases}. \quad (5.51)$$

Here the upper expression refers as usual to the EdS case ($\Omega_{m,0} = 1$), while the lower applies to the general case ($0 < \Omega_{m,0} < 1$). Temperature T can be determined by assuming **thermal equilibrium** between the vacuum-decay photons and those already present. Stefan's law then relates $T(t)$ to the radiation energy density $\rho_r(t)c^2$ as follows:

$$\rho_r(t)c^2 = \frac{4\sigma_{\text{SB}}}{c} [T(t)]^4. \quad (5.52)$$

Substituting equation (5.22) into this expression and expanding the Stefan-Boltzmann constant, we find that

$$\frac{hc}{kT(t)} = \lambda_\nu \times \begin{cases} \left(\frac{t}{t_0} \right)^{2(1-x_m)/3} & (\Omega_{m,0} = 1) \\ \left[\frac{\sinh(t/\tau_0)}{\sinh(t_0/\tau_0)} \right]^{2(1-x_m)/3} & (0 < \Omega_{m,0} < 1) \end{cases} \quad (5.53)$$

where the constant λ_ν is given by

$$\lambda_\nu \equiv \left(\frac{8\pi^5 hc}{15\rho_{r,0}c^2} \right)^{1/4} = 0.46 \text{ cm} \left(\frac{\Omega_{r,0}h_0^2}{4.17 \times 10^{-5}} \right)^{-1/4}. \quad (5.54)$$

This value of λ_ν tells us that the peak of the observed spectrum of decay radiation lies in the microwave region as expected, near that of the CMB ($\lambda_{\text{cmb}} = 0.11 \text{ cm}$). Putting (5.53) back into (5.51), we obtain

$$C(t) = \frac{15\lambda_\nu^4 \mathcal{L}_{\nu,0}V_0}{\pi^4} \times \begin{cases} \left(\frac{t}{t_0} \right) \\ \left[\frac{\cosh(t/\tau_0)}{\cosh(t_0/\tau_0)} \right] \left[\frac{\sinh(t/\tau_0)}{\sinh(t_0/\tau_0)} \right] \end{cases}. \quad (5.55)$$

These two expressions refer to models with $\Omega_{m,0} = 1$ and $0 < \Omega_{m,0} < 1$ respectively. With $C(t)$ thus defined, the SED (5.49) of vacuum decay is completely specified.

5.7 The microwave background

We showed in chapter 3 that the spectral intensity of an arbitrary distribution of sources with comoving number density $n(t)$ and an SED $F(\lambda, t)$ could be expressed as an integral over time by (3.5). Using (5.20), (5.27), (5.49), (5.53) and (5.55), this becomes

$$I_\lambda(\lambda_0) = I_\nu(\lambda_0) \times \begin{cases} \int_{t_{\text{eq}}/t_0}^1 \frac{\tau^{-1} d\tau}{\exp\left[\left(\frac{\lambda_\nu}{\lambda_0}\right) \tau^{-2x_m/3}\right] - 1} \\ \int_{t_{\text{eq}}/\tau_0}^{t_0/\tau_0} \frac{\coth \tau d\tau}{\exp\left[\frac{\lambda_\nu}{\lambda_0} \left(\frac{\sqrt{\Omega_{m,0}} \sinh \tau}{\sqrt{1 - \Omega_{m,0}}}\right)^{-2x_m/3}\right] - 1} \end{cases}. \quad (5.56)$$

Here we have used integration variables $\tau \equiv t/t_0$ in the first case (for models with $\Omega_{m,0} = 1$) and $\tau \equiv t/\tau_0$ in the second (for models with $0 < \Omega_{m,0} < 1$). The dimensional content of both integrals is contained in the prefactor $I_\nu(\lambda_0)$, which is given by

$$I_\nu(\lambda_0) \equiv \frac{5\mathcal{L}_{\nu,0}}{2\pi^5 h H_0} \left(\frac{\lambda_\nu}{\lambda_0}\right)^4 = 15\,500 \text{ CUs} \left(\frac{x_m}{1-x_m}\right) \left(\frac{\lambda_\nu}{\lambda_0}\right)^4. \quad (5.57)$$

We have divided through by photon energy hc/λ_0 so as to express the results in continuum units (CUs) as usual, where $1 \text{ CU} \equiv 1 \text{ photon s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1} \text{ ster}^{-1}$. We will use CUs throughout this book, for the sake of uniformity as well as the fact that these units carry several advantages from the theoretical point of view (section 3.3). The reader who consults the literature, however, will soon find that each part of the electromagnetic spectrum has its own ‘dialect’ of preferred units. In the microwave region it is most common to find background intensities given in terms of the quantity νI_ν , which is the integral of flux per unit frequency *over* frequency, and is usually expressed in units of $\text{nW m}^{-2} \text{ ster}^{-1} = 10^{-6} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ ster}^{-1}$. To translate a given value of νI_ν (in these units) into CUs, one need only multiply by a factor of $10^{-6}/(hc) = 50.34 \text{ erg}^{-1} \text{ \AA}^{-1}$. The Jansky (Jy) is also often encountered, with $1 \text{ Jy} = 10^{-23} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ Hz}^{-1}$. To convert a given value of νI_ν from Jy ster^{-1} into CUs, one multiplies by $10^{-23}/h\lambda = (1509 \text{ Hz erg}^{-1})/\lambda$ with λ in \AA .

Equation (5.56) gives the combined intensity of decay photons which have been emitted at many wavelengths and redshifted by various amounts, but reach us in a waveband centred on λ_0 . The arbitrary volume V_0 has dropped out of the integral as expected and this result is also independent of the uncertainty h_0 in Hubble’s constant since there is a factor of h_0 in both $\mathcal{L}_{\nu,0}$ and H_0 . Results are plotted in figure 5.2 over the waveband 0.01–1 cm, together with observational detections of the background in this part of the spectrum. The most celebrated of these is the COBE detection of the CMB [53] which we have shown as a

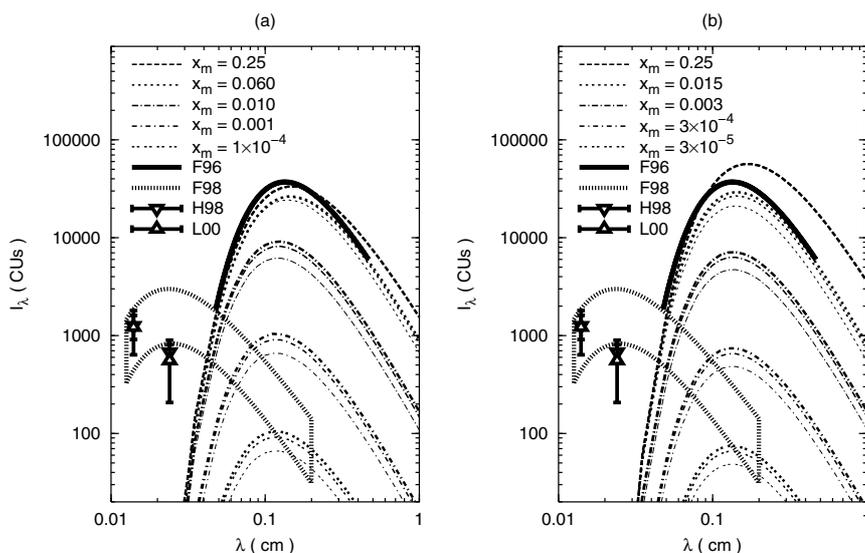


Figure 5.2. The spectral intensity of background radiation due to the decaying vacuum for various values of x_m , compared with observational data in the microwave region (bold unbroken line) and far infrared (bold dotted line and points). For each value of x_m there are three curves representing cosmologies with $\Omega_{m,0} = 1$ (bold lines), $\Omega_{m,0} = 0.3$ (medium lines) and $\Omega_{m,0} = 0.03$ (light lines). The left-hand panel (a) assumes $L = L_{\text{rel}}$ and $\beta = 0.595$, while the right-hand panel (b) assumes $L = L_{\text{th}}$ and $\beta = 1$.

bold unbroken line (F96). The experimental uncertainties in this measurement are far smaller than the thickness of the line. The other observational limits shown in figure 5.2 have been obtained in the far infrared (FIR) region, also from analysis of data from the COBE satellite. These are indicated with bold dotted lines (F98 [59]) and open triangles (H98 [60] and L00 [61]).

Figure 5.2(a) shows the spectral intensity of background radiation from vacuum decay under our standard assumptions, including the relativistic definition (5.31) of vacuum luminosity and the values of $\Omega_{r,0}$ and β given by (5.39) and (5.40) respectively. Five groups of curves are shown, corresponding to values of x_m between 3×10^{-5} and the theoretical maximum of 0.25. For each value of x_m three curves are plotted: one each for the EdS, Λ CDM and Λ BDM cosmologies. As previously noted in connection with the bolometric intensity Q , the choice of cosmological model is less important in determining the background due to vacuum decay than the background due to starlight from galaxies. In fact, the intensities here are actually slightly *lower* in vacuum-dominated models. The reason for this, as before, is that these models have smaller values of $\Omega_{m,0}/\Omega_{r,0}$ and hence larger values of t_{eq} , reducing the size of contributions from the early matter-dominated era when L_v was large.

In figure 5.2(b), we have exchanged the relativistic definition of vacuum luminosity for the thermodynamical one (5.28), and set $\beta = 1$ instead of 0.595. As explained in section 5.4, the increase in β is partly offset by a drop in the present radiation density $\Omega_{r,0}$. There is a net increase in intensity, however, because smaller values of $\Omega_{r,0}$ push t_{eq} back into the past, leading to additional contributions from the early matter-dominated era. These particularly push up the long-wavelength part of the spectrum in figure 5.2(b) relative to figure 5.2(a), as seen most clearly for the case $x_m = 0.25$. Overall, intensities in figure 5.2(b) are higher than those in figure 5.2(a) by about a factor of four.

These figures tell us that *the decaying-vacuum hypothesis is strongly constrained by observations of the microwave background*. The parameter x_m cannot be larger than 0.06 or the intensity of the decaying vacuum would exceed that of the CMB itself under the most conservative assumptions, as represented by figure 5.2(a). This limit tightens to $x_m \leq 0.015$ if different assumptions are made about the luminosity of the vacuum, as shown by figure 5.2(b). These numbers are comparable to the limit of $x \leq 0.07$ obtained from entropy conservation under the assumption that $x = x_r = x_m$ [37]. And insofar as the CMB radiation is usually attributed entirely to relic radiation from the big bang, the real limit on x_m is probably several orders of magnitude stronger than this.

With these upper bounds on x_m , we can finally inquire about the potential of the decaying vacuum as a dark-matter candidate. Since its density is given by (5.16) as a fraction $x/(1-x)$ of that of radiation, we infer that its present density parameter ($\Omega_{v,0}$) satisfies:

$$\Omega_{v,0} = \left(\frac{x_m}{1-x_m} \right) \Omega_{r,0} \leq \begin{cases} 7 \times 10^{-6} & \text{(a)} \\ 1 \times 10^{-6} & \text{(b)} \end{cases} \quad (5.58)$$

Here, (a) and (b) refer to the scenarios represented by figures 5.2(a) and (b), with the corresponding values of $\Omega_{r,0}$ as defined by equations (5.39) and (5.41) respectively. We have assumed that $h_0 \geq 0.6$ as usual. It is clear from the limits (5.58) that a decaying vacuum, at least of the kind we have considered in this chapter, does not contribute significantly to the density of dark matter.

It should be recalled, however, that there are good reasons from quantum theory for expecting some kind of instability for the vacuum in a Universe which progressively cools. (Equivalently, there are good reasons for believing that the cosmological ‘constant’ is not.) Our conclusion is that if the vacuum decays, it either does so very slowly or in a manner that does not upset the isotropy of the cosmic microwave background.

References

- [1] Bronstein M 1933 *Physikalische Zeitschrift der Sowjetunion* **3** 73
- [2] Kragh H 1996 *Cosmology and Controversy* (Princeton, NJ: Princeton University Press) p 36

- [3] Einstein A 1936 *J. Franklin Inst.* **221** 349
- [4] Jordan P 1949 *Nature* **164** 637
- [5] Fierz M 1956 *Helv. Phys. Acta* **29** 128
- [6] Brans C and Dicke R H 1961 *Phys. Rev.* **124** 925
- [7] Bergmann P G 1968 *Int. J. Theor. Phys.* **1** 25
- [8] Endō M and Fukui T 1977 *Gen. Rel. Grav.* **8** 833
- [9] Barrow J D and Maeda K-I 1990 *Nucl. Phys. B* **341** 294
- [10] Fukui T and Overduin J M 2002 *Int. J. Mod. Phys. D* **11** 669
- [11] Madsen M S 1988 *Class. Quantum Grav.* **5** 627
- [12] Maeda K-I 1989 *Phys. Rev. D* **39** 3159
- [13] Barr S M 1987 *Phys. Rev. D* **36** 1691
- [14] Peebles P J E and Ratra B 1988 *Astrophys. J.* **325** L17
- [15] Wetterich C 1988 *Nucl. Phys. B* **302** 668
- [16] Caldwell R R, Dave R and Steinhardt P J 1998 *Phys. Rev. Lett.* **80** 1582
- [17] Zeldovich Y B 1968 *Uspekhi Fiz. Nauk* **95** 209
- [18] Overduin J M and Cooperstock F I 1998 *Phys. Rev. D* **58** 043506
- [19] Chen W and Wu S 1990 *Phys. Rev. D* **41** 695
- [20] Polyakov A M 1982 *Sov. Phys. Usp.* **25** 187
- [21] Adler S L 1982 *Rev. Mod. Phys.* **54** 729
- [22] Dolgov A D 1983 *The Very Early Universe* ed G W Gibbons, S W Hawking and S T C Siklos (Cambridge: Cambridge University Press) p 449
- [23] Banks T 1985 *Nucl. Phys. B* **249** 332
- [24] Peccei R D, Solà J and Wetterich C 1987 *Phys. Lett. B* **195** 183
- [25] Frieman J A *et al* 1995 *Phys. Rev. Lett.* **75** 2077
- [26] Hawking S W 1984 *Phys. Lett. B* **134** 403
- [27] Brown J D and Teitelboim C 1987 *Phys. Lett. B* **195** 177
- [28] Dolgov A D 1997 *Phys. Rev.* **55** 5881
- [29] Mottola E 1985 *Phys. Rev. D* **31** 754
- [30] Tsamis N C and Woodard R P 1993 *Phys. Lett. B* **301** 351
- [31] Mukhanov V F, Abramo L R W and Brandenberger R H 1997 *Phys. Rev. Lett.* **78** 1624
- [32] Banks T 1988 *Nucl. Phys. B* **309** 493
- [33] Coleman S 1988 *Nucl. Phys. B* **310** 643
- [34] Guendelman E I and Kaganovich A B 1997 *Phys. Rev. D* **55** 5970
- [35] Weinberg S 1989 *Rev. Mod. Phys.* **61** 1
- [36] Dolgov A D 1998 *The 4th Paris Cosmology Colloquium* ed N Sanchez and H J de Vega (Singapore: World Scientific)
- [37] Freese K *et al* 1987 *Nucl. Phys. B* **287** 797
- [38] Birkel M and Sarkar S 1997 *Astropart. Phys.* **6** 197
- [39] Olson T S and Jordan T F 1987 *Phys. Rev. D* **35** 3258
- [40] Matyjasek J 1995 *Phys. Rev. D* **51** 4154
- [41] Ratra B and Peebles P J E 1988 *Phys. Rev. D* **37** 3406
- [42] Silveira V and Waga I 1994 *Phys. Rev. D* **50** 4890
- [43] Viana P T P and Liddle A R 1998 *Phys. Rev. D* **57** 674
- [44] Yoshii Y and Sato K 1992 *Astrophys. J.* **387** L7
- [45] Sugiyama N and Sato K 1992 *Astrophys. J.* **387** 439
- [46] Silveira V and Waga I 1997 *Phys. Rev. D* **56** 4625
- [47] Ratra B and Quillen A 1992 *Mon. Not. R. Astron. Soc.* **259** 738

- [48] Bloomfield Torres L F and Waga I 1996 *Mon. Not. R. Astron. Soc.* **279** 712
- [49] Podariu S and Ratra B 2000 *Astrophys. J.* **532** 109
- [50] Overduin J M 1999 *Astrophys. J.* **517** L1
- [51] Pavón D 1991 *Phys. Rev. D* **43** 375
- [52] Lima J A S 1996 *Phys. Rev. D* **54** 2571
- [53] Fixsen D J *et al* 1996 *Astrophys. J.* **473** 576
- [54] Hu W and Silk J 1993 *Phys. Rev. D* **48** 485
- [55] Pollock M D 1980 *Mon. Not. R. Astron. Soc.* **193** 825
- [56] Overduin J M, Wesson P S and Bowyer S 1993 *Astrophys. J.* **404** 1
- [57] Misner C W, Thorne K S and Wheeler J A 1973 *Gravitation* (San Francisco, CA: Freeman) p 859
- [58] Peebles P J E 1993 *Principles of Physical Cosmology* (Princeton, NJ: Princeton University Press) p 164
- [59] Fixsen D J *et al* 1998 *Astrophys. J.* **508** 128
- [60] Hauser M G *et al* 1998 *Astrophys. J.* **508** 25
- [61] Lagache G *et al* 2000 *Astron. Astrophys.* **354** 247

Chapter 6

Axions

6.1 Light axions

Axions are hypothetical particles whose existence would explain what is otherwise a puzzling feature of **quantum chromodynamics** (QCD), the leading theory of strong interactions. QCD contains a dimensionless free parameter (Θ) whose value must be ‘unnaturally’ small in order for the theory not to violate a combination of charge conservation and mirror-symmetry known as charge parity or CP. Upper limits on the electric dipole moment of the neutron currently constrain the value of Θ to be less than about 10^{-9} . The **strong CP problem** is the question: ‘Why is Θ so small?’ This is reminiscent of the cosmological-constant problem which we encountered in chapter 5 (although less severe by many orders of magnitude). Proposed solutions have similarly focused on making Θ , like Λ , a dynamical variable whose value could have been driven toward zero in the early Universe. In the most widely-accepted scenario, due to Peccei and Quinn in 1977 [1], this is accomplished by the spontaneous breaking of a new global symmetry (now called **PQ symmetry**) at energy scales f_{PQ} . As shown by Weinberg [2] and Wilczek [3] in 1978, the symmetry-breaking gives rise to a new particle which eventually acquires a rest energy $m_a c^2 \propto f_{\text{PQ}}^{-1}$. This particle is the axion (a).

Axions, if they exist, meet all the requirements of a successful cold dark-matter or CDM candidate, as listed in section 4.4: they interact weakly with the baryons, leptons and photons of the standard model; they are cold (i.e. non-relativistic during the time when structure begins to form); and they are capable of providing some or even all of the CDM density which is thought to be required, $0 \lesssim \Omega_{\text{cdm}} \lesssim 0.6$. A fourth property, and the one which is of most interest to us here, is that *axions decay generically into photon pairs*. The importance of this process depends on two things: the axion’s rest mass m_a and its two-photon coupling strength $g_{a\gamma\gamma}$. Theoretical and experimental considerations restrict the values of these two parameters but leave open the possibility that decaying axions might contribute strongly to the extragalactic background light (EBL). Our goal

in this chapter will be to calculate the intensity of these contributions, just as we did for those of the decaying vacuum in chapter 5. We will find that observations of the EBL at infrared and optical wavelengths close off part of the parameter space which is left open by other tests. Axions, if they are to make up the CDM, must be either exceedingly light or exceedingly weakly coupled.

6.2 Rest mass

The PQ symmetry-breaking energy scale f_{PQ} is not constrained by the theory and reasonable values for this parameter are such that $m_a c^2$ might, in principle, lie anywhere between 10^{-12} eV and 1 MeV [4]. This broad range of theoretical possibilities has been narrowed down by an impressive combination of cosmological, astrophysical and laboratory-based tests. In summarizing these, it is useful to distinguish between axions with rest energies above and below $m_a c^2 \sim 3 \times 10^{-2}$ eV.

If the axion rest energy lies *below* this value, then most axions arise via processes known as **vacuum misalignment** [5–7] and **axionic string decay** [8]. These are non-thermal mechanisms, meaning that the axions produced in this way were never in thermal equilibrium with the primordial plasma. Their present density would be at least [9]

$$\Omega_a \approx \left(\frac{m_a c^2}{4 \times 10^{-6} \text{ eV}} \right)^{-7/6} h_0^{-2}. \quad (6.1)$$

(This number is currently under debate, and may go up by an order of magnitude or more if string effects play an important role [10].) If we require that axions not provide *too much* CDM ($\Omega_{\text{cdm}} \leq 0.6$) then (6.1) implies a lower limit on the axion rest energy:

$$m_a c^2 \gtrsim 7 \times 10^{-6}. \quad (6.2)$$

This neatly eliminates the lower third of the theoretically-allowed axion mass window. Corresponding upper limits on m_a in this range have come from astrophysics. Prime among these is the fact that the weak couplings of axions to baryons, leptons and photons allow them to stream freely from stellar cores, carrying energy with them. More massive axions could, in principle, cool the core of the Sun, alter the helium-burning phase in red-giant stars, and shorten the duration of the neutrino burst from supernovae such as SN1987a. The last of these effects is particularly sensitive and leads to the upper bound [11, 12]:

$$m_a c^2 \lesssim 6 \times 10^{-3} \text{ eV}. \quad (6.3)$$

Axions with $10^{-5} \lesssim m_a c^2 \lesssim 10^{-2}$ thus remain compatible with both cosmological and astrophysical limits and could provide much or all of the CDM. Moreover, it may be possible to detect these particles in the laboratory

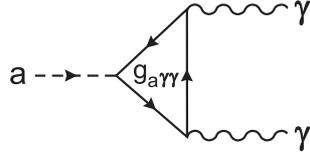


Figure 6.1. The Feynman diagram corresponding to the decay of the axion (a) into two photons (γ) with coupling strength $g_{a\gamma\gamma}$.

by enhancing their conversion into photons with strong magnetic fields, as shown by Sikivie in 1983 [13]. This is now the basis for axion ongoing cavity detector experiments in Japan [14] and the USA [15]. It has also been suggested that the Universe itself might behave like a giant axion cavity detector, since it is threaded by intergalactic magnetic fields. These could stimulate the reverse process, converting photons into a different species of ‘ultralight’ axion and possibly explaining the dimming of high-redshift supernovae without the need for large amounts of vacuum energy [16].

Promising as they are, we will not consider low-mass axions (sometimes known as ‘invisible axions’) further in this chapter. This is because they decay too slowly to have a noticeable impact on the EBL. Axions decay into photon pairs ($a \rightarrow \gamma + \gamma$) via a loop diagram, as illustrated in figure 6.1. The **decay lifetime** of this process is [4]

$$\tau_a = (6.8 \times 10^{24} \text{ s}) m_1^{-5} \zeta^{-2}. \quad (6.4)$$

Here $m_1 \equiv m_a c^2 / (1 \text{ eV})$ is the axion rest energy in units of eV and ζ is a constant which is proportional to the coupling strength $g_{a\gamma\gamma}$ [17]. For our purposes, it is sufficient to treat ζ as a free parameter which depends on the details of the axion theory chosen. Its value has been normalized in equation (6.4) so that $\zeta = 1$ in the simplest **grand unified theories** (GUTs) of strong and electroweak interactions. This could drop to $\zeta = 0.07$ in other theories, however [18], amounting to a strong suppression of the two-photon decay channel. In principle ζ could even vanish altogether, corresponding to a radiatively *stable* axion, although this would require an unlikely cancellation of terms. We will consider values in the range $0.07 \leq \zeta \leq 1$ in what follows. For these values of ζ , and with $m_1 \lesssim 6 \times 10^{-3}$ as given by (6.3), equation (6.4) shows that axions decay on timescales $\tau_a \gtrsim 9 \times 10^{35} \text{ s}$. This is so much longer than the age of the Universe that such particles would truly be invisible.

We therefore shift our attention to axions with rest energies *above* $m_a c^2 \sim 3 \times 10^{-2} \text{ eV}$. Turner showed in 1987 [19] that the vast majority of these would have arisen via **thermal mechanisms** such as Primakoff scattering and photoproduction processes in the early Universe. Application of the Boltzmann equation gives their present comoving number density as $n_a = (830/g_{*F}) \text{ cm}^{-3}$, where g_{*F} counts the number of relativistic degrees of freedom left in the plasma

at the time when axions ‘froze out’ [4]. Ressell [17] has estimated the latter quantity as $g_{*F} = 15$. The present density parameter $\Omega_a = n_a m_a / \rho_{\text{crit},0}$ of thermal axions is thus

$$\Omega_a = 5.2 \times 10^{-3} h_0^{-2} m_1. \quad (6.5)$$

Whether or not this is significant depends on the axion rest mass. The neutrino burst from SN1987a now imposes a *lower* limit on m_a rather than an upper limit as before, because axions in this range are massive enough to interact with nucleons in the supernova core and can no longer stream out freely (this is referred to as the ‘trapped regime’). Only axions at the lower end of the mass-range are able to carry away enough energy to interfere with the duration of the neutrino burst, leading to the limit [20]

$$m_a c^2 \gtrsim 2.2 \text{ eV}. \quad (6.6)$$

Laboratory upper limits on m_a come from the fact that axions cannot carry away too much ‘missing energy’ in processes such as the decay of the kaon [21]. Astrophysical constraints (similar to those mentioned already) are considerably stronger. These now depend critically on whether axions couple only to hadrons at tree-level, or to leptons as well. The former kind are known as **KSVZ axions** after Kim [22] and Shifman, Vainshtein and Zakharov [23]; while the latter take the name **DFSZ axions** after Zhitnitsky [24] and Dine, Fischler and Srednicki [25]. The extra lepton coupling of DFSZ axions allows them to carry so much energy out of the cores of red-giant stars that helium ignition is disrupted unless $m_a c^2 \lesssim 9 \times 10^{-3} \text{ eV}$ [26]. This closes the high-mass window on DFSZ axions, which can consequently exist only with the ‘invisible’ masses discussed earlier. For hadronic (KSVZ) axions, the duration of the helium-burning phase in red giants imposes a weaker bound [27]:

$$m_a c^2 \lesssim 0.7 \zeta^{-1} \text{ eV}. \quad (6.7)$$

This translates into an upper limit $m_a c^2 \lesssim 10 \text{ eV}$ for the simplest axion models with $\zeta \geq 0.07$. And irrespective of the value of ζ , it has been argued that axions with $m_a c^2 \gtrsim 10 \text{ eV}$ can be ruled out because they would interact strongly enough with baryons to produce a detectable signal in existing Čerenkov detectors [28]. In the hadronic case, then, there remains a window of opportunity for **multi-eV axions** with $2 \lesssim m_1 \lesssim 10$. Equation (6.5) shows that these particles would contribute a total density of about $0.03 \lesssim \Omega_a \lesssim 0.15$, where we take $0.6 \leq h_0 \leq 0.9$ as usual. Axions of this kind would not be able to provide the required density of dark matter in the Λ CDM model ($\Omega_{m,0} = 0.3$). They would, however, suffice in low-density models midway between Λ CDM and Λ BDM (section 3.3). Since such models are compatible with current observational data (chapter 4), it is worth proceeding to see whether multi-eV axions can be further constrained by their contributions to the EBL.

6.3 Axion halos

Thermal axions are not as cold as their non-thermal cousins, but will still be found primarily inside gravitational potential wells such as those of galaxies and galaxy clusters [19]. We need not be too specific about the fraction which have settled into galaxies as opposed to larger systems, because we will be concerned primarily with their *combined* contributions to the diffuse background. (Distribution could become an issue if extinction due to dust or gas played a strong role inside the bound regions, but this is not likely to be important for the photon energies under consideration here.) These **axion halos** provide us with a convenient starting-point as cosmological sources of axion-decay radiation, analogous to the galaxies and vacuum source regions of previous chapters. Let us consider to begin with the possibility that axions are cold enough that their fractional contribution (M_h) to the total mass of each halo (M_{tot}) is the same as their fractional contribution to the cosmological matter density:

$$\frac{M_h}{M_{\text{tot}}} = \frac{\Omega_a}{\Omega_{\text{m},0}} = \frac{\Omega_a}{\Omega_a + \Omega_{\text{bar}}}. \quad (6.8)$$

Here we have made the minimal assumption that *all* the non-baryonic dark matter is provided by axions. (As noted earlier, this effectively restricts us to strongly vacuum-dominated cosmological models.) The mass of axions in each halo is, from (6.8):

$$M_h = M_{\text{tot}} \left(1 + \frac{\Omega_{\text{bar}}}{\Omega_a} \right)^{-1}. \quad (6.9)$$

If these regions are distributed with a mean comoving number density n_0 , then the cosmological density of bound axions is

$$\Omega_{\text{a,bound}} = \frac{n_0 M_h}{\rho_{\text{crit},0}} = \frac{n_0 M_{\text{tot}}}{\rho_{\text{crit},0}} \left(1 + \frac{\Omega_{\text{bar}}}{\Omega_a} \right)^{-1}. \quad (6.10)$$

Setting this equal to Ω_a as given by (6.5) fixes the total mass:

$$M_{\text{tot}} = \frac{\Omega_a \rho_{\text{crit},0}}{n_0} \left(1 + \frac{\Omega_{\text{bar}}}{\Omega_a} \right). \quad (6.11)$$

The comoving number density of galaxies at $z = 0$ is commonly taken as [29]

$$n_0 = 0.010 h_0^3 \text{ Mpc}^{-3}. \quad (6.12)$$

Using this together with (6.5) for Ω_a , and setting $\Omega_{\text{bar}} \approx 0.016 h_0^{-2}$ from section 4.3, we find from (6.11) that

$$M_{\text{tot}} = \begin{cases} 9 \times 10^{11} M_{\odot} h_0^{-3} & (m_1 = 3) \\ 1 \times 10^{12} M_{\odot} h_0^{-3} & (m_1 = 5) \\ 2 \times 10^{12} M_{\odot} h_0^{-3} & (m_1 = 8). \end{cases} \quad (6.13)$$

Let us compare these numbers with recent dynamical data on the mass of the Milky Way using the motions of galactic satellites. These assume a **Jaffe profile** [30] for the halo density:

$$\rho_{\text{tot}}(r) = \frac{v_c^2}{4\pi G r^2} \frac{r_j^2}{(r + r_j)^2} \quad (6.14)$$

where v_c is the circular velocity, r_j the Jaffe radius and r the radial distance from the centre of the Galaxy. The data imply that $v_c = 220 \pm 30 \text{ km s}^{-1}$ and $r_j = 180 \pm 60 \text{ kpc}$ [31]. Integrating over r from zero to infinity gives

$$M_{\text{tot}} = \frac{v_c^2 r_j}{G} = (2 \pm 1) \times 10^{12} M_{\odot}. \quad (6.15)$$

This is consistent with (6.13) for a wide range of values of m_1 and h_0 . So axions of this type could, in principle, make up all the dark matter which is required on galactic scales.

Putting (6.11) into (6.9) we obtain for the mass of the axion halos:

$$M_h = \frac{\Omega_a \rho_{\text{crit},0}}{n_0}. \quad (6.16)$$

This could also have been derived as the mass of a region of space of comoving volume $V_0 = n_0^{-1}$ filled with homogeneously-distributed axions of mean density $\rho_a = \Omega_a \rho_{\text{crit},0}$. (This is the approach that we adopted in defining vacuum regions in chapter 5.)

To obtain the halo luminosity, we add up the rest energies of all the decaying axions in the halo and divide by the decay lifetime:

$$L_h = \frac{M_h c^2}{\tau_a}. \quad (6.17)$$

Substituting (6.4) and (6.16) into this equation, we find that

$$\begin{aligned} L_h &= (3.8 \times 10^{40} \text{ erg s}^{-1}) h_0^{-3} \zeta^2 m_1^6 \\ &= \begin{cases} 7 \times 10^9 L_{\odot} h_0^{-3} \zeta^2 & (m_1 = 3) \\ 2 \times 10^{11} L_{\odot} h_0^{-3} \zeta^2 & (m_1 = 5) \\ 3 \times 10^{12} L_{\odot} h_0^{-3} \zeta^2 & (m_1 = 8). \end{cases} \end{aligned} \quad (6.18)$$

These numbers should be compared to the luminosities of the galaxies themselves, which are of order $L_0 = \mathcal{L}_0/n_0 = 2 \times 10^{10} h_0^{-2} L_{\odot}$, where we have used (2.24) for \mathcal{L}_0 . The proposed axion halos are, in principle, capable of outshining their host galaxies unless the axion is either light ($m_1 \lesssim 3$) or weakly-coupled ($\zeta < 1$).

6.4 Intensity

Substituting the halo comoving number density n_0 and luminosity L_h into equation (2.20), we find for the bolometric intensity of the decaying axions:

$$Q = Q_a \int_0^{z_f} \frac{dz}{(1+z)^2 \tilde{H}(z)}. \quad (6.19)$$

Here the dimensional content of the integral is contained in the prefactor Q_a which takes the following numerical values:

$$\begin{aligned} Q_a &= \frac{\Omega_a \rho_{\text{crit},0} c^3}{H_0 \tau_a} = (1.2 \times 10^{-7} \text{ erg s}^{-1} \text{ cm}^{-2}) h_0^{-3} \zeta^2 m_1^6 \quad (6.20) \\ &= \begin{cases} 9 \times 10^{-5} \text{ erg s}^{-1} \text{ cm}^{-2} h_0^{-1} \zeta^2 & (m_1 = 3) \\ 2 \times 10^{-3} \text{ erg s}^{-1} \text{ cm}^{-2} h_0^{-1} \zeta^2 & (m_1 = 5) \\ 3 \times 10^{-2} \text{ erg s}^{-1} \text{ cm}^{-2} h_0^{-1} \zeta^2 & (m_1 = 8). \end{cases} \end{aligned}$$

There are three things to note about this quantity. First, it is comparable in magnitude to the *observed* EBL due to galaxies, $Q_* \approx 3 \times 10^{-4} \text{ erg s}^{-1} \text{ cm}^{-2}$ (chapter 2). Second, unlike Q_* for galaxies or Q_v for decaying vacuum energy, Q_a depends explicitly on the uncertainty h_0 in Hubble's constant. Physically, this reflects the fact that the axion density $\rho_a = \Omega_a \rho_{\text{crit},0}$ in the numerator of (6.20) comes to us from the Boltzmann equation and is independent of h_0 , whereas the density of luminous matter such as that in galaxies is inferred from its luminosity density \mathcal{L}_0 (which is proportional to h_0 , thus cancelling the h_0 -dependence in H_0). The third thing to note about Q_a is that it is independent of n_0 . This is because the collective contribution of decaying axions to the diffuse background is determined by their mean density Ω_a and does not depend on how they are distributed in space.

To evaluate (6.19) we need to specify the cosmological model. Let us assume that the Universe is spatially flat, as is increasingly suggested by the data (chapter 4). Hubble's parameter (2.40) then reduces to

$$\tilde{H}(z) = [\Omega_{m,0}(1+z)^3 + 1 - \Omega_{m,0}]^{1/2} \quad (6.21)$$

where $\Omega_{m,0} = \Omega_a + \Omega_{\text{bar}}$. Putting this into (6.19) along with (6.20) for Q_a , we obtain the plots of $Q(m_1)$ shown in figure 6.2 for $\zeta = 1$. The three bold lines in this plot show the range of intensities obtained by varying h_0 and $\Omega_{\text{bar}} h_0^2$ within the ranges $0.6 \leq h_0 \leq 0.9$ and $0.011 \leq \Omega_{\text{bar}} h_0^2 \leq 0.021$ respectively. We have set $z_f = 30$, since axions were presumably decaying long before they became bound to galaxies. (Results are insensitive to this choice, rising by less than 2% as $z_f \rightarrow 1000$ and dropping by less than 1% for $z_f = 6$.) The axion-decay background is faintest for the largest values of h_0 , as expected from the fact that $Q_a \propto h_0^{-1}$. This is partly offset, however, by the fact that larger values of h_0 also lead to a drop in $\Omega_{m,0}$, extending the age of the Universe and hence the length

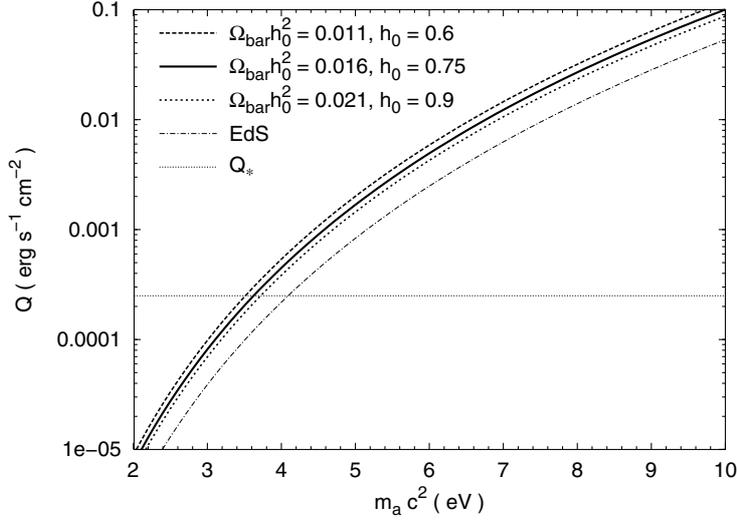


Figure 6.2. The bolometric intensity Q of the background radiation from axions as a function of their rest mass m_a . The faint dash-dotted line shows the equivalent intensity in an EdS model (in which axions alone cannot provide the required CDM). The dotted horizontal line indicates the approximate bolometric intensity (Q_*) of the observed EBL.

of time over which axions have been contributing to the background. (Smaller values of Ω_{bar} raise the intensity slightly for the same reason.) Overall, figure 6.2 confirms that axions with $\zeta = 1$ and $m_a c^2 \gtrsim 3.5$ eV produce a background brighter than that from the galaxies themselves.

