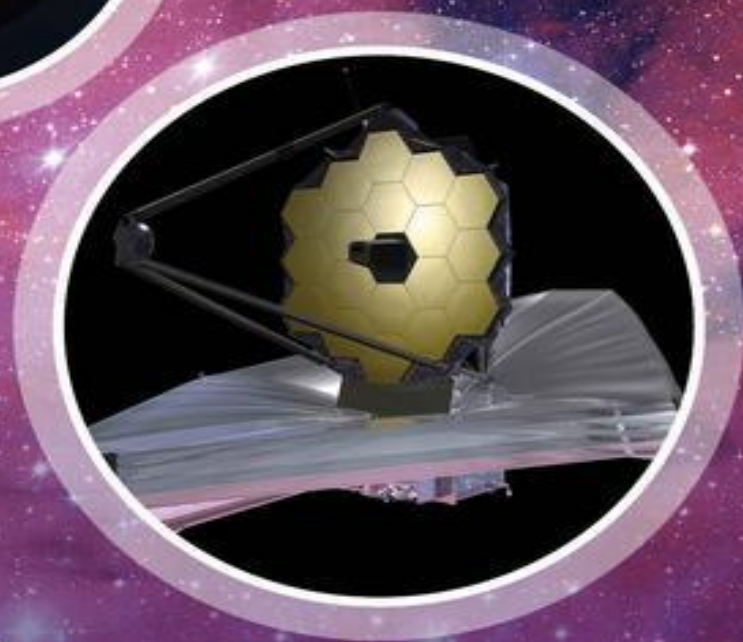


**Marina Cortês
Andrew Liddle**



An Introduction to
MODERN
COSMOLOGY

Fourth Edition

WILEY

Table of Contents

[Cover](#)

[Table of Contents](#)

[Title Page](#)

[Copyright](#)

[Dedication](#)

[Preface](#)

[About the Companion Website](#)

[Chapter 1: A Brief History of Cosmological Ideas](#)

[Chapter 2: Observational Overview](#)

[2.1 In visible light](#)

[2.2 In other wavebands](#)

[2.3 Homogeneity and isotropy](#)

[2.4 The expansion of the Universe](#)

[2.5 Particles in the Universe](#)

[Problems](#)

[Notes](#)

[Chapter 3: Newtonian Gravity](#)

[3.1 The Friedmann equation](#)

[3.2 On the meaning of the expansion](#)

[3.3 Things that go faster than light](#)

[3.4 The fluid equation](#)

[3.5 The acceleration equation](#)

[3.6 On mass, energy, and vanishing factors of \$c^2\$](#)

[Problems](#)

[Note](#)

[Chapter 4: The Geometry of the Universe](#)

[4.1 Flat geometry](#)

[4.2 Spherical geometry](#)

[4.3 Hyperbolic geometry](#)

[4.4 Infinite and observable universes](#)

[4.5 Where did the Big Bang happen?](#)

[4.6 Three values of \$k\$](#)

[Problems](#)

[Notes](#)

[Chapter 5: Simple Cosmological Models](#)

[5.1 Hubble's law](#)

[5.2 Expansion and redshift](#)

[5.3 Solving the equations](#)

[5.4 Particle number densities](#)

[5.5 Evolution including curvature](#)

[Problems](#)

[Note](#)

[Chapter 6: Observational Parameters](#)

[6.1 The expansion rate \$H_0\$](#)

[6.2 The density parameter \$\Omega_0\$](#)

[6.3 The deceleration parameter \$q_0\$](#)

[Problems](#)

[Note](#)

[Chapter 7: The Cosmological Constant](#)

[7.1 Introducing \$\Lambda\$](#)

[7.2 Fluid description of \$\Lambda\$](#)

[7.3 Cosmological models with \$\Lambda\$](#)

[Problems](#)

[Note](#)

[Chapter 8: The Age of the Universe](#)

[Problems](#)

[Chapter 9: The Density of the Universe and Dark Matter](#)

[9.1 Weighing the Universe](#)

[9.2 What might the dark matter be?](#)

[9.3 Dark matter searches](#)

[Problems](#)

[Notes](#)

[Chapter 10: The Cosmic Microwave Background](#)

[10.1 Properties of the microwave background](#)

[10.2 The photon to baryon ratio](#)

[10.3 The origin of the microwave background](#)

[10.4 The origin of the microwave background \(optional advanced treatment\)](#)

[Problems](#)

[Chapter 11: The Early Universe](#)

[Problems](#)

[Chapter 12: Nucleosynthesis: The Origin of the Light Elements](#)

[12.1 Hydrogen and helium](#)

[12.2 Comparing with observations](#)

[12.3 Contrasting decoupling and nucleosynthesis](#)

[Problems](#)