6.5 The infrared and optical backgrounds

To compare our predictions with observational data, it is necessary to calculate the intensity of axionic contributions to the EBL as a function of wavelength. The first step, as usual, is to specify the **spectral energy distribution** or SED of the decay photons in the rest frame. Each axion decays into two photons of energy $\frac{1}{2}m_a c^2$ (figure 6.1), so that the decay photons are emitted at or near a peak wavelength

$$\lambda_a = \frac{2hc}{m_a c^2} = \frac{24\,800 \text{ \AA}}{m_1}. \quad (6.22)$$

Since $2 \lesssim m_1 \lesssim 10$, the value of this parameter tells us that we will be most interested in the infrared and optical bands (roughly 4000–40000 Å). We can

model the decay spectrum with a Gaussian SED as in (3.19) :

$$F(\lambda) = \frac{L_h}{\sqrt{2\pi}\sigma_\lambda} \exp\left[-\frac{1}{2}\left(\frac{\lambda - \lambda_a}{\sigma_\lambda}\right)^2\right]. \quad (6.23)$$

For the standard deviation of the curve, we can use the velocity dispersion v_c of the bound axions [32]. This is 220 km s^{-1} for the Milky Way, giving $\sigma_\lambda \approx 40 \text{ \AA}/m_1$ where we have used $\sigma_\lambda = 2(v_c/c)\lambda_a$ (section 3.4). For axions bound in galaxy clusters, v_c rises to as much as 1300 km s^{-1} [17], implying that $\sigma_\lambda \approx 220 \text{ \AA}/m_1$. For convenience we parametrize σ_λ in terms of a dimensionless quantity $\sigma_{50} \equiv \sigma_\lambda/(50 \text{ \AA}/m_1)$ so that

$$\sigma_\lambda = (50 \text{ \AA}/m_1)\sigma_{50}. \quad (6.24)$$

With the SED $F(\lambda)$ thus specified along with Hubble's parameter (6.21), the spectral intensity of the background radiation produced by axion decays is given by (3.6) as

$$I_\lambda(\lambda_0) = I_a \int_0^{z_f} \frac{\exp\left\{-\frac{1}{2}\left[\frac{\lambda_0/(1+z) - \lambda_a}{\sigma_\lambda}\right]^2\right\} dz}{(1+z)^3[\Omega_{m,0}(1+z)^3 + 1 - \Omega_{m,0}]^{1/2}}. \quad (6.25)$$

The dimensional prefactor in this case reads

$$\begin{aligned} I_a &= \frac{\Omega_a \rho_{\text{crit},0} c^2}{\sqrt{32\pi^3} h H_0 \tau_a} \left(\frac{\lambda_0}{\sigma_\lambda}\right) \\ &= (95 \text{ CUs}) h_0^{-1} \zeta^2 m_1^7 \sigma_{50}^{-1} \left(\frac{\lambda_0}{24\,800 \text{ \AA}}\right). \end{aligned} \quad (6.26)$$

Here we have divided through by the photon energy hc/λ_0 to put results into continuum units or CUs as usual (section 3.3). The source number density in (3.6) cancels out the factor of $1/n_0$ in source luminosity (6.17) so that results are independent of axion distribution, as expected. Evaluating equation (6.25) over $2000 \text{ \AA} \leq \lambda_0 \leq 20\,000 \text{ \AA}$ with $\zeta = 1$ and $z_f = 30$, we obtain the plots of $I_\lambda(\lambda_0)$ in figure 6.3. Three groups of curves are shown, corresponding to $m_a c^2 = 3, 5$ and 8 eV . For each value of m_a there are four curves; these assume $(h_0, \Omega_{\text{bar}} h_0^2) = (0.6, 0.011), (0.75, 0.016)$ and $(0.9, 0.021)$ respectively, with the fourth (faint dash-dotted) curve representing the equivalent intensity in an EdS Universe (as in figure 6.2). Also plotted in figure 6.3 are many of the reported observational constraints on EBL intensity in this waveband. Most have been encountered already in chapter 3. They include data from the OAO-2 satellite (LW76 [33]), several ground-based telescope observations (SS78 [34], D79 [35], BK86 [36]), the Pioneer 10 spacecraft (T83 [37]), sounding rockets (J84 [38], T88 [39]), the Space Shuttle-borne Hopkins UVX experiment (M90 [40]), combined

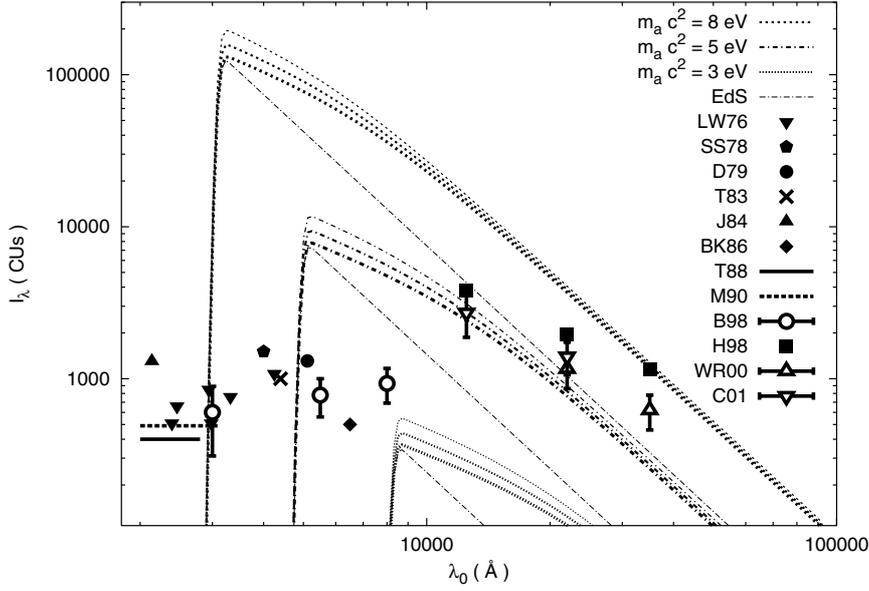


Figure 6.3. The spectral intensity I_λ of the background radiation from decaying axions as a function of observed wavelength λ_0 . The curves for each value of m_a (labelled) correspond to upper, median and lower limits on h_0 and Ω_{bar} together with the equivalent intensity for the EdS model (as in figure 6.2). Also shown are observational upper limits (full symbols and bold lines) and reported detections (empty symbols) over this waveband.

HST/Las Campanas telescope observations (B98 [41]) and the DIRBE instrument aboard the COBE satellite (H98 [42], WR00 [43], C01 [44]).

Figure 6.3 shows that 8 eV axions with $\zeta = 1$ would produce a hundred times more background light at $\sim 3000 \text{ \AA}$ than is actually seen. The background from 5 eV axions would exceed observed levels by a factor of ten at $\sim 5000 \text{ \AA}$. (This, we can note in passing, would colour the night sky green.) Only axions with $m_a c^2 \leq 3 \text{ eV}$ are compatible with observation if $\zeta = 1$. These results are significantly higher than those obtained assuming an EdS cosmology [17, 45] especially at wavelengths longward of the peak. This simply reflects the fact that the background in a low- $\Omega_{\text{m},0}$, high- $\Omega_{\Lambda,0}$ Universe like that considered here receives many more contributions from sources at high redshift.

To obtain more detailed constraints, we can instruct a computer to evaluate the integral (6.25) at more finely-spaced intervals in m_a . Since $I_\lambda \propto \zeta^{-2}$, the value of ζ required to reduce the minimum predicted axion intensity $I_{\text{th}}(m_a, \lambda_0)$ below a given observational limit $I_{\text{obs}}(\lambda_0)$ in figure 6.3 is

$$\zeta(m_a, \lambda_0) \leq \sqrt{\frac{I_{\text{obs}}(\lambda_0)}{I_{\text{th}}(m_a, \lambda_0)}}. \quad (6.27)$$

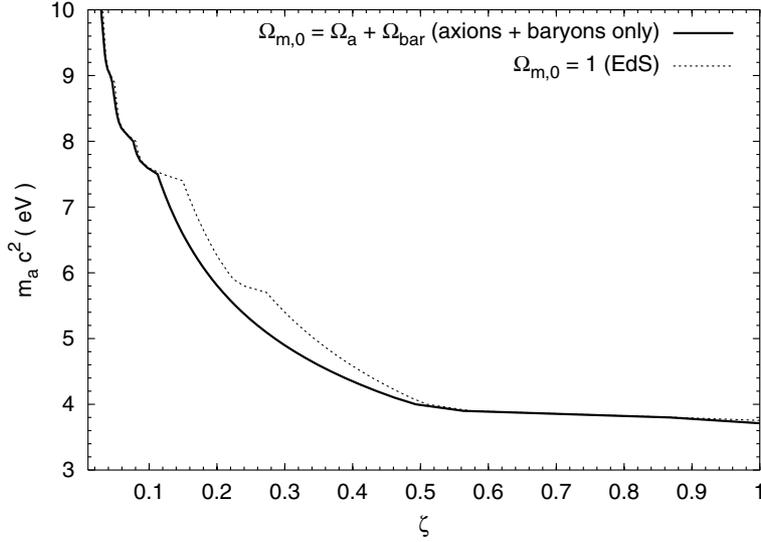


Figure 6.4. The upper limits on the value of $m_a c^2$ as a function of the coupling strength ζ (or vice versa). These are derived by requiring that the minimum predicted axion intensity (as plotted in figure 6.3) be less than or equal to the observational upper limits on the intensity of the EBL.

The upper limit on ζ (for each value of m_a) is then the smallest such value of $\zeta(m_a, \lambda_0)$; i.e. that which brings $I_{\text{th}}(m_a, \lambda_0)$ down to $I_{\text{obs}}(\lambda_0)$ or less for *each* wavelength λ_0 . From this procedure we obtain a function which can be regarded as an upper limit on the axion rest mass m_a as a function of ζ (or vice versa). Results are plotted in figure 6.4 (bold full line). This curve tells us that even in models where the axion–photon coupling is strongly suppressed and $\zeta = 0.07$, the axion cannot be more massive than

$$m_a c^2 \leq 8.0 \text{ eV} \quad (\zeta = 0.07). \quad (6.28)$$

For the simplest axion models with $\zeta = 1$, this tightens to

$$m_a c^2 \leq 3.7 \text{ eV} \quad (\zeta = 1). \quad (6.29)$$

As expected, these limits are stronger than those which would be obtained in an EdS model (faint dotted line in figure 6.4). This is a small effect, however, because the strongest constraints tend to come from the region near the peak wavelength (λ_a), whereas the difference between matter- and vacuum-dominated models is most pronounced at wavelengths longward of the peak where the majority of the radiation originates at high redshift. Figure 6.4 shows that cosmology has the most effect over the range $0.1 \lesssim \zeta \lesssim 0.4$, where upper limits on $m_a c^2$ weakened

by about 10% in the EdS model relative to one in which all the CDM is provided by axions of the kind we have discussed here.

Combining equations (6.6) and (6.29) we conclude that axions in the simplest models are confined to a slender range of viable rest masses:

$$2.2 \text{ eV} \lesssim m_a c^2 \leq 3.7 \text{ eV}. \quad (6.30)$$

Considerations of background radiation thus complement the red-giant bound (6.7) and *close off most, if not all of the multi-eV window for thermal axions*. The range of values (6.30) can be further narrowed by looking for the enhanced signal which might be expected to emanate from concentrations of bound axions associated with galaxies and clusters of galaxies, as first suggested by Kephart and Weiler in 1987 [32]. The most thorough search programme along these lines was that carried out in 1991 by Ressel [17], who found no evidence of such a signal from three selected clusters, further tightening the upper limit on the multi-eV axion window to 3.2 eV in the simplest axion models. Constraints obtained in this way for *non-thermal* axions would be considerably weaker, as noted by several workers [32, 46]. This has no effect on our results, however, since non-thermal axions are vastly outnumbered by thermal ones over the range of rest masses we have considered here.

Let us turn finally to the question of how much dark matter can be provided by axions of this type. With rest energies given by (6.30), equation (6.5) shows that

$$0.014 \lesssim \Omega_a \leq 0.053. \quad (6.31)$$

Here we have taken $0.6 \leq h_0 \leq 0.9$ as usual. This is comparable to the density of baryonic matter (section 4.3), but falls well short of most expectations for CDM density (section 4.4).

Our main conclusions, then, are as follows: thermal axions in the multi-eV window remain (only just) viable at the lightest end of the range of possible rest-masses given by equation (6.30). They may also exist with slightly higher rest-masses, up to the limit given by equation (6.28), but only in certain axion theories where their couplings to photons are weak. In either of these two scenarios, their contributions to the density of dark matter in the Universe are so feeble as to remove much of their motivation as CDM candidates. If they are to provide a significant portion of the dark matter, then axions must have rest masses in the ‘invisible’ range where they do not contribute significantly to the light of the night sky.

References

- [1] Peccei R and Quinn H 1977 *Phys. Rev. Lett.* **38** 1440
- [2] Weinberg S 1978 *Phys. Rev. Lett.* **40** 223
- [3] Wilczek F 1978 *Phys. Rev. Lett.* **40** 279

- [4] Kolb E W and Turner M S 1990 *The Early Universe* (Reading, MA: Addison-Wesley) pp 401–27
- [5] Preskill J, Wise M and Wilczek F 1983 *Phys. Lett. B* **120** 127
- [6] Abbott L and Sikivie P 1983 *Phys. Lett. B* **120** 133
- [7] Dine M and Fischler W 1983 *Phys. Lett. B* **120** 137
- [8] Davis R 1986 *Phys. Lett. B* **180** 255
- [9] Sikivie P 2000 *Beyond the Desert 1999* ed H V Klapdor-Kleingrothaus and I V Krivosheina (Oxford: Institute of Physics Press) p 547
- [10] Battye R A and Shellard E P S 2000 *Beyond the Desert 1999* ed H V Klapdor-Kleingrothaus and I V Krivosheina (Oxford: Institute of Physics Press) p 565
- [11] Janka H-T *et al* 1996 *Phys. Rev. Lett.* **76** 2621
- [12] Keil W *et al* 1997 *Phys. Rev. D* **56** 2419
- [13] Sikivie P 1983 *Phys. Lett.* **51** 1415
- [14] Yamamoto K *et al* 2000 *Dark Matter in Astro- and Particle Physics* ed H V Klapdor-Kleingrothaus (Heidelberg: Springer) p 638
- [15] Asztalos S J and Kinion D 2000 *Dark Matter in Astro- and Particle Physics* ed H V Klapdor-Kleingrothaus (Heidelberg: Springer) p 630
- [16] Csáki C, Kaloper N and Terning J 2002 *Phys. Rev. Lett.* **88** 161302
- [17] Ressel M T 1991 *Phys. Rev. D* **44** 3001
- [18] Kaplan D B 1985 *Nucl. Phys. B* **260** 215
- [19] Turner M S 1987 *Phys. Rev. Lett.* **59** 2489
- [20] Turner M S 1988 *Phys. Rev. Lett.* **60** 1797
- [21] Kim J-E 1987 *Phys. Rep.* **150** 1
- [22] Kim J E 1979 *Phys. Rev. Lett.* **43** 103
- [23] Shifman M A, Vainshtein A I and Zakharov V I 1980 *Nucl. Phys. B* **166** 493
- [24] Zhitnitsky A R 1980 *Sov. J. Nucl. Phys.* **31** 260
- [25] Dine M, Fischler W and Srednicki M 1981 *Phys. Lett. B* **104** 199
- [26] Raffelt G and Weiss A 1995 *Phys. Rev. D* **51** 1495
- [27] Raffelt G G and Dearborn D S P 1987 *Phys. Rev. D* **36** 2211
- [28] Engel J, Seckel D and Hayes A C 1990 *Phys. Rev. D* **65** 960
- [29] Peebles P J E 1993 *Principles of Physical Cosmology* (Princeton, NJ: Princeton University Press) p 122
- [30] Jaffe W 1983 *Mon. Not. R. Astron. Soc.* **202** 995
- [31] Kochanek C S 1996 *Astrophys. J.* **457** 228
- [32] Kephart T W and Weiler T J 1987 *Phys. Rev. Lett.* **58** 171
- [33] Lillie C F and Witt A N 1976 *Astrophys. J.* **208** 64
- [34] Spinrad H and Stone R P S 1978 *Astrophys. J.* **226** 609
- [35] Dube R R, Wickes W C and Wilkinson D T 1979 *Astrophys. J.* **232** 333
- [36] Boughn S P and Kuhn J R 1986 *Astrophys. J.* **309** 33
- [37] Toller G N 1983 *Astrophys. J.* **266** L79
- [38] Jakobsen P *et al* 1984 *Astron. Astrophys.* **139** 481
- [39] Tennyson P D *et al* 1988 *Astrophys. J.* **330** 435
- [40] Murthy J *et al* 1990 *Astron. Astrophys.* **231** 187
- [41] Bernstein R A 1999 *The Low Surface Brightness Universe* (Astronomical Society of the Pacific Conference Series, Volume 170) ed J I Davies, C Impey and S Phillipps (San Francisco, CA: ASP) p 341
- [42] Hauser M G *et al* 1998 *Astrophys. J.* **508** 25
- [43] Wright E L and Reese E D 2000 *Astrophys. J.* **545** 43

- [44] Cambr sy L *et al* 2001 *Astrophys. J.* **555** 563
- [45] Overduin J M and Wesson P S 1993 *Astrophys. J.* **414** 449
- [46] Lodenquai J F and Dixit V V 1987 *Phys. Lett. B* **194** 350

Chapter 7

Neutrinos

7.1 The decaying-neutrino hypothesis

Experiments now indicate that neutrinos possess non-zero rest mass and make up at least part of the dark matter. If different neutrino species have different rest masses, then heavier species can decay into lighter ones plus a photon. The neutrino number density, moreover, is high enough that their decay photons might be observable, as first appreciated by Cowsik in 1977 [1] and de Rujula and Glashow in 1980 [2]. The strength of the expected signal depends on the way in which neutrino rest masses are incorporated into the standard model of particle physics. In minimal extensions of this model, radiative neutrino decays are characterized by lifetimes on the order of 10^{29} yr or more [3]. This is so much longer than the age of the Universe that neutrinos are effectively *stable* and would not produce a detectable signal. In other theories such as those involving supersymmetry, however, their decay lifetime can drop to 10^{15} yr [4]. This is within five orders of magnitude of the age of the Universe and opens up the possibility of significant contributions to the extragalactic background light (EBL).

Decay photons from neutrinos with lifetimes this short are also interesting because of their potential for resolving a number of longstanding astrophysical puzzles related to the ionization of hydrogen and nitrogen in the interstellar and intergalactic medium [5, 6]. As was first pointed out by Melott *et al* in 1988 [7], these point in particular to a neutrino which decays on timescales of order $\tau_\nu \sim 10^{24}$ s and has a rest energy $m_\nu \sim 30$ eV. The latter value fits awkwardly with current thinking on large-scale structure formation in the early Universe (section 4.5). Neutrinos of this kind could help with so many other problems, however, that the prospect of their existence bears further scrutiny. Sciama and others have been led on this basis to develop a detailed scenario, now known as the **decaying-neutrino hypothesis** [8–10]. In this, the rest energy and decay lifetime of the massive τ neutrino are specified as

$$m_{\nu_\tau} c^2 = 28.9 \pm 1.1 \text{ eV} \quad \tau_\nu = (2 \pm 1) \times 10^{23} \text{ s}. \quad (7.1)$$

The τ neutrino decays into a μ neutrino plus a photon (figure 7.1). Assuming that $m_{\nu_\tau} \gg m_{\nu_\mu}$, conservation of energy and momentum require this photon to have an energy $E_\gamma = \frac{1}{2}m_{\nu_\tau}c^2 = 14.4 \pm 0.5$ eV.

The concreteness of this proposal has made it eminently testable. Some of the strongest bounds come from searches for the line emission near 14 eV that would be expected from concentrations of decaying dark matter in clusters of galaxies. No such signal has been seen in the direction of the galaxy cluster surrounding the quasar 3C 263 [11] or in the direction of the rich cluster Abell 665 which was observed using the Hopkins Ultraviolet Telescope in 1991 [12]. It may be, however, that absorption plays a stronger role than expected along the lines of sight to these clusters or that most of their dark matter is in another form [4, 13]. A potentially more robust test of the decaying-neutrino hypothesis comes from the *diffuse* background light. This has been looked at in a number of studies [14–20]. The task is a challenging one for several reasons. Photons whose energies lie near 14 eV are strongly absorbed by both dust and neutral hydrogen, and the distribution of these quantities in intergalactic space is uncertain. It is also notoriously difficult, perhaps more so in this part of the spectrum than any other, to distinguish between those parts of the background which are truly extragalactic and those which are due to a complex mixture of competing foreground signals [21, 22].

Here we reconsider the problem with the help of the formalism developed in preceding chapters, adapting it to allow for absorption by dust and neutral hydrogen. Despite all the uncertainties, we will come to a firm conclusion: existing observations are sufficient to rule out the decaying-neutrino hypothesis, unless either the rest mass or decay lifetime of the neutrino lie outside the bounds specified by the theory.

7.2 Bound neutrinos

Neutrinos, the original hot dark-matter candidates, are likely to be found predominantly outside gravitational potential wells such as those of individual galaxies. However, some at least will have tended to collect into galactic dark-matter halos since the epoch of galaxy formation. These **neutrino halos** comprise our sources of background radiation in this section. Their comoving number density is just that of the galaxies themselves, given by (6.12) as $n_0 = 0.010h_0^3 \text{ Mpc}^{-3}$. The decay photon wavelengths at emission can be taken as distributed normally about the peak wavelength corresponding to E_γ :

$$\lambda_\nu = \frac{hc}{E_\gamma} = 860 \pm 30 \text{ \AA}. \quad (7.2)$$

This lies in the extreme ultraviolet (EUV) portion of the spectrum, although the redshifted tail of the observed photon spectrum will stretch across the far ultraviolet (FUV) and near ultraviolet (NUV) bands as well. [For definiteness,

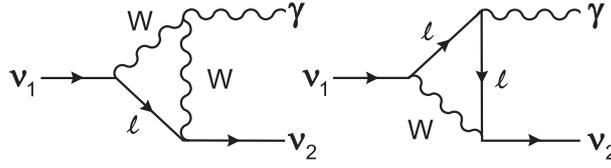


Figure 7.1. Feynman diagrams corresponding to the decay of a massive neutrino (ν_1) into a second, lighter neutrino species (ν_2) together with a photon (γ). The process is mediated by charged leptons (ℓ) and the W boson (W).

we take these wavebands to extend over 100–912 Å (EUV), 912–2000 Å (FUV) and 2000–4000 Å (NUV). It should, however, be noted that universal standards have yet to be established in this area and individual conventions vary.] The spectral energy distribution (SED) of the decaying-neutrino halos is then given by equation (3.19) :

$$F(\lambda) = \frac{L_h}{\sqrt{2\pi}\sigma_\lambda} \exp\left[-\frac{1}{2}\left(\frac{\lambda - \lambda_\gamma}{\sigma_\lambda}\right)^2\right] \quad (7.3)$$

where L_h , the halo luminosity, has yet to be determined. For the standard deviation σ_λ we can follow the same procedure as in the preceding chapter and use the velocity dispersion of the bound neutrinos, giving $\sigma_\lambda = 2\lambda_\nu v_c/c$. Let us parametrize this for convenience using the range of uncertainty given by (7.2) for the value of λ_ν , so that $\sigma_{30} \equiv \sigma_\lambda/(30 \text{ \AA})$.

The halo luminosity is just the ratio of the number of decaying neutrinos (N_τ) to their decay lifetime (τ_ν), multiplied by the energy of each decay photon (E_γ). Because the latter is just above the hydrogen-ionizing energy of 13.6 eV, we also need to multiply the result by an **efficiency factor** ϵ (between zero and one), to reflect the fact that some of the decay photons are absorbed by neutral hydrogen in their host galaxy before they can leave the halo and contribute to its luminosity. Altogether, then:

$$L_h = \frac{\epsilon N_\tau E_\gamma}{\tau_\nu} = \frac{\epsilon M_h c^2}{2\tau_\nu}. \quad (7.4)$$

Here we have expressed N_τ as the number of neutrinos with rest mass $m_{\nu_\tau} = 2E_\gamma/c^2$ per halo mass M_h .

To calculate the mass of the halo, let us follow reasoning similar to that adopted for axion halos in section 6.3 and assume that the ratio of baryonic-to-total mass in the halo is comparable to the ratio of baryonic-to-total matter density in the Universe at large:

$$\frac{M_{\text{tot}} - M_h}{M_{\text{tot}}} = \frac{M_{\text{bar}}}{M_{\text{tot}}} = \frac{\Omega_{\text{bar}}}{\Omega_{\text{bar}} + \Omega_\nu}. \quad (7.5)$$

Here we have made the economical assumption that there are no *other* contributions to the matter density, apart from those of baryons and massive neutrinos. It follows from equation (7.5) that

$$M_h = M_{\text{tot}} \left(1 + \frac{\Omega_{\text{bar}}}{\Omega_\nu} \right)^{-1}. \quad (7.6)$$

We take $M_{\text{tot}} = (2 \pm 1) \times 10^{12} M_\odot$ following equation (6.15). For Ω_{bar} we use the value $(0.016 \pm 0.005) h_0^{-2}$ quoted in section 4.3. And to calculate Ω_ν we put the neutrino rest mass m_{ν_τ} into equation (4.7), giving

$$\Omega_\nu = (0.31 \pm 0.01) h_0^{-2}. \quad (7.7)$$

Inserting these values of M_{tot} , Ω_{bar} and Ω_ν into (7.6), we obtain

$$M_h = (0.95 \pm 0.01) M_{\text{tot}} = (1.9 \pm 0.9) \times 10^{12} M_\odot. \quad (7.8)$$

This leaves a baryonic mass $M_{\text{bar}} = M_{\text{tot}} - M_h \approx 1 \times 10^{11} M_\odot$, in good agreement with the observed sum of contributions from stars in the disc, bulge and halo of our own Galaxy, together with the matter making up the interstellar medium.

The neutrino density (7.7), when combined with that of baryons, leads to a total present-day matter density of

$$\Omega_{\text{m},0} = \Omega_{\text{bar}} + \Omega_\nu = (0.32 \pm 0.01) h_0^{-2}. \quad (7.9)$$

As pointed out by Sciamia [9], massive neutrinos are thus consistent with a critical-density EdS Universe ($\Omega_{\text{m},0} = 1$) if

$$h_0 = 0.57 \pm 0.01. \quad (7.10)$$

This is just below the range of values which many workers now consider observationally viable for Hubble's constant (section 4.3). But it is a striking fact that the same neutrino rest mass which resolves several unrelated astrophysical problems also implies a reasonable expansion rate in the simplest cosmological model. In the interests of testing the decaying-neutrino hypothesis in a self-consistent way, we will follow Sciamia in adopting the narrow range of values (7.10) for this chapter only.

7.3 Luminosity

To evaluate the halo luminosity (7.4), it remains to find the fraction ϵ of decay photons which escape from the halo. A derivation of ϵ is given in appendix C, based on the hydrogen photoionization cross section with respect to 14 eV photons, and taking into account the distribution of decaying neutrinos relative to that of neutral hydrogen in the galactic disc. Here we show that the results can

be simply understood once it is appreciated that the photoionization cross section and hydrogen density are such that effectively all of the decay photons striking the disc are absorbed. The probability of absorption for a single decay photon is then proportional to the solid angle subtended by the galactic disc, as seen from the point where the photon is released.

We model the distribution of τ neutrinos (and their decay photons) in the halo with a **flattened ellipsoidal profile** which has been advocated in the context of the decaying-neutrino scenario by Salucci and Sciama [23]. This has

$$\rho_\nu(r, z) = 4n_\odot m_{\nu\tau} \mathcal{N}_\nu(r, \theta) \quad (7.11)$$

with

$$\mathcal{N}_\nu(r, \theta) \equiv \left[1 + \sqrt{(r/r_\odot)^2 \sin^2 \theta + (r/h)^2 \cos^2 \theta} \right]^{-2}.$$

Here r and θ are spherical coordinates, $n_\odot = 5 \times 10^7 \text{ cm}^{-3}$ is the number density of halo neutrinos in the vicinity of the Sun, $r_\odot = 8 \text{ kpc}$ is the distance of the Sun from the centre of the Galaxy, and $h = 3 \text{ kpc}$ is the scale height of the halo. Although this function has essentially been constructed to account for the ionization structure of the Milky Way, it agrees reasonably well with dark-matter halo distributions which have derived on strictly dynamical grounds [24].

Defining $x \equiv r/r_\odot$, one can use (7.11) to express the mass M_h of the halo in terms of the halo radius r_h as

$$M_h(r_h) = M_\nu \int_{\theta=0}^{\pi/2} \int_{x=0}^{x_{\max}(r_h, \theta)} \mathcal{N}_\nu(x, \theta) x^2 \sin \theta \, dx \, d\theta \quad (7.12)$$

where

$$M_\nu \equiv 16\pi n_\odot m_{\nu\tau} r_\odot^3 = 9.8 \times 10^{11} M_\odot$$

$$x_{\max}(r_h, \theta) = (r_h/r_\odot) / \sqrt{\sin^2 \theta + (r_\odot/h)^2 \cos^2 \theta}.$$

Outside $x > x_{\max}$, we assume that the halo density drops off exponentially and can be ignored. Using (7.12) it can be shown that halos whose masses M_h are given by (7.8) have scale radii $r_h = (70 \pm 25) \text{ kpc}$. This is consistent with evidence from the motion of galactic satellites [25].

We now put ourselves in the position of a decay photon released at cylindrical coordinates (y_ν, z_ν) inside the halo, as shown in figure 7.2(a). It may be seen that the disc of the Galaxy presents an approximately elliptical figure with maximum angular width 2α and angular height 2β , where

$$\alpha = \tan^{-1} \sqrt{(r_d^2 - d^2) / [(y_\nu - d)^2 + z_\nu^2]}$$

$$\beta = \frac{1}{2} \left[\tan^{-1} \left(\frac{y_\nu + r_d}{z_\nu} \right) - \tan^{-1} \left(\frac{y_\nu - r_d}{z_\nu} \right) \right]. \quad (7.13)$$

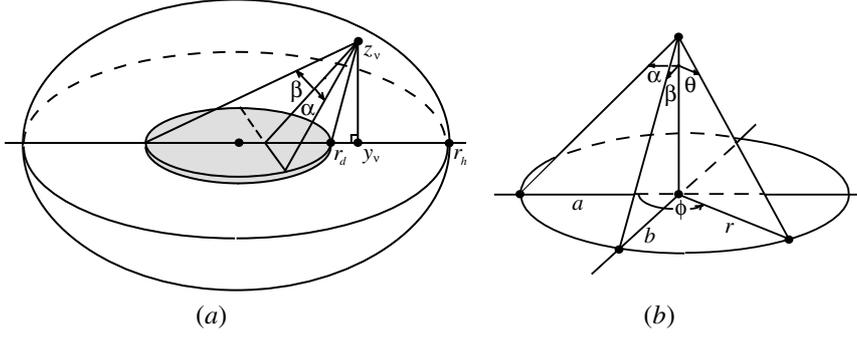


Figure 7.2. Left-hand side (a): absorption of decay photons inside the halo. For neutrinos decaying at (y_v, z_v) , the probability that decay photons will be absorbed inside the halo (radius r_h) is essentially the same as the probability that they will strike the galactic disc (radius r_d , shaded). Right-hand side (b): an ellipse of semi-major axis a , semi-minor axis b and radial arm $r(\phi)$ subtends corresponding angles of α , β and $\theta(\phi)$ respectively.

Here $d = [(y_v^2 + z_v^2 + r_d^2) - \sqrt{(y_v^2 + z_v^2 + r_d^2)^2 - 4y_v^2 r_d^2}]/2y_v$, $y_v = r \sin \theta$, $z_v = r \cos \theta$ and $r = r_\odot x$. In spherical coordinates centred on the photon, the solid angle subtended by an ellipse is

$$\Omega_e = \int_0^{2\pi} d\phi \int_0^{\theta(\phi)} \sin \theta' d\theta' = \int_0^{2\pi} [1 - \cos \theta(\phi)] d\phi. \quad (7.14)$$

Here $\theta(\phi)$ is the angle subtended by a radial arm of the ellipse as depicted in figure 7.2(b). The cosine of this angle can be expressed in terms of α and β by

$$\cos \theta(\phi) = \left[1 + \frac{\tan^2 \alpha \tan^2 \beta}{\tan^2 \alpha \sin^2 \phi + \tan^2 \beta \cos^2 \phi} \right]^{-1/2}. \quad (7.15)$$

The single-point probability that a photon released at (x, θ) will escape from the halo is then

$$\mathcal{P}_e = 1 - \frac{\Omega_e(\alpha, \beta)}{4\pi}. \quad (7.16)$$

For a given halo size r_h and disc size r_d , we obtain a good approximation to ϵ by averaging \mathcal{P}_e over all locations (x, θ) in the halo and weighting by the neutrino number density \mathcal{N}_v . Trial and error shows that, for the range of halo sizes considered here, a disc radius of $r_d = 36$ kpc leads to values which are within 1% of those found in appendix C. The latter read:

$$\epsilon = \begin{cases} 0.63 & (r_h = 45 \text{ kpc}) \\ 0.77 & (r_h = 70 \text{ kpc}) \\ 0.84 & (r_h = 95 \text{ kpc}). \end{cases} \quad (7.17)$$

As expected, the escape fraction of decay photons goes up as the scale size of the halo increases relative to that of the disc. As $r_h \gg r_d$ one gets $\epsilon \rightarrow 1$, while a small halo with $r_h \lesssim r_d$ leads to $\epsilon \approx 0.5$.

With the decay lifetime τ_ν , halo mass M_h and efficiency factor ϵ all known, equation (7.4) gives for the luminosity of the halo:

$$L_h = (6.5 \times 10^{42} \text{ erg s}^{-1}) f_h f_\tau^{-1}. \quad (7.18)$$

Here we have introduced two dimensionless constants f_h and f_τ in order to parametrize the uncertainties in ϵM_h and τ_ν . For the ranges of values given earlier, these take the values $f_h = 1.0 \pm 0.6$ and $f_\tau = 1.0 \pm 0.5$ respectively. Setting $f_h = f_\tau = 1$ gives a halo luminosity of about $2 \times 10^9 L_\odot$, or less than 5% of the optical luminosity of the Milky Way ($L_0 = 2 \times 10^{10} h_0^{-2} L_\odot$), with h_0 as specified by equation (7.10).

The combined bolometric intensity of decay photons from all the bound neutrinos out to a redshift z_f is given by (2.21) as usual:

$$Q_{\text{bound}} = Q_h \int_0^{z_f} \frac{dz}{(1+z)^2 \tilde{H}(z)} \quad (7.19)$$

where

$$Q_h \equiv \frac{cn_0 L_h}{H_0} = (2.0 \times 10^{-5} \text{ erg s}^{-1} \text{ cm}^{-2}) h_0^2 f_h f_\tau^{-1}.$$

The h_0 -dependence in this quantity comes from the fact that we have so far considered only neutrinos in galaxy halos, whose number density n_0 goes as h_0^3 . Since we follow Sciamia in adopting the EdS cosmology in this chapter, the Hubble expansion rate (2.40) is

$$\tilde{H}(z) = (1+z)^{3/2}. \quad (7.20)$$

Putting this into (7.19), we find

$$Q_{\text{bound}} \approx \frac{2}{5} Q_h = (8.2 \times 10^{-6} \text{ erg s}^{-1} \text{ cm}^{-2}) h_0^2 f_h f_\tau^{-1}. \quad (7.21)$$

(The approximation is good to better than 1% if $z_f \geq 8$.) Here we have neglected absorption *between* the galaxies, an issue we will return to shortly. Despite their size, dark-matter halos in the decaying-neutrino hypothesis are not very bright. Their combined intensity is about 1% of that of the EBL due to galaxies, $Q_* \approx 3 \times 10^{-4} \text{ erg s}^{-1} \text{ cm}^{-2}$ (section 2.3). The main reason for this is the long neutrino decay lifetime, five orders of magnitude longer than the age of the galaxies. The ultraviolet decay photons have not had time to contribute as much to the EBL as their counterparts in the optical part of the spectrum.

7.4 Free-streaming neutrinos

The cosmological density of decaying τ -neutrinos in dark-matter halos is small: $\Omega_{\nu, \text{bound}} = n_0 M_h / \rho_{\text{crit}} = (0.068 \pm 0.032) h_0$. With h_0 as given by (7.10), this

amounts to less than 6% of the total neutrino density, equation (7.7). Therefore, as expected for hot dark-matter particles, the bulk of the EBL contributions in the decaying-neutrino scenario come from neutrinos which are distributed on larger scales. We will refer to these collectively as **free-streaming neutrinos**, though some of them may actually be associated with larger scale systems such as clusters of galaxies. (The distinction is not critical for our purposes, since we are concerned with the summed contributions of all such neutrinos to the diffuse background.) Their cosmological density is found using (7.7) as $\Omega_{\nu,\text{free}} = \Omega_{\nu} - \Omega_{\nu,\text{bound}} = 0.30h_0^{-2}f_f$, where the dimensionless constant f_f parametrizes the uncertainties in this quantity and takes the value $f_f = 1.00 \pm 0.05$.

To identify ‘sources’ of radiation in this section we follow the same procedure as with vacuum regions (section 5.4) and axions (section 6.3) and divide up the Universe into regions of comoving volume $V_0 = n_0^{-1}$. The mass of each region is

$$M_f = \Omega_{\nu,\text{free}}\rho_{\text{crit},0}V_0 = \Omega_{\nu,\text{free}}\rho_{\text{crit},0}/n_0. \quad (7.22)$$

The luminosity of these sources has the same form as equation (7.4) except that we put $M_h \rightarrow M_f$ and drop the efficiency factor ϵ since the density of intergalactic hydrogen is too low to absorb a significant fraction of the decay photons within each region. Thus,

$$L_f = \frac{\Omega_{\nu,\text{free}}\rho_{\text{crit},0}c^2}{2n_0\tau_\nu}. \quad (7.23)$$

With the values already mentioned for $\Omega_{\nu,\text{free}}$ and τ_ν , and with $\rho_{\text{crit},0}$ and n_0 given by (2.36) and (6.12) respectively, equation (7.23) implies a comoving luminosity density due to free-streaming neutrinos of

$$\mathcal{L}_f = n_0L_f = (1.2 \times 10^{-32} \text{ erg s}^{-1} \text{ cm}^{-3})f_f f_\tau^{-1}. \quad (7.24)$$

This is high: $0.5h_0^{-1}$ times the luminosity density of the Universe, as given by equation (2.24). Based on this, we may anticipate strong constraints on the decaying-neutrino hypothesis from observation. Similar conclusions follow from the bolometric intensity due to free-streaming neutrinos. Replacing L_h with L_f in (7.19), we find

$$Q_{\text{free}} = \frac{2cn_0L_f}{5H_0} = (1.2 \times 10^{-4} \text{ erg s}^{-1} \text{ cm}^{-2})h_0^{-1}f_f f_\tau^{-1}. \quad (7.25)$$

This is of the same order of magnitude as Q_* and goes as h_0^{-1} rather than h_0^2 . Taking into account the uncertainties in h_0 , f_h , f_f and f_τ , the bolometric intensity of bound and free-streaming neutrinos together is

$$Q = Q_{\text{bound}} + Q_{\text{free}} = (0.33 \pm 0.17)Q_*. \quad (7.26)$$

In principle, then, these particles are capable of producing a background as bright as that from the galaxies themselves, equation (2.49). The vast majority of their light comes from free-streaming neutrinos. These are more numerous than their halo-bound counterparts, and are not appreciably affected by absorption at source.

7.5 Intergalactic absorption

To constrain neutrinos in a more quantitative way, it is necessary to determine their EBL contributions as a function of wavelength. In the absence of absorption, this would be done as in previous chapters by putting the source luminosity (which may be either L_h for the galaxy halos or L_f for the free-streaming neutrinos) into the SED (7.3) and substituting the latter into equation (3.6). This gives

$$I_\lambda(\lambda_0) = I_\nu \int_0^{z_f} (1+z)^{-9/2} \exp \left\{ -\frac{1}{2} \left[\frac{\lambda_0/(1+z) - \lambda_\nu}{\sigma_\lambda} \right]^2 \right\} dz. \quad (7.27)$$

Here we have used (7.20) for $\tilde{H}(z)$. The numerical value of the prefactor I_ν is given with the help of (7.18) for bound neutrinos and (7.23) for free-streaming ones as follows:

$$\begin{aligned} I_\nu &= \frac{cn_0}{\sqrt{32\pi^3 H_0 \sigma_\lambda}} \times \begin{cases} L_h \\ L_f \end{cases} \quad (7.28) \\ &= \begin{cases} (940 \text{ CUs}) h_0^2 f_h f_\tau^{-1} \sigma_{30}^{-1} (\lambda_0/\lambda_\nu) & \text{(bound)} \\ (5280 \text{ CUs}) h_0^{-1} f_f f_\tau^{-1} \sigma_{30}^{-1} (\lambda_0/\lambda_\nu) & \text{(free)}. \end{cases} \end{aligned}$$

Now, however, we must take into account the fact that decay photons (from both bound and free-streaming neutrinos) encounter significant amounts of absorbing material as they travel through the **intergalactic medium** (IGM) on the way to our detectors. The wavelength of neutrino decay photons, $\lambda_\nu = 860 \pm 30 \text{ \AA}$, is just shortward of the Lyman- α line at 912 \AA , which means that these photons are absorbed almost as strongly as they can be by neutral hydrogen (this is, in fact, one of the prime motivations of the theory). It is also very close to the waveband of peak extinction by dust. The simplest way to handle both these types of absorption is to include an opacity term $\tau(\lambda_0, z)$ inside the argument of the exponential, so that (7.27) is modified to read

$$I_\lambda(\lambda_0) = I_\nu \int_0^{z_f} (1+z)^{-9/2} \exp \left\{ -\frac{1}{2} \left[\frac{\lambda_0/(1+z) - \lambda_\nu}{\sigma_\lambda} \right]^2 - \tau(\lambda_0, z) \right\} dz. \quad (7.29)$$

The **optical depth** $\tau(\lambda_0, z)$ can be broken into separate terms corresponding to hydrogen gas and dust along the line of sight:

$$\tau(\lambda_0, z) = \tau_{\text{gas}}(\lambda_0, z) + \tau_{\text{dust}}(\lambda_0, z). \quad (7.30)$$

Our best information about both of these quantities comes from high-redshift quasar spectra. The fact that these can be seen at all already puts limits on the degree of attenuation due to intervening matter. And the shape of the observed spectra can give us clues about the effects of the absorbing medium in specific wavebands.

The gas component contains fewer uncertainties and is better understood at present. Zuo and Phinney [26] have developed a formalism to describe the absorption due to randomly distributed clouds such as **quasar absorption-line systems** and normalized this to the number of Lyman-limit systems at $z = 3$. We use their model 1, which gives the highest absorption below $\lambda_0 \lesssim 2000 \text{ \AA}$ and is thus conservative for our purposes. Assuming an EdS cosmology, the optical depth at λ_0 due to neutral hydrogen out to a redshift z is given by

$$\tau_{\text{gas}}(\lambda_0, z) = \begin{cases} \tau_{\text{ZP}} \left(\frac{\lambda_0}{\lambda_{\text{L}}} \right)^{3/2} \ln(1+z) & (\lambda_0 \leq \lambda_{\text{L}}) \\ \tau_{\text{ZP}} \left(\frac{\lambda_0}{\lambda_{\text{L}}} \right)^{3/2} \ln \left(\frac{1+z}{\lambda_0/\lambda_{\text{L}}} \right) & [\lambda_{\text{L}} < \lambda_0 < \lambda_{\text{L}}(1+z)] \\ 0 & [\lambda_0 \geq \lambda_{\text{L}}(1+z)] \end{cases} \quad (7.31)$$

where $\lambda_{\text{L}} = 912 \text{ \AA}$ and $\tau_{\text{ZP}} = 2.0$.

Dust is a more complicated and potentially more important issue, and we pause to discuss this critically before presenting our model. The simplest possibility, and the one which should be most effective in obscuring a diffuse signal like that considered here, would be for the dust to be spread uniformly through intergalactic space. A quantitative estimate of opacity due to a uniform dusty intergalactic medium has, in fact, been suggested [27] but is regarded as an extreme upper limit because it would lead to excessive reddening of quasar spectra [28]. Subsequent discussions have tended to treat intergalactic dust as clumpy [29], with significant debate about the extent to which such clumps would redden and/or hide background quasars, possibly helping to explain the observed ‘turnoff’ in quasar population at around $z \sim 3$ [30–33]. Most of these models assume an EdS cosmology. The effects of dust extinction could be enhanced if $\Omega_{\text{m},0} < 1$ and/or $\Omega_{\Lambda,0} > 0$ [32], but we ignore this possibility here because neutrinos (not vacuum energy) are assumed to make up the critical density in the decaying-neutrino scenario.

We will use a formalism due to Fall and Pei [34] in which dust is associated with damped Ly α absorbers whose numbers and density profiles are sufficient to obscure a portion of the light reaching us from $z \sim 3$, but not to account fully for the turnoff in quasar population. Obscuration is calculated based on the column density of hydrogen in these systems, together with estimates of the dust-to-gas ratio, and is normalized to the observed quasar luminosity function. The resulting mean optical depth at λ_0 out to redshift z is

$$\tau_{\text{dust}}(\lambda_0, z) = \int_0^z \frac{\tau_{\text{FP}}(z')(1+z')}{(1+\Omega_{\text{m},0}z')^{1/2}} \xi \left(\frac{\lambda_0}{1+z'} \right) dz'. \quad (7.32)$$

Here $\xi(\lambda)$ is the **extinction** of light by dust at wavelength λ relative to that in the B-band (4400 \AA). If $\tau_{\text{FP}}(z) = \text{constant}$ and $\xi(\lambda) \propto \lambda^{-1}$, then τ_{dust} is proportional to $\lambda_0^{-1}[(1+z)^3 - 1]$ or $\lambda_0^{-1}[(1+z)^{2.5} - 1]$, depending on cosmology [27, 29].

In the more general treatment of Fall and Pei [34], $\tau_{\text{FP}}(z)$ is parametrized as a function of redshift so that

$$\tau_{\text{FP}}(z) = \tau_{\text{FP}}(0)(1+z)^\delta \quad (7.33)$$

where $\tau_{\text{FP}}(0)$ and δ are adjustable parameters. Assuming an EdS cosmology ($\Omega_{\text{m},0} = 1$), the observational data are consistent with lower limits of $\tau_*(0) = 0.005$, $\delta = 0.275$ (**model A**); best-fit values of $\tau_*(0) = 0.016$, $\delta = 1.240$ (**model B**); or upper limits of $\tau_*(0) = 0.050$, $\delta = 2.063$ (**model C**). We will use all three models in what follows.

To calculate the extinction $\xi(\lambda)$ in the 300–2000 Å range, we use numerical Mie scattering routines in conjunction with various dust populations. In performing these calculations, we tacitly assume that intergalactic and interstellar dust are similar in nature, which is a reasonable assumption that is, of course, very difficult to test. Many people have constructed dust-grain models that reproduce the average extinction curve for the **diffuse interstellar medium** (DISM) at $\lambda > 912$ Å [35] but there have been fewer studies at shorter wavelengths. One such study is that of Martin and Rouleau [36], who extended earlier silicate/graphite synthetic extinction curves due to Draine and Lee [37] assuming:

- (1) two populations of homogeneous spherical dust grains composed of graphite and silicates respectively;
- (2) a power-law size distribution of the form $a^{-3.5}$ where a is the grain radius;
- (3) a range of grain radii from 50–2500 Å; and
- (4) solar abundances of carbon and silicon relative to hydrogen [38].

The last of these assumptions is questionable in light of new work which shows that heavy elements may be far less abundant in the DISM than they are in the Sun. Snow and Witt [39], for example, report interstellar abundances of $214 \times 10^{-6}/\text{H}$ and $18.6 \times 10^{-6}/\text{H}$ for carbon and silicon respectively. This cuts earlier values in half and actually makes it difficult for a simple silicate/graphite model to reproduce the observed DISM extinction curve. We therefore derive new dust-extinction curves based on the revised abundances. In the interests of obtaining conservative bounds on the decaying-neutrino hypothesis, we also consider a range of modified grain populations, looking, in particular, for those which provide optimal extinction efficiency in the FUV without drifting too far from the average DISM curve in the optical and NUV bands. We describe the general characteristics of these models here and show the resulting extinction curves in figure 7.3; more details are found in [20]).

Our **population 1** grain model (figure 7.3, dash-dotted line) assumes the standard grain model employed by other workers, but uses the new, lower abundance numbers together with dielectric functions due to Draine [40]. The shape of the extinction curve provides a reasonable fit to observation at longer wavelengths (reproducing, for example, the absorption bump at 2175 Å); but its magnitude is too low, confirming the inadequacies of the old dust model.

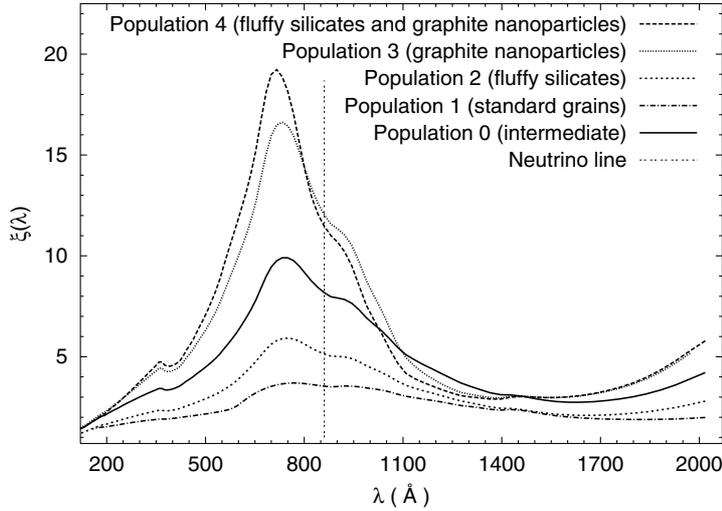


Figure 7.3. The FUV extinction (relative to that in the B-band) produced by five different dust-grain populations. Standard grains (population 1) produce the least extinction, while PAH-like carbon nanoparticles (population 3) produce the most extinction near the decaying neutrino line at 860 Å (vertical line).

Extinction in the vicinity of 860 Å is also weak (with a peak value of $\tau_{\max} \sim 1.0 \times 10^{-21} \text{ cm}^2 \text{ H}^{-1}$ at 770 Å), so that this model is able to ‘hide’ very little of the light from decaying neutrinos. Insofar as it almost certainly underestimates the true extent of extinction by dust, this grain model provides a good *lower limit* on absorption in the context of the decaying-neutrino hypothesis.

The silicate component of our **population 2** grain model is modified along the lines of the ‘fluffy silicate’ model which has been suggested as a resolution of the heavy-element abundance crisis in the DISM [41]. We replace the standard silicates of population 1 by silicate grains with a 45% void fraction, assuming a silicon abundance of $32.5 \times 10^{-6}/\text{H}$ [20]. We also decrease the size of the graphite grains ($a = 50\text{--}250 \text{ Å}$) and reduce the carbon depletion to 60% to provide a better match to the DISM curve. This mixture provides a better match to the interstellar data at optical wavelengths, and also shows significantly more FUV extinction than population 1, with a peak of $\tau_{\max} \sim 2.0 \times 10^{-21} \text{ cm}^2 \text{ H}^{-1}$ at 750 Å. Results are shown in figure 7.3 as a short-dashed line.

For **population 3**, we retain the standard silicates of population 1 but modify the graphite component as an approximation to the polycyclic aromatic hydrocarbon (PAH) nanostructures which have recently been proposed as carriers of the 2175 Å absorption bump [42]. PAH nanostructures are thought to consist of stacks of molecules such as coronene ($\text{C}_{24}\text{H}_{12}$), circumcoronene ($\text{C}_{54}\text{H}_{18}$) and

larger species in various states of edge hydrogenation. They have been linked to the $3.4 \mu\text{m}$ absorption feature in the DISM [43] as well as the extended red emission in nebular environments [44]. With sizes in the range $7\text{--}30 \text{ \AA}$, these structures are much smaller than the canonical graphite grains. Their dielectric functions, however, go over to that of graphite in the high-frequency limit [42]. So as an approximation to these particles, we use spherical graphite grains with extremely small radii ($3\text{--}150 \text{ \AA}$). This greatly increases extinction near the neutrino-decay peak, giving $\tau_{\text{max}} \sim 3.9 \times 10^{-21} \text{ cm}^2 \text{ H}^{-1}$ at 730 \AA . Results are shown in figure 7.3 as a dotted line.

Our **population 4** grain model, finally, combines the distinctive features of populations 2 and 3. It has a graphite component made up of nanoparticles (as in population 3) and a fluffy silicate component with a 45% porosity (like that of population 2). Results (plotted as a long-dashed line in figure 7.3) are close to those obtained with population 3. This is because extinction in the FUV waveband is dominated by small-particle contributions, so that silicates (whatever their void fraction) are of secondary importance. Absolute extinction rises slightly, with a peak of $\tau_{\text{max}} \sim 4.1 \times 10^{-21} \text{ cm}^2 \text{ H}^{-1}$ at 720 \AA . The quantity which is of most interest to us, extinction relative to that in the B-band, drops slightly longward of 800 \AA . Therefore it is the population 3 grains which provide us with the highest value of $\xi(\lambda_0)$ near 860 \AA , and hence a conservative *upper limit* on dust absorption in the context of the decaying-neutrino scenario. Neither the population 3 nor the population 4 grains fit the average DISM curve as well as those of population 2, because the Mie scattering formalism cannot accurately reproduce the behaviour of nanoparticles near the 2175 \AA resonance. Their high levels of extinction in the FUV region, however, suit these grain models for our purpose, which is to set the strongest possible limits on the decaying-neutrino hypothesis.

7.6 The ultraviolet background

We are now ready to specify the total optical depth (7.30) and hence to evaluate the intensity integral (7.29). We will do this using three combinations of the dust models just described, with a view to establishing lower and upper bounds on the EBL intensity predicted by the theory. A *minimum-absorption* model is obtained by combining Fall and Pei's model A with the extinction curve of the population 1 (standard) dust grains. At the other end of the spectrum, model C of Fall and Pei together with the population 3 (nanoparticle) grains provides the most conservative *maximum-absorption* model (for $\lambda_0 \gtrsim 800 \text{ \AA}$). Finally, as an intermediate model, we combine model B of Fall and Pei with the 'middle-of-the-road' extinction curve labelled as population 0 in figure 7.3.

The resulting predictions for the spectral intensity of the FUV background due to decaying neutrinos are plotted in figure 7.4 (light lines) and compared with observational limits (bold lines and points). The curves in the bottom half of this figure refer to EBL contributions from bound neutrinos only, while those

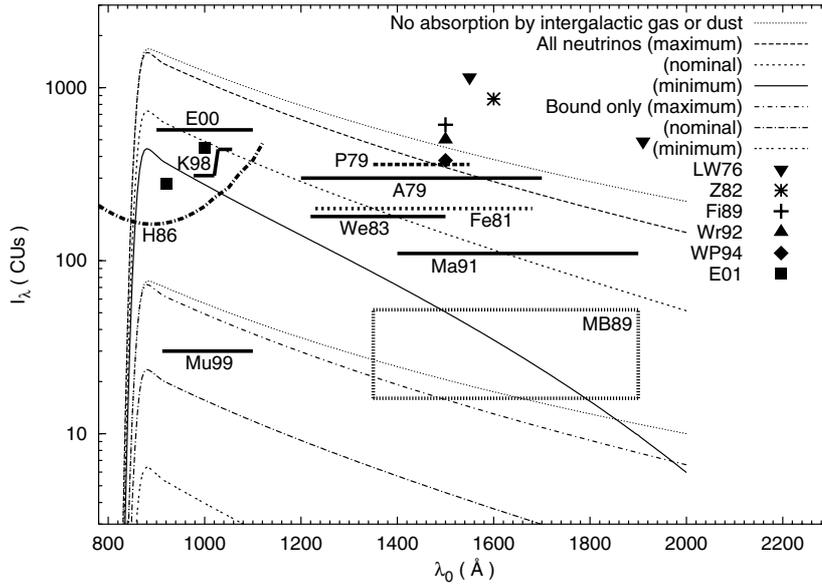


Figure 7.4. The spectral intensity I_λ of background radiation from decaying neutrinos as a function of observed wavelength λ_0 (light curves), plotted together with observational upper limits on EBL intensity in the far ultraviolet (points and bold curves). The bottom four theoretical curves refer to bound neutrinos only, while the top four refer to bound and free-streaming neutrinos together. The minimum predicted signals consistent with the theory are combined with the highest possible extinction in the intergalactic medium and vice versa. The faint dotted lines show the signal that would be seen (in the maximum-intensity case) if there were no intergalactic extinction at all.

in the top half correspond to contributions from both bound and free-streaming neutrinos together.

We begin our discussion with the bound neutrinos. The key results are the three widely-spaced curves in the lower half of the figure, with peak intensities of about 6, 20 and 80 CUs at $\lambda_0 \approx 900 \text{ \AA}$. These are obtained by letting h_0 and f_h take their minimum, nominal and maximum values respectively in (7.29), with the reverse order applying to f_τ . Simultaneously we have adopted the maximum, intermediate and minimum-absorption models for intergalactic dust, as described earlier. Thus the highest-intensity model is paired with the lowest possible dust extinction and vice versa. (The faint dotted line appended to the highest-intensity curve is included for comparison purposes and shows the effects of neglecting absorption by dust and gas altogether.) These curves should be seen as extreme upper and lower bounds on the theoretical intensity of EBL contributions from decaying neutrinos in galaxy halos.

They are best compared with an experimental measurement by Martin and Bowyer in 1989 [45], labelled ‘MB89’ in figure 7.4. These authors used data from a rocket-borne imaging camera to search for small-scale fluctuations in the FUV EBL, and deduced from this that the combined light of external galaxies (and their associated halos) reaches the Milky Way with an intensity of 16–52 CUs over 1350–1900 Å. There is now some doubt as to whether this was really an extragalactic signal, and indeed whether it is feasible to detect such a signal at all, given the brightness and fluctuations of the galactic foreground in this waveband [46]. Viable or not, however, it is of interest to see what a detection of this order would mean for the decaying-neutrino hypothesis. Figure 7.4 shows that it would constrain the theory only weakly. The expected signal in this waveband lies below 20 CUs in even the most optimistic scenario where signal strength is highest and absorption is weakest. In the nominal ‘best-fit’ scenario this drops to less than 7 CUs. As noted already (section 7.3), the low intensity of the background light from decaying neutrinos in galactic halos (as compared to that from the galaxies themselves) is due primarily to the long neutrino decay lifetime. In order to place significant constraints on the theory, one needs the stronger signal which comes from free-streaming, as well as bound neutrinos. This, in turn, requires limits on the intensity of the total background rather than that associated with fluctuations.

The curves in the upper half of figure 7.4 (with peak intensities of about 300, 700 and 2000 CUs at $\lambda_0 \approx 900$ Å) represent the combined EBL contributions from all decaying neutrinos. These are obtained by letting f_h and f_f take their minimum, nominal and maximum values respectively in (7.29), with the reverse order applying to f_τ as well as h_0 (the latter change being due to the fact that the dominant free-streaming contribution goes as h_0^{-1} rather than h_0^2). Simultaneously we have adopted the maximum-, intermediate- and minimum-absorption models for intergalactic dust, as before. Intensity is greatly reduced in the maximum-absorption case (unbroken line): by 11% at 900 Å, 53% at 1400 Å and 86% at 1900 Å. The bulk of this reduction is due to dust, especially at longer wavelengths where most of the light originates at high redshifts. Comparable reduction factors in the intermediate-absorption case (short-dashed line) are 9% at 900 Å, 28% at 1400 Å and 45% at 1900 Å. In the minimum-absorption case (long-dashed line), dust becomes less important than gas at shorter wavelengths and the intensity is reduced by a total of 9% at 900 Å, 21% at 1400 Å and 31% at 1900 Å. These three curves cover the full range of theoretical EBL intensities in the context of the decaying-neutrino scenario. For comparison we show also the intensity that would be observed in the maximum-intensity case if there were no absorption by intergalactic dust or neutral hydrogen at all (faint dotted line).

The most conservative constraints on the theory are obtained by comparing the minimum-intensity theoretical EBL contributions (unbroken line) with observational upper limits on total EBL intensity. Most of the limits which have been reported to date over the FUV waveband are represented in figure 7.4, and we describe these briefly here. They can be usefully divided into two groups: those above and below the Lyman α -line at 1216 Å. At the longest wavelengths are two

more points from the analysis of OAO-2 satellite data by Lillie and Witt ([47]; labelled ‘LW76’ in figure 7.4) which we have already encountered in chapter 3. Nearby is an upper limit from the Russian Prognoz satellite by Zvereva *et al* ([48]; ‘Z82’). Considerably stronger broadband limits have come from rocket experiments by Paresce *et al* ([49]; ‘P79’), Anderson *et al* ([50]; ‘A79’) and Feldman *et al* ([51]; ‘Fe81’), as well as an analysis of data from the Solrad-11 spacecraft by Weller ([52]; ‘We83’).

A number of studies have proceeded by establishing a correlation between background intensity and the column density of neutral hydrogen inside the Milky Way, and then extrapolating this out to zero column density to obtain the presumed extragalactic component. Martin *et al* [53] applied this method to data taken by the Berkeley UVX experiment, setting an upper limit of 110 CUs on the intensity of any unidentified EBL contributions over 1400–1900 Å (‘Ma91’). The correlation method is subject to uncertainties involving the true extent of scattering by dust, as well as absorption by ionized and molecular hydrogen at high galactic latitudes. Henry [22] and Henry and Murthy [55] approach these issues differently and raise the upper limit on background intensity to 400 CUs over 1216–3200 Å. A good indication of the complexity of the problem is found at 1500 Å, where Fix *et al* [56] used data from the DE-1 satellite to identify an isotropic background flux of 530 ± 80 CUs (‘Fi89’), the highest value reported so far. The same data were subsequently reanalysed by Wright [54] who found a much lower best-fit value of 45 CUs, with a conservative upper limit of 500 CUs (‘Wr92’). The former would rule out the decaying-neutrino hypothesis, while the latter does not constrain it at all. A third treatment of the same data has led to an intermediate result of 300 ± 80 CUs ([57]; ‘WP94’).

Limits on the FUV background shortward of $\text{Ly}\alpha$ have been even more controversial. Several studies have been based on data from the Voyager 2 ultraviolet spectrograph, beginning with that of Holberg [58], who obtained limits between 100 and 200 CUs over 500–1100 Å (labelled ‘H86’ in figure 7.4). A reanalysis of the data over 912–1100 Å by Murthy *et al* [59] led to similar numbers. In a subsequent reanalysis, however, Murthy *et al* [60] tightened this bound to 30 CUs over the same waveband (‘Mu99’). The statistical validity of these results has been vigorously debated [61, 62], with a second group asserting that the original data do not justify a limit smaller than 570 CUs (‘E00’). Of these Voyager-based limits, the strongest (‘Mu99’) is wholly incompatible with the decaying-neutrino hypothesis, while the weakest (‘E00’) puts only a modest constraint on the theory. Two new experiments have yielded results midway between these extremes: the DUVE orbital spectrometer [63] and the EURD spectrograph aboard the Spanish MINISAT 01 [64]. Upper limits on continuum emission from the former instrument are 310 CUs over 980–1020 Å and 440 CUs over 1030–1060 Å (‘K98’), while the latter has produced upper bounds of 280 CUs at 920 Å and 450 CUs at 1000 Å (‘E01’).

What do these observational data imply for the decaying-neutrino hypothesis? Longward of $\text{Ly}\alpha$, figure 7.4 shows that they span very nearly the

same parameter space as the minimum and maximum-intensity predictions of the theory (unbroken and long-dashed lines). Most stringent are Weller's Solrad-11 result ('We83') and the correlation-method constraint of Martin *et al* ('Ma91'). Taken on their own, these data constrain the decaying-neutrino hypothesis rather severely but do not rule it out. Absorption (by dust in particular) plays a critical role in reducing the strength of the signal.

Shortward of $\text{Ly}\alpha$, most of the signal originates nearby and intergalactic absorption is far less important. Ambiguity here comes rather from the spread in reported limits which, in turn, reflects the formidable experimental challenges in this part of the spectrum. Nevertheless it is clear that both the Voyager-based limits of Holberg ('H86') and Murthy *et al* ('Mu99'), as well as the new EURD measurement at 920 \AA ('E01') are incompatible with the theory. These upper bounds are violated by the weakest predicted signal, which assumes the strongest possible extinction (unbroken line). The easiest way to reconcile theory with observation is to increase the neutrino decay lifetime. If we require that $I_{\text{th}} < I_{\text{obs}}$, then the previously mentioned EURD measurement ('E01') implies a lower bound of $\tau_\nu > 3 \times 10^{23} \text{ s}$. This rises to $(5 \pm 3) \times 10^{23} \text{ s}$ and $(26 \pm 10) \times 10^{23} \text{ s}$ for the Voyager limits ('H86' and 'Mu99' respectively). All these numbers lie outside the range of lifetimes required in the decaying-neutrino scenario, $\tau_\nu = (2 \pm 1) \times 10^{23} \text{ s}$. The DUVE constraint ('K98') is more forgiving but still pushes the theory to the edge of its available parameter space. *Taken together, these data may safely be said to exclude the decaying-neutrino hypothesis.* This conclusion is in accord with current thinking on the value of Hubble's constant (section 4.3) and structure formation (section 4.5), as well as more detailed analysis of the EURD data [65].

These limits would be weakened (by a factor of up to nearly one-third) if the value of Hubble's constant h_0 were allowed to exceed 0.57 ± 0.01 , since the dominant free-streaming contributions to $I_\lambda(\lambda_0)$ go as h_0^{-1} . (This would be only partly offset by the fact that the bound ones are proportional to h_0^2 .) A higher expansion rate would, however, exacerbate problems with structure formation and the age of the Universe, the more so because the dark matter in this theory is hot. It would also mean sacrificing the critical density of neutrinos. Another possibility would be to consider neutrinos of lower rest mass. This would, however, entail a proportionate drop in the energy of the decay photons. The latter would, in fact, have to drop below the Lyman or hydrogen-ionizing limit, thus removing the whole motivation for the proposed neutrinos in the first place. Similar considerations apply to neutrinos with longer decay lifetimes.

Our conclusions, then, are as follows. Neutrinos with rest masses and decay lifetimes as specified by the decaying-neutrino scenario produce levels of ultraviolet background radiation very close to and, in several cases, above experimental upper limits on the intensity of the EBL. At wavelengths longer than 1200 \AA , where intergalactic absorption is most effective, the theory is marginally compatible with observation—if one adopts the upper limits on dust density consistent with quasar obscuration; and *if* the dust grains are extremely small. At

wavelengths in the range 900–1200 Å, predicted intensities are either comparable to or higher than those actually seen. Thus, while there is now good experimental evidence that some of the dark matter is provided by massive neutrinos, the light of the night sky tells us that these particles cannot have the rest masses and decay lifetimes attributed to them in the decaying-neutrino hypothesis.

References

- [1] Cowsik R 1977 *Phys. Rev. Lett.* **39** 784
- [2] de Rujula A and Glashow S L 1980 *Phys. Rev. Lett.* **45** 942
- [3] Pal P B and Wolfenstein L 1982 *Phys. Rev. D* **25** 766
- [4] Bowyer S *et al* 1995 *Phys. Rev. D* **52** 3214
- [5] Melott A L and Sciama D W 1981 *Phys. Rev. Lett.* **46** 1369
- [6] Sciama D W and Melott A L 1982 *Phys. Rev. D* **25** 2214
- [7] Melott A L, McKay D W and Ralston J P 1988 *Astrophys. J.* **324** L43
- [8] Sciama D W 1990 *Astrophys. J.* **364** 549
- [9] Sciama D W 1993 *Modern Cosmology and the Dark Matter Problem* (Cambridge: Cambridge University Press)
- [10] Sciama D W 1997 *Mon. Not. R. Astron. Soc.* **289** 945
- [11] Fabian A, Naylor T and Sciama D 1991 *Mon. Not. R. Astron. Soc.* **249** 21
- [12] Davidsen A F *et al* 1991 *Nature* **351** 128
- [13] Sciama D W 1993 *Pub. Astron. Soc. Pac.* **105** 102
- [14] Stecker F W 1980 *Phys. Rev. Lett.* **45** 1460
- [15] Kimble R, Bowyer S and Jakobsen P 1981 *Phys. Rev. Lett.* **46** 80
- [16] Sciama D W 1991 *The Early Observable Universe from Diffuse Backgrounds* ed B Rocca-Volmerange, J M Deharveng and J Van Tran Thanh (Paris: Edition Frontières) p 127
- [17] Overduin J M, Wesson P S and Bowyer S 1993 *Astrophys. J.* **404** 460
- [18] Dodelson S and Jubas J M 1994 *Mon. Not. R. Astron. Soc.* **266** 886
- [19] Overduin J M and Wesson P S 1997 *Astrophys. J.* **483** 77
- [20] Overduin J M, Seahra S S, Duley W W and Wesson P S 1999 *Astron. Astrophys.* **349** 317
- [21] Bowyer S 1991 *Ann. Rev. Astron. Astrophys.* **29** 59
- [22] Henry R C 1991 *Ann. Rev. Astron. Astrophys.* **29** 89
- [23] Salucci P and Sciama D W 1990 *Mon. Not. R. Astron. Soc.* **244** 9P
- [24] Gates E I, Gyuk G and Turner M S 1995 *Astrophys. J.* **449** L123
- [25] Kochanek C S 1996 *Astrophys. J.* **457** 228
- [26] Zuo L and Phinney E S 1993 *Astrophys. J.* **418** 28
- [27] Ostriker J P and Cowie L L 1981 *Astrophys. J.* **243** L127
- [28] Wright E L 1981 *Astrophys. J.* **250** 1
- [29] Ostriker J P and Heisler J 1984 *Astrophys. J.* **278** 1
- [30] Wright E L 1986 *Astrophys. J.* **311** 156
- [31] Wright E L and Malkan M A 1987 *Bull. Am. Astron. Soc.* **19** 699
- [32] Heisler J and Ostriker J P 1988 *Astrophys. J.* **332** 543
- [33] Wright E L 1990 *Astrophys. J.* **353** 413
- [34] Fall S M and Pei Y C 1993 *Astrophys. J.* **402** 479
- [35] Mathis J S 1990 *Ann. Rev. Astron. Astrophys.* **28** 37

- [36] Martin P G and Rouleau F 1991 *Extreme Ultraviolet Astronomy* ed R F Malina and S Bowyer (New York: Pergamon) p 341
- [37] Draine B T and Lee H M 1984 *Astrophys. J.* **285** 89
- [38] Meyer J-P 1979 *Les Elements et leurs Isotopes dan L'univers* (Université de Liège) p 153
- [39] Snow T P and Witt A N 1996 *Astrophys. J.* **468** L65
- [40] Draine B T 1995 <http://www.astro.princeton.edu/~draine/dust/>
- [41] Mathis J S 1996 *Astrophys. J.* **472** 643
- [42] Duley W W and Seahra S 1998 *Astrophys. J.* **507** 874
- [43] Duley W W and Seahra S S 1999 *Astrophys. J.* **522** L129
- [44] Seahra S S and Duley W W 1999 *Astrophys. J.* **520** 719
- [45] Martin C and Bowyer S 1989 *Astrophys. J.* **338** 677
- [46] Sassee T P, Lampton M, Bowyer S and Wu X 1995 *Astrophys. J.* **447** 630
- [47] Lillie C F and Witt A N 1976 *Astrophys. J.* **208** 64
- [48] Zvereva A M *et al* 1982 *Astron. Astrophys.* **116** 312
- [49] Paresce F, Margon B, Bowyer S and Lampton M 1979 *Astrophys. J.* **230** 304
- [50] Anderson R C *et al* 1979 *Astrophys. J.* **234** 415
- [51] Feldman P D, Brune W H and Henry R C 1991 *Astrophys. J.* **249** L51
- [52] Weller C S 1983 *Astrophys. J.* **268** 899
- [53] Martin C, Hurwitz M and Bowyer S 1991 *Astrophys. J.* **379** 549
- [54] Wright E L 1992 *Astrophys. J.* **391** 34
- [55] Henry R C and Murthy J 1993 *Astrophys. J.* **418** L17
- [56] Fix J D, Craven J D and Frank L A 1989 *Astrophys. J.* **345** 203
- [57] Witt A N and Petersohn J K 1994 *The First Symposium on the Infrared Cirrus and Diffuse Interstellar Clouds* (Astronomical Society of the Pacific Conference Series, Volume 58) ed R M Cutri and W B Latter (San Francisco, CA: ASP) p 91
- [58] Holberg J B 1986 *Astrophys. J.* **311** 969
- [59] Murthy J, Henry R C and Holberg J B 1991 *Astrophys. J.* **383** 198
- [60] Murthy J *et al* 1999 *Astrophys. J.* **522** 904
- [61] Edelstein J, Bowyer S and Lampton M 2000 *Astrophys. J.* **539** 187
- [62] Murthy J *et al* 2001 *Astrophys. J.* **557** L47
- [63] Korpela E J, Bowyer S and Edelstein J 1998 *Astrophys. J.* **495** 317
- [64] Edelstein J *et al* 2001 *Astrophys. Space Sci.* **276** 177
- [65] Bowyer S *et al* 1999 *Astrophys. J.* **526** 10

Chapter 8

Supersymmetric weakly interacting particles

8.1 The lightest supersymmetric particle

Weakly interacting massive particles (WIMPs) are hypothetical particles whose interaction strengths are comparable to those of neutrinos, but whose rest masses exceed those of the baryons. The most widely-discussed examples arise in the context of *supersymmetry* (SUSY), which is motivated quite independently of the dark-matter problem as a theoretical framework for many attempts to unify the forces of nature. SUSY predicts that, for every known fermion in the standard model, there exists a new bosonic ‘superpartner’ and vice versa (more than doubling the number of fundamental degrees of freedom in the simplest models; see [1] for a review). These superpartners were recognized as potential dark-matter candidates in the early 1980s by Cabibbo *et al* [2], Pagels and Primack [3], Weinberg [4] and others [5–9], with the generic term ‘WIMP’ being coined shortly thereafter [10].

There is, as yet, no firm experimental evidence for SUSY WIMPs. This means that their rest energies, if they exist, lie beyond the range currently probed by accelerators (and, in particular, beyond the rest energies of their standard-model counterparts). Supersymmetry is, therefore, not an exact symmetry of nature. The masses of the superpartners, like that of the axion (chapter 6), are thought to have been generated by a symmetry-breaking event in the early Universe. Subsequently, as the temperature of the expanding fireball dropped below their rest energies, heavier species would have dropped out of equilibrium and begun to disappear by pair annihilation, leaving progressively lighter ones behind. Eventually, only one massive superpartner would have remained: the **lightest supersymmetric particle** (LSP). It is this particle which plays the role of the WIMP in SUSY theories. Calculations using the Boltzmann equation show that the collective density of relic LSPs today lies within one or two orders of magnitude of the suspected cold dark-matter (CDM) density across much of the parameter space of the simplest SUSY theories [11].

There are at least three ways in which SUSY WIMPs could contribute to the extragalactic background light (EBL). The first is by **pair annihilation** to photons. This occurs in even the simplest or minimal SUSY model (MSSM) but it is a very slow process because it takes place via intermediate loops of charged particles such as leptons and quarks and their antiparticles. Processes involving intermediate bound states contribute as well, but these are also suppressed by loop diagrams and are equally suppressed. (Each vertex in such a diagram weakens the interaction by one factor of the coupling parameter.) The reason for the stability of the LSP in the MSSM is an additional new symmetry of nature, known as **R-parity**, which is necessary (among other things) to protect the proton from decaying via intermediate SUSY states. There are also *non*-minimal SUSY theories which do not conserve R-parity (and in which the proton can decay). In these theories, LSPs can contribute to the EBL in two more ways: by direct decay into photons via **loop diagrams**; and also indirectly via **tree-level decays** to secondary particles which then scatter off pre-existing background photons to produce a signal.

In this chapter, we will estimate the strength of EBL contributions from all three kinds of processes (annihilations, one-loop and tree-level decays). The strongest signals come from decaying LSPs in non-minimal SUSY theories, and we will find that these particles are strongly constrained by data on background radiation in the x-ray and γ -ray bands. While our results depend on a number of theoretical parameters, they imply in general that the LSP, if it decays, must do so on timescales which are considerably longer than the lifetime of the Universe.

8.2 Neutralinos

The first step is to choose an LSP from the lineup of SUSY candidate particles. Early workers variously identified this as the **photino** $\tilde{\gamma}$ [2], the **gravitino** \tilde{g} [3], the **sneutrino** $\tilde{\nu}$ [7] or the **selectron** \tilde{e} [8]. (SUSY superpartners are denoted by a tilde and take the same names as their standard-model counterparts, with a prefix ‘s’ for superpartners of fermions and a suffix ‘ino’ for those of bosons.) Many of these contenders were eliminated by Ellis *et al*, who showed in 1984 that the LSP is, in fact, most likely to be a **neutralino** $\tilde{\chi}$, a linear superposition of the photino $\tilde{\gamma}$, the zino \tilde{Z} and two neutral higgsinos \tilde{h}_1^0 and \tilde{h}_2^0 [9]. (These are the SUSY spin- $\frac{1}{2}$ counterparts of the photon, Z^0 and Higgs bosons respectively.) There are four neutralinos, each a mass eigenstate made up of (in general) different amounts of photino, zino, etc, although in special cases a neutralino could be ‘mostly photino’, say, or ‘pure zino’. The LSP is, by definition, the lightest such eigenstate. Accelerator searches place a lower limit on its rest energy which currently stands at $m_{\tilde{\chi}}c^2 > 46$ GeV [12].

In minimal SUSY, the density of neutralinos drops only by way of the (slow) pair-annihilation process, and it is quite possible for these particles to ‘overclose’ the Universe if their rest energy is too high. This does not change the geometry of

the Universe but rather speeds up its expansion rate, whose square is proportional to the total matter density from equation (2.33). In such a situation, the Universe would have reached its present size in too short a time. Lower bounds on the age of the Universe thus impose an upper bound on the neutralino rest energy which has been set at $m_{\tilde{\chi}}c^2 \lesssim 3200$ GeV [13]. Detailed exploration of the parameter space of minimal SUSY theory shows that this upper limit tightens in most cases to $m_{\tilde{\chi}}c^2 \lesssim 600$ GeV [14]. Much recent work is focused on a slimmed-down version of the MSSM known as the constrained minimal SUSY model (CMSSM), in which all existing experimental bounds and cosmological requirements are comfortably met by neutralinos with rest energies in the range $90 \text{ GeV} \lesssim m_{\tilde{\chi}}c^2 \lesssim 400 \text{ GeV}$ [15].

Even in its constrained minimal version, SUSY theory contains at least five adjustable input parameters, making the neutralino a considerably harder proposition to test than the axion or the massive neutrino. Fortunately, there are several other ways (besides accelerator searches) to look for these particles. Because their rest energies are above the temperature at which they decoupled from the primordial fireball, WIMPs have non-relativistic velocities and ought to be found predominantly in gravitational potential wells like those of our own galaxy. They will occasionally scatter against target nuclei in terrestrial laboratories as the Earth follows the Sun around the Milky Way. Efforts to detect such scattering events are hampered by both low event rates and the relatively small amounts of energy deposited. But they could be given away by the directional dependence of nuclear recoils, or by the seasonal modulation of the event rate due to the Earth's motion around the Sun. Both effects are now the basis for **direct detection** experiments around the world [16]. Great excitement followed the announcement by the DAMA team in 2000 of a WIMP signature with rest energy $m_{\tilde{\chi}}c^2 = 59^{+17}_{-14}$ GeV [17]. Two other groups (CDMS [18] and EDELWEISS [19]) have, however, been unable to reproduce this result; and it is generally regarded as a false start but one which is stimulating a great deal of productive follow-up work.

A second, **indirect search** strategy is to look for annihilation byproducts from neutralinos which have collected inside massive bodies. Most attention has been directed at the possibility of detecting antiprotons from the galactic halo [20] or neutrinos from the Sun [21] or Earth [22]. The heat generated in the cores of gas giants like Jupiter or Uranus has also been considered as a potential annihilation signature [23]. The main challenge in each case lies in separating the signal from the background noise. Neutrinos from WIMP annihilations in the Sun or Earth provide perhaps the best chance for a detection because backgrounds for these cases are small and relatively well understood. In the case of the Earth, one looks for a flux of neutrino-induced muons which can be distinguished from the atmospheric background by the fact that they are travelling *straight up*. The AMANDA team, whose detectors are buried deep in Antarctic ice, is one of several which have reported preliminary results in experiments of this kind [24]. After various filters were applied to the first year's worth of raw data, 15 candidate

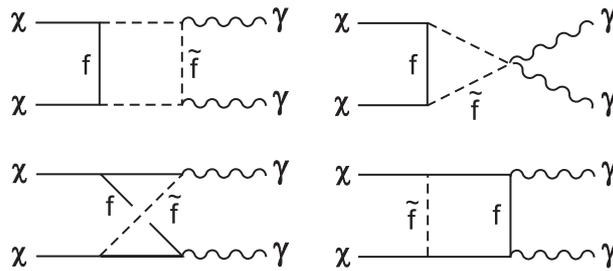


Figure 8.1. Some Feynman diagrams corresponding to the annihilation of two neutralinos ($\tilde{\chi}$), producing a pair of photons (γ). The process is mediated by fermions (f) and their supersymmetric counterparts, the sfermions (\tilde{f}).

events remained. This was consistent with the expected statistical background of 17 events, leaving only an upper limit on the size of a possible WIMP population inside the Earth at present.