[Note](#)

[Chapter 13: The Inflationary Universe](#)

[13.1 Problems with the Hot Big Bang](#)

[13.2 Inflationary expansion](#)

[13.3 Solving the Big Bang problems](#)

[13.4 How much inflation?](#)

[13.5 Inflation and particle physics](#)

[Problems](#)

[Chapter 14: The Initial Singularity](#)

[Problem](#)

[Chapter 15: Overview: The Standard Cosmological Model](#)

[15.1 Expansion](#)

[15.2 Geometry](#)

[15.3 Age](#)

[15.4 Fate](#)

[15.5 Contents](#)

[15.6 Early history](#)

[15.7 Outlook](#)

[Advanced Topic 1: General Relativistic Cosmology](#)

[A1.1 The metric of space-time](#)

[A1.2 The Einstein equations](#)

[A1.3 Aside: Topology of the Universe](#)

[Problems](#)

[Advanced Topic 2: Classic Cosmology: Distances and Luminosities](#)

[A2.1 Light propagation and redshift](#)

- [A2.2 The observable Universe](#)
- [A2.3 Luminosity distance](#)
- [A2.4 Angular diameter distance](#)
- [A2.5 Source counts](#)
- [Problems](#)
- [Advanced Topic 3: Neutrino Cosmology](#)
 - [A3.1 The massless case](#)
 - [A3.2 Massive neutrinos](#)
 - [A3.3 Neutrinos and structure formation](#)
 - [Problems](#)
- [Advanced Topic 4: Baryogenesis](#)
 - [Problem](#)
- [Advanced Topic 5: Structures in the Universe](#)
 - [A5.1 The observed structures](#)
 - [A5.2 Gravitational instability](#)
 - [A5.3 The clustering of galaxies](#)
 - [A5.4 Cosmic microwave background anisotropies](#)
 - [A5.5 The origin of structure](#)
 - [Problems](#)
- [Advanced Topic 6 Constraining Cosmological Models](#)
 - [A6.1 Cosmological models and parameters](#)
 - [A6.2 Key cosmological observations](#)
 - [A6.3 Cosmological data analysis](#)
 - [A6.4 The Standard Cosmological Model: 2025 edition](#)
 - [A6.5 Tensions and hints of new physics](#)
 - [A6.6 The future](#)
 - [Problems](#)
- [Bibliography](#)
- [Numerical Answers and Hints to Problems](#)
- [Index](#)
- [End User License Agreement](#)

List of Illustrations

Chapter 2

- [Figure 2.1 If viewed from above the disk, our own Milky Way galaxy would probably resemble ...](#)
- [Figure 2.2 A map of galaxy positions in a narrow slice of the Universe, as measured by the...](#)
- [Figure 2.3 Images of the Coma cluster of galaxies in visible/infrared light \(top\) and in X-...](#)
- [Figure 2.4 A computer simulation showing the predicted distribution of matter in the Univer...](#)
- [Figure 2.5 The cosmic microwave background spectrum as measured by the FIRAS experiment...](#)
- [Figure 2.6 The cosmic microwave background temperature variations on the sky as measured b...](#)
- [Figure 2.7 A plot of velocity versus estimated distance for a set of 1355 galaxies. A straight-line relation...](#)
- [Figure 2.8 According to Hubble's law, the further away from us a galaxy is, the faster it i...](#)
- [Figure 2.9 The Planck function of Equation \(2.7\). There are far more photons with very low...](#)
- [Figure 2.10 The energy density distribution of a blackbody spectrum, given by Equation \(2.8\). Most of...](#)

Chapter 3

- [Figure 3.1 The particle at radius \$r\$ only feels gravitational attraction from the shaded region. Any...](#)
- [Figure 3.2 The comoving coordinate system is carried along with the expansion, so that any...](#)

Chapter 4

- [Figure 4.1 A sketch of a spherical surface, representing positive \$k\$. A triangle is shown, which has...](#)

[Figure 4.2 A sketch of a saddle surface, representing the hyperbolic geometry obtained when \$k\$ is...](#)

Chapter 5

[Figure 5.1 A photon travels a distance \$dr\$ between two galaxies A and B.](#)

[Figure 5.2 A schematic illustration of the evolution of a universe containing radiation and...](#)

[Figure 5.3 Three possible evolutions for the Universe, corresponding to the different signs of \$k\$. The...](#)

Chapter 7

[Figure 7.1 Different models for the Universe can be identified by their location in the plane showing the...](#)

Chapter 8

[Figure 8.1 Predicted ages as fractions of the Hubble time \$H_0^{-1}\$ for open universes and for universes with...](#)

Chapter 9

[Figure 9.1 The rotation curve of the spiral galaxy NGC3198. We see that it remains roughly...](#)