8.3 Pair annihilation

Pair annihilation into *photons* (figure 8.1) provides a complementary indirect search technique. For the range of neutralino rest energies under consideration here ($50 \text{ GeV} \lesssim m_{\tilde{\chi}} c^2 \lesssim 1000 \text{ GeV}$), the photons so produced lie in the γ -ray portion of the spectrum. Many workers, beginning with Sciama [8] and Silk and Srednicki [20], have studied the possibility of γ -rays from SUSY WIMP annihilations in the halo of the Milky Way, which gives the strongest signal. Prognoses for detection have ranged from very optimistic [25] to very pessimistic [26]; converging gradually to the conclusion that neutralino-annihilation contributions would be at or somewhat below the level of the galactic background, and possibly distinguishable from it by their spectral shape [27–29]. Uncertainties are considerable, arising not only from weak signal-to-noise levels but also from factors such as the assumed distribution of WIMPs in the halo and scattering processes along the line of sight. Recent studies have focused on possible enhancements of the signal such as might occur in the presence of a high-density core [30], a flattened halo [31], a very extended singular halo [32] or a massive black hole at the galactic centre [33]. It has also been pointed out that the large effective area of atmospheric Čerenkov detectors could compensate for the weaker signal from neutralinos at the upper end of the mass range [34]. This, however, appears unlikely to raise the expected signal above the galactic background [35]. Other authors have carried the search away from the halo of the Milky Way toward systems such as dwarf spheroidal galaxies [36], the Large Magellanic Cloud [37] and the giant elliptical galaxy M87 in Virgo [38].

The possibility of neutralino-annihilation contributions to the *diffuse extragalactic* background, rather than the signal from localized concentrations

of dark matter, has received far less attention. First to apply the problem to SUSY WIMPs were Cabibbo *et al* [2] who assumed a WIMP rest energy (10–30 eV) which we now know is far too low. Like the decaying neutrino (chapter 7), this would have produced a background in the ultraviolet. It is excluded, however, by an argument (sometimes known as the Lee–Weinberg bound) which restricts WIMPs to rest energies above 2 GeV [6]. EBL contributions from SUSY WIMPs in this range were first estimated by Silk and Srednicki [20]. Their conclusion, and those of most workers who have followed them [39–41], is that neutralino annihilations would be responsible for no more than a small fraction of the observed γ -ray background. Here we review this argument, reversing our usual procedure and attempting to set a reasonably conservative *upper* limit on neutralino contributions to the EBL.

As usual, we take for our sources of background radiation the galactic dark-matter halos with comoving number density n_0 . Each neutralino pair annihilates into two monochromatic photons of energy $E_\gamma \approx m_{\tilde{\chi}} c^2$ [29]. We model this with the Gaussian SED (3.19)

$$F(\lambda) = \frac{L_{\text{h,ann}}}{\sqrt{2\pi}\sigma_\lambda} \exp\left[-\frac{1}{2}\left(\frac{\lambda - \lambda_{\text{ann}}}{\sigma_\lambda}\right)^2\right]. \quad (8.1)$$

Photon wavelengths are distributed normally about the peak wavelength:

$$\lambda_{\text{ann}} = hc/m_{\tilde{\chi}}c^2 = (1.2 \times 10^{-6} \text{ \AA})m_{10}^{-1} \quad (8.2)$$

where $m_{10} \equiv m_{\tilde{\chi}}c^2/(10 \text{ GeV})$ is the neutralino rest energy in units of 10 GeV. The standard deviation σ_γ can be related to the velocity dispersion of bound dark-matter particles as in previous chapters, so that $\sigma_\lambda = 2(v_c/c)\lambda_{\text{ann}}$. With $v_c \sim 220 \text{ km s}^{-1}$ and $m_{10} \sim 1$ this is of order $\sim 10^{-9} \text{ \AA}$. For convenience we specify this from now on in terms of a dimensionless parameter $\sigma_9 \equiv \sigma_\lambda/(10^{-9} \text{ \AA})$.

The luminosity $L_{\text{h,ann}}$ of the halo is given by

$$L_{\text{h,ann}} = 2m_{\tilde{\chi}}c^2 \langle \sigma v \rangle_{\gamma\gamma} \mathcal{P} \quad (8.3)$$

where $\langle \sigma v \rangle_{\gamma\gamma}$ is the **annihilation cross section**, $\mathcal{P} \equiv m_{\tilde{\chi}}^{-2} \int \rho_{\tilde{\chi}}^2(r) 4\pi r^2 dr$ and $\rho_{\tilde{\chi}}(r)$ is the mass density distribution of neutralinos in the halo [32]. For the cross section we turn to the particle-physics literature. Berezhinsky *et al* [42] have calculated this quantity as $\langle \sigma v \rangle_{\gamma\gamma} \approx a_{\gamma\gamma}$ (for non-relativistic neutralinos) where

$$a_{\gamma\gamma} = \frac{\hbar^2 c^3 \alpha^4 (m_{\tilde{\chi}} c^2)^2}{3^6 \pi (m_{\tilde{f}} c^2)^4} \left(\frac{Z_{11}}{\sin \theta_W} \right)^4 (45 + 48y + 139y^2)^2. \quad (8.4)$$

Here α is the fine structure constant, $m_{\tilde{f}}$ is the mass of an intermediate sfermion, $y \equiv (Z_{12}/Z_{11}) \tan \theta_W$, θ_W is the weak mixing angle and Z_{ij} are elements of the real orthogonal matrix which diagonalizes the neutralino mass matrix. In particular, the ‘pure photino’ case is specified by $Z_{11} = \sin \theta_W$, $y = 1$ and

the ‘pure zino’ by $Z_{11} = \cos \theta_W$, $y = -\tan^2 \theta_W$. Collecting these expressions together and parametrizing the sfermion rest energy by $\tilde{m}_{10} \equiv m_{\tilde{f}}/10 \text{ GeV}$, we obtain:

$$\langle \sigma v \rangle_{\gamma\gamma} = (8 \times 10^{-27} \text{ cm}^3 \text{ s}^{-1}) f_\chi m_{10}^2 \tilde{m}_{10}^{-4}. \quad (8.5)$$

Here f_χ ($= 1$ for photinos, 0.4 for zinos) is a dimensionless quantity whose value parametrizes the makeup of the neutralino.

Since we attempt in this section to set an upper limit on EBL contributions from neutralino annihilations, we take $f_\chi \approx 1$ (the photino case). In the same spirit, we would like to use lower limits for the sfermion mass \tilde{m}_{10} . It is important to estimate this quantity accurately since the cross section goes as \tilde{m}_{10}^{-4} . Giudice and Griest [43] have made a detailed study of photino annihilations and find a lower limit on \tilde{m}_{10} as a function of m_{10} , assuming that photinos provide at least $0.025 h_0^{-2}$ of the critical density. Over the range $0.1 \leq m_{10} \leq 4$, this lower limit is empirically well fit by a function of the form $\tilde{m}_{10} \approx 4 m_{10}^{0.3}$. If this holds over our broader range of masses, then we obtain an upper limit on the neutralino annihilation cross section of $\langle \sigma v \rangle_{\gamma\gamma} \lesssim (3 \times 10^{-29} \text{ cm}^3 \text{ s}^{-1}) m_{10}^{0.8}$. This expression gives results which are about an order of magnitude higher than the cross sections quoted by Gao *et al* [40] for photinos with $m_{10} \approx 1$.

For the WIMP density distribution $\rho_\chi(r)$ we adopt the simple and widely-used *isothermal model* [29]:

$$\rho_\chi(r) = \rho_\odot \left(\frac{a^2 + r_\odot^2}{a^2 + r^2} \right). \quad (8.6)$$

Here $\rho_\odot = 5 \times 10^{-25} \text{ g cm}^{-3}$ is the approximate dark-matter density in the Solar vicinity, assuming a spherical halo [44], $r_\odot = 8 \text{ kpc}$ is the distance of the Sun from the galactic centre [45] and $a = (2-20) \text{ kpc}$ is a core radius. To fix this latter parameter, we can integrate (8.6) over volume to obtain total halo mass $M_h(r)$ inside radius r :

$$M_h(r) = 4\pi\rho_\odot \int_0^r r' (a^2 + r_\odot^2) \left[1 - \left(\frac{a}{r'} \right) \tan^{-1} \left(\frac{r'}{a} \right) \right] dr'. \quad (8.7)$$

Observations of the motions of galactic satellites imply that the total mass inside 50 kpc is about $5 \times 10^{11} M_\odot$ [46]. This in (8.7) implies $a = 9 \text{ kpc}$, which we consequently adopt. The maximum extent of the halo is not well constrained observationally but can be specified if we suppose that $M_h = (2 \pm 1) \times 10^{12} M_\odot$ as in (6.15). Equation (8.7) then gives a radius $r_h = (170 \pm 80) \text{ kpc}$. The cosmological density of WIMPs in galactic dark-matter halos adds up to $\Omega_h = n_0 M_h / \rho_{\text{crit},0} = (0.07 \pm 0.04) h_0$.

If there are no other sources of CDM, then the total matter density is $\Omega_{\text{m},0} = \Omega_h + \Omega_{\text{bar}} \approx 0.1 h_0$ and the observed flatness of the Universe (section 4.6) implies a strongly vacuum-dominated cosmology. While we use this as a lower limit on WIMP contributions to the dark matter in subsequent sections, it is quite possible that CDM also exists in larger-scale regions such as galaxy clusters. To

take this into account in a general way, we define a **cosmological enhancement factor** $f_c \equiv (\Omega_{m,0} - \Omega_{\text{bar}})/\Omega_h$ representing the added contributions from WIMPs outside galactic halos (or perhaps in halos which extend far enough to fill the space between galaxies [47]). This takes the value $f_c = 1$ for the most conservative case just described but rises to $f_c = (4 \pm 2)h_0^{-1}$ in the Λ CDM model with $\Omega_{m,0} = 0.3$, and $(14 \pm 7)h_0^{-1}$ in the EdS model with $\Omega_{m,0} = 1$.

With $\rho_{\tilde{\chi}}(r)$ known, we are in a position to calculate the quantity \mathcal{P} , which becomes:

$$\mathcal{P} = \frac{2\pi\rho_{\odot}^2(a^2 + r_{\odot}^2)^2}{m_{\tilde{\chi}}^2 a} \left[\tan^{-1}\left(\frac{r_h}{a}\right) - \frac{(r_h/a)}{1 + (r_h/a)^2} \right]. \quad (8.8)$$

Using the values for ρ_{\odot} , r_{\odot} and a specified before and setting $r_h = 250$ kpc to get an upper limit, we find that $\mathcal{P} \leq (5 \times 10^{65} \text{ cm}^{-3})m_{10}^{-2}$. Putting this result together with the cross section (8.5) into (8.3), we obtain:

$$L_{h,\text{ann}} \leq (1 \times 10^{38} \text{ erg s}^{-1})f_{\chi}m_{10}\tilde{m}_{10}^{-4}. \quad (8.9)$$

Inserting Giudice and Griest's [43] lower limit on the sfermion mass \tilde{m}_{10} (as empirically fit before), we find that (8.9) gives an upper limit on halo luminosity of $L_{h,\text{ann}} \leq (5 \times 10^{35} \text{ erg s}^{-1})f_{\chi}m_{10}^{-0.2}$. Higher estimates can be found in the literature [48] but these assume a singular halo whose density drops off as only $\rho_{\tilde{\chi}}(r) \propto r^{-1.8}$ and extends out to a very large halo radius $r_h = 4.2h_0^{-1}$ Mpc. For a standard isothermal distribution of the form (8.6), our results confirm that halo luminosity due to neutralino annihilations alone is very low, amounting to less than 10^{-8} times the total bolometric luminosity of the Milky Way.

The combined bolometric intensity of neutralino annihilations between redshift z_f and the present is given by substituting the comoving number density n_0 and luminosity $L_{h,\text{ann}}$ into equation (2.20) to give

$$Q = Q_{\tilde{\chi},\text{ann}} \int_0^{z_f} \frac{dz}{(1+z)^2[\Omega_{m,0}(1+z)^3 + (1-\Omega_{m,0})]^{1/2}} \quad (8.10)$$

where $Q_{\tilde{\chi},\text{ann}} = (cn_0L_{h,\text{ann}}f_c)/H_0$ and we have assumed spatial flatness. With values for all these parameters as specified earlier, we find

$$Q = \begin{cases} (1 \times 10^{-12} \text{ erg s}^{-1} \text{ cm}^{-2})h_0^2 f_{\chi} m_{10}^{-0.2} & (\Omega_{m,0} = 0.1h_0) \\ (3 \times 10^{-12} \text{ erg s}^{-1} \text{ cm}^{-2})h_0 f_{\chi} m_{10}^{-0.2} & (\Omega_{m,0} = 0.3) \\ (1 \times 10^{-11} \text{ erg s}^{-1} \text{ cm}^{-2})h_0 f_{\chi} m_{10}^{-0.2} & (\Omega_{m,0} = 1). \end{cases} \quad (8.11)$$

Here we have set $z_f = 30$ (larger values do not substantially increase the value of Q) and used values of $f_c = 1$, $4h_0^{-1}$ and $20h_0^{-1}$ respectively. The effects of the cosmological enhancement factor f_c are partially offset in (8.10) by the fact that a Universe with higher matter density $\Omega_{m,0}$ is younger, and hence contains less background light in general. Even the highest value of Q given in (8.11) is

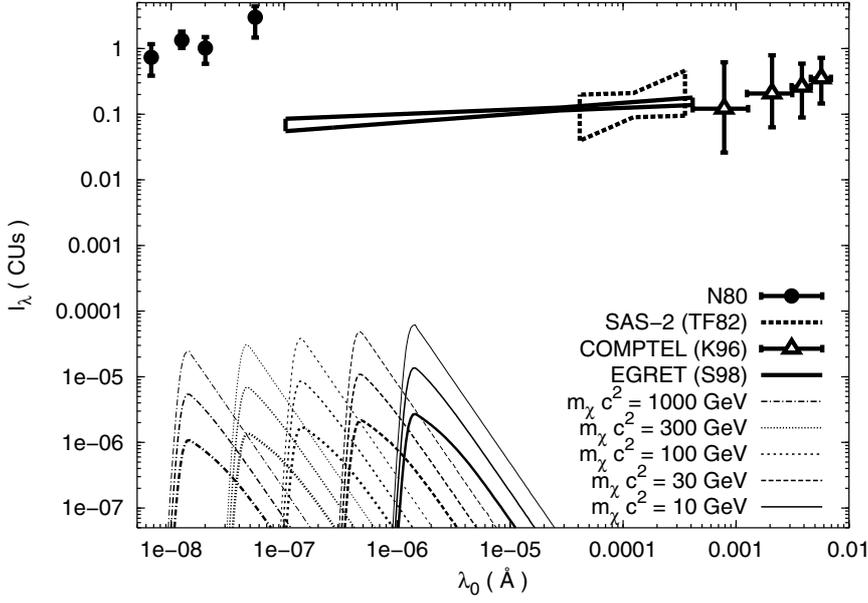


Figure 8.2. The spectral intensity of the diffuse γ -ray background due to neutralino annihilations (lower left), compared with observational limits from high-altitude balloon experiments (N80), as the SAS-2 spacecraft and the COMPTEL and EGRET instruments aboard the Compton Gamma-Ray Observatory. The three plotted curves for each value of $m_{\tilde{\chi}} c^2$ depend on the total density of neutralinos: galaxy halos only ($\Omega_{m,0} = 0.1h_0$; bold lines), Λ CDM model ($\Omega_{m,0} = 0.3$; medium lines) or EdS model ($\Omega_{m,0} = 1$; light lines).

negligible in comparison to the approximate intensity (2.25) of the EBL due to ordinary galaxies.

The total spectral intensity of annihilating neutralinos is found by substituting the SED (8.1) into (3.6) to give

$$I_{\lambda}(\lambda_0) = I_{\tilde{\chi},\text{ann}} \int_0^{z_f} \frac{\exp\left\{-\frac{1}{2}\left[\frac{\lambda_0/(1+z) - \lambda_{\text{ann}}}{\sigma_{\lambda}}\right]^2\right\}}{(1+z)^3[\Omega_{m,0}(1+z)^3 + 1 - \Omega_{m,0}]^{1/2}} dz. \quad (8.12)$$

For a typical neutralino with $m_{10} \approx 10$ the annihilation spectrum peaks near $\lambda_0 \approx 10^{-7}$ Å. The dimensional prefactor reads as

$$\begin{aligned} I_{\tilde{\chi},\text{ann}} &= \frac{n_0 L_{\text{h,ann}} f_c}{\sqrt{32\pi^3} h H_0} \left(\frac{\lambda_0}{\sigma_{\lambda}}\right) \\ &= (0.0002 \text{ CUs}) h_0^2 f_{\chi} f_c m_{10}^{0.8} \sigma_9^{-1} \left(\frac{\lambda_0}{10^{-7} \text{ \AA}}\right). \end{aligned} \quad (8.13)$$

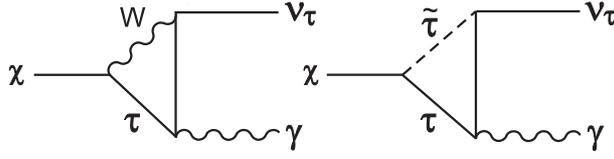


Figure 8.3. Feynman diagrams corresponding to one-loop decays of the neutralino ($\tilde{\chi}$) into a neutrino (here ν_τ) and a photon (γ). The process can be mediated by the W boson and a τ muon, or by the τ and its supersymmetric counterpart ($\tilde{\tau}$).

Here we have divided through by the photon energy hc/λ_0 to put results into continuum units or CUs as usual (section 3.3). Equation (8.12) gives the combined intensity of radiation from neutralino annihilations, emitted at various wavelengths and redshifted by various amounts but observed at wavelength λ_0 . Results are plotted in figure 8.2 together with observational constraints. We defer discussion of this plot (and the data) to section 8.7, where they are compared with the other WIMP results for this chapter.

8.4 One-loop decays

Figure 8.2 confirms that neutralino annihilations do not produce enough background radiation to lead to useful constraints on these particles. We therefore turn in the rest of this chapter to *non*-minimal SUSY theories in which R-parity is not necessarily conserved and the LSP (in this case the neutralino) can decay. The cosmological consequences of R-parity breaking have been reviewed by Bouquet and Salati [49]. There is only one *direct* decay mode into photons:

$$\tilde{\chi} \rightarrow \nu + \gamma. \quad (8.14)$$

Feynman diagrams for this process are shown in figure 8.3. Because these decays occur via loop diagrams, they are again subdominant. We consider theories in which R-parity breaking occurs spontaneously. This introduces a scalar sneutrino with a non-zero vacuum expectation value $v_R \equiv \langle \tilde{\nu}_{\tau_R} \rangle$, as discussed by Masiero and Valle [50]. Neutralino decays into photons could be detectable if $m_{\tilde{\chi}}$ and v_R are large [51].

The photons produced in this way are again monochromatic, with $E_\gamma = \frac{1}{2}m_{\tilde{\chi}}c^2$. In fact the SED here is the same as (8.1) except that the peak wavelength is doubled, $\lambda_{\text{loop}} = 2hc/m_{\tilde{\chi}}c^2 = (2.5 \times 10^{-6} \text{ \AA}) m_{10}^{-1}$. The only parameter that needs to be recalculated is the halo luminosity L_h . For one-loop neutralino decays of lifetime $\tau_{\tilde{\chi}}$, this takes the form:

$$L_{h,\text{loop}} = \frac{N_{\tilde{\chi}} b_\gamma E_\gamma}{\tau_{\tilde{\chi}}} = \frac{b_\gamma M_h c^2}{2\tau_{\tilde{\chi}}}. \quad (8.15)$$

Here $N_{\tilde{\chi}} = M_h/m_{\tilde{\chi}}$ is the number of neutralinos in the halo and b_γ is the **branching ratio** or fraction of neutralinos that decay into photons. This is estimated by Berezhinsky *et al* [51] as

$$b_\gamma \approx 10^{-9} f_R^2 m_{10}^2 \quad (8.16)$$

where the new parameter $f_R \equiv v_R/(100 \text{ GeV})$. The requirement that SUSY WIMPs do not carry too much energy out of stellar cores implies that f_R is of order ten or more [50]. We take $f_R > 1$ as a lower limit.

For halo mass we adopt $M_h = (2 \pm 1) \times 10^{12} M_\odot$ as usual, with $r_h = (170 \pm 80) \text{ kpc}$ from the discussion following (8.7). As in the previous section, we parametrize our lack of certainty about the distribution of neutralinos on larger scales with the cosmological enhancement factor f_c . Collecting these results together and expressing the decay lifetime in dimensionless form as $f_\tau \equiv \tau_{\tilde{\chi}}/(1 \text{ Gyr})$, we obtain for the luminosity of one-loop neutralino decays in the halo:

$$L_{\text{h,loop}} = (6 \times 10^{40} \text{ erg s}^{-1}) m_{10}^2 f_R^2 f_\tau^{-1}. \quad (8.17)$$

With $m_{10} \sim f_R \sim f_\tau \sim 1$, equation (8.17) gives $L_{\text{h,loop}} \sim 2 \times 10^7 L_\odot$. This is considerably brighter than the halo luminosity due to neutralino annihilations in minimal SUSY models but still amounts to less than 10^{-3} times the bolometric luminosity of the Milky Way.

The combined bolometric intensity is found as in the previous section, but with $L_{\text{h,ann}}$ in (8.10) replaced by $L_{\text{h,loop}}$ so that

$$Q = \begin{cases} (1 \times 10^{-7} \text{ erg s}^{-1} \text{ cm}^{-2}) h_0^2 m_{10}^2 f_R^2 f_\tau^{-1} & (\Omega_{\text{m},0} = 0.1 h_0) \\ (4 \times 10^{-7} \text{ erg s}^{-1} \text{ cm}^{-2}) h_0 m_{10}^2 f_R^2 f_\tau^{-1} & (\Omega_{\text{m},0} = 0.3) \\ (2 \times 10^{-6} \text{ erg s}^{-1} \text{ cm}^{-2}) h_0 m_{10}^2 f_R^2 f_\tau^{-1} & (\Omega_{\text{m},0} = 1). \end{cases} \quad (8.18)$$

This is again small. However, we see that massive ($m_{10} \gtrsim 10$) neutralinos which provide close to the critical density ($\Omega_{\text{m},0} \sim 1$) and decay on timescales of order 1 Gyr or less ($f_\tau \lesssim 1$) could, in principle, rival the intensity of the conventional EBL.

To obtain more quantitative constraints, we turn to spectral intensity. This is given by equation (8.12) as before, except that the dimensional prefactor $I_{\tilde{\chi},\text{ann}}$ must be replaced by

$$\begin{aligned} I_{\tilde{\chi},\text{loop}} &= \frac{n_0 L_{\text{h,loop}} f_c}{\sqrt{32\pi^3} h H_0} \left(\frac{\lambda_0}{\sigma_\lambda} \right) \\ &= (30 \text{ CUs}) h_0^2 m_{10}^3 f_R^2 f_\tau^{-1} f_c \sigma_9^{-1} \left(\frac{\lambda_0}{10^{-7} \text{ \AA}} \right). \end{aligned} \quad (8.19)$$

Results are plotted in figure 8.4 for neutralino rest energies $1 \leq m_{10} \leq 100$. While their bolometric intensity is low, these particles are capable of significant EBL contributions in narrow portions of the γ -ray background. To keep the

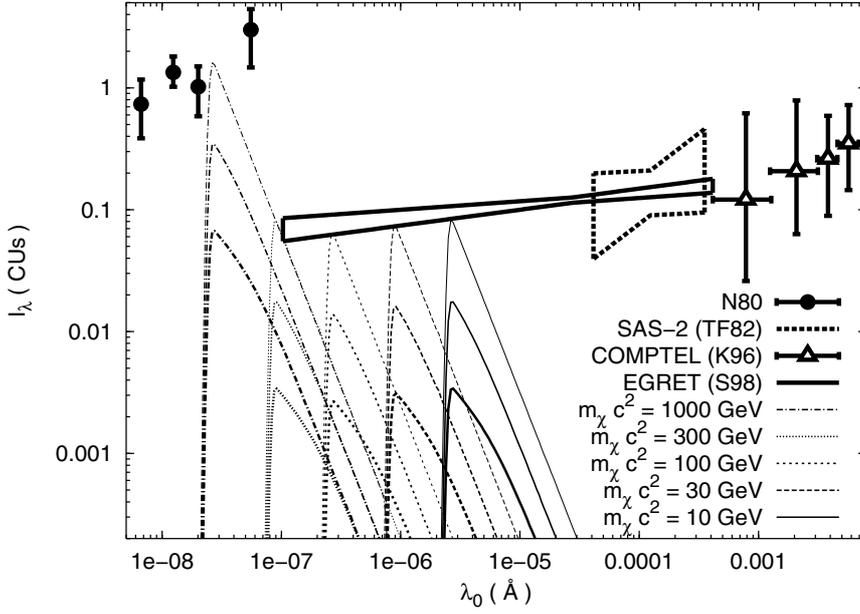


Figure 8.4. The spectral intensity of the diffuse γ -ray background due to neutralino one-loop decays (lower left), compared with observational upper limits from high-altitude balloon experiments (filled dots), SAS-2, EGRET and COMPTEL. The three plotted curves for each value of $m_{\tilde{\chi}}c^2$ correspond to models with $\Omega_{m,0} = 0.1h_0$ (heavy lines), $\Omega_{m,0} = 0.3$ (medium lines) and $\Omega_{m,0} = 1$ (light lines). For clarity we have assumed decay lifetimes in each case such that the highest theoretical intensities lie just under the observational constraints.

diagram from becoming too cluttered, we have assumed values of f_{τ} such that the highest predicted intensity in each case stays just below the EGRET limits. Numerically, this corresponds to lower bounds on the decay lifetime $\tau_{\tilde{\chi}}$ between 100 Gyr (for $m_{\tilde{\chi}}c^2 = 10$ GeV) and 10^5 Gyr (for $m_{\tilde{\chi}}c^2 = 300$ GeV). For rest energies at the upper end of this range, these limits are probably optimistic because the decay photons are energetic enough to undergo **pair production** on cosmic microwave background (CMB) photons. Some of the decay photons, in other words, do not reach us from cosmological distances but are instead reprocessed into lower-energy radiation along the way. As we show in the next section, however, stronger limits arise from a different process in any case. We defer further discussion of figure 8.4 (including the observational data and the limits on neutralino lifetime that follow from them) to section 8.7.

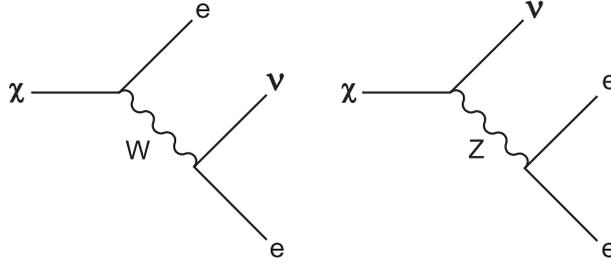


Figure 8.5. The Feynman diagrams corresponding to tree-level decays of the neutralino ($\tilde{\chi}$) into a neutrino (ν) and a lepton–antilepton pair (here, the electron and positron). The process can be mediated by the W or Z boson.

8.5 Tree-level decays

The dominant decay processes for the LSP neutralino in non-minimal SUSY (assuming spontaneously broken R-parity) are *tree-level* decays to leptons and neutrinos, as follows:

$$\tilde{\chi} \rightarrow \ell^+ + \ell^- + \nu_\ell. \quad (8.20)$$

Of particular interest is the case in which $\ell = e$; Feynman diagrams for this decay are shown in figure 8.5. Although these processes do not contribute directly to the EBL, they do so indirectly, because the high-energy electrons undergo **inverse Compton scattering** (ICS) off the CMB photons via $e + \gamma_{\text{cmb}} \rightarrow e + \gamma$. This gives rise to a flux of high-energy photons which can be at least as important as those from the direct (one-loop) neutralino decays considered in the previous section [52].

The spectrum of photons produced in this way depends on the rest energy of the original neutralino. We consider first the case $m_{10} \lesssim 10$, which is more or less pure ICS. The input (‘zero-generation’) electrons are monoenergetic, but after multiple scatterings they are distributed as E^{-2} [53]. From this the spectrum of outgoing photons can be calculated [54] and shown to take the following form:

$$N_{\text{ics}}(E) \propto \begin{cases} E^{-3/2} & (E \leq E_{\text{max}}) \\ 0 & (E > E_{\text{max}}) \end{cases} \quad (8.21)$$

where

$$E_{\text{max}} = \frac{4}{3} \left(\frac{E_e}{m_e c^2} \right)^2 E_{\text{cmb}}.$$

Here $E_e = \frac{1}{3} m_{\tilde{\chi}} c^2 = (3.3 \text{ GeV}) m_{10}$ is the energy of the input electrons, m_e is their rest mass and $E_{\text{cmb}} = 2.7 k T_{\text{cmb}}$ is the mean energy of the CMB photons. Using $m_e c^2 = 0.51 \text{ MeV}$ and $T_{\text{cmb}} = 2.7 \text{ K}$, and allowing for decays at arbitrary redshift z after Berezhinsky [54], we obtain the expression $E_{\text{max}}(z) = (36 \text{ keV}) m_{10}^2 (1+z)^{-1}$.

The halo SED may be determined as a function of wavelength by setting $F(\lambda) d\lambda = EN(E) dE$ where $E = hc/\lambda$. Normalizing the spectrum so that $\int_0^\infty F(\lambda) d\lambda = L_{\text{h,tree}}$ then gives:

$$F_{\text{ics}}(\lambda) = \frac{L_{\text{h,tree}}}{2} \times \begin{cases} \sqrt{\lambda_\gamma} \lambda^{-3/2} & (\lambda \geq \lambda_\gamma) \\ 0 & (\lambda < \lambda_\gamma) \end{cases} \quad (8.22)$$

where $\lambda_\gamma = hc/E_{\text{max}} = (0.34 \text{ \AA}) m_{10}^{-2} (1+z)$ and $L_{\text{h,tree}}$ is the halo luminosity due to tree-level decays.

In the case of more massive neutralinos with $m_{10} \gtrsim 10$, the situation is complicated by the fact that outgoing photons become energetic enough to initiate pair production via $\gamma + \gamma_{\text{cmb}} \rightarrow e^+ + e^-$. This injects new electrons into the ICS process, resulting in **electromagnetic cascades**. For particles which decay at high redshifts ($z \gtrsim 100$), other processes such as photon–photon scattering must also be taken into account [55]. Cascades on *non*-CMB background photons may also be important [56]. A full treatment of these effects requires detailed numerical analysis as carried out for example in [57]. Here we simplify the problem by assuming that the LSP is stable enough to survive into the late matter-dominated (or vacuum-dominated) era. The primary effect of cascades is to steepen the decay spectrum at high energies, so that [54]

$$N_{\text{casc}}(E) \propto \begin{cases} E^{-3/2} & (E \leq E_x) \\ E^{-2} & (E_x < E \leq E_c) \\ 0 & (E > E_c) \end{cases} \quad (8.23)$$

where

$$E_x = \frac{1}{3} \left(\frac{E_0}{m_e c^2} \right)^2 E_{\text{cmb}} (1+z)^{-1} \quad E_c = E_0 (1+z)^{-1}.$$

Here E_0 is a minimum absorption energy. We adopt the numerical expressions $E_x = (1.8 \times 10^3 \text{ GeV})(1+z)^{-1}$ and $E_c = (4.5 \times 10^4 \text{ GeV})(1+z)^{-1}$ after Protheroe *et al* [58]. Employing the relation $F(\lambda) d\lambda = EN(E) dE$ and normalizing as before, we find:

$$F_{\text{casc}}(\lambda) = \frac{L_{\text{h,tree}}}{[2 + \ln(\lambda_x/\lambda_c)]} \times \begin{cases} \sqrt{\lambda_x} \lambda^{-3/2} & (\lambda \geq \lambda_x) \\ \lambda^{-1} & (\lambda_x > \lambda \geq \lambda_c) \\ 0 & (\lambda < \lambda_c) \end{cases} \quad (8.24)$$

where the new parameters are $\lambda_x = hc/E_x = (7 \times 10^{-9} \text{ \AA})(1+z)$ and $\lambda_c = hc/E_c = (3 \times 10^{-10} \text{ \AA})(1+z)$.

The luminosity $L_{\text{h,tree}}$ is given by

$$L_{\text{h,tree}} = \frac{N_{\tilde{\chi}} b_e E_e}{\tau_{\tilde{\chi}}} = \frac{2b_e M_{\text{h}} c^2}{3\tau_{\tilde{\chi}}}. \quad (8.25)$$

Here b_e is now the branching ratio for all processes of the form $\tilde{\chi} \rightarrow e + \text{all}$, and $E_e = \frac{2}{3} m_{\tilde{\chi}} c^2$ is the total energy lost to the electrons. We assume that all of this

eventually finds its way into the EBL. Berezhinsky *et al* [51] supply the following branching ratio:

$$b_e \approx 10^{-6} f_\chi f_R^2 m_{10}^2. \quad (8.26)$$

Here f_χ parametrizes the composition of the neutralino, taking the value 0.4 for the pure higgsino case. With the halo mass specified by (6.15) and $f_\tau \equiv \tau_{\tilde{\chi}}/(1 \text{ Gyr})$ as usual, we obtain:

$$L_{\text{h,tree}} = (8 \times 10^{43} \text{ erg s}^{-1}) m_{10}^2 f_\chi f_R^2 f_\tau^{-1}. \quad (8.27)$$

This is approximately four orders of magnitude higher than the halo luminosity due to one-loop decays, and provides for the first time the possibility of significant EBL contributions. With all adjustable parameters taking values of order unity, we find that $L_{\text{h,tree}} \sim 2 \times 10^{10} L_\odot$, which is comparable to the bolometric luminosity of the Milky Way.

The combined bolometric intensity of all neutralino halos is computed as in the previous two sections. Replacing $L_{\text{h,loop}}$ in (8.10) with $L_{\text{h,tree}}$ leads to the values

$$Q = \begin{cases} (2 \times 10^{-4} \text{ erg s}^{-1} \text{ cm}^{-2}) h_0^2 m_{10}^2 f_\chi f_R^2 f_\tau^{-1} & (\Omega_{\text{m},0} = 0.1 h_0) \\ (5 \times 10^{-4} \text{ erg s}^{-1} \text{ cm}^{-2}) h_0 m_{10}^2 f_\chi f_R^2 f_\tau^{-1} & (\Omega_{\text{m},0} = 0.3) \\ (2 \times 10^{-3} \text{ erg s}^{-1} \text{ cm}^{-2}) h_0 m_{10}^2 f_\chi f_R^2 f_\tau^{-1} & (\Omega_{\text{m},0} = 1). \end{cases} \quad (8.28)$$

These are of the same order as or higher than the bolometric intensity of the EBL from ordinary galaxies, equation (2.25).

To obtain the spectral intensity, we substitute the SEDs $F_{\text{ics}}(\lambda)$ and $F_{\text{casc}}(\lambda)$ into equation (3.6). The results can be written:

$$I_\lambda(\lambda_0) = I_{\tilde{\chi},\text{tree}} \int_0^{z_f} \frac{\mathcal{F}(z) dz}{(1+z)[\Omega_{\text{m},0}(1+z)^3 + (1-\Omega_{\text{m},0})]^{1/2}} \quad (8.29)$$

where the quantities $I_{\tilde{\chi},\text{tree}}$ and $\mathcal{F}(z)$ are defined as follows. For neutralino rest energies $m_{10} \lesssim 10$ (ICS):

$$\begin{aligned} I_{\tilde{\chi},\text{tree}} &= \frac{n_0 L_{\text{h,tree}} f_c}{8\pi h H_0 m_{10}} \left(\frac{0.34 \text{ \AA}}{\lambda_0} \right)^{1/2} \\ &= (300 \text{ CUs}) h_0^2 m_{10} f_\chi f_R^2 f_\tau^{-1} f_c \left(\frac{\lambda_0}{\text{\AA}} \right)^{-1/2} \\ \mathcal{F}(z) &= \begin{cases} 1 & [\lambda_0 \geq \lambda_\gamma(1+z)] \\ 0 & [\lambda_0 < \lambda_\gamma(1+z)]. \end{cases} \end{aligned} \quad (8.30)$$

Conversely, for $m_{10} \gtrsim 10$ (cascades):

$$I_{\tilde{\chi},\text{tree}} = \frac{n_0 L_{\text{h,tree}} f_c}{4\pi h H_0 [2 + \ln(\lambda_x/\lambda_c)]} \left(\frac{7 \times 10^{-9} \text{ \AA}}{\lambda_0} \right)^{1/2}$$

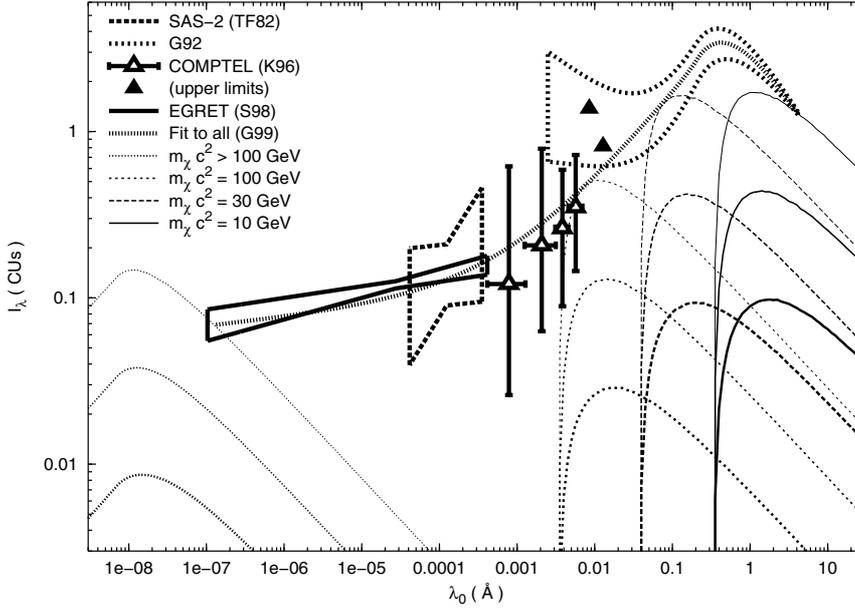


Figure 8.6. The spectral intensity of the diffuse γ -ray and x-ray backgrounds due to neutralino tree-level decays, compared with observational upper limits from SAS-2, EGRET and COMPTEL in the γ -ray region, and from Gruber's fits to the experimental data in the x-ray region (G92,G99). The three plotted curves for each value of $m_{\tilde{\chi}} c^2$ correspond to models with $\Omega_{m,0} = 0.1 h_0$ (bold lines), $\Omega_{m,0} = 0.3$ (medium lines) and $\Omega_{m,0} = 1$ (light lines). For clarity we have assumed decay lifetimes in each case such that the highest theoretical intensities lie just under the observational constraints.

$$\begin{aligned}
 &= (0.02 \text{ CU s}) h_0^2 m_{10}^2 f_{\tilde{\chi}} f_R^2 f_{\tau}^{-1} f_c \left(\frac{\lambda_0}{\text{\AA}} \right)^{-1/2} \quad (8.31) \\
 \mathcal{F}(z) &= \begin{cases} 1 & [\lambda_0 \geq \lambda_x(1+z)] \\ \frac{\lambda_0}{\lambda_x(1+z)} & [\lambda_x(1+z) > \lambda_0 \geq \lambda_c(1+z)] \\ 0 & [\lambda_0 < \lambda_c(1+z)]. \end{cases}
 \end{aligned}$$

Numerical integration of equation (8.29) leads to the plots shown in figure 8.6. Cascades (like the pair annihilations we have considered already) dominate the γ -ray part of the spectrum. The ICS process, however, is most important at lower energies, in the x-ray region. We discuss the observational limits and the constraints that can be drawn from them in more detail in section 8.7.

8.6 Gravitinos

Gravitinos (\tilde{g}) are the SUSY spin- $\frac{3}{2}$ counterparts of gravitons. Although often discussed along with neutralinos, they are not favoured as dark-matter candidates, at least in the simplest SUSY theories. The reason for this, often called the **gravitino problem** [5], boils down to the fact that they interact *too* weakly, not only with other particles but with themselves as well. Hence they annihilate too slowly and survive long enough to ‘overclose’ the Universe unless some other way is found to reduce their numbers. Decays are one possibility, but not if the gravitino is a stable LSP. Gravitino decay products must also not be allowed to interfere with processes such as primordial nucleosynthesis [4]. Inflation, followed by a judicious period of reheating, can thin out their numbers to almost any desired level. But the **reheat temperature** T_R must satisfy $kT_R \lesssim 10^{12}$ GeV or gravitinos will once again become too numerous [9]. Related arguments based on entropy production, primordial nucleosynthesis and the CMB power spectrum force this number down to $kT_R \lesssim (10^9\text{--}10^{10})$ GeV [59] or even $kT_R \lesssim (10^6\text{--}10^9)$ GeV [60]. These temperatures are incompatible with the generation of **baryon asymmetry** in the Universe, a process which is usually taken to require $kT_R \sim 10^{14}$ GeV or higher [61].

Recent developments are, however, beginning to loosen the baryogenesis requirement [62], and there are alternative models in which baryon asymmetry is generated at energies as low as ~ 10 TeV [63] or even 10 MeV–1 GeV [64]. With this in mind we include a brief look at gravitinos here. There are two possibilities: (1) If the gravitino is *not* the LSP, then it decays early in the history of the Universe, well before the onset of the matter-dominated era. In the model of Dimopoulos *et al* [65], for example, the gravitino decays both radiatively and hadronically and is, in fact, ‘long-lived for its mass’ with a lifetime of $\tau_{\tilde{g}} = (2 - 9) \times 10^5$ s. Particles of this kind have important consequences for nucleosynthesis, and might affect the shape of the CMB, if $\tau_{\tilde{g}}$ were to exceed $\sim 10^7$ s. However, they are irrelevant as far as the EBL is concerned. We therefore restrict our attention to the case (2), in which the gravitino is the LSP. Moreover, in light of the results we have already obtained for the neutralino, we disregard annihilations and consider only models in which the LSP can decay.

The decay mode depends on the specific mechanism of R-parity violation. We follow Berezhinsky [66] and concentrate on dominant *tree-level* processes. In particular we consider the decay:

$$\tilde{g} \rightarrow e^+ + \text{all} \quad (8.32)$$

followed by ICS off the CMB, as in the previous section on neutralinos. The spectrum of photons produced by this process is identical to that in section 8.5, except that the monoenergetic electrons have energy $E_e = \frac{1}{2}m_{\tilde{g}}c^2 = (5 \text{ GeV})m_{10}$ [66], where $m_{\tilde{g}}c^2$ is the rest energy of the gravitino and $m_{10} \equiv m_{\tilde{g}}c^2/(10 \text{ GeV})$ as before. This, in turn, implies that $E_{\text{max}} = (81 \text{ keV})m_{10}^2(1+z)^{-1}$ and

$\lambda_\gamma = hc/E_{\max} = (0.15 \text{ \AA})m_{10}^{-2}(1+z)$. The values of λ_x and λ_c are unchanged.

The SED comprises equations (8.22) for ICS and (8.24) for cascades, as before. Only the halo luminosity needs to be recalculated. This is similar to equation (8.25) for neutralinos, except that the factor of $\frac{2}{3}$ becomes $\frac{1}{2}$, and the branching ratio can be estimated at [66]

$$b_e \sim \left(\frac{\alpha}{\pi}\right)^2 = 5 \times 10^{-6}. \quad (8.33)$$

Using our standard value for the halo mass M_h , and parametrizing the gravitino decay lifetime by $f_\tau \equiv \tau_{\tilde{g}}/(1 \text{ Gyr})$ as before, we obtain the following halo luminosity due to gravitino decays:

$$L_{h,\text{grav}} = (3 \times 10^{44} \text{ erg s}^{-1})f_\tau^{-1}. \quad (8.34)$$

This is higher than the luminosity due to neutralino decays and exceeds the luminosity of the Milky Way by several times if $f_\tau \sim 1$.

The bolometric intensity of all gravitino halos is computed exactly as before. Replacing $L_{h,\text{tree}}$ in (8.10) with $L_{h,\text{grav}}$, we find:

$$Q = \begin{cases} (7 \times 10^{-4} \text{ erg s}^{-1} \text{ cm}^{-2})h_0^2 f_\tau^{-1} & (\Omega_{m,0} = 0.1h_0) \\ (2 \times 10^{-3} \text{ erg s}^{-1} \text{ cm}^{-2})h_0 f_\tau^{-1} & (\Omega_{m,0} = 0.3) \\ (8 \times 10^{-3} \text{ erg s}^{-1} \text{ cm}^{-2})h_0 f_\tau^{-1} & (\Omega_{m,0} = 1). \end{cases} \quad (8.35)$$

It is clear that gravitinos must decay on timescales longer than the lifetime of the Universe ($f_\tau \gtrsim 16$), or they would produce a background brighter than that of the galaxies.

The spectral intensity is the same as before, equation (8.29), but with the new numbers for λ_γ and L_h . This results in

$$I_\lambda(\lambda_0) = I_{\tilde{g}} \int_0^{z_f} \frac{\mathcal{F}(z) dz}{(1+z)[\Omega_{m,0}(1+z)^3 + (1-\Omega_{m,0})]^{1/2}} \quad (8.36)$$

where the prefactor $I_{\tilde{g}}$ is defined as follows. For $m_{10} \lesssim 10$ (ICS):

$$\begin{aligned} I_{\tilde{g}} &= \frac{n_0 L_{h,\text{grav}} f_c}{8\pi h H_0 m_{10}} \left(\frac{0.15 \text{ \AA}}{\lambda_0}\right)^{1/2} \\ &= (800 \text{ CUs}) h_0^2 m_{10}^{-1} f_\tau^{-1} f_c \left(\frac{\lambda_0}{\text{\AA}}\right)^{-1/2}. \end{aligned} \quad (8.37)$$

Conversely, for $m_{10} \gtrsim 10$ (cascades):

$$\begin{aligned} I_{\tilde{g}} &= \frac{n_0 L_{h,\text{grav}} f_c}{4\pi h H_0 [2 + \ln(\lambda_x/\lambda_c)]} \left(\frac{7 \times 10^{-9} \text{ \AA}}{\lambda_0}\right)^{1/2} \\ &= (0.06 \text{ CUs}) h_0^2 f_\tau^{-1} f_c \left(\frac{\lambda_0}{\text{\AA}}\right)^{-1/2}. \end{aligned} \quad (8.38)$$

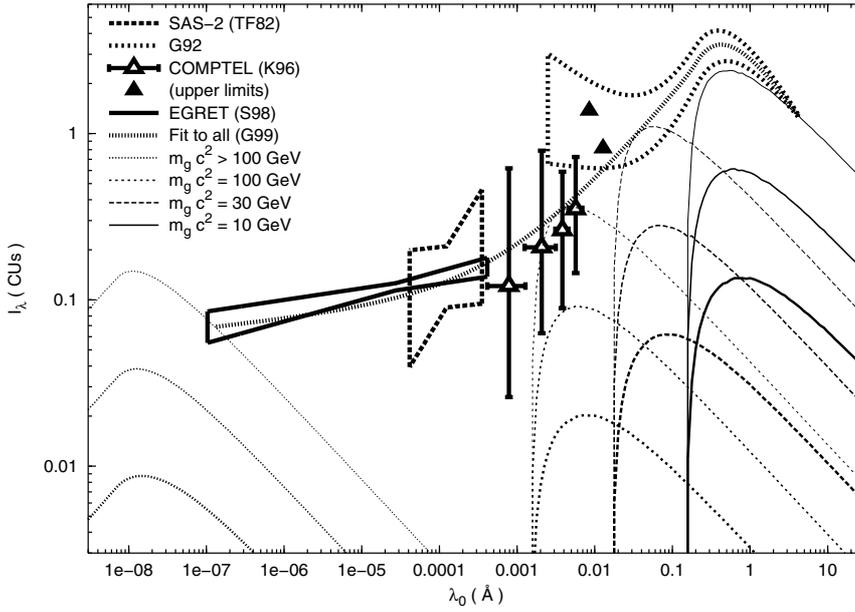


Figure 8.7. The spectral intensity of the diffuse γ -ray and x-ray backgrounds due to gravitino tree-level decays, compared with observational upper limits. These come from SAS-2, COMPTEL and EGRET in the γ -ray region, and from Gruber's fits to the experimental data in the x-ray region. The three plotted curves for each value of $m_{\tilde{g}} c^2$ correspond to models with $\Omega_{m,0} = 0.1 h_0$ (bold lines), $\Omega_{m,0} = 0.3$ (medium lines) and $\Omega_{m,0} = 1$ (light lines). For clarity we have assumed decay lifetimes in each case such that the highest theoretical intensities lie just under the observational constraints.

The function $\mathcal{F}(z)$ has the same form as in equations (8.30) and (8.31) and does not need to be redefined (requiring only the new value for the cut-off wavelength λ_γ). Because the branching ratio b_e in (8.33) is independent of the gravitino rest mass, m_{10} appears in these results only through λ_γ . Thus the ICS part of the spectrum goes as m_{10}^{-1} while the cascade part does not depend on m_{10} at all. As with neutralinos, cascades dominate the γ -ray part of the spectrum and the ICS process is most important in the x-ray region. Numerical integration of equation (8.36) leads to the results plotted in figure 8.7. We proceed to discuss these next, beginning with the observational data.

8.7 The x-ray and γ -ray backgrounds

The experimental situation as regards EBL intensity in the x-ray and γ -ray regions is clearer than that in the optical and ultraviolet. Firm detections (as opposed to upper limits) have been made in both bands and these are consistent with

expectations based on known astrophysical sources. The constraints that we derive here are thus conservative ones, in the sense that the EBL flux which could plausibly be due to decaying WIMPs is almost certainly smaller than the levels actually measured.

We have not used any data on the lowest-energy, or soft x-ray background, which lies just beyond the ultraviolet and extends over approximately 0.1–3 keV or 4–100 Å. The hard x-ray background (3–800 keV or 0.02–4 Å) is, however, crucial in constraining the decays of low-mass neutralinos and gravitinos via the ICS process, as can be seen in figures 8.6 and 8.7. (The high-energy spectrum is conventionally divided into wavebands whose precise definition, however, varies from author to author. The definitions we use here coincide roughly with energy ranges in which different detection techniques must be used.) We have included two compilations of observations in the hard x-ray band, both by Gruber [67, 68]. The first (labelled ‘G92’ in figures 8.6 and 8.7) is an empirical fit to various pre-1992 measurements, including those from the Kosmos and Apollo spacecraft, HEAO-1 and balloon experiments. In plotting the uncertainty for this curve, we have used the fact that the relative scatter of the data increases logarithmically from 2% at 3 keV to 60% at 3 MeV. The second compilation (‘G99’) is a revision of this fit in light of new data at higher energies, and has been extended deep into the γ -ray region. The prominent peak in the range 3–300 keV (0.04–4 Å) is widely attributed to integrated light from **active galactic nuclei** (AGN) [69].

In the low-energy γ -ray region (0.8–30 MeV or 0.0004–0.02 Å) we have used results from the COMPTEL instrument on the Compton Gamma-Ray Observatory (CGRO), which was operational from 1990–2000 [70]. Four data points are plotted in figures 8.2, 8.4, 8.6 and 8.7, and two more (upper limits only) appear at low energies in figures 8.6 and 8.7. These experimental results, which interpolate smoothly between the backgrounds at both lower and higher energies, played a key role in the demise of the ‘MeV bump’ (visible in figures 8.6 and 8.7 as a significant upturn in Gruber’s fit to the pre-1992 data from about 0.002–0.02 Å). This apparent feature in the background had previously attracted a great deal of attention from theoretical cosmologists as a possible signature of new physics. Figures 8.6 and 8.7 suggest that it could also have been interpreted as evidence for a long-lived non-minimal SUSY WIMP with a rest energy near 100 GeV. The MeV bump is, however, no longer believed to be real, as the new fit (‘G99’) makes clear. Most of the background in this region is now suspected to be due to **Type Ia supernovae** (SNIa) [71].

There are a number of good measurements in the high-energy γ -ray band (30 MeV–30 GeV or 4×10^{-7} – 4×10^{-4} Å). We have plotted two of these: one from the SAS-2 satellite which flew in 1972–73 [72], and one from the EGRET instrument which was part of the CGRO mission along with COMPTEL [73]. As may be seen in figures 8.2, 8.4, 8.6 and 8.7, the new results essentially confirm the old ones and extend them to energies as high as 120 GeV ($\lambda_0 = 10^{-7}$ Å), with error bars which have been reduced in size by a factor of about ten. Most of this extragalactic background is thought to arise from unresolved **blazars**, highly

variable AGN whose relativistic jets are pointed in nearly our direction [74].

Finally, we have made use of some observations in the very high-energy (VHE) γ -ray region (30 GeV–30 TeV or 4×10^{-10} – 4×10^{-7} Å). Because the extragalactic component of the background has not yet been measured beyond 120 GeV, we have fallen back on measurements of *total* γ -ray flux, obtained by Nishimura *et al* in 1980 [75] using a series of high-altitude balloon experiments. These are shown as filled dots (‘N80’) in figures 8.2 and 8.4. They constitute a robust upper limit on EBL flux, since the majority of the observed signals are due to cosmic-ray interactions in the atmosphere of the Earth.

Some comments are in order here about notation and units in this part of the spectrum. For experimental reasons, measurements are often expressed in terms of **integral flux** $E I_E(>E_0)$ or number of particles with energies above E_0 . This presents no difficulties in the high-energy γ -ray region where the differential spectrum is well approximated with a single power-law component, $I_E(E_0) = I_*(E_0/E_*)^{-\alpha}$. The conversion to integral form is then given by

$$E I_E(>E_0) = \int_{E_0}^{\infty} I_E(E) dE = \frac{E_* I_*}{\alpha - 1} \left(\frac{E_0}{E_*} \right)^{1-\alpha}. \quad (8.39)$$

The spectrum in either case can be specified by its index α , together with the values of E_* and I_* (or E_0 and $E I_E$ the integral case). Thus the final SAS-2 results were reported as $\alpha = 2.35^{+0.4}_{-0.3}$ with $E I_E = (5.5 \pm 1.3) \times 10^{-5} \text{ s}^{-1} \text{ cm}^{-2} \text{ ster}^{-1}$ for $E_0 = 100 \text{ MeV}$ [72]. The EGRET spectrum is instead specified by $\alpha = 2.10 \pm 0.03$ with $I_* = (7.32 \pm 0.34) \times 10^{-9} \text{ s}^{-1} \text{ cm}^{-2} \text{ ster}^{-1} \text{ MeV}^{-1}$ for $E_* = 451 \text{ MeV}$ [73]. To convert a differential flux in these units to I_λ in CUs, we multiply by $E_0/\lambda_0 = E_0^2/hc = 80.66 E_0^2$ where E_0 is photon energy in MeV.

We now discuss our results, beginning with the neutralino annihilation fluxes plotted in figure 8.2. These are at least three orders of magnitude fainter than the background detected by EGRET [73] (and four orders of magnitude below the upper limit set by the data of Nishimura *et al* [75] at shorter wavelengths). This agrees with previous studies assuming a critical density of neutralinos [20, 40]. Figure 8.2 shows that EBL contributions would drop by another order of magnitude in the favoured scenario with $\Omega_{m,0} \approx 0.3$, and by another if neutralinos are confined to galaxy halos ($\Omega_{m,0} \approx 0.1h_0$). Unfortunately, the same stability that makes minimal SUSY WIMPs so compelling as dark-matter candidates also makes them hard to detect. As discussed in section 8.3, prospects for observing these particles can improve substantially if one looks for the enhanced flux from high-density regions like the Galactic centre.

Figure 8.4 shows the EBL contributions from one-loop neutralino decays in *non-minimal* SUSY. We have put $h_0 = 0.75$, $z_f = 30$ and $f_R = 1$. Depending on their decay lifetime (here parametrized by f_τ), these particles are capable, in principle, of producing a background comparable to (or even in excess of) the EGRET limits. The plots in figure 8.4 correspond to the smallest values of f_τ which are consistent with the data for $m_{10} = 1, 3, 10, 30$ and 100.

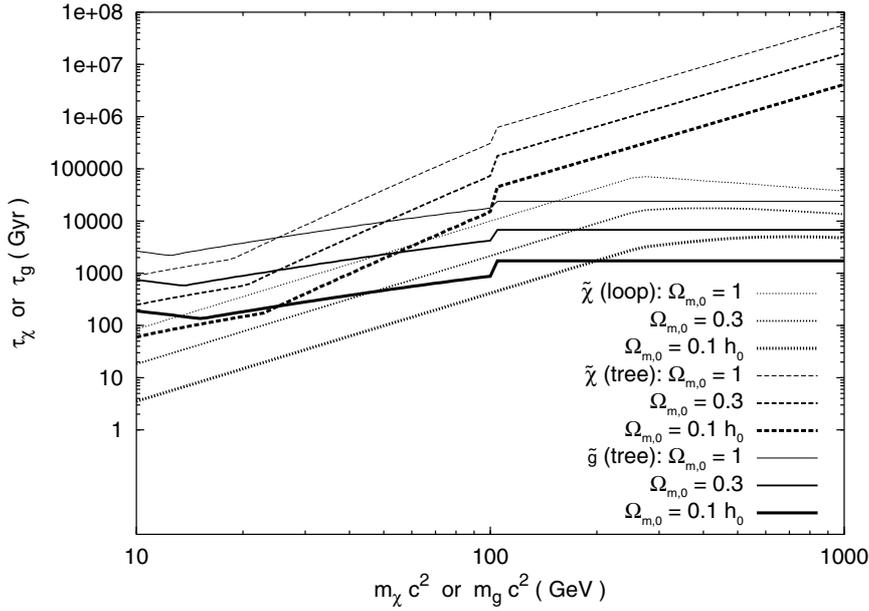


Figure 8.8. The lower limits on WIMP decay lifetime derived from observations of the x-ray and γ -ray backgrounds. Neutralino bounds are shown for both one-loop decays (dotted lines) and tree-level decays (dashed lines). For gravitinos we show only the tree-level constraints (unbroken lines). For each process there are three curves corresponding to models with $\Omega_{m,0} = 1$ (light lines), $\Omega_{m,0} = 0.3$ (medium lines) and $\Omega_{m,0} = 0.1h_0$ (bold lines).

Following the same procedure here as we did for axions in chapter 6, we can repeat this calculation over more finely-spaced intervals in neutralino rest mass, obtaining a lower limit on decay lifetime $\tau_{\tilde{\chi}}$ as a function of $m_{\tilde{\chi}}$. Results are shown in figure 8.8 (dotted lines). The lower limits obtained in this way range from 4 Gyr for the lightest neutralinos (assumed to be confined to galaxy halos with a total matter density of $\Omega_{m,0} = 0.1h_0$) to 70 000 Gyr for the heaviest (if these provide enough CDM to put $\Omega_{m,0} = 1$). A typical intermediate limit (for $m_{\tilde{\chi}}c^2 = 100$ GeV and $\Omega_{m,0} = 0.3$) is $\tau_{\tilde{\chi}} > 2000$ Gyr.

Figure 8.6 is a plot of EBL flux from *indirect* neutralino decays via the tree-level, ICS and cascade processes described in section 8.5. These provide us with our strongest constraints on non-minimal SUSY WIMPs. We have set $h_0 = 0.75$, $z_f = 30$ and $f_{\tilde{\chi}} = f_R = 1$, and assumed values of f_{τ} such that the highest predicted intensities lie just under observational limits, as before. Contributions from neutralinos at the light end of the mass range are constrained by x-ray data, while those at the heavy end are checked by the EGRET measurements. Both the shape and absolute intensity of the ICS spectra depend on the neutralino rest

mass. The cascade spectra, however, depend on m_{10} through intensity alone (via the prefactor $I_{\tilde{\chi},\text{tree}}$). When normalized to the observational upper limit, all curves for $m_{10} > 10$ therefore overlap. Normalizing across the full range of neutralino rest masses (as for one-loop decays), we obtain the lower bounds on lifetime $\tau_{\tilde{\chi}}$ plotted in figure 8.8 (dashed lines). These range from 60 Gyr to 6×10^7 Gyr, depending on the values of $m_{\tilde{\chi}}c^2$ and $\Omega_{m,0}$. A typical result (for $m_{\tilde{\chi}}c^2 = 100$ GeV and $\Omega_{m,0} = 0.3$) is $\tau_{\tilde{\chi}} > 70\,000$ Gyr.

Figure 8.7, finally, shows the EBL contributions from tree-level *gravitino* decays. These follow the same pattern as the neutralino decays, and need little additional comment. Requiring that the predicted signal not exceed the x-ray and γ -ray observations, we obtain the lower limits on decay lifetime plotted in figure 8.8 (unbroken lines). These curves are seen to be flatter for gravitinos than they are for neutralinos. This is a consequence of the fact that the branching ratio (8.33) is independent of m_{10} . Our lower limits on $\tau_{\tilde{g}}$ range from 200 to 20 000 Gyr, with a typical value (for $m_{\tilde{g}}c^2 = 100$ GeV and $\Omega_{m,0} = 0.3$) of $\tau_{\tilde{g}} > 4000$ Gyr.

These results agree with estimates in the literature [51, 66] and confirm that, whether it is a neutralino or gravitino, *the LSP in non-minimal SUSY theories must be very nearly stable*. Neither particle can be excluded as a dark-matter candidate on this basis; the LSP may after all be perfectly stable, which is the simplest (minimal) case. However, an ‘almost-stable’ LSP is difficult to understand in the context of non-minimal SUSY theory, because it requires that R-parity conservation be violated, but violated at improbably low levels. Equivalently, it introduces a very small dimensionless quantity into the theory, analogous to the parameter Θ associated with the strong-CP problem (chapter 6). To this extent, then, our constraints on decay lifetime suggest that the SUSY WIMP either exists in the context of minimal SUSY theory, or not at all.

Let us sum up our findings in this chapter. We have considered neutralinos and gravitinos, either of which could be the LSP and hence make up the ‘missing mass’ in SUSY theories. Both particles remain viable as dark-matter candidates but their properties are constrained by ever-improving data on the intensity of the extragalactic x-ray and γ -ray backgrounds. In the minimal SUSY model, where R-parity is strictly conserved, neutralinos annihilate so slowly that no useful limits can be set at this time. (This would be even more true of gravitinos.) In the broader class of *non-minimal* SUSY theories, however, R-parity conservation is no longer guaranteed and the LSP (whether neutralino or gravitino) can decay. We have shown that any such decay must occur on timescales longer than 10^2 – 10^8 Gyr (for neutralinos) or 10^2 – 10^4 Gyr (for gravitinos), depending on their rest masses and various theoretical input parameters. For either case it is clear that these particles must be almost perfectly stable. This, in turn, implies that R-parity must be broken very gently, if it is broken at all.

References

- [1] Jungman G, Kamionkowski M and Griest K 1996 *Phys. Rep.* **267** 195
- [2] Cabibbo N, Farrar G R and Maiani L 1981 *Phys. Lett. B* **105** 155
- [3] Pagels H and Primack J R 1982 *Phys. Rev. Lett.* **48** 223
- [4] Weinberg S 1982 *Phys. Rev. Lett.* **48** 1303
- [5] Ellis J, Linde A D and Nanopoulos D V 1982 *Phys. Lett. B* **118** 59
- [6] Goldberg H 1983 *Phys. Rev. Lett.* **50** 1419
- [7] Ibáñez L E 1984 *Phys. Lett. B* **137** 160
- [8] Sciama D W 1984 *Phys. Lett. B* **137** 169
- [9] Ellis J *et al* 1984 *Nucl. Phys. B* **238** 453
- [10] Steigman G and Turner M S 1985 *Nucl. Phys. B* **253** 375
- [11] Ellis J and Olive K A 2001 *Phys. Lett. B* **514** 114
- [12] Ellis J, Falk T, Ganis G, Olive K A 2000 *Phys. Rev. D* **62** 075010
- [13] Griest K, Kamionkowski M and Turner M S 1990 *Phys. Rev. D* **41** 3565
- [14] Ellis J, Falk T and Olive K A 1998 *Phys. Lett. B* **444** 367
- [15] Roszkowski L, de Austri R R and Nihei T 2001 *J. High Energy Phys.* **08** 024
- [16] Baudis L and Klapdor-Kleingrothaus H V 2000 *Beyond the Desert 1999* ed H V Klapdor-Kleingrothaus and I V Krivosheina (Oxford: Institute of Physics Press) p 881
- [17] Belli P *et al* 2000 *Beyond the Desert 1999* ed H V Klapdor-Kleingrothaus and I V Krivosheina (Oxford: Institute of Physics Press) p 869
- [18] Abusaidi R *et al* 2000 *Phys. Rev. Lett.* **84** 5699
- [19] Benoit A *et al* 2001 *Phys. Lett. B* **513** 15
- [20] Silk J and Srednicki M 1984 *Phys. Rev. Lett.* **53** 624
- [21] Silk J, Olive K and Srednicki M 1985 *Phys. Rev. Lett.* **55** 257
- [22] Freese K 1986 *Phys. Lett. B* **167** 295
- [23] Krauss L M, Srednicki M and Wilczek F 1986 *Phys. Rev. D* **33** 2079
- [24] Bai X *et al* 2001 *Dark Matter in Astro- and Particle Physics* ed H V Klapdor-Kleingrothaus (Heidelberg: Springer) p 699
- [25] Silk J and Bloemen H 1987 *Astrophys. J.* **313** L47
- [26] Stecker F W 1988 *Phys. Lett. B* **201** 529
- [27] Rudaz S 1989 *Phys. Rev. D* **39** 3549
- [28] Stecker F W and Tylka A J 1989 *Astrophys. J.* **343**
- [29] Bouquet A, Salati P and Silk J 1989 *Phys. Rev. D* **40** 3168
- [30] Bergström L 1989 *Nucl. Phys. B* **325** 647
- [31] Freese K and Silk J 1989 *Phys. Rev. D* **40** 3828
- [32] Berezhinsky V, Bottino A and Mignola G 1994 *Phys. Lett. B* **325** 136
- [33] Gondolo P and Silk J 1999 *Phys. Rev. Lett.* **83** 1719
- [34] Urban M *et al* 1992 *Phys. Lett. B* **293** 149
- [35] Chardonnet P *et al* 1995 *Astrophys. J.* **454** 774
- [36] Lake G 1990 *Nature* **346** 39
- [37] Gondolo P 1994 *Nucl. Phys. Proc. Suppl. B* **B35** 148
- [38] Baltz E A *et al* 1999 *Phys. Rev. D* **61** 023514
- [39] Cline D B and Gao Y-T 1990 *Astron. Astrophys.* **231** L23
- [40] Gao Y-T, Stecker F W and Cline D B 1991 *Astron. Astrophys.* **249** 1
- [41] Overduin J M and Wesson P S 1997 *Astrophys. J.* **480** 470
- [42] Berezhinsky V S, Bottino A and de Alfaro V 1992 *Phys. Lett. B* **274** 122

- [43] Giudice G F and Griest K 1989 *Phys. Rev. D* **40** 2549
- [44] Gates E I, Gyuk G and Turner M S 1995 *Astrophys. J.* **449** L123
- [45] Metzger M R, Caldwell J A R and Schechter P L 1998 *Astron. J.* **115** 635
- [46] Kochanek C S 1996 *Astrophys. J.* **457** 228
- [47] Peebles P J E 1993 *Principles of Physical Cosmology* (Princeton, NJ: Princeton University Press) p 421
- [48] Berezhinsky V S, Gurevich A V and Zybin K P 1992 *Phys. Lett. B* **294** 221
- [49] Bouquet A and Salati P 1987 *Nucl. Phys. B* **284** 557
- [50] Masiero A and Valle J W F 1990 *Phys. Lett. B* **251** 273
- [51] Berezhinsky V, Masiero A and Valle J W F 1991 *Phys. Lett. B* **266** 382
- [52] Barbieri R and Berezhinsky V 1988 *Phys. Lett. B* **205** 559
- [53] Blumenthal G R and Gould R J 1970 *Rev. Mod. Phys.* **42** 237
- [54] Berezhinsky V S 1992 *Nucl. Phys. B* **380** 478
- [55] Svensson R and Zdziarski A A 1990 *Astrophys. J.* **349** 415
- [56] Coppi P S and Aharonian F A 1997 *Astrophys. J.* **487** L9
- [57] Kribs G D and Rothstein I Z 1997 *Phys. Rev. D* **55** 4435
- [58] Protheroe R J, Stanev T and Berezhinsky V S 1995 *Phys. Rev. D* **51** 4134
- [59] Ellis J, Kim E and Nanopoulos D V 1984 *Phys. Lett. B* **145** 181
- [60] Kawasaki M and Moroi T 1995 *Prog. Theor. Phys.* **93** 879
- [61] Kolb E W and Turner M S 1990 *The Early Universe* (Reading: Addison-Wesley) p 157
- [62] Giudice G F, Kolb E W and Riotto A 2001 *Phys. Rev. D* **64** 023508
- [63] Kuzmin V A, Rubakov V A and Shaposhnikov M E 1985 *Phys. Lett. B* **155** 36
- [64] Dimopoulos S and Hall L J 1987 *Phys. Lett. B* **196** 135
- [65] Dimopoulos S *et al* 1988 *Astrophys. J.* **330** 545
- [66] Berezhinsky V S 1991 *Phys. Lett. B* **261** 71
- [67] Gruber D E 1992 *The X-Ray Background* ed X Barcon and A C Fabian (Cambridge: Cambridge University Press) p 44
- [68] Gruber D E 1999 *Astrophys. J.* **520** 124
- [69] Zdziarski A A 1996 *Mon. Not. R. Astron. Soc.* **281** L9
- [70] Kappadath S C *et al* 1996 *Astron. Astrophys. Suppl. Ser.* **120** 619
- [71] The L-S, Leising M D and Clayton D D 1993 *Astrophys. J.* **403** 32
- [72] Thompson D J and Fichtel C E 1982 *Astron. Astrophys.* **109** 352
- [73] Sreekumar P *et al* 1998 *Astrophys. J.* **494** 523
- [74] McNaron-Brown K *et al* 1995 *Astrophys. J.* **451** 575
- [75] Nishimura J *et al* 1980 *Astrophys. J.* **238** 394

Chapter 9

Black holes

9.1 Primordial black holes

Black holes are regions of space from which light cannot escape. It might therefore appear that little could be learned about these objects from measurements of the extragalactic background light (EBL). In fact, experimental data on EBL intensity constrain black holes *more strongly* than any of the other dark-matter candidates we have discussed so far. Before explaining how this comes about, we distinguish between ‘ordinary’ black holes (which form via the gravitational collapse of massive stars at the end of their lives) and **primordial black holes** (PBHs) which could have arisen from the collapse of overdense regions in the early Universe. The existence of the former is very nearly an established astronomical fact, while the latter remain hypothetical. However, it is the latter (PBHs) which are of most interest to us as potential dark-matter candidates.

The reason for this is as follows. Ordinary black holes come from baryonic progenitors (i.e. stars) and can hence be classified with the baryonic dark matter of the Universe. (They are, of course, not ‘baryonic’ in all respects, since among other things their baryon number is not defined.) Ordinary black holes are therefore subject to the nucleosynthesis bound (4.4) on the density of baryonic matter, which limits them to less than 5% of the critical density. PBHs are *not* subject to this bound because they form during the radiation-dominated era, before nucleosynthesis begins. *A priori*, nothing prevents them from making up most of the density in the Universe. Moreover, they constitute *cold* dark matter because their velocities are low. (That is, they collectively obey a dustlike equation of state, even though they might individually be better described as ‘radiation-like’ than baryonic.) PBHs were first proposed as dark-matter candidates by Zeldovich and Novikov in 1966 [1] and Hawking in 1971 [2].

Black holes contribute to the EBL via a process discovered by Hawking in 1974 and often called **Hawking evaporation** [3]. Photons cannot escape from inside the black hole, but they are produced *at or near the horizon* by

quantum fluctuations in the surrounding curved spacetime. These give rise to a net flux of particles which propagates outward from a black hole (of mass M) at a rate proportional to M^{-2} (with the black-hole mass itself dropping at the same rate). For ordinary, stellar-mass black holes, this process occurs so slowly that contributions to the EBL are insignificant and the designation ‘black’ remains perfectly appropriate over the lifetime of the Universe. PBHs, however, can have masses far smaller than those of a star, leading to correspondingly higher luminosities. Those with $M \lesssim 10^{15}$ g (equivalent to the mass of a small asteroid) would, in fact, evaporate quickly enough to shed all their mass over less than ~ 10 Gyr. They would already have expired in a blaze of high-energy photons and other elementary particles as $M \rightarrow 0$.

Our goal in this chapter will be to estimate the impact of this process on the intensity of the EBL, and also to consider the corresponding problem for the higher-dimensional analogues of black holes known as **solitons**. In four dimensions this is a calculation that has been refined by many people and has entered into the ‘standard lore’ of dark-matter astrophysics. For the most part we will be content to review these well-established results, commenting as appropriate on the ways in which they have been extended and modified by later research. We will confirm that observations of the γ -ray background exclude a significant role for four-dimensional black holes under the most natural assumptions. Solitons are not as easy to constrain but we will find that the bolometric intensity of the background allows us to put useful limits on these objects as well.

9.2 Initial mass distribution

Our sources of background radiation in this chapter are the PBHs themselves. We will take these to be distributed homogeneously throughout space. This is not necessarily realistic since their low velocities mean that PBHs will tend to collect inside the potential wells of galaxies and galaxy clusters. Clustering is not of great concern to us here, however, for the same reason already discussed in connection with axions and WIMPs (i.e. we are interested primarily in the combined contributions of all PBHs to the diffuse background). A complication is introduced, however, by the fact that PBHs cover such a wide range of masses and luminosities that we can no longer treat all sources in the same way (as we did for the galaxy halos in previous chapters). Instead we must define quantities like number density and energy spectrum as functions of *PBH mass* as well as time, and integrate our final results over both parameters.

The first step is to identify the distribution of PBH masses at the time when they formed. There is little prospect of probing the time before nucleosynthesis experimentally, so any theory of PBH formation is necessarily speculative to some degree. However, the scenario requiring the least extrapolation from known physics would be one in which PBHs arose via the gravitational collapse of

small initial density fluctuations on a standard Robertson–Walker background. It is furthermore reasonable to assume that the equation of state had the usual form (2.27) and that the initial density fluctuations were distributed as

$$\delta = \epsilon(M_i/M_f)^{-n}. \quad (9.1)$$

Here M_i is the initial mass of the PBH, M_f is the mass lying inside the **particle horizon** (or causally connected Universe) at PBH formation time, and ϵ is a proportionality constant.

An investigation of PBH formation under these conditions was carried out in 1975 by Carr [4], who showed that the process is favoured over an extended range of masses only if $n = \frac{2}{3}$. Proceeding on this assumption, he found that the **initial mass distribution** of PBHs formed with masses between M_i and $M_i + dM_i$ per unit comoving volume is

$$n(M_i) dM_i = \rho_f M_f^{-2} \zeta \left(\frac{M_i}{M_f} \right)^{-\beta} dM_i \quad (9.2)$$

where ρ_f is the mean density at formation time. The parameters β and ζ are formally given by $2(2\gamma - 1)/\gamma$ and $\epsilon \exp[-(\gamma - 1)^2/2\epsilon^2]$ respectively, where γ is the equation-of-state parameter in (2.27). However, in the interests of lifting every possible restriction on conditions prevailing in the early Universe, we follow Carr [5] in treating β and ζ as free parameters, not necessarily related to γ and ϵ . Insofar as the early Universe was governed by the equation of state (2.27), β will take values between 2 (dustlike or ‘soft’) and 3 (stiff or ‘hard’), with $\beta = \frac{5}{2}$ corresponding to the most natural situation (i.e. $\gamma = \frac{4}{3}$ as for a non-interacting relativistic gas). But we will allow β to take values up to 4, corresponding to ‘superhard’ early conditions. The parameter ζ can be understood physically as the fraction of the Universe which goes into PBHs of mass M_f at time t_f . It is a measure of the initial inhomogeneity of the Universe.

The fact that equation (9.2) has no exponential cut-off at high mass is important because it allows us (at least in principle) to obtain a substantial cosmological density of PBHs. Since $2 \leq \beta \leq 4$, however, the power-law distribution is dominated by PBHs of *low* mass. This is the primary reason why PBHs turn out to be so tightly constrained by data on background radiation. It is the low-mass PBHs whose contributions to the EBL via Hawking evaporation are the strongest.

Much subsequent effort has gone into the identification of alternative formation mechanisms which could give rise to a more favourable distribution of PBH masses (i.e. one peaked at sufficiently high mass to provide the requisite CDM density without the unwanted background radiation from the low-mass tail). We pause here to survey some of these developments. Perhaps the least speculative possibility is that PBHs arise from a post-inflationary spectrum of density fluctuations which is not perfectly scale-invariant but has a characteristic length of some kind [6]. The parameter ζ in (9.2) would then depend explicitly on

the inflationary potential (or analogous quantities). This kind of dependence has been discussed, for example, in the context of two-stage inflation [7], extended inflation [8], chaotic inflation [9], ‘plateau’ inflation [10], hybrid inflation [11] and inflation via isocurvature fluctuations [12].

A narrow spectrum of masses might also be expected if PBHs formed during a spontaneous **phase transition** rather than arising from primordial fluctuations. The quark–hadron transition [13], grand unified symmetry-breaking transition [14] and Weinberg–Salam phase transition [15] have all been considered in this regard. The initial mass distribution in each case would be peaked near the horizon mass M_f at transition time. The quark–hadron transition has attracted particular attention because PBH formation at this time would be enhanced by a temporary softening of the equation of state; and because M_f for this case is coincidentally close to M_\odot , so that PBHs might be responsible for MACHO observations of microlensing in the halo [16]. **Cosmic string loops** have also been explored as possible seeds for PBHs with a peaked mass spectrum [17, 18]. And considerable interest has recently been generated by the discovery that PBHs could provide a physical realization of the theoretical phenomenon known as **critical collapse** [19]. If this is so, then initial PBH masses would no longer necessarily be clustered near M_f .

While any of these proposals can, in principle, concentrate the PBH population within a narrow mass range, all of them face the same problem of *fine-tuning* if they are to produce the desired present-day density of PBHs. In the case of inflationary mechanisms it is the form of the potential which must be adjusted. In others it is the bubble nucleation rate, the string mass per unit length or the fraction of the Universe going into PBHs at formation time. The degree of fine-tuning required is typically of order one part in 10^9 . Thus, while modifications of the initial mass distribution can weaken the ‘standard’ constraints on PBH properties (which we derive later), they do not as yet have a compelling physical basis. Similar comments apply to PBH-based explanations for specific classes of observational phenomena. It has been suggested, for instance, that PBHs with the right mass could be responsible for certain kinds of **γ -ray bursts** [20–22] or for long-term **quasar variability** via microlensing [23, 24]. Other connections have been drawn to diffuse γ -ray emission from the galactic halo [25, 26] as well as the MACHO events mentioned earlier [27, 28]. These suggestions are intriguing, but experimental confirmation of any one (or more) of them would raise almost as many questions as it answers.

9.3 Evolution and number density

In order to obtain the comoving number density of PBHs from their initial mass distribution, we use the fact that PBHs evaporate at a rate which is inversely

proportional to the square of their masses:

$$\frac{dM}{dt} = -\frac{\alpha}{M^2}. \quad (9.3)$$

This applies to uncharged, non-rotating black holes, which is a reasonable approximation in the case of PBHs since these objects discharge quickly relative to their lifetimes [29] and also give up angular momentum by preferentially emitting particles with spin [30]. The parameter α depends in general on the PBH mass M and its behaviour was worked out in detail by Page in 1976 [31]. The PBHs which are of most importance for our purposes are those with $4.5 \times 10^{14} \text{ g} \leq M \leq 9.4 \times 10^{16} \text{ g}$. Black holes in this range are light enough (and therefore ‘hot’ enough) to emit massless particles (including photons) as well as ultrarelativistic electrons and positrons. The corresponding value of α is

$$\alpha = 6.9 \times 10^{25} \text{ g}^3 \text{ s}^{-1}. \quad (9.4)$$

For $M > 9.4 \times 10^{16} \text{ g}$, the value of α drops to $3.8 \times 10^{25} \text{ g}^3 \text{ s}^{-1}$ because the larger black hole is ‘cooler’ and no longer able to emit electrons and positrons. EBL contributions from PBHs of this mass are of lesser importance because of the shape of the mass distribution.