[Figure 9.2 A schematic illustration of a galactic disk, with a few globular clusters, embed...](#)

Chapter 10

[Figure 10.1 The evolution of the blackbody spectrum as the Universe expands. The expansion ...](#)

[Figure 10.2 Because they have traveled toward us uninterrupted since the Universe was at 3000 K, the...](#)

Chapter 11

[Figure 11.1 A schematic illustration of the temperature–time relation, assuming \$\Omega_0 = 0.3\$ and \$h = 0.7\$...](#)

Chapter 12

[Figure 12.1 The predicted abundances of light nuclei, as a function of \$\Omega_B h^2\$ along the top and of the...](#)

Chapter 13

[Figure 13.1 An illustration of the horizon problem. We receive microwave radiation from poi...](#)

[Figure 13.2 Possible evolution of the density parameter \$\Omega_{\text{tot}}\$. There might or might not be a period before inflation,...](#)

[Figure 13.3 A schematic illustration of the inflationary solution to the horizon problem, w...](#)

Chapter 14

[Figure 14.1 The solid line shows the true \(decelerating\) scale factor. The dotted line extr...](#)

List of Tables

Chapter 4

[Table 4.1 A summary of possible geometries.](#)

Chapter 11

[Table 11.1 Different stages of the Universe's evolution \(taking \$\Omega_0 = 0.3\$ and \$h = 0.7\$ \). Some nu...](#)

Chapter 12

[Table 12.1 A comparison of nucleosynthesis and decoupling.](#)

An Introduction to Modern Cosmology

Marina Cortês and Andrew Liddle

Institute for Astrophysics and Space Sciences

Faculty of Sciences

University of Lisbon

Campo Grande

Lisbon, Portugal

Fourth Edition

WILEY

This fourth edition first published 2026

© 2026 John Wiley & Sons Ltd

Edition History

John Wiley & Sons Ltd (1e, 1998); John Wiley & Sons Ltd (2e, 2003); John Wiley & Sons Ltd (3e, 2015)

All rights reserved, including rights for text and data mining and training of artificial intelligence technologies or similar technologies. 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, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

The right of Marina Cortês and Andrew Liddle to be identified as the author(s) of this work has been asserted in accordance with law.

Registered Office(s)

John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA

John Wiley & Sons Ltd, New Era House, 8 Oldlands Way, Bognor Regis, West Sussex, PO22 0NO, UK

For details of our global editorial offices, customer services, and more information about Wiley products, visit us at www.wiley.com.

The manufacturer's authorized representative according to the EU General Product Safety Regulation is Wiley-VCH GmbH, Boschstr. 12, 69469 Weinheim, Germany, e-mail: Product_Safety@wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Trademarks: Wiley and the Wiley logo are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates in the United States and other countries and may not be used without written permission. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc. is not associated with any product or vendor mentioned in this book.

Limit of Liability/Disclaimer of Warranty

While the publisher and the authors have used their best efforts in preparing this work, including a review of the content of the work, neither the publisher nor the authors make any representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

Library of Congress Cataloging-in-Publication Data:

Names: Liddle, Andrew R. author | Cortês, Marina V. author

Title: An introduction to modern cosmology / Marina V. Cortês, Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal, Andrew Liddle, Institute for Astrophysics and Space Sciences, Faculty of Sciences, University of Lisbon, Campo Grande, Lisbon, Portugal.

Description: Fourth edition. | Hoboken, NJ, USA : John Wiley & Sons Ltd, 2026. | Includes index.

Identifiers: LCCN 2025045575 | ISBN 9781394278435 print | ISBN 9781394278503 ePDF | ISBN 9781394278442 ePUB

Subjects: LCSH: Cosmology

Classification: LCC QB981 .L567 2026

LC record available at <https://lcn.loc.gov/2025045575>

Cover Design: Wiley

Cover Images: Courtesy of NASA, © Nihit RPG/Shutterstock, © Outer Space/Shutterstock, © Vadim Sadovski/Shutterstock

To Dawa and Kylo

Preface

The development of cosmology will no doubt be seen as one of the scientific triumphs of the twentieth century. At its beginning, cosmology hardly existed as a scientific discipline. By its end, the Hot Big Bang cosmology stood secure as the accepted description of the universe as a whole. Telescopes such as the Hubble and James Webb Space Telescopes are capable of seeing light from galaxies so distant that the light has been traveling toward us for most of the lifetime of the universe. The cosmic microwave background, a fossil relic of a time when the universe was both denser and hotter, is routinely detected and its properties examined. That our universe is presently expanding is established without doubt, and its material composition is accurately determined.