As the PBH mass drops below $4.5 \times 10^{14} \text{ g}$, its energy kT climbs above the rest energies of successively heavier particles, beginning with muons and pions. As each mass threshold is passed, the PBH is able to emit more particles and the value of α increases further. At temperatures above the **quark–hadron transition** ($kT \approx 200 \text{ MeV}$), MacGibbon and Webber have shown that relativistic quark and gluon jets are likely to be emitted rather than massive elementary particles [32]. These jets subsequently fragment into stable particles, and the photons produced in this way are actually more important (at these energies) than the primary photon flux. The precise behaviour of α in this regime depends on one’s choice of particle physics. A plot of $\alpha(M)$ for the standard model is found in the review by Halzen *et al* [33], who also note that α climbs to $7.8 \times 10^{26} \text{ g}^3 \text{ s}^{-1}$ at $kT = 100 \text{ GeV}$, and that its value would be at least three times higher in supersymmetric extensions of the standard model where there are many more particle states to be emitted.

As we will shortly see, however, EBL contributions from PBHs at these temperatures are suppressed by the fact that the latter have already evaporated away all their mass and vanished. If we assume for the moment that PBH evolution is adequately described by (9.3) with $\alpha = \text{constant}$ as given by (9.4), then integration gives

$$M(t) = (M_i^3 - 3\alpha t)^{1/3}. \quad (9.5)$$

The lifetime t_{pbh} of a PBH is found by setting $M(t_{\text{pbh}}) = 0$, giving $t_{\text{pbh}} = M_i^3/3\alpha$. Therefore the initial mass of a PBH which is just disappearing today ($t_{\text{pbh}} = t_0$) is given by

$$M_* = (3\alpha t_0)^{1/3}. \quad (9.6)$$

Taking $t_0 = 16$ Gyr and using (9.4) for α , we find that $M_* = 4.7 \times 10^{14}$ g. A numerical analysis allowing for changes in the value of α over the full range of PBH masses with $0.06 \leq \Omega_{\text{m},0} \leq 1$ and $0.4 \leq h_0 \leq 1$ leads to a somewhat larger result [33]:

$$M_* = (5.7 \pm 1.4) \times 10^{14} \text{ g.} \quad (9.7)$$

PBHs with $M \approx M_*$ are exploding at redshift $z \approx 0$ and consequently dominate the spectrum of EBL contributions. The parameter M_* is therefore of central importance in what follows.

We now obtain the comoving number density of PBHs with masses between M and $M + dM$ at any time t . Since this is the same as the comoving number density of PBHs with *initial* masses between M_i and $M_i + dM_i$ at formation time, we can write

$$n(M, t) dM = n(M_i) dM_i. \quad (9.8)$$

Inverting equation (9.5) to get $M_i = (M^3 + 3\alpha t)^{1/3}$ and differentiating, we find from equations (9.2) and (9.8) that

$$n(M, t) dM = \mathcal{N} \left(\frac{M}{M_*} \right)^2 \left[\left(\frac{M}{M_*} \right)^3 + \left(\frac{t}{t_0} \right) \right]^{-(\beta+2)/3} d \left(\frac{M}{M_*} \right) \quad (9.9)$$

where we have used (9.6) to replace M_*^3 with $3\alpha t_0$. Here the parameter \mathcal{N} is formally given in terms of the various parameters at PBH formation time by $\mathcal{N} = (\zeta \rho_f / M_f) (M_f / M_*)^{\beta-1}$ and has the dimensions of a number density. As we will see, it corresponds roughly to the comoving number density of PBHs of mass M_* . Following Page and Hawking [34], we allow \mathcal{N} to move up or down as required by observational constraints. The theory to this point is thus specified by two adjustable input parameters: the **PBH normalization** \mathcal{N} and the **equation-of-state parameter** β . In terms of the dimensionless variables $\mathcal{M} \equiv M/M_*$ and $\tau \equiv t/t_0$ (9.9) reads

$$n(\mathcal{M}, \tau) d\mathcal{M} = \mathcal{N} \mathcal{M}^2 (\mathcal{M}^3 + \tau)^{-(\beta+2)/3} d\mathcal{M}. \quad (9.10)$$

This is the comoving number density of PBHs with mass ratios between \mathcal{M} and $\mathcal{M} + d\mathcal{M}$ at time $t = t_0 \tau$.

9.4 Cosmological density

To obtain the present mass density of PBHs with mass ratios between \mathcal{M} and $\mathcal{M} + d\mathcal{M}$, we multiply equation (9.10) by the PBH mass $M = M_* \mathcal{M}$ and put $\tau = 1$ so that

$$\rho_{\text{pbh}}(\mathcal{M}, 1) d\mathcal{M} = \mathcal{N} M_* \mathcal{M}^{1-\beta} (1 + \mathcal{M}^{-3})^{-(\beta+2)/3} d\mathcal{M}. \quad (9.11)$$

The total PBH density is then found by integrating over \mathcal{M} from zero to infinity. The integral can be solved by changing variables to $x \equiv \mathcal{M}^{-3}$, whereupon

$$\rho_{\text{pbh}} = \frac{1}{3} \mathcal{N} M_* \int_0^\infty x^{a-1} (1-x)^{-(a+b)} dx. \quad (9.12)$$

Here $a \equiv \frac{1}{3}(\beta - 2)$ and $b \equiv \frac{4}{3}$. The solution is

$$\rho_{\text{pbh}} = k_\beta \mathcal{N} M_* \quad k_\beta \equiv \frac{\Gamma(a)\Gamma(b)}{3\Gamma(a+b)} \quad (9.13)$$

where $\Gamma(x)$ is the gamma function. Allowing β to take values from 2 through $\frac{5}{2}$ (the most natural situation) and up to 4, we obtain:

$$k_\beta = \begin{cases} \infty & (\beta = 2) \\ 1.87 & (\beta = 2.5) \\ 0.88 & (\beta = 3) \\ 0.56 & (\beta = 3.5) \\ 0.40 & (\beta = 4). \end{cases} \quad (9.14)$$

The combined mass density of all PBHs in the Universe is thus of order $\rho_{\text{pbh}} \approx \mathcal{N} M_*$. That is, \mathcal{N} can be thought of as the present number density of PBHs, if the latter are imagined to consist predominantly of objects with $M \approx M_*$. Equation (9.13) can be recast as a relation between \mathcal{N} and the PBH density parameter $\Omega_{\text{pbh}} = \rho_{\text{pbh}}/\rho_{\text{crit},0}$:

$$\Omega_{\text{pbh}} = \frac{k_\beta \mathcal{N} M_*}{\rho_{\text{crit},0}}. \quad (9.15)$$

PBH number density \mathcal{N} and density Ω_{pbh} are thus interchangeable as free parameters. If we adopt the most natural value for β ($=2.5$) together with an upper limit due to Page and Hawking of $\mathcal{N} \lesssim 10^4 \text{ pc}^{-3}$ [34], then equations (2.36), (9.7), (9.14) and (9.15) together imply that Ω_{pbh} is, at most, of order $\sim 10^{-8} h_0^{-2}$. If this upper limit on \mathcal{N} holds (as we will confirm under the simplest assumptions), then there is little hope for PBHs to make up the dark matter.

Equation (9.14) shows that one way to boost their importance would be to assume a soft equation of state at formation time (i.e. a value of β close to 2 as for dustlike matter, rather than 2.5 as for radiation). Physically this is related to the fact that low-pressure matter offers little resistance to gravitational collapse. Such a softening has, in fact, been shown to occur during the quark-hadron transition [16], leading to increases in Ω_{pbh} for PBHs which form at that time (subject to the fine-tuning problem noted in section 9.2). For PBHs which arise from primordial density fluctuations, however, conditions of this kind are unlikely to hold throughout the formation epoch. In the limit $\beta \rightarrow 2$ equation (9.2) breaks down in any case because it becomes possible for PBHs to form on scales smaller than the horizon [4].

9.5 Spectral energy distribution

Hawking [35] proved that an uncharged, non-rotating black hole emits bosons (such as photons) in any given quantum state with energies between E and $E + dE$ at the rate

$$d\dot{N} = \frac{\Gamma_s dE}{2\pi\hbar[\exp(E/kT) - 1]}. \quad (9.16)$$

Here T is the effective **black-hole temperature**, and Γ_s is the **absorption coefficient** or probability that the same particle would be absorbed by the black hole if incident upon it in this state. The function $d\dot{N}$ is related to the spectral energy distribution (SED) of the black hole by $d\dot{N} = F(\lambda, \mathcal{M}) d\lambda/E$, since we have defined $F(\lambda, \mathcal{M}) d\lambda$ as the energy emitted between wavelengths λ and $\lambda + d\lambda$. Here we anticipate the fact that F will depend explicitly on the PBH mass \mathcal{M} as well as wavelength. The PBH SED thus satisfies

$$F(\lambda, \mathcal{M}) d\lambda = \frac{\Gamma_s E dE}{2\pi\hbar[\exp(E/kT) - 1]}. \quad (9.17)$$

The absorption coefficient Γ_s is, in general, a complicated function of \mathcal{M} and E as well as the quantum numbers s (spin), ℓ (total angular momentum) and m (axial angular momentum) of the emitted particles. Its form was first calculated by Page [31]. At high energies, and in the vicinity of the peak of the emitted spectrum, a good approximation is given by [36]

$$\Gamma_s \propto M^2 E^2. \quad (9.18)$$

This approximation breaks down at low energies, where it gives rise to errors of order 50% for $(GME/\hbar c^3) \sim 0.05$ [37] or (with $E = 2\pi\hbar c/\lambda$ and $M \sim M_*$) for $\lambda \sim 10^{-3}$ Å. This is adequate for our purposes, as we will find that the strongest constraints on PBHs come from those with masses $M \sim M_*$ at wavelengths $\lambda \sim 10^{-4}$ Å.

Putting (9.18) into (9.17) and making the change of variable to wavelength $\lambda = \hbar c/E$, we obtain the SED

$$F(\lambda, \mathcal{M}) d\lambda = \frac{C \mathcal{M}^2 \lambda^{-5} d\lambda}{\exp(\hbar c/kT\lambda) - 1} \quad (9.19)$$

where C is a proportionality constant. This has the same form as the blackbody spectrum, equation (3.22). One should, however, keep in mind that we have made three simplifying assumptions in arriving at this equation. First, we have neglected the black-hole charge and spin (as justified in section 9.3). Second, we have used an approximation for the absorption coefficient Γ_s . And third, we have treated all the emitted photons as if they are in the same quantum state whereas, in fact, the emission rate (9.16) applies separately to the $\ell = s$ ($= 1$), $\ell = s + 1$ and $\ell = s + 2$ modes. There are thus actually *three distinct* quasi-blackbody photon

spectra with different characteristic temperatures for any single PBH. However, Page [31] has demonstrated that the $\ell = s$ mode is overwhelmingly dominant, with the $\ell = s+1$ and $\ell = s+2$ modes contributing less than 1% and 0.01% of the total photon flux respectively. Equation (9.19) is thus a reasonable approximation to the SED of the PBH as a whole.

To fix the value of C we use the fact that the total flux of photons (in all modes) radiated by a black hole of mass M is given [31]

$$\dot{N} = \int d\dot{N} = \int_{\lambda=0}^{\infty} \frac{F(\lambda, \mathcal{M}) d\lambda}{hc/\lambda} = 5.97 \times 10^{34} \text{ s}^{-1} \left(\frac{M}{1 \text{ g}} \right)^{-1}. \quad (9.20)$$

Inserting (9.19) and recalling that $M = M_* \mathcal{M}$, we find that

$$C \int_0^{\infty} \frac{\lambda^{-4} d\lambda}{\exp(hc/kT\lambda) - 1} = (5.97 \times 10^{34} \text{ g s}^{-1}) \frac{hc}{M_* \mathcal{M}^3}. \quad (9.21)$$

The definite integral on the left-hand side of this equation can be solved by switching variables to $\nu = c/\lambda$:

$$\int_0^{\infty} \frac{\nu^2 d\nu/c^3}{\exp(h\nu/kT) - 1} = \left(\frac{hc}{kT} \right)^{-3} \Gamma(3)\zeta(3) \quad (9.22)$$

where $\Gamma(n)$ and $\zeta(n)$ are the gamma function and Riemann zeta function respectively. We then apply the fact that the temperature T of an uncharged, non-rotating black hole is given by

$$T = \frac{\hbar c^3}{8\pi\kappa GM}. \quad (9.23)$$

Putting (9.22) and (9.23) into (9.21) and rearranging terms leads to

$$C = (5.97 \times 10^{34} \text{ g s}^{-1}) \frac{(4\pi)^6 h G^3 M_*^2}{c^5 \Gamma(3)\zeta(3)}. \quad (9.24)$$

Using $\Gamma(3) = 2! = 2$ and $\zeta(3) = \sum_{k=1}^{\infty} k^{-3} = 1.201$ along with (9.7) for M_* , we find

$$C = (270 \pm 120) \text{ erg } \text{\AA}^4 \text{ s}^{-1}. \quad (9.25)$$

We can also use the definitions (9.23) to define a useful new quantity:

$$\lambda_{\text{pbh}} \equiv \frac{hc}{kT\mathcal{M}} = \left(\frac{4\pi}{c} \right)^2 GM_* = (6.6 \pm 1.6) \times 10^{-4} \text{ \AA}. \quad (9.26)$$

The size of this characteristic wavelength tells us that we will be concerned primarily with the high-energy γ -ray portion of the spectrum. In terms of C and λ_{pbh} the SED (9.19) now reads

$$F(\lambda, \mathcal{M}) = \frac{C \mathcal{M}^2 / \lambda^5}{\exp(\mathcal{M} \lambda_{\text{pbh}} / \lambda) - 1}. \quad (9.27)$$

While this contains no explicit time-dependence, the spectrum does of course depend on time through the PBH mass ratio \mathcal{M} .

9.6 Luminosity

To compute the PBH luminosity we employ equation (3.1), integrating the SED $F(\lambda, \mathcal{M})$ over all wavelengths to obtain:

$$L(\mathcal{M}) = C\mathcal{M}^2 \int_0^\infty \frac{\lambda^{-5} d\lambda}{\exp(\mathcal{M}\lambda_{\text{pbh}}/\lambda) - 1}. \quad (9.28)$$

This definite integral is also solved by means of a change of variable to frequency ν , with the result that

$$L(\mathcal{M}) = C\mathcal{M}^2 (\mathcal{M}\lambda_{\text{pbh}})^{-4} \Gamma(4)\zeta(4). \quad (9.29)$$

Using equations (9.7), (9.24) and (9.26) along with the values $\Gamma(4) = 3! = 6$ and $\zeta(4) = \pi^4/96$, we can put this into the form

$$L(\mathcal{M}) = L_{\text{pbh}}\mathcal{M}^{-2} \quad (9.30)$$

where

$$L_{\text{pbh}} = \frac{(5.97 \times 10^{34} \text{ g s}^{-1})\pi^2 hc^3}{512\zeta(3)GM_*^2} = (1.0 \pm 0.4) \times 10^{16} \text{ erg s}^{-1}.$$

Compared to the luminosity of an ordinary star, the typical PBH (of mass ratio $\mathcal{M} \approx 1$) is not very luminous. A PBH of 900 kg or so might theoretically be expected to reach the Sun's luminosity; however, in practice it would already have exploded, having long since reached an effective temperature high enough to emit a wide range of massive particles as well as photons. The low luminosity of black holes in general can be emphasized by using the relation $\mathcal{M} \equiv M/M_*$ to recast equation (9.30) in the form

$$\frac{L}{L_\odot} = 1.7 \times 10^{-55} \left(\frac{M}{M_\odot} \right)^{-2}. \quad (9.31)$$

This expression is not strictly valid for PBHs of masses near M_\odot , having been derived for those with $M \sim M_* \sim 10^{15}$ g. For more massive PBHs, the luminosity is if anything *lower*. (In fact, a black hole of Solar mass would be colder than the CMB and would absorb radiation faster than it could emit it.) So, Hawking evaporation or not, most black holes are indeed very black.

9.7 Bolometric intensity

To obtain the total bolometric intensity of PBHs out to a look-back time $t_0 - t_f$, we substitute the PBH number density (9.10) and luminosity (9.28) into the integral (2.14) as usual. This time, however, number density $n(t)$ is to be replaced by $n(\mathcal{M}, \tau) d\mathcal{M}$, the number density of PBHs with mass ratios between \mathcal{M} and

$\mathcal{M} + d\mathcal{M}$ at time $t = t_0\tau$. Similarly, $L(\mathcal{M})$ takes the place of $L(t)$. We then integrate over PBH mass \mathcal{M} as well as time t in order to obtain the total intensity. Employing τ as an integration variable, we find:

$$Q = ct_0\mathcal{N}L_{\text{pbh}} \int_{\tau_f}^1 \tilde{R}(\tau) d\tau \int_{\mathcal{M}_c(\tau)}^{\infty} \frac{d\mathcal{M}}{(\mathcal{M}^3 + \tau)^\varepsilon}. \quad (9.32)$$

Here $\varepsilon \equiv (\beta + 2)/3$ and $\mathcal{M}_c(\tau)$ is a minimum **cut-off mass**, equal to the mass of the lightest PBH which has not yet evaporated at time τ . This arises because the initial PBH mass distribution (9.2) must, in general, have a non-zero minimum M_{min} in order to avoid divergences at low mass. As soon as the lightest PBHs have formed and started to evaporate, however, the cut-off begins to drop from its initial value of $\mathcal{M}_c(0) = M_{\text{min}}/M_*$. If M_{min} is of the order of the Planck mass as suggested by Barrow *et al* [38], then one finds that $\mathcal{M}_c(\tau)$ drops to zero well before the end of the radiation-dominated era. Since we are concerned with times later than this, we can safely set $\mathcal{M}_c(\tau) = 0$ in what follows.

Using (9.15) for \mathcal{N} , we therefore rewrite equation (9.32) as

$$Q = Q_{\text{pbh}}\Omega_{\text{pbh}} \int_{\tau_f}^1 \tilde{R}(\tau) d\tau \int_0^{\infty} \frac{d\mathcal{M}}{(\mathcal{M}^3 + \tau)^\varepsilon} \quad (9.33)$$

where

$$Q_{\text{pbh}} = \frac{ct_0\rho_{\text{crit},0}L_{\text{pbh}}}{k_\beta M_*}. \quad (9.34)$$

In this form, equation (9.33) can be used to put a rough upper limit on Ω_{pbh} from the bolometric intensity of the background light [39]. Let us assume that the Universe is flat, as suggested by most observations (chapter 4). Then its age t_0 can be obtained from equation (2.70) as

$$t_0 = \frac{2\tilde{t}_0}{3H_0}. \quad (9.35)$$

Here $\tilde{t}_0 \equiv \tilde{t}_m(0)$ where $\tilde{t}_m(z)$ is the dimensionless function

$$\tilde{t}_m(z) \equiv \frac{1}{\sqrt{1 - \Omega_{m,0}}} \sinh^{-1} \sqrt{\frac{1 - \Omega_{m,0}}{\Omega_{m,0}(1+z)^3}}. \quad (9.36)$$

Equation (9.35) is necessary if we are to obtain the right dependence on Hubble's constant h_0 in our results, and equation (9.36) will prove useful in specifying not only the value of t_0 but τ_f as well. Putting (9.35) into (9.34) and using equations (2.19), (2.36), (9.7) and (9.30), we find that

$$\begin{aligned} Q_{\text{pbh}} &= \frac{(5.97 \times 10^{34} \text{ g s}^{-1})\pi^2 hc^4 \rho_{\text{crit},0} \tilde{t}_0}{768\zeta(3)GH_0 k_\beta M_*^3} \\ &= (2.3 \pm 1.4)h_0 \tilde{t}_0 k_\beta^{-1} \text{ erg s}^{-1} \text{ cm}^{-2}. \end{aligned} \quad (9.37)$$

We are now ready to evaluate equation (9.33). To begin with we note that the integral over mass has an analytic solution:

$$\int_0^\infty \frac{d\mathcal{M}}{(\mathcal{M}^3 + \tau)^\varepsilon} = k_\varepsilon \tau^{\frac{1}{3}-\varepsilon} \quad k_\varepsilon \equiv \frac{\Gamma(\frac{1}{3})\Gamma(\varepsilon - \frac{1}{3})}{3\Gamma(\varepsilon)}. \quad (9.38)$$

If we further specialize to the EdS case ($\Omega_{m,0} = 1$) then $\tilde{t}_0 = 1$ and equation (2.61) implies:

$$\tilde{R}(\tau) = \tau^{2/3}. \quad (9.39)$$

Putting equations (9.38) and (9.39) into (9.33), we find that

$$Q = Q_{\text{pbh}} \Omega_{\text{pbh}} k_\varepsilon \int_{\tau_f}^1 \tau^{1-\varepsilon} d\tau = Q_{\text{pbh}} \Omega_{\text{pbh}} k_\varepsilon \left(\frac{1 - \tau_f^{2-\varepsilon}}{2 - \varepsilon} \right). \quad (9.40)$$

The parameter τ_f is simply obtained for the EdS case by inverting (9.39) to give $\tau_f = (1 + z_f)^{-3/2}$. The subscript ‘f’ (by which we usually mean ‘formation’) is here a misnomer since we do not integrate back to PBH formation time, which occurred in the early stages of the radiation-dominated era. Rather we integrate back to the redshift at which processes like pair production become significant enough to render the Universe approximately opaque to the (primarily γ -ray) photons from PBH evaporation. Following Kribs *et al* [37] this is $z_f \approx 700$.

Using this value of z_f and substituting equations (9.37) and (9.38) into (9.40), we find that the bolometric intensity of background radiation due to evaporating PBHs in an EdS Universe is

$$Q = h_0 \Omega_{\text{pbh}} \times \begin{cases} 0 & (\beta = 2) \\ 2.3 \pm 1.4 \text{ erg s}^{-1} \text{ cm}^{-2} & (\beta = 2.5) \\ 6.6 \pm 4.2 \text{ erg s}^{-1} \text{ cm}^{-2} & (\beta = 3) \\ 17 \pm 10 \text{ erg s}^{-1} \text{ cm}^{-2} & (\beta = 3.5) \\ 45 \pm 28 \text{ erg s}^{-1} \text{ cm}^{-2} & (\beta = 4). \end{cases} \quad (9.41)$$

This vanishes for $\beta = 2$ because $k_\beta \rightarrow \infty$ in this limit, as discussed in section 9.4. The case $\beta = 4$ (i.e. $\varepsilon = 2$) is evaluated with the help of L’Hôpital’s rule, which gives $\lim_{\varepsilon \rightarrow 2} (1 - \tau_f^{2-\varepsilon}) / (2 - \varepsilon) = -\ln \tau_f$.

All the values of Q in (9.41) are far higher than the actual bolometric EBL intensity in an EdS Universe, which is $\frac{2}{5} Q_* = 1.0 \times 10^{-4} \text{ erg s}^{-1} \text{ cm}^{-2}$ from (2.49). Moreover this background is already well accounted for by the integrated light from galaxies. A firm upper bound on Ω_{pbh} (for the most natural situation with $\beta = 2.5$) is therefore

$$\Omega_{\text{pbh}} < (4.4 \pm 2.8) \times 10^{-5} h_0^{-1}. \quad (9.42)$$

For harder initial equations of state ($\beta > 2.5$) the PBH density would have to be even lower. PBHs in the simplest formation scenario are thus eliminated as important dark-matter candidates, even without reference to the γ -ray spectrum.

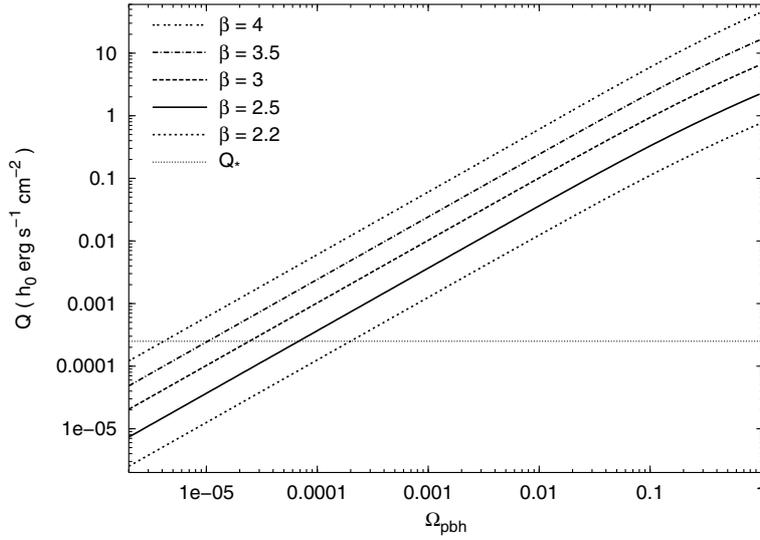


Figure 9.1. The bolometric intensity due to evaporating primordial black holes as a function of their collective density Ω_{pbh} and the equation-of-state parameter β . We have assumed that $\Omega_{\text{m},0} = \Omega_{\text{bar}} + \Omega_{\text{pbh}}$ with $\Omega_{\text{bar}} = 0.016h_0^{-2}$, $h_0 = 0.75$ and $\Omega_{\Lambda,0} = 1 - \Omega_{\text{m},0}$. The horizontal dotted line indicates the approximate bolometric intensity (Q_*) of the observed EBL.

This conclusion is based on equation (9.30) for PBH luminosity, which is valid over the most important part of the integral. Corrections for higher and lower masses will not alter the bound (9.42) by more than an order of magnitude.

The results in (9.41) assume an EdS cosmology. For more general flat models containing both matter and vacuum energy, integrated background intensity would go up, because the Universe is older, and down because $Q \propto \Omega_{\text{pbh}}$. The latter effect is the stronger one, so that our constraints on Ω_{pbh} will be weakened in a model such as Λ CDM (with $\Omega_{\text{m},0} = 0.3$, $\Omega_{\Lambda,0} = 0.7$). To determine the importance of this effect, we can re-evaluate the integral (9.33) using the general formula (2.68) for $\tilde{R}(\tau)$ in place of (9.39). We will make the minimal assumption that PBHs constitute the *only* CDM, so that $\Omega_{\text{m},0} = \Omega_{\text{pbh}} + \Omega_{\text{bar}}$ with Ω_{bar} given by (4.4) as usual. Equation (2.70) shows that the parameter τ_f is given for arbitrary values of $\Omega_{\text{m},0}$ by $\tau_f = \tilde{t}_m(z_f)/\tilde{t}_0$ where the function $\tilde{t}_m(z)$ is defined as before by (9.36).

Carrying out the integration in (9.33), we obtain the plots of bolometric intensity Q as a function of Ω_{pbh} shown in figure 9.1. As before Q is proportional to h_0 because it goes as both $\rho_{\text{pbh}} = \Omega_{\text{pbh}}\rho_{\text{crit},0} \propto h_0^2$ and $t_0 \propto h_0^{-1}$. Since $Q \rightarrow 0$ for $\beta \rightarrow 2$ we have chosen a minimum value of $\beta = 2.2$ as being representative of ‘soft’ conditions.

Figure 9.1 confirms that, regardless of cosmological model, PBH contributions to the background light are too high unless $\Omega_{\text{pbh}} \ll 1$. The values in equation (9.41) are recovered at the right-hand edge of the figure where Ω_{pbh} approaches one, as expected. For all other models, if we impose a conservative upper bound $Q < Q_*$ (as indicated by the faint dotted line) then it follows that $\Omega_{\text{pbh}} < (6.9 \pm 4.2) \times 10^{-5} h_0^{-1}$ for $\beta = 2.5$. This is about 60% higher than the limit (9.42) for the EdS case.

9.8 Spectral intensity

Stronger limits on PBH density can be obtained from the γ -ray background, where these objects contribute most strongly to the EBL and where we have good data (as summarized in section 8.7). Spectral intensity is found as usual by substituting the comoving PBH number density (9.10) and SED (9.27) into equation (3.5). As in the bolometric case, we now integrate over PBH mass \mathcal{M} as well as time $\tau = t/t_0$, so that

$$I_\lambda(\lambda_0) = \frac{ct_0}{4\pi} \int_{\tau_f}^1 \tilde{R}^2(\tau) d\tau \int_{\mathcal{M}_c(\tau)}^\infty n(\mathcal{M}, \tau) F(\tilde{R}\lambda_0, \mathcal{M}) d\mathcal{M}. \quad (9.43)$$

Following the discussion in section 9.7 we set $\mathcal{M}_c(\tau) = 0$. In light of our bolometric results it is unlikely that PBHs make up a significant part of the dark matter, so we no longer tie the value of $\Omega_{\text{m},0}$ to Ω_{pbh} . Models with $\Omega_{\text{m},0} > \Omega_{\text{bar}}\Omega_{\text{pbh}}$ must therefore contain a second species of cold dark matter (other than PBHs) to provide the required matter density. Putting (9.10) and (9.27) into (9.43) and using (9.7), (9.15) and (9.24), we find that

$$I_\lambda(\lambda_0) = I_{\text{pbh}}\Omega_{\text{pbh}} \int_{\tau_f}^1 \tilde{R}^{-3}(\tau) d\tau \int_0^\infty \frac{\mathcal{M}^4 (\mathcal{M}^3 + \tau)^{-\epsilon} d\mathcal{M}}{\exp[\lambda_{\text{pbh}}\mathcal{M}/\tilde{R}(\tau)\lambda_0] - 1}. \quad (9.44)$$

Here the dimensional prefactor is a function of both β and λ_0 and reads

$$\begin{aligned} I_{\text{pbh}} &= \frac{(5.97 \times 10^{34} \text{ g s}^{-1})(4\pi)^5 G^3 M_* \rho_{\text{crit},0} \tilde{t}_0}{3\zeta(3)c^5 k_\beta H_0 \lambda_0^4} \\ &= [(2.1 \pm 0.5) \times 10^{-7} \text{ CUs}] h_0 k_\beta^{-1} \tilde{t}_0 \left(\frac{\lambda_0}{\text{\AA}}\right)^{-4}. \end{aligned} \quad (9.45)$$

We have divided through by the photon energy hc/λ_0 to put the results in units of CUs as usual. The range of uncertainty in $I_\lambda(\lambda_0)$ is noticeably smaller than that in Q , equation (9.37). This is because $I_\lambda(\lambda_0)$ depends only linearly on M_* whereas Q is proportional to M_*^{-3} . (This, in turn, results from the fact that $I_\lambda \propto C \propto M_*^2$ whereas $Q \propto L_{\text{pbh}} \propto M_*^{-2}$. One more factor of M_*^{-1} comes from $\mathcal{N} \propto \rho_{\text{pbh}}/M_*$ in both cases.) Like Q , I_λ depends linearly on h_0 since integrated intensity in either case is proportional to both $\rho_{\text{pbh}} \propto \rho_{\text{crit},0} \propto h_0^2$ and $t_0 \propto h_0^{-1}$.

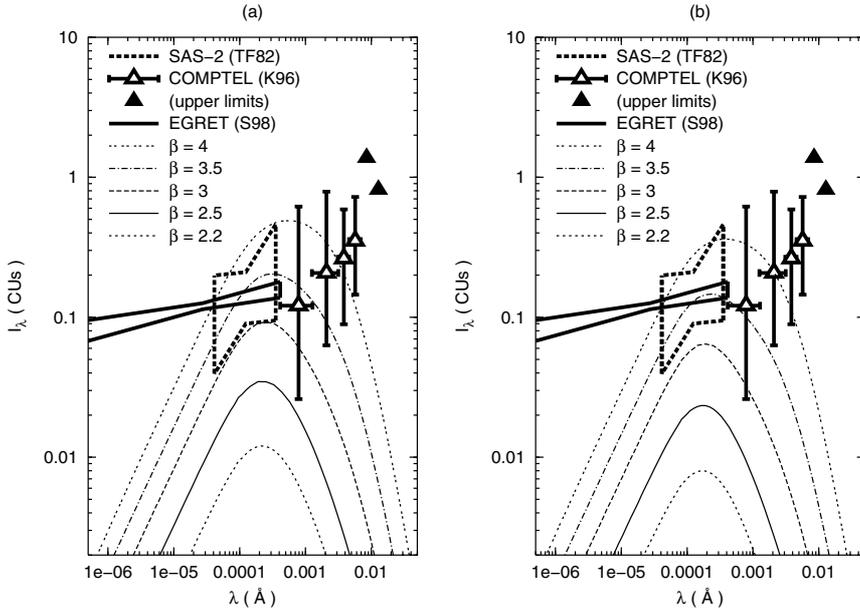


Figure 9.2. The spectral intensity of the diffuse γ -ray background from evaporating primordial black holes in flat models, as compared with experimental limits from the SAS-2, COMPTEL and EGRET instruments. The left-hand panel (a) assumes $\Omega_{m,0} = 0.06$, while the right-hand panel (b) is plotted for $\Omega_{m,0} = 1$ (the EdS case). All curves assume $\Omega_{\text{pbh}} = 10^{-8}$ and $h_0 = 0.75$.

Numerical integration of equation (9.44) leads to the plots in figure 9.2, where we have set $\Omega_{\text{pbh}} = 10^{-8}$. Following Page and Hawking [34] we have chosen values of $\Omega_{m,0} = 0.06$ in the left-hand panel (a) and $\Omega_{m,0} = 1$ in the right-hand panel (b). (Results are not strictly comparable in the former case, however, since we assume that $\Omega_{\Lambda,0} = 1 - \Omega_{m,0}$ rather than $\Omega_{\Lambda,0} = 0$.) Our results are in good agreement with the earlier ones except at the longest wavelengths (lowest energies), where PBH evaporation is no longer well described by a simple blackbody SED, and where the spectrum begins to be affected by pair production on nuclei. As expected the spectra peak near 10^{-4} Å in the γ -ray region. Also plotted in figure 9.2 are the data from SAS-2 [40], COMPTEL [41] and EGRET [42] (bold lines and points).

By adjusting the value of Ω_{pbh} up or down from its value of 10^{-8} in figure 9.2, we can match the theoretical PBH spectra to those measured (e.g. by EGRET), thereby obtaining the maximum value of Ω_{pbh} consistent with

observation. For $\beta = 2.5$ this results in

$$\Omega_{\text{pbh}} < \begin{cases} (4.2 \pm 1.1) \times 10^{-8} h_0^{-1} & (\Omega_{\text{m},0} = 0.06) \\ (6.2 \pm 1.6) \times 10^{-8} h_0^{-1} & (\Omega_{\text{m},0} = 1). \end{cases} \quad (9.46)$$

These limits are three orders of magnitude stronger than the one from bolometric intensity, again confirming that PBHs in the simplest formation scenario cannot be significant contributors to the dark matter. Equation (9.46) can be put into the form of an upper limit on the PBH number-density normalization \mathcal{N} with the help of (9.15), giving

$$\mathcal{N} < \begin{cases} (2.2 \pm 0.8) \times 10^4 h_0 \text{ pc}^{-3} & (\Omega_{\text{m},0} = 0.06) \\ (3.2 \pm 1.1) \times 10^4 h_0 \text{ pc}^{-3} & (\Omega_{\text{m},0} = 1). \end{cases} \quad (9.47)$$

These numbers are in good agreement with the original Page–Hawking bound of $\mathcal{N} < 1 \times 10^4 \text{ pc}^{-3}$, which was obtained for $h_0 = 0.6$ [34].

Subsequent workers have refined the γ -ray background constraints on Ω_{pbh} and \mathcal{N} in a number of ways. An important development was the realization by MacGibbon and Webber [32] that PBHs whose effective temperatures have climbed above the rest energy of hadrons are likely to give off more photons by *indirect* processes than by direct emission. This is because hadrons are composite particles, made up of quarks and gluons. It is these elementary constituents (rather than their composite bound states) which should be emitted from a very hot PBH in the form of relativistic **quark and gluon jets**. Numerical simulations and accelerator experiments indicate that these jets subsequently fragment into secondary particles whose decays (especially those of the pions) produce a far greater flux of photons than that emitted directly from the PBH. The net effect is to increase the PBH luminosity, particularly in low-energy γ -rays, strengthening the constraint on Ω_{pbh} by about an order of magnitude [36]. The most recent upper limit obtained in this way using EGRET data (assuming $\Omega_{\text{m},0} = 1$) is $\Omega_{\text{pbh}} < (5.1 \pm 1.3) \times 10^{-9} h_0^{-2}$ [43].

Complementary bounds on PBH contributions to the dark matter have come from *direct* searches for those evaporating within a few kpc of the Earth. Such limits are subject to more uncertainty than ones based on the EBL because they depend on assumptions about the degree to which PBHs are clustered. If there is *no* clustering then (9.42) can be converted into a stringent upper bound on the **local PBH evaporation rate**, $\dot{\mathcal{N}} < 10^{-7} \text{ pc}^{-3} \text{ yr}^{-1}$. This, however, relaxes to $\dot{\mathcal{N}} \lesssim 10 \text{ pc}^{-3} \text{ yr}^{-1}$ if PBHs are strongly clustered [33], in which case limits from direct searches could potentially become competitive with those based on the EBL. Data taken at energies near 50 TeV with the CYGNUS air-shower array have led to a bound of $\dot{\mathcal{N}} < 8.5 \times 10^5 \text{ pc}^{-3} \text{ yr}^{-1}$ [44]; and a comparable limit of $\dot{\mathcal{N}} < (3.0 \pm 1.0) \times 10^6 \text{ pc}^{-3} \text{ yr}^{-1}$ has been obtained at 400 GeV using an imaging atmospheric Čerenkov technique developed by the Whipple collaboration [45]. The strongest constraint yet reported has come from balloon measurements of the

cosmic-ray antiproton flux below 0.5 GeV, from which it has been inferred that $\dot{N} < 0.017 \text{ pc}^{-3} \text{ yr}^{-1}$ [46].

Other ideas have been advanced which could weaken the bounds on PBHs as dark-matter candidates. It might be, for instance, that these objects leave behind **stable relics** rather than evaporating completely [47]. This, however, raises a new problem (similar to the ‘gravitino problem’ discussed in section 8.6) because such relics would have been overproduced by quantum and thermal fluctuations in the early Universe. Inflation can be invoked to reduce their density but must be finely tuned if the same relics are to make up an interesting fraction of the dark matter today [48]. A more promising possibility has been opened up by the suggestion of Heckler [49, 50] that particles emitted from the surface interact strongly enough above a critical temperature to form a **black-hole photosphere**. This would make the PBH appear cooler as seen from a distance than its actual surface temperature, just as the Solar photosphere makes the Sun appear cooler than its core. (In the case of the black hole, however, one has not only an electromagnetic photosphere but a QCD ‘gluosphere’.) The reality of this effect is still under debate [43] but preliminary calculations indicate that it could reduce the intensity of PBH contributions to the γ -ray background by 60% at 100 MeV and by as much as two orders of magnitude at 1 GeV [51].

Finally, as discussed already in section 9.2, the limits obtained here can be weakened or evaded if PBH formation occurs in such a way as to produce fewer low-mass objects. The challenge faced in such proposals is to explain how a distribution of this kind comes about in a natural way. A common procedure is to turn the question around and use observational data on the present intensity of the γ -ray background as a probe of the original PBH formation mechanism. Such an approach has been applied, for example, to put constraints on the spectral index of density fluctuations in the context of PBHs which form via critical collapse [37], or inflation with a ‘blue’ or tilted spectrum [52]. Even if they do not exist, in other words, primordial black holes provide a valuable window on conditions in the early Universe, where information is otherwise scarce.

9.9 Higher-dimensional ‘black holes’

In view of the fact that conventional black holes are disfavoured as dark-matter candidates, it is worthwhile to consider alternatives. One of the simplest of these is the extension of the black-hole concept from the four-dimensional (4D) spacetime of general relativity to higher dimensions. Higher-dimensional relativity, also known as **Kaluza–Klein gravity**, has a long history and underlies modern attempts to unify gravity with the standard model of particle physics (see [53] for a review). The extra dimensions have traditionally been assumed to be compact, in order to explain their non-appearance in low-energy physics. The past few years, however, have witnessed a surge of interest in *non-compactified* theories of higher-dimensional gravity [54–56]. In such theories

the dimensionality of spacetime can manifest itself at experimentally accessible energies. We focus on the prototypical five-dimensional (5D) case, although the extension to higher dimensions is straightforward in principle.