We are now in an era where understanding of cosmology has shifted from the qualitative to the quantitative. The turn of the millennium saw the establishment of what has come to be known as the Standard Cosmological Model, representing an almost universal consensus among cosmologists as to the best description of our universe. Nevertheless, it is a model with a major surprise – the belief that our universe is presently experiencing accelerated expansion. Add to that ongoing mysteries such as the properties of the so-called dark matter, which is believed to be the dominant form of matter in the universe, and it is clear that we have some way to go before we can say that a full picture of the physics of the universe is in our grasp. The first decades of the twenty-first century have seen dramatic leaps in the precision of cosmological probes, but this accuracy has yet to yield any substantial improvements to our overall understanding.

Such a bold endeavor as cosmology easily captures the imagination, and there has been increasing demand for cosmology to be taught at university in an accessible manner. Traditionally, cosmology was taught as the tail end of a general relativity course, with a derivation of the metric for an expanding universe and a few solutions. Such a course fails to capture the flavor of modern cosmology, which takes classic physical sciences like thermodynamics, atomic physics, and gravitation and apply them on a grand scale.

In fact, introductory modern cosmology can be tackled in a different way, by avoiding general relativity altogether. By a lucky chance and a subtle bit of cheating, the correct equations describing an expanding universe can be obtained from Newtonian gravity. From this basis, one can study all the triumphs of the Hot Big Bang cosmology – the expansion of the Universe, the prediction of its age, the existence of the cosmic microwave background, and the abundances of light elements such as helium and deuterium – and even go on to discuss more speculative ideas such as the inflationary cosmology.

The origin of this book, first published in 1998, was a short lecture course at the University of Sussex, around 20 lectures, taught by Andrew to students in the final year of a bachelor's degree or the penultimate year of a master's degree. The prerequisites are all very standard physics, and the emphasis is aimed at physical intuition rather than mathematical rigor. It has been further developed through his experiences of teaching cosmology at Imperial College London and at the University of Edinburgh, plus much useful feedback from readers.

The structure of the book is a central “spine,” the main chapters from 1 to 15, which provide a self-contained introduction to modern cosmology. In addition, there are six Advanced Topic chapters, each with prerequisites, which can be added to extend the course as desired. Ordinarily, the best time to tackle those Advanced Topics is immediately after their prerequisites have been attained, though they could also be included at any later stage.

Cosmology is an interesting course to teach, as it is not like most of the other subjects taught in undergraduate physics courses. There is no perceived wisdom, built up over a century or more, which provides an unquestionable foundation, as in thermodynamics, electromagnetism, and even quantum mechanics and general relativity. Each successive edition of this book has introduced new discoveries, and lecturers can expect opportunities to crop up during their course to discuss new results which impact on cosmologists' views of the universe.

You can follow our own evolving views by checking out this book's WWW Home Page. There you may find some updates on observations, a detailed solution set for instructors, and also a list of any errors in the book that we are aware of. If you are confident you've found one yourself, and it's not on the list, we'd be very pleased to hear of it.

Marina Cortês and Andrew Liddle
Lisbon
July 2025

Some fundamental constants

Newton's constant	G	$6.672 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ sec}^{-2}$
Speed of light	c	$2.998 \times 10^8 \text{ m sec}^{-1}$ or $3.076 \times 10^{-7} \text{ Mpc yr}^{-1}$
Reduced Planck constant	$\hbar = h/2\pi$	$1.055 \times 10^{-34} \text{ m}^2 \text{ kg sec}^{-1}$
Boltzmann constant	k_B	$1.381 \times 10^{-23} \text{ J K}^{-1}$ or $8.619 \times 10^{-5} \text{ eV K}^{-1}$
Radiation constant	$\alpha = \pi^2 k_B^4 / 15 \hbar^3 c^3$	$7.565 \times 10^{-16} \text{ J m}^{-3} \text{ K}^{-4}$
Electron mass-energy	$m_e c^2$	0.511 MeV
Proton mass-energy	$m_p c^2$	938.3 MeV
Neutron mass-energy	$m_n c^2$	939.6 MeV
Thomson cross-section	σ_e	$6.652 \times 10^{-29} \text{ m}^2$
Free neutron half-life	t_{half}	609 sec

Some conversion factors

$$\begin{aligned}
 1 \text{ pc} &= 3.261 \text{ light years} = 3.086 \times 10^{16} \text{ m} \\
 1 \text{ yr} &= 3.156 \times 10^7 \text{ sec} \\
 1 \text{ eV} &= 1.602 \times 10^{-19} \text{ J} \\
 1 M_{\odot} &= 1.989 \times 10^{30} \text{ kg} \\
 1 \text{ J} &= 1 \text{ kg m}^2 \text{ sec}^{-2} \\
 1 \text{ Hz} &= 1 \text{ sec}^{-1}
 \end{aligned}$$