Let us begin by recalling that black holes in standard 4D relativity are described geometrically by the **Schwarzschild metric**. In isotropic coordinates this reads

$$ds^2 = \left(\frac{1 - GM_s/2c^2r}{1 + GM_s/2c^2r} \right)^2 c^2 dt^2 - \left(1 + \frac{GM_s}{2c^2r} \right)^4 (dr^2 + r^2 d\Omega^2) \quad (9.48)$$

where $d\Omega^2 \equiv d\theta^2 + \sin^2\theta d\phi^2$. This is a description of the static, spherically-symmetric spacetime around a pointlike object (such as a collapsed star or primordial density fluctuation) with **Schwarzschild mass** M_s . As we have seen, it is unlikely that such objects can make up the dark matter.

If the Universe has *more* than four dimensions, then the same object must be modelled with a higher-dimensional analogue of the Schwarzschild metric. Various possibilities have been explored over the years, with most attention focusing on a 5D solution discussed in detail by Gross and Perry [57], Sorkin [58] and Davidson and Owen [59] in the early 1980s. This is now generally known as the **soliton metric** and reads:

$$ds^2 = \left(\frac{ar - 1}{ar + 1} \right)^{2\xi\kappa} c^2 dt^2 - \left(\frac{a^2r^2 - 1}{a^2r^2} \right)^2 \left(\frac{ar + 1}{ar - 1} \right)^{2\xi(\kappa-1)} (dr^2 + r^2 d\Omega^2) \\ - \left(\frac{ar + 1}{ar - 1} \right)^{2\xi} dy^2. \quad (9.49)$$

Here y is the new coordinate and there are three metric parameters (a, ξ, κ) rather than just one (M_s) as in equation (9.48). Only two of these are independent, however, because a **consistency condition** (which follows from the field equations) requires that $\xi^2(\kappa^2 - \kappa + 1) = 1$. In the limit where $\xi \rightarrow 0$, $\kappa \rightarrow \infty$ and $\xi\kappa \rightarrow 1$, equation (9.49) reduces to (9.48) on 4D hypersurfaces $y = \text{constant}$. In this limit we can also identify the parameter a as $a = 2c^2/GM_s$ where M_s is the Schwarzschild mass.

We wish to understand the physical properties of this solution in four dimensions. To accomplish this we do two things. First, we assume that Einstein’s field equations in their usual form hold in the full *five*-dimensional spacetime. Second, we assume that the Universe in five dimensions is *empty*, with no 5D matter fields or cosmological constant. The field equations then simplify to

$$\mathcal{R}_{AB} = 0. \quad (9.50)$$

Here \mathcal{R}_{AB} is the 5D Ricci tensor, defined in exactly the same way as the 4D one except that spacetime indices A, B run over 0–4 instead of 0–3. Putting a 5D metric such as (9.49) into the vacuum 5D field equations (9.50), we find that we

recover the 4D field equations (2.1) with a non-zero energy–momentum tensor $\mathcal{T}_{\mu\nu}$. *Matter and energy, in other words, are induced in 4D by pure geometry in 5D.* It is by studying the properties of this **induced-matter energy–momentum tensor** ($\mathcal{T}_{\mu\nu}$) that we learn what the soliton looks like in four dimensions.

The details of the mechanism just outlined [60] and its application to solitons in particular [61, 62] have been well studied and we do not review this material here. It is important to note, however, that the Kaluza–Klein soliton differs from an ordinary black hole in several key respects. It contains a singularity at its centre; but this centre is located at $r = 1/a$ rather than $r = 0$. (The point $r = 0$ is, in fact, not even part of the manifold, which ends at $r = 1/a$.) And its event horizon, insofar as it has one, *also* shrinks to a point at $r = 1/a$. For these reasons the soliton is better classified as a naked singularity than a black hole.

Solitons in the induced-matter picture are further distinguished from conventional black holes by the fact that they have an extended matter distribution rather than having all their mass compressed into the singularity. It is this feature which proves to be of most use to us in putting constraints on solitons as dark-matter candidates [63]. The time–time component of the induced-matter energy–momentum tensor gives us the density of the solitonic fluid as a function of radial distance:

$$\rho_s(r) = \frac{c^2 \xi^2 \kappa a^6 r^4}{2\pi G (ar - 1)^4 (ar + 1)^4} \left(\frac{ar - 1}{ar + 1} \right)^{2\xi(\kappa - 1)}. \quad (9.51)$$

From the other elements of $\mathcal{T}_{\mu\nu}$ one finds that pressure can be written $p_s = \frac{1}{3}\rho_s c^2$, so that the soliton has a radiation-like equation of state. In this respect it more closely resembles a primordial black hole (which forms during the radiation-dominated era) than one which arises as the endpoint of stellar collapse. The elements of $\mathcal{T}_{\mu\nu}$ can also be used to calculate the **gravitational mass** of the fluid inside r :

$$M_g(r) = \frac{2c^2 \xi \kappa}{Ga} \left(\frac{ar - 1}{ar + 1} \right)^\xi. \quad (9.52)$$

At large distances $r \gg 1/a$ from the centre the soliton’s density (9.51) and gravitational mass (9.52) go over to

$$\rho_s(r) \rightarrow \frac{c^2 \xi^2 \kappa}{2\pi G a^2 r^4} \quad M_g(r) \rightarrow M_g(\infty) = \frac{2c^2 \xi \kappa}{Ga}. \quad (9.53)$$

The second of these expressions shows that the asymptotic value of M_g is, in general, not the same as M_s [$M_g(\infty) = \xi \kappa M_s$ for $r \gg 1/a$], but reduces to it in the limit $\xi \kappa \rightarrow 1$. Viewed in four dimensions, the soliton resembles a hole in the geometry surrounded by a spherically-symmetric ball of ultrarelativistic matter whose density falls off at large distances as $1/r^4$. If the Universe does have more than four dimensions, then objects like this should be common, being generic to 5D Kaluza–Klein gravity in exactly the same way black holes are to 4D general relativity.

Let us, therefore, assess their impact on the background radiation, using the same methods as usual, and assuming that the fluid making up the soliton is in fact composed of photons (although one might also consider ultrarelativistic particles such as neutrinos in principle). We do not have spectral information on these so we proceed bolometrically. Putting the second of equations (9.53) into the first gives

$$\rho_s(r) \approx \frac{GM_g^2}{8\pi c^2 \kappa r^4}. \quad (9.54)$$

Numbers can be attached to the quantities κ , r and M_g as follows. The first (κ) is technically a free parameter. However, a natural choice from the physical point of view is $\kappa \sim 1$. For this case the consistency relation implies $\xi \sim 1$ also, guaranteeing that the asymptotic gravitational mass of the soliton is close to its Schwarzschild one. To obtain a value for r , let us assume that solitons are distributed homogeneously through space with average separation d and mean density $\bar{\rho}_s = \Omega_s \rho_{\text{crit},0} = M_s/d^3$. Since ρ_s drops as r^{-4} whereas the number of solitons at a distance r climbs only as r^3 , the local density of solitons is largely determined by the nearest one. We can therefore replace r by $d = (M_s/\Omega_s \rho_{\text{crit},0})^{1/3}$. The last unknown in (9.54) is the soliton mass M_g ($= M_s$ if $\kappa = 1$). The fact that $\rho_s \propto r^{-4}$ is reminiscent of the density profile of the galactic dark-matter halo, equation (6.14). Theoretical work on the classical tests of 5D general relativity [64] and limits on violations of the **equivalence principle** [65] also suggests that solitons are likely to be associated with dark matter on galactic or larger scales. Let us therefore express M_s in units of the mass of the Galaxy, which from (6.15) is $M_{\text{gal}} \approx 2 \times 10^{12} M_\odot$. Equation (9.54) then gives the local energy density of solitonic fluid as

$$\rho_s c^2 \approx (3 \times 10^{-17} \text{ erg cm}^{-3}) h_0^{8/3} \Omega_s^{4/3} \left(\frac{M_s}{M_{\text{gal}}} \right)^{2/3}. \quad (9.55)$$

To get a characteristic value, we take $M_s = M_{\text{gal}}$ and adopt our usual values $h_0 = 0.75$ and $\Omega_s = \Omega_{\text{cdm}} = 0.3$. Let us, moreover, compare our result to the average energy density of the CMB, which dominates the spectrum of background radiation (figure 1.2). The latter is found from (5.37) as $\rho_{\text{cmb}} c^2 = \Omega_\gamma \rho_{\text{crit},0} c^2 = 4 \times 10^{-13} \text{ erg cm}^{-3}$. We therefore obtain

$$\frac{\rho_s}{\rho_{\text{cmb}}} \approx 7 \times 10^{-6}. \quad (9.56)$$

This is of the same order of magnitude as the limit set on anomalous contributions to the CMB by COBE and other experiments. We infer, therefore, that the dark matter may consist of solitons but that they are probably not more massive than galaxies. Similar arguments can be made on the basis of tidal effects and gravitational lensing [63]. To go further and put more detailed constraints on these candidates from background radiation or other considerations will require a more detailed investigation of their microphysical properties.

Let us summarize our results for this chapter. We have noted that standard (stellar) black holes cannot provide the dark matter insofar as their contributions to the density of the Universe are effectively baryonic. Primordial black holes evade this constraint but we have reconfirmed the classic results of Page, Hawking and others: the collective density of such objects must be negligible, for otherwise their presence would have been obvious in the γ -ray background. In fact, we have shown that their bolometric intensity alone is sufficient to rule them out as important dark-matter candidates. These constraints may be relaxed if primordial black holes form in such a way as to favour objects of higher mass, but it is not clear that such a distribution can be shown to arise in a natural way. As an alternative, we have considered black-hole like objects in higher-dimensional gravity. If the world does have more than four dimensions, as suggested by modern unification theories, then these should exist and could be the dark matter. This is consistent with the bolometric intensity of the EBL, but there are a number of theoretical issues to be worked out before a more definitive assessment of their potential can be made.

References

- [1] Y B Zeldovich and I D Novikov 1966 *Sov. Astron.* **10** 602
- [2] Hawking S W 1971 *Mon. Not. R. Astron. Soc.* **152** 75
- [3] Hawking S W 1974 *Nature* **248** 30
- [4] Carr B J 1975 *Astrophys. J.* **201** 1
- [5] Carr B J 1976 *Astrophys. J.* **206** 8
- [6] Khlopov M Y, Malomed B A and Zeldovich Y B 1985 *Mon. Not. R. Astron. Soc.* **215** 575
- [7] Nasel'skii P D and Polnarëv A G 1985 *Sov. Astron.* **29** 487
- [8] Hsu S D H 1990 *Phys. Lett. B* **251** 343
- [9] Carr B J and Lidsey J E 1993 *Phys. Rev. D* **48** 543
- [10] Ivanov P, Naselsky P and Novikov I 1994 *Phys. Rev. D* **50** 7173
- [11] García-Bellido J, Linde A and Wands D 1996 *Phys. Rev. D* **54** 6040
- [12] Yokoyama J 1997 *Astron. Astrophys.* **318** 673
- [13] Crawford M and Schramm D N 1982 *Nature* **298** 538
- [14] Kodama H, Sasaki M and Sato K 1982 *Prog. Theor. Phys.* **68** 1979
- [15] Hawking S W, Moss I G and Stewart J M 1982 *Phys. Rev. D* **26** 2681
- [16] Jedamzik K 1997 *Phys. Rev. D* **55** R5871
- [17] Hawking S W 1989 *Phys. Lett. B* **231** 237
- [18] Polnarëv A and Zembowicz R 1991 *Phys. Rev. D* **43** 1106
- [19] Niemeyer J C and Jedamzik K 1998 *Phys. Rev. Lett.* **80** 5481
- [20] Cline D B and Hong W 1992 *Astrophys. J.* **401** L57
- [21] Cline D B, Sanders D A and Hong W 1997 *Astrophys. J.* **486** 169
- [22] Green A M 2002 *Phys. Rev.* **65** 027301
- [23] Hawkins M R S 1993 *Nature* **366** 242
- [24] Hawkins M R S 1996 *Mon. Not. R. Astron. Soc.* **278** 787
- [25] Wright E L 1996 *Astrophys. J.* **459** 487
- [26] Cline D B 1998 *Astrophys. J.* **501** L1

- [27] Green A M and Liddle A R 1999 *Phys. Rev.* **60** 063509
- [28] Green A M 2000 *Astrophys. J.* **537** 708
- [29] Carter B 1974 *Phys. Rev. Lett.* **33** 558
- [30] Page D N 1976 *Phys. Rev. D* **14** 3260
- [31] Page D N 1976 *Phys. Rev. D* **13** 198
- [32] MacGibbon J H and Webber B R 1990 *Phys. Rev. D* **41** 3052
- [33] Halzen F, Zas E, MacGibbon J H and Weekes T C 1991 *Nature* **353** 807
- [34] Page D N and Hawking S W 1976 *Astrophys. J.* **206** 1
- [35] Hawking S W 1975 *Commun. Math. Phys.* **43** 199
- [36] MacGibbon J H and Carr B J 1991 *Astrophys. J.* **371** 447
- [37] Kribs G D, Leibovich A K and Rothstein I Z 1999 *Phys. Rev. D* **60** 103510
- [38] Barrow J D, Copeland E J and Liddle A R 1991 *Mon. Not. R. Astron. Soc.* **253** 675
- [39] Overduin J M and Wesson P S 1992 *Vistas Astron.* **35** 439
- [40] Thompson D J and Fichtel C E 1982 *Astron. Astrophys.* **109** 352
- [41] Kappadath S C *et al* 1996 *Astron. Astrophys. Suppl. Ser.* **120** 619
- [42] Sreekumar P *et al* 1998 *Astrophys. J.* **494** 523
- [43] Carr B J and MacGibbon J H 1998 *Phys. Rep.* **307** 141
- [44] Alexandreas D E *et al* 1993 *Phys. Rev. Lett.* **71** 2524
- [45] Connaughton V *et al* 1998 *Astropart. Phys.* **8** 178
- [46] Maki K, Mitsui T and Orito S 1996 *Phys. Rev. Lett.* **76** 3473
- [47] MacGibbon J H 1987 *Nature* **329** 308
- [48] Barrow J D, Copeland E J and Liddle A R 1992 *Phys. Rev. D* **46** 645
- [49] Heckler A F 1997 *Phys. Rev. D* **55** 480
- [50] Heckler A F 1997 *Phys. Rev. Lett.* **78** 3430
- [51] Cline J M, Mostoslavsky M and Servant G 1999 *Phys. Rev. D* **59** 063009
- [52] Kim H I, Lee C H and MacGibbon J H 1999 *Phys. Rev. D* **59** 063004
- [53] Overduin J M and Wesson P S 1997 *Phys. Rep.* **283** 303
- [54] Wesson P S *et al* 1996 *Int. J. Mod. Phys.* **A11** 3247
- [55] Arkani-Hamed N, Dimopoulos S and Dvali G 1998 *Phys. Lett. B* **429** 263
- [56] Randall L and Sundrum R 1999 *Phys. Rev. Lett.* **83** 4690
- [57] Gross D J and Perry M J 1983 *Nucl. Phys. B* **226** 29
- [58] Sorkin R D 1983 *Phys. Rev. Lett.* **51** 87
- [59] Davidson A and Owen D A 1985 *Phys. Lett. B* **155** 247
- [60] Wesson P S and Ponce de Leon J 1992 *J. Math. Phys.* **33** 3883
- [61] Liu H and Wesson P S 1992 *J. Math. Phys.* **33** 3888
- [62] Wesson P S and Ponce de Leon J 1994 *Class. Quantum Grav.* **11** 1341
- [63] Wesson P S 1994 *Astrophys. J.* **420** L49
- [64] Liu H and Overduin J M 2000 *Astrophys. J.* **538** 386
- [65] Overduin J M 2000 *Phys. Rev. D* **62** 102001

Chapter 10

Conclusions

Why is the sky dark at night, rather than being filled with the light of distant galaxies? We have answered this question both qualitatively and quantitatively. The brightness of the night sky is set by the finite age of the Universe, which limits the number of photons that galaxies have been able to contribute to the extragalactic background light. Expansion further darkens an already-black sky by stretching and dimming the light from distant galaxies. This is, however, a secondary effect. If we could freeze the expansion without altering the lifetime of the galaxies, then the night sky would brighten by no more than a factor of two to three at optical wavelengths, depending on the evolutionary history of the galaxies and the makeup of the cosmological fluid.

What makes up the dark matter? This question still awaits a definitive answer. However, we have seen that the machinery which settles the first question goes some way toward resolving this one as well. Most dark-matter candidates are not, in fact, perfectly black. Like the galaxies, they contribute to the extragalactic background light with characteristic signatures at specific wavelengths. Experimental data on the spectral intensity of this light therefore tell us what the dark matter can (or cannot) be. It cannot be vacuum energy decaying into photons, because this would lead to levels of microwave background radiation in excess of those observed. It cannot consist of axions or neutrinos with rest energies in the eV range, because these would produce too much infrared, optical or ultraviolet background light, depending on their lifetimes and coupling parameters. It *could* consist of supersymmetric weakly interacting massive particles (WIMPs) such as neutralinos or gravitinos, but data on the x-ray and γ -ray backgrounds imply that these must be very nearly stable. The same data exclude a significant role for primordial black holes, whose Hawking evaporation produces too much light at γ -ray wavelengths. Higher-dimensional analogues of black holes known as solitons are more difficult to constrain but an analysis based on the integrated intensity of the background radiation at all wavelengths suggests that they could be dark-matter objects if their masses are not larger than those of galaxies.

Table 10.1. Constraints on dark-matter candidates from background radiation.

Candidate	Relevant waveband	Prognosis	Caveat(s)
Vacuum energy	Microwave	Ruled out	Other decay modes
Multi-eV axions	IR, optical	Unlikely	Rest mass/coupling
Light neutrinos	UV	Ruled out	Rest mass/lifetime
WIMPs	X-ray to γ -ray	Plausible	Lifetime/parameters
Black holes	γ -ray	Ruled out	Formation mechanisms
Solitons	(bolometric)	Possible	Theoretical parameters

The relevant waveband and overall prognosis for these dark-matter candidates are summarized in table 10.1 along with possible caveats for each case. Thus, theories in which the vacuum does not couple explicitly to matter or radiation cannot be ruled out in this way. ‘Invisible’ axions and neutrinos with very low rest masses would evade our bounds, although neutrinos in this case could not provide a significant fraction of the dark matter. Very weakly-coupled axions or long-lived neutrinos are also possible in principle, but face other problems. Stable or extremely long-lived supersymmetric WIMPs remain viable, subject to a number of theoretical parameters (such as the supersymmetry-breaking energy scale) whose values have yet to be worked out. Primordial black holes could conceivably play a larger role on the cosmic stage if they form in such a way as to give them a more sharply-peaked mass distribution. Solitons are the proverbial dark-horse candidates: they cannot be excluded, but a decisive test must await more detailed investigation of their theoretical properties.

Table 10.1, of course, does not exhaust the list of possibilities. Many other kinds of dark matter have been proposed, and most can be similarly constrained by means of their contributions to the background light. We mention a few of these without going into details. Very massive objects (VMOs) are dark-matter clumps with masses in the range $100\text{--}10^5 M_\odot$. They might, for instance, be black holes which formed from the collapse of even more massive progenitors. These high-redshift progenitors would also have been extremely luminous and it is their light (redshifted into the infrared band) which puts the strongest limits on VMOs as dark-matter candidates [1]. Warm-dark-matter (WDM) candidates furnish another example. These are particles with rest energies in the keV range whose existence would help resolve certain problems with conventional theories of structure formation. While they are not part of most extensions of the standard model, two possible WDM candidates are light gravitinos and a fourth generation of ‘sterile’ neutrinos which interact more weakly than those of the standard model. These could leave a mark in the x-ray background [2]. A number of other dark-matter particles which have been proposed are very massive and would affect primarily the γ -ray background. Among these are WIMPzillas, non-thermal

relics with rest energies in the range 10^{12} – 10^{16} GeV [3]. Gluinos [4] and axinos [5], the supersymmetric counterparts of gluons and axions, are other examples. High-energy γ -rays also provide us with a crucial experimental window on some of the dark-matter candidates predicted by higher-dimensional unified-field theories such as string and M theory [6] and brane-world scenarios [7, 8].

The dark night sky is thus our most versatile dark-matter detector. Its potential in this regard is summed up nicely by a quote from the eighteenth-century English amateur astronomer and cosmologist Thomas Wright. In his 1750 book titled *An original theory or new hypothesis of the Universe*, he wrote as follows:

We may justly suppose that so many radiant bodies were not created barely to enlighten an infinite void, but to make their much more numerous attendants visible. [9]

Of the attendants we have discussed, which is most likely to make up the dark matter? It would be foolhardy to answer this too confidently, but we will close this book with some opinions of our own. The results we have obtained here support the view that the dark matter cannot involve ‘ordinary’ particles, by which we mean those of the standard model of particle physics (including those which may have fallen into black holes). We think it unlikely that baryons, neutrinos or black holes will be found to form the bulk of the dark matter. Vacuum energy, similarly, is a possibility only insofar as it does not decay into baryons or photons. This leaves us with the other ‘exotic’ candidates. There is no doubt that the frontrunners here are axions and supersymmetric WIMPs. These particles are theoretically well motivated and physically well defined enough to be testable. Both, however, arise in the context of particle-physics theories which do not yet make room for gravity. It may be that the nature of the dark matter will continue to resist understanding until a more fully unified theory of all the interactions is at hand. In our view such a theory is likely to involve more than four dimensions. Thus, while we are justified at present in directing most of our theoretical and experimental efforts at WIMPs and axions, experience also suggests to us that the dark matter may involve higher-dimensional objects such as solitons and branes.

A final word is in order about the phrase ‘dark matter’, which while natural is perhaps semantically unfortunate. It has been known since Einstein that mass and energy are equivalent: in recent years particle physicists have found it increasingly difficult to define the mass of a particle (as opposed, for example, to the energy of a resonance); it is now acknowledged by workers in general relativity that the energy of the gravitational field in that theory cannot be localized; and in higher-dimensional extensions of general relativity, energy in spacetime is derived from the geometry of higher dimensions in a way that intimately mixes the properties of ‘matter’ and ‘vacuum’. The darkness of the night sky is consistent with other cosmological data in telling us that there is something out there which gravitates or curves spacetime. We are of the opinion that this substance, whatever we call it,

will eventually be identified. However, while some of it may be in a form which is familiar, some may also be in a form which is truly unexpected.

References

- [1] Carr B J and Sakellariadou M 1999 *Astrophys. J.* **516** 195
- [2] Abazajian K, Fuller G M and Tucker W H 2001 *Astrophys. J.* **562** 593
- [3] Ziaeeepour H 2001 *Astropart. Phys.* **16** 101
- [4] Berezhinsky V, Kachelrieß M and Ostapchenko S 2002 *Phys. Rev. D* **65** 083004
- [5] Kim H B and Kim J E 2002 *Phys. Lett. B* **527** 18
- [6] Benakli K, Ellis J and Nanopoulos D V 1999 *Phys. Rev. D* **59** 047301
- [7] Hall L J and Smith D 1999 *Phys. Rev. D* **60** 085008
- [8] Hannestad S and Raffelt G G 2001 *Phys. Rev. Lett.* **87** 051301
- [9] Jaki S L 2001 *The Paradox of Olbers' Paradox* 2nd edn (Pinckney, MI: Real View Books) p 112

Appendix A

Bolometric intensity integrals

A.1 Radiation-dominated models

This appendix is a summary of analytic solutions to the bolometric intensity integrals (2.42) and (2.45) in chapter 2. We begin with the situation in which the cosmic fluid is primarily composed of radiation-like matter (such as photons or light neutrinos), so that $\Omega_{r,0} \neq 0$ but $\Omega_{m,0} = \Omega_{\Lambda,0} = 0$. Equations (2.42) and (2.45) give

$$\frac{Q}{Q_*} = \int_1^{1+z_f} \frac{dx}{x^3 \sqrt{(1 - \Omega_{r,0}) + \Omega_{r,0} x^2}} \quad (\text{A.1})$$

$$\frac{Q_{\text{stat}}}{Q_*} = \int_1^{1+z_f} \frac{dx}{x^2 \sqrt{(1 - \Omega_{r,0}) + \Omega_{r,0} x^2}}. \quad (\text{A.2})$$

We assume that the density of relativistic matter is positive. Equation (A.1) then has two solutions, depending on whether the Universe is open or closed. For open models ($\Omega_{r,0} < 1$),

$$\begin{aligned} \frac{Q}{Q_*} = & \frac{1}{2(1 - \Omega_{r,0})} - \frac{\sqrt{1 + \Omega_{r,0} z_f (2 + z_f)}}{2(1 - \Omega_{r,0})(1 + z_f)^2} + \frac{\Omega_{r,0}}{4(1 - \Omega_{r,0})^{3/2}} \\ & \times \ln \left[\left(\frac{\sqrt{1 + \Omega_{r,0} z_f (2 + z_f)} + \sqrt{1 - \Omega_{r,0}}}{\sqrt{1 + \Omega_{r,0} z_f (2 + z_f)} - \sqrt{1 - \Omega_{r,0}}} \right) \right. \\ & \left. \times \left(\frac{1 - \sqrt{1 - \Omega_{r,0}}}{1 + \sqrt{1 - \Omega_{r,0}}} \right) \right]. \end{aligned} \quad (\text{A.3})$$

This simplifies considerably in the limit $z_f \rightarrow \infty$, for which

$$\frac{Q}{Q_*} \rightarrow \frac{1}{2(1 - \Omega_{r,0})} + \frac{\Omega_{r,0}}{4(1 - \Omega_{r,0})^{3/2}} \ln \left(\frac{1 - \sqrt{1 - \Omega_{r,0}}}{1 + \sqrt{1 - \Omega_{r,0}}} \right). \quad (\text{A.4})$$

For closed models ($\Omega_{r,0} > 1$), the solution is different:

$$\begin{aligned} \frac{Q}{Q_*} &= \frac{-1}{2(\Omega_{r,0} - 1)} + \frac{\sqrt{1 + \Omega_{r,0}z_f(2 + z_f)}}{2(\Omega_{r,0} - 1)(1 + z_f)^2} + \frac{\Omega_{r,0}}{2(\Omega_{r,0} - 1)^{3/2}} \\ &\times \left[\cos^{-1} \left(\sqrt{\frac{\Omega_{r,0} - 1}{\Omega_{r,0}}} \frac{1}{(1 + z_f)} \right) \right. \\ &\left. - \cos^{-1} \left(\sqrt{\frac{\Omega_{r,0} - 1}{\Omega_{r,0}}} \right) \right]. \end{aligned} \quad (\text{A.5})$$

This also simplifies in the limit $z_f \rightarrow \infty$, as follows:

$$\frac{Q}{Q_*} \rightarrow \frac{-1}{2(\Omega_{r,0} - 1)} + \frac{\Omega_{r,0}}{2(\Omega_{r,0} - 1)^{3/2}} \left[\frac{\pi}{2} - \cos^{-1} \left(\sqrt{\frac{\Omega_{r,0} - 1}{\Omega_{r,0}}} \right) \right]. \quad (\text{A.6})$$

For the static case, (A.2) has a single solution for both open and closed models, as follows:

$$\frac{Q_{\text{stat}}}{Q_*} = \frac{1}{1 - \Omega_{r,0}} - \frac{\sqrt{1 + \Omega_{r,0}z_f(2 + z_f)}}{(1 - \Omega_{r,0})(1 + z_f)}. \quad (\text{A.7})$$

In the limit $z_f \rightarrow \infty$, this reduces to

$$\frac{Q_{\text{stat}}}{Q_*} \rightarrow \frac{1 - \sqrt{\Omega_{r,0}}}{1 - \Omega_{r,0}}. \quad (\text{A.8})$$

These solutions are plotted in figure 2.2 of the main text.

A.2 Matter-dominated models

We proceed to models in which the cosmic fluid is dominated by dustlike rather than relativistic matter, so that $\Omega_{m,0} \neq 0$ but $\Omega_{r,0} = \Omega_{\Lambda,0} = 0$. If vacuum energy can be neglected (a supposition made by most cosmologists until quite recently), then this is a good description of conditions in the present Universe. The case $\Omega_{m,0} \approx 0.3$ (i.e. the Λ CDM model) has received special attention in the literature. (The other important special case is of course the Einstein–de Sitter model with $\Omega_{m,0} = 1$, treated in section 2.7.) Equations (2.42) and (2.45) read

$$\frac{Q}{Q_*} = \int_1^{1+z_f} \frac{dx}{x^3 \sqrt{(1 - \Omega_{m,0}) + \Omega_{m,0}x}} \quad (\text{A.9})$$

$$\frac{Q_{\text{stat}}}{Q_*} = \int_1^{1+z_f} \frac{dx}{x^2 \sqrt{(1 - \Omega_{m,0}) + \Omega_{m,0}x}}. \quad (\text{A.10})$$

We assume that the matter density is positive, as before. Equation (A.9) then has two solutions, depending on whether the Universe is open or closed. For open

models ($\Omega_{m,0} < 1$),

$$\begin{aligned} \frac{Q}{Q_*} = & \frac{1}{2(1 - \Omega_{m,0})} \left[1 - \frac{3\Omega_{m,0}}{2(1 - \Omega_{m,0})} \right] \\ & - \frac{\sqrt{1 + \Omega_{m,0}z_f}}{2(1 - \Omega_{m,0})(1 + z_f)} \left[\frac{1}{1 + z_f} - \frac{3\Omega_{m,0}}{2(1 - \Omega_{m,0})} \right] \\ & + \frac{3\Omega_{m,0}^2}{8(1 - \Omega_{m,0})^{5/2}} \ln \left[\left(\frac{\sqrt{1 + \Omega_{m,0}z_f} - \sqrt{1 - \Omega_{m,0}}}{\sqrt{1 + \Omega_{m,0}z_f} + \sqrt{1 - \Omega_{m,0}}} \right) \right. \\ & \left. \times \left(\frac{1 + \sqrt{1 - \Omega_{m,0}}}{1 - \sqrt{1 - \Omega_{m,0}}} \right) \right]. \end{aligned} \quad (\text{A.11})$$

In the limit $z_f \rightarrow \infty$ this simplifies to

$$\begin{aligned} \frac{Q}{Q_*} \rightarrow & \frac{1}{2(1 - \Omega_{m,0})} \left[1 - \frac{3\Omega_{m,0}}{2(1 - \Omega_{m,0})} \right] \\ & + \frac{3\Omega_{m,0}^2}{8(1 - \Omega_{m,0})^{5/2}} \ln \left(\frac{1 + \sqrt{1 - \Omega_{m,0}}}{1 - \sqrt{1 - \Omega_{m,0}}} \right). \end{aligned} \quad (\text{A.12})$$

For closed models ($\Omega_{m,0} > 1$), the solution is somewhat different:

$$\begin{aligned} \frac{Q}{Q_*} = & \frac{-1}{2(\Omega_{m,0} - 1)} \left[1 + \frac{3\Omega_{m,0}}{2(\Omega_{m,0} - 1)} \right] + \frac{\sqrt{1 + \Omega_{m,0}z_f}}{2(\Omega_{m,0} - 1)(1 + z_f)} \\ & \times \left[\frac{1}{1 + z_f} + \frac{3\Omega_{m,0}}{2(\Omega_{m,0} - 1)} \right] + \frac{3\Omega_{m,0}^2}{4(\Omega_{m,0} - 1)^{5/2}} \\ & \times \left[\tan^{-1} \left(\frac{\sqrt{1 + \Omega_{m,0}z_f}}{\sqrt{\Omega_{m,0} - 1}} \right) - \tan^{-1} \left(\frac{1}{\sqrt{\Omega_{m,0} - 1}} \right) \right]. \end{aligned} \quad (\text{A.13})$$

This simplifies, as follows, for $z_f \rightarrow \infty$:

$$\begin{aligned} \frac{Q}{Q_*} \rightarrow & \frac{-1}{2(\Omega_{m,0} - 1)} \left[1 + \frac{3\Omega_{m,0}}{2(\Omega_{m,0} - 1)} \right] \\ & + \frac{3\Omega_{m,0}^2}{4(\Omega_{m,0} - 1)^{5/2}} \left[\frac{\pi}{2} - \tan^{-1} \left(\frac{1}{\sqrt{\Omega_{m,0} - 1}} \right) \right]. \end{aligned} \quad (\text{A.14})$$

The solution of (A.10) for a static Universe also depends on the spatial curvature. For open models,

$$\begin{aligned} \frac{Q_{\text{stat}}}{Q_*} = & \frac{1}{1 - \Omega_{m,0}} - \frac{\sqrt{1 + \Omega_{m,0}z_f}}{(1 - \Omega_{m,0})(1 + z_f)} + \frac{\Omega_{m,0}}{2(1 - \Omega_{m,0})^{3/2}} \\ & \times \ln \left[\left(\frac{\sqrt{1 + \Omega_{m,0}z_f} + \sqrt{1 - \Omega_{m,0}}}{\sqrt{1 + \Omega_{m,0}z_f} - \sqrt{1 - \Omega_{m,0}}} \right) \left(\frac{1 - \sqrt{1 - \Omega_{m,0}}}{1 + \sqrt{1 - \Omega_{m,0}}} \right) \right]. \end{aligned} \quad (\text{A.15})$$

In the limit $z_f \rightarrow \infty$ this reduces to

$$\frac{Q_{\text{stat}}}{Q_*} \rightarrow \frac{1}{1 - \Omega_{m,0}} + \frac{\Omega_{m,0}}{2(1 - \Omega_{m,0})^{3/2}} \ln \left(\frac{1 - \sqrt{1 - \Omega_{m,0}}}{1 + \sqrt{1 - \Omega_{m,0}}} \right). \quad (\text{A.16})$$

For closed models, the solution is

$$\begin{aligned} \frac{Q_{\text{stat}}}{Q_*} &= \frac{-1}{\Omega_{m,0} - 1} + \frac{\sqrt{1 + \Omega_{m,0}z_f}}{(\Omega_{m,0} - 1)(1 + z_f)} + \frac{\Omega_{m,0}}{(\Omega_{m,0} - 1)^{3/2}} \\ &\quad \times \left[\tan^{-1} \left(\frac{\sqrt{1 + \Omega_{m,0}z_f}}{\sqrt{\Omega_{m,0} - 1}} \right) - \tan^{-1} \left(\frac{1}{\sqrt{\Omega_{m,0} - 1}} \right) \right] \end{aligned} \quad (\text{A.17})$$

with the following limit as $z_f \rightarrow \infty$:

$$\frac{Q_{\text{stat}}}{Q_*} \rightarrow \frac{-1}{\Omega_{m,0} - 1} + \frac{\Omega_{m,0}}{(\Omega_{m,0} - 1)^{3/2}} \left[\frac{\pi}{2} - \tan^{-1} \left(\frac{1}{\sqrt{\Omega_{m,0} - 1}} \right) \right]. \quad (\text{A.18})$$

These results are displayed in figure 2.3 of the main text.

A.3 Vacuum-dominated models

We move on finally to the models filled with vacuum energy, so that $\Omega_{r,0} = \Omega_{m,0} = 0$ but $\Omega_{\Lambda,0} \neq 0$. If evidence continues to build for a vacuum-dominated Universe, then these may provide a better description of large-scale dynamics than the matter-dominated models of the previous section, despite the fact that they contain no matter at all. They will, moreover, approximate the real Universe more and more closely as matter density falls with time, while that of the vacuum energy stays constant.

We consider values in the range $-0.5 \leq \Omega_{\Lambda,0} \leq 1.5$ in this section, although a negative cosmological term is somewhat disfavoured both theoretically (through the interpretation of $\Omega_{\Lambda,0}$ as a large-scale energy density) and observationally (since negative values of $\Omega_{\Lambda,0}$ imply very short lifetimes for the Universe). We exclude the special cases $\Omega_{\Lambda,0} = 0$ (Milne model) and $\Omega_{\Lambda,0} = 1$ (de Sitter model) which are treated in section 2.7. Equations (2.42) and (2.45) take the forms

$$\frac{Q}{Q_*} = \int_1^{1+z_f} \frac{dx}{x^2 \sqrt{\Omega_{\Lambda,0} + (1 - \Omega_{\Lambda,0})x^2}} \quad (\text{A.19})$$

$$\frac{Q_{\text{stat}}}{Q_*} = \int_1^{1+z_f} \frac{dx}{x \sqrt{\Omega_{\Lambda,0} + (1 - \Omega_{\Lambda,0})x^2}}. \quad (\text{A.20})$$

Equation (A.19) for the expanding model has the same form as (A.2) for the static model with relativistic matter, and the solution (good for both open and closed models) is

$$\frac{Q}{Q_*} = \frac{1}{\Omega_{\Lambda,0}} - \frac{\sqrt{1 + (1 - \Omega_{\Lambda,0})(2 + z_f)z_f}}{\Omega_{\Lambda,0}(1 + z_f)}. \quad (\text{A.21})$$

As $z_f \rightarrow \infty$, this reduces (for models with $\Omega_{\Lambda,0} \leq 1$) to

$$\frac{Q}{Q_*} \rightarrow \frac{1 - \sqrt{1 - \Omega_{\Lambda,0}}}{\Omega_{\Lambda,0}} \quad (\Omega_{\Lambda,0} \leq 1). \quad (\text{A.22})$$

For models with $\Omega_{\Lambda,0} > 1$, limiting values of EBL intensity occur as $z_f \rightarrow z_{\max}(\Omega_{\Lambda,0}) < \infty$, as discussed in section 2.7 in the main text. In the expanding case equation (A.21) gives simply

$$\frac{Q}{Q_*} \rightarrow \frac{1}{\Omega_{\Lambda,0}} \quad (\Omega_{\Lambda,0} > 1) \quad (\text{A.23})$$

where we have used (2.58) for z_{\max} . For the static case, equation (A.20) has three possible solutions, depending whether $\Omega_{\Lambda,0} < 0$, $0 < \Omega_{\Lambda,0} < 1$, or $\Omega_{\Lambda,0} > 1$. For models of the first kind,

$$\frac{Q_{\text{stat}}}{Q_*} = \frac{1}{\sqrt{-\Omega_{\Lambda,0}}} \left[\cos^{-1} \left(\sqrt{\frac{-\Omega_{\Lambda,0}}{1 - \Omega_{\Lambda,0}}} \frac{1}{1 + z_f} \right) - \cos^{-1} \left(\sqrt{\frac{-\Omega_{\Lambda,0}}{1 - \Omega_{\Lambda,0}}} \right) \right] \quad (\text{A.24})$$

which has the following limit as $z_f \rightarrow \infty$:

$$\frac{Q_{\text{stat}}}{Q_*} \rightarrow \frac{1}{\sqrt{-\Omega_{\Lambda,0}}} \left[\frac{\pi}{2} - \cos^{-1} \left(\sqrt{\frac{-\Omega_{\Lambda,0}}{1 - \Omega_{\Lambda,0}}} \right) \right]. \quad (\text{A.25})$$

For models in which $0 < \Omega_{\Lambda,0} < 1$, the solution of (A.20) reads

$$\begin{aligned} \frac{Q_{\text{stat}}}{Q_*} &= \frac{1}{2\sqrt{\Omega_{\Lambda,0}}} \ln \left[\left(\frac{\sqrt{1 + (1 - \Omega_{\Lambda,0})(2 + z_f)z_f} - \sqrt{\Omega_{\Lambda,0}}}{\sqrt{1 + (1 - \Omega_{\Lambda,0})(2 + z_f)z_f} + \sqrt{\Omega_{\Lambda,0}}} \right) \right. \\ &\quad \left. \times \left(\frac{1 + \sqrt{\Omega_{\Lambda,0}}}{1 - \sqrt{\Omega_{\Lambda,0}}} \right) \right] \end{aligned} \quad (\text{A.26})$$

which reduces to the following as $z_f \rightarrow \infty$:

$$\frac{Q_{\text{stat}}}{Q_*} \rightarrow \frac{1}{2\sqrt{\Omega_{\Lambda,0}}} \ln \left(\frac{1 + \sqrt{\Omega_{\Lambda,0}}}{1 - \sqrt{\Omega_{\Lambda,0}}} \right). \quad (\text{A.27})$$

When $\Omega_{\Lambda,0} > 1$, finally, we obtain

$$\begin{aligned} \frac{Q_{\text{stat}}}{Q_*} &= \frac{1}{2\sqrt{\Omega_{\Lambda,0}}} \ln \left[\left(\frac{\sqrt{\Omega_{\Lambda,0}} - \sqrt{1 + (1 - \Omega_{\Lambda,0})(2 + z_f)z_f}}{\sqrt{\Omega_{\Lambda,0}} + \sqrt{1 + (1 - \Omega_{\Lambda,0})(2 + z_f)z_f}} \right) \right. \\ &\quad \left. \times \left(\frac{\sqrt{\Omega_{\Lambda,0}} + 1}{\sqrt{\Omega_{\Lambda,0}} - 1} \right) \right]. \end{aligned} \quad (\text{A.28})$$

In the limit $z_f \rightarrow z_{\max}$, where the latter is defined as in (2.58), this expression reduces to

$$\frac{Q_{\text{stat}}}{Q_*} \rightarrow \frac{1}{2\sqrt{\Omega_{\Lambda,0}}} \ln \left(\frac{\sqrt{\Omega_{\Lambda,0}} + 1}{\sqrt{\Omega_{\Lambda,0}} - 1} \right) \quad (\text{A.29})$$

a result close to that in (A.27). Figure 2.4 in the main text summarizes these results.