Commonly used symbols

z	redshift	<i>defined on page</i> 12 , 48
H_0	Hubble constant	13 , 59
r	physical distance	13
v	velocity	13
f	frequency	16
T	temperature	17
k_B	Boltzmann constant	17
ϵ	energy density	19
α	radiation constant	19
G	Newton's gravitational constant	23
ρ	mass density	24
a	scale factor	26
x	comoving distance	26
k	curvature	27
p	pressure	28
	(or occasionally momentum)	15
n	number density	53
H	Hubble parameter	60
h	Hubble constant	60
	(or Planck's constant)	16
Ω_0	present density parameter	61
ρ_c	critical density	61
Ω	density parameter	62
Ω_k	curvature "density parameter"	63
q_0	deceleration parameter	63
Λ	cosmological constant	65
Ω_Λ	cosmological constant density parameter	66
t	time	73
t_0	present age	73
Ω_B	baryon density parameter	80
Y_4	helium abundance	114
d_{lum}	luminosity distance	151
d_{diam}	angular diameter distance	156
$\Delta T/T$, C_ℓ	cosmic microwave background anisotropies	181 , 182

About the Companion Website

This book is accompanied by a companion website for instructors:



<http://www.wiley.com/go/AnIntroductiontoModernCosmology4e>

This website includes:

- Full model solutions

Chapter 1

A Brief History of Cosmological Ideas

The cornerstone of modern cosmology is the belief that the place that we occupy in the Universe is in no way special. This is known as the **cosmological principle**, and it is an idea that is both powerful and simple. It is intriguing, then, that for the bulk of the history of civilization it was believed that we occupy a very special location, usually the center, in the scheme of things.

The ancient Greeks, in a model further developed by the Alexandrian Ptolemy, believed that the Earth must lie at the center of the cosmos. It would be circled by the Moon, the Sun, and the planets, and then the “fixed” stars would be yet further away. A complex combination of circular motions, Ptolemy’s epicycles, was devised in order to explain the motions of the planets, especially the phenomenon of retrograde motion where planets appear to temporarily reverse their direction of motion. It was not until the early 1500s that Copernicus stated forcefully the view, initiated nearly 2000 years before by Aristarchus, that one should regard the Earth, and the other planets, as going around the Sun. By ensuring that the planets moved at different speeds, retrograde motion could easily be explained by this theory. However, although Copernicus is credited with removing the anthropocentric view of the Universe, which placed the Earth at its center, he still believed that the Sun was at the center.

Newton’s theory of gravity put what had been an empirical science (Kepler’s discovery that the planets moved on elliptical orbits) on a solid footing, and it appears that Newton believed that the stars were also suns pretty much like our own, distributed evenly throughout infinite space, in a static configuration. However, it seems that Newton was aware that such a static configuration is unstable.

Over the next 200 years, it became increasingly understood that the nearby stars are not evenly distributed, but rather are located in a disk-shaped assembly, which we now know as the Milky Way galaxy. The Herschels were able to identify the disk structure in the late 1700s, but their observations were not perfect, and they wrongly concluded that the Solar System lay at its center. Only in the early 1900s was this convincingly overturned, by Shapley, who realized that we are some two-thirds of the radius away from the center of the Galaxy. Even then, he apparently still believed our galaxy to be at the center of the Universe. It was only in 1952 that Baade finally demonstrated that the Milky Way is a fairly typical galaxy, leading to the modern view, known as the **cosmological principle** (or sometimes the Copernican principle), that the Universe looks the same wherever you are.

It is important to stress that the cosmological principle isn’t exact. For example, no one thinks that sitting in a lecture theater is exactly the same as relaxing at the beach, and the interior of the Sun is a very different environment from the interstellar regions. Rather, it is an approximation that we believe holds better and better the larger the length scales we consider. Even on the scale of individual galaxies, it is not very good, but once we take very large regions (though still much smaller than the Universe itself), containing say a million galaxies, we expect every such region to look more or less like every other one. The cosmological principle is therefore a property of the global Universe, breaking down if one looks at local phenomena.

The cosmological principle is the basis of Big Bang cosmology. The Big Bang is the best description we have of our Universe, and the aim of this book is to explain why. The Big Bang is a picture of our Universe as an evolving entity, which was very different in the past as compared to the present. Originally, it was forced to compete with a rival idea, the Steady State Universe, which holds that the Universe does not evolve but rather has looked the same forever, with new material being created to fill the gaps as the Universe expands. However, the observations we will describe now support the Big Bang so strongly that the Steady State theory is almost never considered.

Chapter 2

Observational Overview

For most of history, astronomers have had to rely on light in the visible part of the spectrum in order to study the Universe. One of the great astronomical achievements of the twentieth century was the exploitation of the full electromagnetic spectrum for astronomical measurements. We now have instruments capable of making observations of radio waves, microwaves, infrared light, visible light, ultraviolet light, X-rays, and gamma rays, which all correspond to light waves of different (in this case increasing) frequency. We have also entered an epoch where we can go beyond the electromagnetic spectrum and receive information of other types. A remarkable feature of observations of a nearby supernova in 1987 was that it was also seen through the detection of neutrinos, an extraordinarily weakly interacting type of particle normally associated with radioactive decay. Very high-energy cosmic rays, consisting of highly relativistic elementary particles, are now routinely detected, though as yet there is no clear understanding of their astronomical origin. And, most excitingly of all, in 2015, the Advanced Laser Interferometer Gravitational-wave Observatory (LIGO) made the first direct detection of gravitational waves, ripples in space-time itself, which has become a technique to observe astronomical events such as colliding stars and black hole mergers.

The advent of large ground- and satellite-based telescopes operating in all parts of the electromagnetic spectrum has revolutionized our picture of the Universe. While there are probably gaps in our knowledge, some of which may be important for all we know, we do seem to have a consistent picture, based on the cosmological principle, of how material is distributed in the Universe. Our discussion here is brief; for a more detailed discussion of the observed Universe, see some of the more advanced textbooks listed in the Bibliography. A huge array of images can be found online; the ones shown here are just a taster for what you can find for yourself.

2.1 In visible light

Historically, our picture of the Universe was built up through ever more careful observations using visible light.

Stars: The main source of visible light in the Universe is nuclear fusion within stars. The Sun is a fairly typical star, with a mass of about 2×10^{30} kg. This is known as a solar mass, indicated $M_{\odot}$, and is a convenient unit for measuring masses.

The nearest stars to us are a few light years away, a light year being the distance (about 10^{16} m) that light can travel in a year. For historical reasons, an alternative unit, known as the **parsec** and denoted “pc,” ¹ is more commonly used in cosmology. A parsec equals 3.261 light years. In cosmology, one seldom considers individual stars, instead preferring to adopt as the smallest considered unit the conglomerations of stars known as ...

Galaxies: Our Solar System lies some way off-center in a giant disk structure known as the Milky Way galaxy. It contains a staggering hundred thousand million (10^{11}) or so stars, with masses ranging from about a tenth that of our Sun to tens of times larger. It consists of a central bulge, plus a disk of radius 12.5 kiloparsecs (kpc, equal to 10^3 pc), and a thickness of only about 0.3 kpc. We are located in the disk about 8 kpc from the center. The disk rotates slowly (and also differentially, with the outer edges moving more slowly, just as more distant planets in the Solar System orbit more slowly). At our radius, the galaxy rotates over a period of 200 million years. Because we are within it, we can't get an image of our own galaxy, but it probably looks not unlike the M100 galaxy shown in [Figure 2.1](#).