Appendix B

Dynamics with a decaying vacuum

This appendix outlines the solution of equations (5.15), (5.16), (5.18) and (5.19) in the flat decaying-vacuum cosmology which is the subject of chapter 5. These equations specify the scale factor R and densities ρ_m , ρ_r and ρ_v of matter, radiation and decaying vacuum energy respectively. We apply the boundary conditions $R(0) = 0$, $\rho_m(t_0) = \rho_{m,0}$ and $\rho_r(t_0) = \rho_{r,0}$.

The key is the differential equation (5.19), which gives $R(t)$. This may be solved analytically in three different regimes:

$$\begin{aligned}
 \text{Regime 1:} & \quad \rho_r + \rho_v \gg \rho_m + \rho_c \\
 \text{Regime 2:} & \quad \rho_r + \rho_v \ll \rho_m, \quad \rho_c = 0 \\
 \text{Regime 3:} & \quad \rho_r + \rho_v \ll \rho_m + \rho_c, \quad \rho_c \neq 0
 \end{aligned} \tag{B.1}$$

Here we distinguish between the time-varying part of the vacuum energy density (ρ_v) and its constant part (ρ_c). Regime 1 describes the *radiation*-dominated era. Regime 2 is a good approximation to the *matter*-dominated era in the Einstein–de Sitter (EdS) model, with $\rho_{m,0} = \rho_{\text{crit},0}$. Regime 3 should be used instead when $\rho_{m,0} < \rho_{\text{crit},0}$, as in the Λ CDM model (with $\rho_{m,0} = 0.3\rho_{\text{crit},0}$) or the Λ BDM model (with $\rho_{m,0} = 0.03\rho_{\text{crit},0}$). For brevity we refer to this as the *vacuum*-dominated regime.

B.1 Radiation-dominated regime

Equation (5.19) takes the following form in regime 1 (with $x = x_r$):

$$\left(\frac{\dot{R}}{R} \right)^2 = \frac{8\pi G\rho_v}{3x_r}. \tag{B.2}$$

Inserting (5.18) for $\rho_v(R)$ then produces a differential equation for $R(t)$, which can be solved to yield:

$$R_1(t) = \left[\frac{32\pi G\alpha_v(1-x_r)^2 t^2}{3x_r} \right]^{1/4(1-x_r)}. \tag{B.3}$$

Substitution back into (5.18) then gives the vacuum energy density:

$$\rho_{v,1}(t) = \frac{\alpha x_r}{(1-x_r)^2} t^{-2} \quad (\text{B.4})$$

where $\alpha \equiv 3/(32\pi G) = 4.47 \times 10^5 \text{ g cm}^{-2} \text{ s}^2$. The corresponding radiation density is, from (5.16):

$$\rho_{r,1}(t) = \left(\frac{1-x_r}{x_r} \right) \rho_{v,1}(t). \quad (\text{B.5})$$

To check these results one can imagine that regime 1 extends to the present day. Dividing through by the critical density $\rho_{\text{crit}} = 3H^2/8\pi G$, we obtain

$$\Omega_{v+r,1}(t) = \frac{\rho_{v,1} + \rho_{r,1}}{\rho_{\text{crit}}} = \frac{\rho_{v,1}/x_r}{\rho_{\text{crit}}} = \frac{1}{[2(1-x_r)H]^2} t^{-2}. \quad (\text{B.6})$$

Differentiating (B.3) with respect to time, we find that $H_0 = 1/[2(1-x_r)t_0]$. Thus $\Omega_{v+r,1} = 1$ at the present time, which is consistent with our assumption of flatness.

B.2 Matter-dominated regime

For regime 2, equation (5.19) takes the form (with $x = x_m$):

$$\left(\frac{\dot{R}}{R} \right)^2 = \frac{8\pi G \rho_m}{3}. \quad (\text{B.7})$$

From the conservation equation (5.15), matter density is given by

$$\rho_m(R) = \alpha_m R^{-3} \quad (\text{B.8})$$

where $\alpha_m = \text{constant}$. Inserting this into (B.7) produces a differential equation for $R(t)$ whose solution is:

$$R_2(t) = (6\pi G \alpha_m)^{1/3} t^{2/3}. \quad (\text{B.9})$$

Substitution back into (5.18) then gives the vacuum energy in terms of the constants α_v and α_m as

$$\rho_{v,2}(t) = \alpha_v (6\pi G \alpha_m t^2)^{-4(1-x_m)/3}. \quad (\text{B.10})$$

One does not need to solve for α_v and α_m in this expression. The dependence on t is enough, as we show by using (5.16) to write down the corresponding radiation density:

$$\rho_{r,2}(t) = \left(\frac{1-x_m}{x_m} \right) \rho_{v,2}(t) = \alpha_r t^{-8(1-x_m)/3} \quad (\text{B.11})$$

where α_r is a new constant. This is fixed in terms of the present radiation density $\rho_{r,0}$ by the boundary condition $\rho_{r,2}(t_0) = \rho_{r,0}$, so that $\alpha_r = \rho_{r,0} t_0^{8(1-x_m)/3}$. The vacuum density can then be written in terms of α_r as well. Using (5.16) again one has

$$\rho_{v,2}(t) = \left(\frac{x_m}{1-x_m} \right) \rho_{r,2}(t) = \left(\frac{\alpha_r x_m}{1-x_m} \right) t^{-8(1-x_m)/3}. \quad (\text{B.12})$$

The quantities ρ_r and ρ_v will not necessarily be continuous across the phase transition $t = t_{\text{eq}}$, but $R^{4(1-x)} \rho_v$ will be, since it is conserved by (5.17).

Matter density during regime 2 is found by substituting (B.9) directly into (B.8). This gives

$$\rho_{m,2} = \frac{1}{6\pi G} t^{-2}. \quad (\text{B.13})$$

To check this one can divide through by the critical density to obtain $\Omega_{m,2}(t)$. Recalling that the lifetime of a flat, matter-dominated Universe is $t_0 = 2/3H_0$, one finds simply

$$\Omega_{m,2}(t) = \left(\frac{t}{t_0} \right)^{-2}. \quad (\text{B.14})$$

Thus $\Omega_{m,2} = 1$ in the limit $t \rightarrow t_0$ as required.

We do not need an expression for matter density in the radiation-dominated era. For the sake of the completeness of the plots in figure 5.1, however, we can extend $\rho_m(t)$ into regime 1 by imposing continuity across the phase transition $t = t_{\text{eq}}$. (This is an extra condition on the theory, which conserves particle number $R^3 \rho_m$ rather than density ρ_m .) Using equations (B.3), (B.8) and (B.13), we obtain

$$\rho_{m,1}(t) = \frac{1}{6\pi G t_{\text{eq}}^2} \left(\frac{t}{t_{\text{eq}}} \right)^{-3/2(1-x_r)}. \quad (\text{B.15})$$

This expression joins smoothly onto (B.13) at $t = t_{\text{eq}}$ as desired.

B.3 Vacuum-dominated regime

The derivation in this section follows the same logic as in the previous section but involves a little more work. Equation (5.19) takes the following form in regime 3 (with $x = x_m$):

$$\left(\frac{\dot{R}}{R} \right)^2 = \frac{8\pi G}{3} (\alpha_m R^{-3} + \rho_c) \quad (\text{B.16})$$

where we have used equation (B.8). This is the same as the differential equation (2.64) and can be solved by the methods of section 2.8. Assuming $\rho_c > 0$, the result is

$$R_3(t) = \left[\sqrt{\frac{\alpha_m}{\rho_c}} \sinh(\sqrt{6\pi G \rho_c} t) \right]^{2/3}. \quad (\text{B.17})$$

We then proceed as for regime 2, substituting (B.17) into (5.18) to obtain an expression for $\rho_{v,3}(t)$ in terms of α_m and α_v . This gives the time-dependence of the radiation density $\rho_{r,3}$ (as well as $\rho_{v,3}$). Application of the boundary condition $\rho_{r,3}(t_0) = \rho_{r,0}$ then gives

$$\rho_{r,3}(t) = \rho_{r,0} \left[\frac{\sinh(\sqrt{6\pi G\rho_c}t)}{\sinh(\sqrt{6\pi G\rho_c}t_0)} \right]^{-8(1-x_m)/3}. \quad (\text{B.18})$$

The corresponding vacuum density is, from (5.16),

$$\rho_{v,3}(t) = \left(\frac{x_m}{1-x_m} \right) \rho_{r,3}(t). \quad (\text{B.19})$$

The matter density is similarly found by putting (B.17) into (B.8) and applying the boundary condition $\rho_{m,3}(t_0) = \rho_{m,0}$ to obtain

$$\rho_{m,3}(t) = \rho_{m,0} \left[\frac{\sinh(\sqrt{6\pi G\rho_c}t)}{\sinh(\sqrt{6\pi G\rho_c}t_0)} \right]^{-2}. \quad (\text{B.20})$$

For plotting purposes, this can be extended into regime 1 by imposing continuity on ρ_m at t_{eq} , as in the preceding section. The result is

$$\rho_{m,1}(t) = \rho_{m,0} \left[\frac{\sinh(\sqrt{6\pi G\rho_c}t_{\text{eq}})}{\sinh(\sqrt{6\pi G\rho_c}t_0)} \right]^{-2} \left(\frac{t}{t_{\text{eq}}} \right)^{-3/2(1-x_r)}. \quad (\text{B.21})$$

This joins smoothly onto (B.20) at $t = t_{\text{eq}}$ as desired.

The structural similarity between the results of this section and those of the preceding one allows us to combine them into a single set of equations if we define a function $\mathcal{S}_m(t)$ such that

$$\mathcal{S}_m(t) \equiv \begin{cases} t & (\rho_c = 0) \\ \sinh(\sqrt{6\pi G\rho_c}t) & (\rho_c \neq 0). \end{cases} \quad (\text{B.22})$$

The results of these two sections, finally, can be expressed in terms of $\Omega_{m,0}$ by recalling that $8\pi G\rho_{\text{crit},0} = 3H_0^2$ so that $\sqrt{6\pi G\rho_c} = \frac{3}{2}\sqrt{1-\Omega_{m,0}}H_0$ for flat models. Equations (5.20) through (5.26) and figure 5.1 in the main text summarize these results.

Appendix C

Absorption by galactic hydrogen

In this appendix, we take up the problem (section 7.3) of calculating the ‘efficiency factor’ ϵ , or fraction of neutrino decay photons of energy $E_\gamma = 14.4 \pm 0.5$ eV which are absorbed by neutral hydrogen inside the galaxy before they can escape and contribute to its luminosity. The number density of decay photons is proportional to \mathcal{N}_ν , equation (7.11).

Consider a decay photon emitted along the direction (θ, ϕ) at some point $P = (r_p, z_p)$ in the halo, where r, z are cylindrical coordinates. We wish to find out how much absorbing material lies between P and the edge of the halo, which is described by the ellipsoidal surface:

$$\left(\frac{x}{r_\odot}\right)^2 + \left(\frac{y}{r_\odot}\right)^2 + \left(\frac{z}{h}\right)^2 = \xi^2. \quad (\text{C.1})$$

Here x, y, z are rectangular coordinates, $r_\odot = 8$ kpc is the distance of the Sun from the galactic centre, $h = 3$ kpc is the scale height of the halo and $\xi = r_h/r_\odot$ is the halo scale factor. Outside the ellipsoidal surface we assume that the neutrino number density drops off exponentially and can be neglected. The value of ξ is fixed by specifying the total mass of the halo. We require this to be $M_h = (2 \pm 1) \times 10^{12} M_\odot$, which from section 7.3 corresponds to $r_h = 70 \pm 25$ kpc. Thus $\xi = 8.8 \pm 3.1$.

If d is the distance to the edge of the halo, then $x = r_p + d \sin \theta \cos \phi$, $y = d \sin \theta \sin \phi$ and $z = z_p + d \cos \theta$. Substitution into equation (C.1) yields two roots:

$$d^\pm = -\frac{B}{A} \pm \sqrt{\left(\frac{B}{A}\right)^2 - \left(\frac{C}{A}\right)^2} \quad (\text{C.2})$$

where $A = h^2 \sin^2 \theta + r_\odot^2 \cos^2 \theta$, $B = h^2 r_p \sin \theta \cos \phi + r_\odot^2 z_p \cos \theta$ and $C = h^2 r_p^2 + r_\odot^2 z_p^2 - r_\odot^2 h^2 \xi$. The quantities d^\pm are the column lengths in the directions $Q^+ = (\theta, \phi)$ and $Q^- = (\theta + \pi, \phi + \pi)$.

In general, the fraction of (monochromatic) photons transmitted across a distance d through an absorbing medium of ionization cross section σ and

number density n is given by $f = e^{-\tau}$ where τ , the optical depth, is defined by $\tau = \int_0^d \sigma n dr'$. For neutral hydrogen the photoionization cross section is $\sigma = \sigma_H (13.6 \text{ eV}/E_\gamma)^3$ where $\sigma_H = 6.3 \times 10^{-18} \text{ cm}^2$ [1]. The number density of neutral hydrogen atoms in the Milky Way is roughly described by the following function [2]:

$$n_H(r, z) = n_d f_d(r) \exp \left[-\frac{1}{2} \left(\frac{z}{z_d(r)} \right)^2 \right] \quad (\text{C.3})$$

where:

$$f_d(r) = \begin{cases} r/r_1 & \text{if } r < r_1 \\ \exp[(r_2 - r)/r_3] & \text{if } r > r_2 \\ 1 & \text{otherwise} \end{cases}$$

$$z_d(r) = \begin{cases} 0.12 \text{ kpc} + 0.023(r - r_0) & \text{if } r > r_0 \\ 0.12 \text{ kpc} & \text{otherwise.} \end{cases}$$

Here $n_d = 0.35 \text{ cm}^{-3}$, $r_1 = 5 \text{ kpc}$, $r_2 = 13 \text{ kpc}$, $r_3 = 1.9 \text{ kpc}$ and $r_0 = 9.5 \text{ kpc}$. The constants $\sigma = 5.3 \times 10^{-18} \text{ cm}^2$ (for 14.4 eV photons) and $n_d = 0.35 \text{ cm}^{-3}$ may be usefully combined into a single absorption coefficient $k_d \equiv \sigma n_d = 5730 \text{ kpc}^{-1}$. The large size of this number indicates that regions close to the galactic plane (where $n_H \approx n_d$) will be totally opaque to decay photons.

The optical depths of columns PQ^+ and PQ^- are:

$$\tau_p^\pm(\theta, \phi) = k_d \int_0^{d^\pm} f_d(r) \exp \left[-\frac{1}{2} \left(\frac{z}{z_d(r)} \right)^2 \right] dr' \quad (\text{C.4})$$

where $x = r_p + r' \sin \theta \cos \phi$, $y = r' \sin \theta \sin \phi$, $z = z_p + r' \cos \theta$ and $r = (x^2 + y^2)^{1/2}$. The probability that a photon emitted at P will reach the edge of the halo (averaged over all decay directions) is then:

$$\mathcal{P}_e(r_p, z_p) = \frac{1}{2\pi} \int_{\phi=0}^{\pi} \int_{\theta=0}^{\pi/2} [e^{-\tau_{p,+}(\theta, \phi)} + e^{-\tau_{p,-}(\theta, \phi)}] \sin \theta d\theta d\phi. \quad (\text{C.5})$$

Thanks to cylindrical symmetry, we need integrate only over one quadrant of the sky. It remains to average \mathcal{P}_e over all points P in the halo. The average must be weighted by the number density of decay photons \mathcal{N}_v to reflect the fact that more photons are emitted near the centre of the halo than at its edges. Using (7.11) this may be written as $\mathcal{N}_v(r_p, z_p) = [1 + \sqrt{(r_p/r_\odot)^2 (z/h)^2}]^{-2}$. So the final efficiency factor is

$$\epsilon = \frac{\int_0^{r_h} \int_0^{z_{\max}(r_p)} \mathcal{P}_e(r_p, z_p) \mathcal{N}_v(r_p, z_p) r_p dr_p dz_p}{\int_0^{r_h} \int_0^{z_{\max}(r_p)} \mathcal{N}_v(r_p, z_p) r_p dr_p dz_p} \quad (\text{C.6})$$

where $z_{\max}(r_p) = h \sqrt{\xi^2 - (r_p/r_\odot)^2}$. Numerical integration of (C.6) produces the values quoted in equation (7.17) of the main text. Figure C.1 illustrates the shape

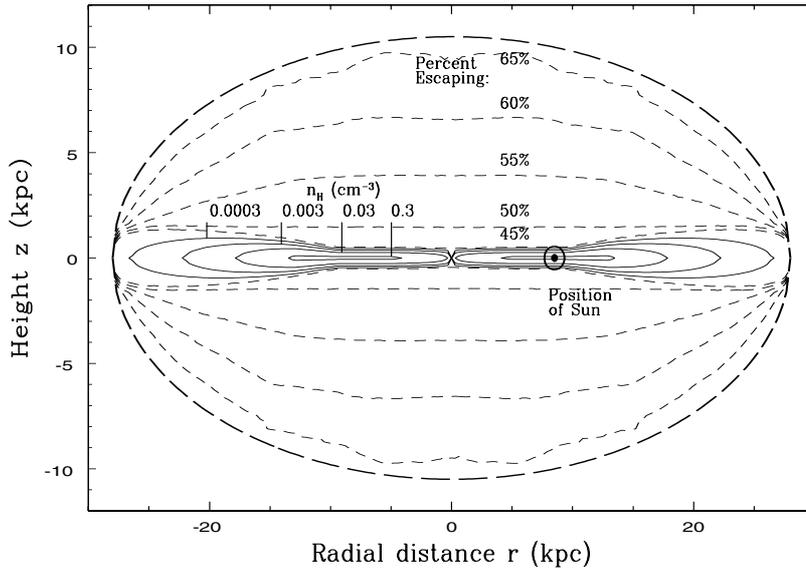


Figure C.1. Iso-absorption contours (light short-dashed lines) indicating the percentage of decay photons produced inside the halo which escape absorption by neutral hydrogen (solid lines) and reach the edge of the halo (heavy long-dashed lines). This plot assumes a small dark-matter halo of radial size $r_h = 28$ kpc.

of several ‘iso-absorption’ contours for a dark-matter halo with $r_h = 28$ kpc [3], considerably smaller than the ones we consider in chapter 7. In this case, the scale sizes of the disc and halo are close enough that nearly 50% of the decay photons are absorbed, cutting the luminosity of the neutrino halo in half.

References

- [1] Bowers R and Deeming T 1984 *Astrophysics II: Interstellar Matter and Galaxies* (Boston, MA: Jones and Bartlett) pp 383–91
- [2] Scheffler H and Elsässer H 1988 *Physics of the Galaxy and Interstellar Medium* (Berlin: Springer) pp 350–64
- [3] Overduin J M and Wesson P S 1997 *Astrophys. J.* **483** 77

Index

- Absorption
 - by Earth's atmosphere, 10
 - by intergalactic dust, 63, 128, 135
 - by neutral hydrogen, 135
 - in the disc, 129, 130, 208
 - intergalactic, 63, 128, 136
- Active galactic nuclei (AGN), 164
- Age
 - of metal-poor halo stars, 70
 - of the galaxies, 7, 28
 - of the Universe (t_0), 39, 70, 94, 99, 180
- Age crisis, 70
- AMANDA, 148
- Anthropic principle, 80
- Antipodal redshift, 80
- Apollo, 164
- Arrow of time, 91
- Asteroids, 2
- Axinos, 194
- Axion decay, 113
 - and cavity detectors, 115
 - bolometric intensity due to, 119
 - Feynman diagram, 115
 - halo luminosity due to, 118
 - lifetime (τ_a), 115
 - spectral intensity due to, 121
- Axions, 13, 76, 113, 146, 192
 - and red giants, 114, 116
 - and supernovae, 114, 116
 - comoving number density (n_a), 115
 - coupling parameter (ζ), 115, 122
 - density parameter (Ω_a), 116, 124
 - DFSZ, 116
 - in clusters of galaxies, 124
 - in galaxy halos, 117
 - invisible, 115
 - KSVZ (hadronic), 116
 - multi-eV, 116
 - production mechanisms
 - axionic string decay, 114
 - thermal, 115
 - vacuum misalignment, 114
 - rest mass (m_a), 114, 116, 124
- B-band, 21, 50, 136
- Baade–Wesselink method, 68
- Baryogenesis, 161
- Baryon-to-photon ratio, 84, 98
- Baryonic dark matter (BDM), 67
 - density parameter (Ω_{bar}), 67
- Baryonic matter
 - dark fraction (f_{bdm}), 72
 - density parameter (Ω_{bar}), 72, 130
- BDM, *see* Baryonic dark matter
- Beige dwarfs, 73
- Berkeley UVX, 142
- Bianchi identities, 91
- Big-bang theory, 6
- Big bounce
 - in vacuum-dominated models, 32
 - with variable Λ -term, 94
- Black dwarfs, 72
- Black holes, 13, 72, 170
 - higher-dimensional, *see* Solitons

- primordial, *see* Primordial black holes
 - temperature, 178
- Blazars, 164
- Boltzmann equation, 115, 146
- BOOMERANG, 83
- Brane-world scenario, 194
- Brans–Dicke theory, 92
- Brown dwarfs, 72

- CDM, *see* Cold dark matter
- CDMS, 148
- Cepheid variables, 68
- Čerenkov detectors, 116, 149, 185
- CGRO, 164
- Charge-parity (CP), 113
- Circular velocity (v_c), 50, 118
- CMB, *see* Cosmic microwave background
- CNOC collaboration, 74
- COBE, 11, 49, 83, 95, 108, 122
- Coincidence problem, 86
- Cold dark matter (CDM), 67, 70, 73, 85, 113, 146, 170
 - density parameter (Ω_{cdm}), 67, 76
- Comoving luminosity density (\mathcal{L}), 45
 - at $z = 0$ (\mathcal{L}_0), 21
 - relative to $z = 0$ ($\tilde{\mathcal{L}}$), 46
- Comoving number density (n), 19
 - at $z = 0$ (n_0), 117
 - relative to $z = 0$ (\tilde{n}), 59
- COMPTEL, 153, 156, 160, 163, 164, 184
- Conformal rescaling, 92
- Conservation
 - of energy, 90, 91, 93
 - of entropy, 103
 - of galaxies, 23
 - of particle number, 97
- Constrained minimal supersymmetric model (CMSSM), 148
- Continuum units (CUs), 46, 108

- Cosmic microwave background (CMB), 7, 13, 41
 - angular power spectrum, 83, 94
 - temperature (T_{cmb}), 103, 157
- Cosmic string loops, 78, 173
- Cosmological constant (Λ), 68, 78
 - variable, *see* Variable
 - cosmological term
- Cosmological-constant problem, 79, 85, 90, 113
- Cosmological models
 - closed (spherical), 25, 83
 - de Sitter, 30
 - Eddington–Lemaître, 80
 - Einstein–de Sitter (EdS), 27, 30, 47, 70, 73, 81, 98, 122, 133, 152, 181
 - flat (Euclidean), 24
 - Λ BDM, 48, 70, 72, 84, 98, 116
 - Λ CDM, 48, 70, 81, 85, 98, 116, 152
 - Milne, 31
 - OCDM, 48, 70
 - open (hyperbolic), 25, 85
 - radiation, 29
 - SCDM, *see* EdS above
- Cosmological scale factor (R), 17
 - as function of t , 37, 38, 71, 99, 181
 - at present (R_0), 17
 - relative to present (\tilde{R}), 19
- Critical density (ρ_{crit}), 24
 - present value of ($\rho_{\text{crit},0}$), 24
- Curvature constant (k), 17, 26
- Curvature scalar (\mathcal{R}), 17, 92
- CUs, *see* Continuum units
- CYGNUS air-shower array, 185

- DAMA, 148
- Dark energy, *see* Vacuum energy
- Dark matter, 13, 66, 192
 - baryonic, *see* Baryonic dark matter
 - cold, *see* Cold dark matter

- ‘four elements’ of, 69
- hot, *see* Hot dark matter
- warm, *see* Warm dark matter
- DASI, 83
- de Sitter model, *see* Cosmological models
- DE-1, 142
- Decaying-neutrino hypothesis, 77, 127, 142
- Dicke coincidences, 27
- DISM, *see* Interstellar medium, diffuse
- Dust
 - and dimming of SNIa, 82
 - and galaxy SEDs, 56
 - and lensing statistics, 81
 - extinction (ξ), 136, 137
- Dust grains
 - fluffy silicate, 138
 - Mie scattering, 137
 - PAH nanostructures, 138
 - silicate/graphite, 137
- DUVE, 142
- EBL, *see* Extragalactic background light
- Eclipsing binaries, 69
- EDELWEISS, 148
- EGRET, 153, 156, 160, 163, 164, 184
- Einstein–de Sitter (EdS) model, *see* Cosmological models
- Einstein limit ($\Omega_{\Lambda, E}$), 80
- Einstein tensor, 92
- Einstein’s ‘biggest blunder’, 79
- Einstein’s field equations
 - in 4 dimensions, 16
 - in 5 dimensions, 187
- Electromagnetic cascades, 158, 160
- Electron–positron annihilation, 103
- Electroweak theory, 79
- Energy–momentum tensor ($\mathcal{T}_{\mu\nu}$), 17
 - for induced matter, 188
 - for perfect fluids, 22, 96
 - for scalar fields, 93
- Equation of state, 22, 172
 - dustlike, 22, 96, 172, 176
 - radiation, 22, 96, 172
 - scalar field, 23
 - stiff, 23, 172
 - superhard, 172
 - vacuum, 22, 102
- Equivalence principle, 189
- EURD, 142
- Extragalactic background light (EBL), 11
 - bolometric intensity of (Q), 19
 - in de Sitter model, 30
 - in EdS model, 30
 - in Milne model, 31
 - in models with matter and vacuum energy, 33
 - in radiation model, 30
 - typical value (Q_*), 22
 - spectral intensity of (I_λ), 43
 - with blackbody SED, 53
 - with delta-function SED, 48
 - with Gaussian SED, 50
 - with two-component SED, 59
- Free-streaming neutrinos
 - bolometric intensity due to, 134
 - comoving luminosity density due to, 134
- Friedmann–Lemaître equation, 26
- Galaxies
 - faint blue excess, 58
 - faint number counts, 81, 94
 - low surface brightness, 61
 - luminosity (L_0), 21
 - luminosity density of, *see* Comoving luminosity density
 - luminosity ratio (ℓ_0), 57
 - merger parameter (η), 57
 - morphological types, 55

- normal (quiescent), 55
 - luminosity (L_n), 57
 - number density (n_n), 57
- number density of, *see*
 - Comoving number density
- peculiar velocities, 75
- spectral energy distributions, 56
- starburst (active), 44, 55
 - luminosity (L_s), 57
 - number density (n_s), 57
- starburst fraction (f), 57
 - at $z = 0$ (f_0), 57
- Galaxy clusters
 - baryon fraction, 74
 - rich, 74
- Galaxy current (J_g^μ), 18
- Galaxy distribution
 - bias in, 81
 - power spectrum of, 75
- Galaxy evolution
 - luminosity density, 45
 - number density, 59, 75
 - starburst fraction, 59
- Gallex, 78
- γ -ray band, 11, 147, 160, 163, 183
 - high-energy, 164
 - low-energy, 164
 - very-high-energy (VHE), 165
- γ -ray bursts (GRBs), 173
- General relativity, 16
 - classical tests of, 189
 - in higher dimensions, *see*
 - Kaluza–Klein gravity
- GI theory, *see* Gravitational instability theory
- Gluinos, 194
- Grand unified theories (GUTs), 79, 115, 173
- Gravitational instability (GI) theory, 75–77
- Gravitational lensing, 70, 80
 - statistics of, 81, 94
- Gravitino problem, 161, 186
- Gravitinos, 147, 193
 - bolometric intensity due to, 162
 - decay, *see* WIMP decay
 - decay lifetime ($\tau_{\tilde{g}}$), 162, 166
 - rest mass ($m_{\tilde{g}}$), 161
 - spectral intensity due to, 162
- Hawking evaporation, 13, 170
- HDM, *see* Hot dark matter
- HEAO-1, 164
- Higgsinos (\tilde{h}), 147, 159
- Homestake, 78
- Hopkins Ultraviolet Telescope, 128
- Hopkins UVX, 49, 121
- Hot dark matter (HDM), 67, 77, 128
- Hubble Deep Field (HDF), 46, 72
- Hubble Key Project (HKP), 68
- Hubble Space Telescope (HST), 11, 49, 122
- Hubble’s energy and number effects, 17
- Hubble’s law, 6
- Hubble’s parameter (H), 20
 - as function of t , 38
 - as function of z , 26, 119, 133
 - at present (H_0), 20
 - normalized (h_0), 20, 70, 130
 - relative to present (\tilde{H}), 21, 26
- HZT, 81
- ICS, *see* Inverse Compton scattering
- Inflation, 23, 27, 73, 161, 186
 - and reheat temperature, 161
- Infrared (IR) band, 11, 12, 120
 - far infrared (FIR), 109
 - near infrared (NIR), 48
- Integral flux, conversion to, 165
- Intensity
 - of a single galaxy, 17
 - of all galaxies, *see* Extragalactic background light
 - of galaxies in a shell
 - bolometric, 19

- spectral, 42
- ‘reading-room’, 9, 39
- Intergalactic medium (IGM), 135
- Interstellar medium, diffuse (DISM), 137
- Inverse Compton scattering (ICS), 157, 160, 161
- Isothermal density profile, 151
- J-band, 49
- Jaffe profile, 118
- K-band, 49, 81
- Kaluza–Klein gravity, 186
- Kitt Peak telescope, 12
- Kosmos, 164
- L-band, 49
- Large Magellanic Cloud (LMC), 68, 72, 149
- Large-scale structure formation, 73, 81, 94, 127
- Las Campanas telescope, 49, 122
- Lee–Weinberg bound, 150
- LEP collider, 77
- Light-travel time, 7
- Lightest supersymmetric particle (LSP), 146, 161
- Long-baseline radio interferometry, 69
- Look-back time, 19
- LSND, 78
- Luminous matter
 - average density, 6
 - density parameter (Ω_{lum}), 68, 71
 - rate of energy output, 5
- Lyman- α absorbers, 72, 75, 84, 136
- M-theory, 194
- Mach’s principle, 92
- MACHO survey, 72, 73, 173
- MACHOs, *see* Massive compact halo objects
- Magnitude-redshift relation, 81
- MAP, 84
- Mass-to-light ratio, 74
- Mass-to-light ratio (M/L), 72
- Massive compact halo objects (MACHOs), 72, 74
- Matter density (ρ_m), 23, 99
- Matter density parameter ($\Omega_{m,0}$), 25, 67, 85, 130
 - evolution of (Ω_m), 85
- Matter–antimatter annihilation, 95
- Matter–radiation equality, epoch of (t_{eq}), 97, 100
- MAXIMA, 83
- Metric tensor ($g_{\mu\nu}$), 17
- MeV bump, demise of, 164
- Microlensing, 72
- Milky Way
 - dynamical mass, 74, 118, 130, 151, 155
 - luminosity (B-band), 74
- Milne model, *see* Cosmological models
- Minimal supersymmetric model (MSSM), 147
- Naked singularities, 188
- Neutralinos, 147
 - annihilation cross section, 150, 151
 - bolometric intensity due to, 152, 155
 - decay, *see* WIMP decay
 - decay lifetime ($\tau_{\tilde{\chi}}$), 154, 156, 166
 - in halos, *see* WIMP halos
 - rest mass ($m_{\tilde{\chi}}$), 147
 - spectral intensity due to, 153, 155, 159
- Neutrino decay, 127, 150
 - Feynman diagrams, 129
 - lifetime (τ_ν), 127
 - spectral intensity due to, 141
- Neutrino halos, 128
 - bolometric intensity due to, 133

- efficiency factor (ϵ), 129, 132, 206
- ellipsoidal density profile, 131, 206
- luminosity (L_h), 129, 133
- mass (M_h), 130
- spectral intensity due to, 140
- Neutrino oscillation, 78
- Neutrinoless double-beta decay, 78
- Neutrinos, 13, 67, 192
 - blackbody temperature (T_ν), 103
 - bound, *see* Neutrino halos
 - density parameter (Ω_ν), 67, 77, 78, 103, 130
 - free-streaming, *see* Free-streaming neutrinos
 - in clusters of galaxies, 128
 - number density (n_ν), 76
 - rest mass (m_ν), 77, 127
 - sterile, 193
- Neutron stars, 72
- Non-minimal coupling, 92
- Nucleosynthesis
 - primordial, 71, 94, 98, 161
 - stellar, 5
- OA0-2, 49, 121, 142
- OGLE microlensing survey, 69
- Olbers' paradox, 1
 - and absorption, 3, 64
 - and astronomy textbooks, 4
 - and expansion, 4, 8, 35, 64, 192
 - and finite age, 4, 7, 35, 64, 192
 - and galaxy distribution, 3
 - and spatial curvature, 4
 - and the laws of physics, 4
 - bolometric resolution of, 35
 - in de Sitter model, 31
 - spectral resolution of, 63
- Optical band, 10, 120
- Optical depth, 135, 207
 - in dust, 136
 - in neutral hydrogen, 136
- Pair production, 156, 158, 181
- PBHs, *see* Primordial black holes
- Peccei-Quinn (PQ) symmetry, 113
- Phase transition, 97, 173
- Photinos, 147, 150, 151
- Pioneer 10, 46, 49, 121
- Planck, 84
- Planck function, 51
- Planck's law, 8
- Primordial black holes (PBHs), 170, 192
 - and γ -ray bursts, 173
 - and γ -ray emission from halo, 173
 - and MACHO events, 173
 - and quasar variability, 173
 - bolometric intensity due to, 180–182
 - comoving number density, 175
 - cut-off mass, 180
 - density (ρ_{pbh}), 176
 - density parameter (Ω_{pbh}), 176, 181, 185
 - direct detection, 185
 - equation-of-state parameter (β), 175
 - formation of
 - during phase transition, 173
 - via collapse of small density fluctuations, 172
 - via cosmic string loops, 173
 - via critical collapse, 173, 186
 - via inflationary mechanisms, 172, 186
 - initial mass distribution, 172
 - lifetime (t_{pbh}), 174
 - local evaporation rate, 185
 - luminosity, 179
 - normalization (\mathcal{N}), 175, 185
 - photosphere, 186
 - spectral intensity due to, 183, 184
 - stable relics, 186
 - typical mass (M_*), 175
- Prognoz, 142

- Pure luminosity evolution (PLE), 52
- Quantum chromodynamics (QCD), 79, 113
- Quark and gluon jets, 185
- Quark–hadron transition, 173, 176
- Quasars, 3
 - lensing frequency, 81
 - spectra, *see* Lyman- α absorbers
 - turnoff, 136
 - variability, 72, 173
- Quintessence, 23, 67, 92
- R-parity, 147, 154, 161, 167
- Radiation density (ρ_r), 23, 99
- Radiation density parameter ($\Omega_{r,0}$), 25, 67, 104
- Radio galaxy lobes, 75
- Red clump stars, 69
- Red dwarfs, 72
- Redshift (z), 6, 20
- Ricci tensor
 - in 4 dimensions ($\mathcal{R}_{\mu\nu}$), 17
 - in 5 dimensions (\mathcal{R}_{AB}), 187
- Robertson–Walker metric, 17
- SAGE, 78
- SAS-2, 153, 156, 160, 163, 164, 184
- Scalar field, 23, 90, 91
- Scalar–tensor theories, 92
- Schwarzschild metric, 187
- SCP, 81
- SED, *see* Spectral energy distribution
- Selectrons, 147
- Sloan Digital Sky Survey (SDSS), 21, 61
- Sneutrinos, 147
- Soliton metric, 187
- Solitons, 171, 192
 - density profile, 188, 189
 - gravitational mass, 188
 - local energy density (ρ_s), 189
- Solrad-11, 142
- Spectral energy distribution, 42
 - blackbody, 52, 106, 177
 - delta-function, 44
 - Gaussian, 50, 121, 129, 150
 - inverse Compton scattering, 158
 - with cascades, 158
 - normalization, 42
 - primordial black hole, 177
- Static analogue
 - bolometric intensity, 28
 - spectral intensity, 43
- Steady-state theory, 3
- Stefan’s law, 52, 103, 107
- String theory, 194
- Strong-CP problem, 113, 167
- Subaru Deep Field, 61
- Sunyaev–Zeldovich (SZ) effect, 68
- Super-Kamiokande, 78
- Supernovae, Type Ia (SNIa), 81, 94, 114, 164
- Supersymmetry, 127, 146, 174
- SUSY, *see* Supersymmetry
- Total density parameter ($\Omega_{\text{tot},0}$), 25, 26
- Two Degree Field (2dF), 21
- Ultraviolet (UV) band, 11, 150
 - extreme ultraviolet (EUV), 128
 - far ultraviolet (FUV), 128
 - near ultraviolet (NUV), 48, 128
- Unified field theory, 16
- Vacuum decay, 13, 90
 - and cosmic microwave background, 95
 - bolometric intensity of, 106
 - branching ratio (β), 103, 104
 - comoving luminosity density of ($\mathcal{L}_{\nu,0}$), 105
 - coupling parameter (x), 97
 - into baryons, 95

- into radiation, 95, 97
- luminosity (L_v), 104
 - relativistic, 102
 - thermodynamical, 102
- phenomenological model, 95
- source regions, 101
- spectral intensity of, 108
- Vacuum density (ρ_v), 23, 99
 - in terms of Λ (ρ_Λ), 24, 79, 96
- Vacuum density parameter ($\Omega_{\Lambda,0}$), 25, 78, 82, 83, 85
 - evolution of (Ω_Λ), 85
- Vacuum energy, 67, 78, 192
- Variable cosmological term, 80, 90
- Velocity dispersion, 121, 129, 150
- Very massive objects (VMOs), 193
- VIRGO consortium, 81
- Voyager 2, 142

- Warm dark matter (WDM), 193
- Weakly-interacting massive particles (WIMPs), 13, 76, 146, 192
 - cosmological enhancement factor (f_c), 152
 - decay, *see* WIMP decay
 - direct detection of, 148
 - in halos, *see* WIMP halos
 - indirect detection of, 148
 - pair annihilation of, 146
 - Feynman diagrams, 149
- Whipple collaboration, 185
- White dwarfs, 72
- Wien's law, 44
- WIMP decay
 - branching ratio, 155, 158, 162
 - Feynman diagrams, 154, 157
 - one-loop, 147, 154, 165
 - tree-level, 147, 157, 161
- WIMP halos, 150
 - cosmological density of, 151
 - density profile, *see* Isothermal density profile
 - luminosity of, 150, 152, 154, 155, 158, 159, 162
- WIMPs, *see* Weakly-interacting massive particles
- WIMPzillas, 193

- X-ray band, 11, 147, 160, 163, 193
 - hard, 164
 - soft, 164

- Zero-point field (ZPF), 67
- Zinos, 147, 151
- Zodiacal light, 46