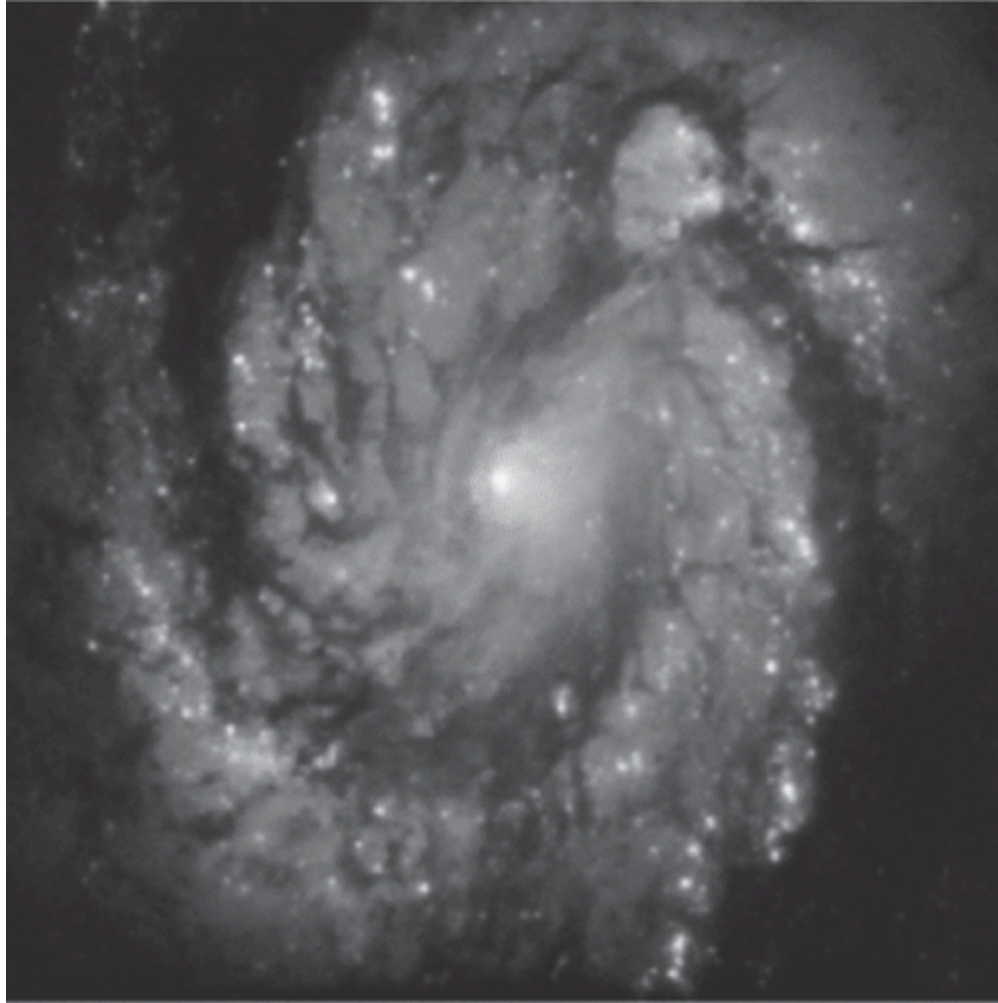


Figure 2.1 If viewed from above the disk, our own Milky Way galaxy would probably resemble the M100 galaxy, imaged here by the Hubble Space telescope. (Courtesy of NASA.)

Our galaxy is surrounded by smaller collections of stars, known as globular clusters. These are distributed more or less symmetrically about the bulge, at distances of 5–30 kpc. Typically, they each contain a million stars, and are thought to be remnants of the formation of the galaxy. As we shall discuss later, it is believed that the entire disk and globular cluster system are embedded in a larger spherical structure known as the Galactic halo.

Galaxies are the most visually striking and beautiful astronomical objects in the Universe, exhibiting a wide range of properties. However, in cosmology, the detailed structure of a galaxy is usually irrelevant, and galaxies are normally thought of as point-like objects emitting light, often broken into subclasses according to colors, luminosities, and morphologies.

The Local Group: Our galaxy resides within a small concentrated group of galaxies known as the local group. The nearest galaxy is a small irregular galaxy known as the Large Magellanic Cloud (LMC), which is 50 kpc away from the Sun. The nearest galaxy of similar size to our own is the Andromeda Galaxy, at a distance of 770 kpc. The Milky Way is one of the largest galaxies in the local group. A typical galaxy group occupies a volume of a few cubic **megaparsecs**. The megaparsec, denoted Mpc and equal to a million parsecs, is the cosmologist's favorite unit for measuring distances, because it is roughly the separation between neighboring galaxies. It equals 3.086×10^{22} m.

Clusters of Galaxies, Superclusters, and Voids: Surveying larger regions of the Universe, say on a scale of hundreds of Mpc, one sees a variety of large-scale structures, as shown in [Figure 2.2](#). This figure is not a photograph, but rather a carefully constructed map of the nearby region of our Universe, on a scale of about $1:10^{27}$! In some places, galaxies are clearly grouped into clusters of galaxies.



Circular map of the Universe details galaxy positions forming filamentary structure across angular slices, labeled with redshift and hours.

Figure 2.2 A map of galaxy positions in a narrow slice of the Universe, as measured by the Sloan Digital Sky Survey. Our galaxy is located at the center, and the survey radius is around 600 Mpc. The galaxy positions were obtained by measurement of the shift of spectral lines, as described in [Section 2.4](#). (Figure courtesy M. Blanton and the Sloan Digital Sky Survey, www.sdss3.org.)

A famous example of a cluster of galaxies is the Coma cluster, which is about 100 Mpc away from our own Galaxy. The upper panel of [Figure 2.3](#) shows a combined optical/infrared image of Coma; although the image resembles a star field, almost every source is a distinct galaxy (the main exception being two bright stars in the upper right quadrant). Coma contains perhaps 10 000 galaxies, mostly too faint to show in this image, orbiting in their common gravitational field.



Two panels detail infrared light and a labeled extended X-ray of the Coma cluster, with a scale of 10 arcmin.

Figure 2.3 Images of the Coma cluster of galaxies in visible/infrared light (top) and in X-rays (bottom), the latter being on a larger angular scale. (Courtesy of NASA/Spitzer satellite and ESA/U. Briel/MPE Garching/XMM-Newton satellite.)

However, most galaxies, sometimes called field galaxies, are not part of a cluster. Galaxy clusters are the largest gravitationally collapsed objects in the Universe, and they themselves are grouped into superclusters, joined by filaments and walls of galaxies. In between this “foamlike” structure lie large voids, some as large as 50 Mpc across. Structures in the Universe will be further described in [Advanced Topic 5](#). [Figure 2.4](#) shows an example of a computer simulation aiming to model the distribution of material within the Universe.



Simulated large scale structure of the universe details a dense web of filaments and nodes with a scale bar for 125 Mpc per hour.

Figure 2.4 A computer simulation showing the predicted distribution of matter in the Universe on large scales. (Courtesy of V. Springel and the Virgo Consortium.)

Large-scale Smoothness: Only once we get to scales of hundreds of megaparsecs or more does the Universe begin to appear smooth, as revealed by extremely large galaxy surveys such as the 2dF galaxy redshift survey and the Sloan Digital Sky Survey. Such surveys do not find any huge structures on scales greater than those described above; the galaxy superclusters and voids are believed to be the biggest structures in the present Universe. The observation that the Universe does indeed become smooth on the largest scales, the cosmological principle, is the underpinning of modern cosmology. It is interesting that, while the smoothness of the matter distribution on large scales has been a key assumption of cosmology for decades, it is only fairly recently that a convincing observational demonstration has become possible.

2.2 In other wavebands

Observations using visible light provide us with a good picture of what’s going on in the present-day Universe. However, many other wavebands make vital contributions to our understanding, and in particular our best knowledge of cosmology comes not from visible light but from microwaves.

Microwaves: For cosmology, this is by far the most important waveband. Penzias & Wilson’s accidental discovery in 1965 that the Earth is bathed in microwave radiation, with a blackbody spectrum at a temperature of around 3 K, was and is one of the most powerful pieces of information in support of the Big Bang theory, around which cosmology is based. This is now known as the cosmic microwave background or simply CMB. Observations by the Far InfraRed Absolute Spectrometer (FIRAS) experiment on board the COsmic Background Explorer (COBE) satellite have confirmed that the radiation is extremely close to the blackbody form at a temperature 2.725 ± 0.001 K. This data is shown in [Figure 2.5](#). Furthermore, the temperature coming from different parts of the sky is astonishingly uniform, and this presents the best evidence that we can use the cosmological principle as the foundation of cosmology. Even more importantly, it is possible to identify tiny variations, only one part in a hundred thousand, between the intensities of the microwaves coming from different directions. These were first discovered by the COBE satellite in 1992, and [Figure 2.6](#) shows precision measurements published by the *Planck* Satellite collaboration in 2018. It is believed that these patterns, known as cosmic microwave background anisotropies, are intimately related to the origin of structure in the Universe, and at present such measurements give us by far the most accurate understanding of cosmology. This topic is explored further in [Advanced Topics 5](#) and [6](#).



Graph of Intensity versus Waves per centimeter, plotting cosmic microwave background curve data points with error bars multiplied by 400.

Figure 2.5 The cosmic microwave background spectrum as measured by the FIRAS experiment on the COBE satellite. The error bars are so small that they have been multiplied by 400 to make them visible on this plot, and the best-fit blackbody spectrum at **2.725 K**, shown by the line, is an excellent fit.



Map of cosmic microwave temperature variations across an oval projection scaled from negative 300 to 300 microkelvin.

Figure 2.6 The cosmic microwave background temperature variations on the sky as measured by the *Planck* Satellite, with scale in microKelvin. (Courtesy of ESA/Planck collaboration.)

Infrared: Carrying out surveys in the infrared part of the spectrum, as was first done by the highly successful InfraRed Astronomical Satellite (IRAS) in the 1980s, is an excellent way of spotting young galaxies, in which star formation is at an early stage. Infrared surveys pick up a somewhat different population of galaxies to surveys carried out in optical light, though obviously the brightest galaxies are seen in both. Infrared is particularly good for looking through the dust in our own Galaxy to see distant objects, as it is absorbed and scattered much less strongly than visible radiation. Accordingly, it is best for studying the region close to our Galactic plane, where obscuration by dust is strongest. Infrared is also vital for studying the most distant galaxies, whose emission has been shifted predominantly into the infrared part of the spectrum by the expansion of the Universe after the light was emitted; NASA's James Webb Space Telescope is the current state-of-the-art for such studies. The far infrared can probe emission from interstellar dust grains to learn about dust-enshrouded environments that are opaque at visible and near-infrared wavelengths.

X-rays: These are a vital probe of clusters of galaxies; in between the galaxies lies gas so hot that it emits in the X-ray part of the spectrum, corresponding to temperatures of tens of millions of Kelvin. This gas is thought to be a remnant material from the formation of the galaxies, which failed to collapse to form stars. X-ray emission from the Coma galaxy cluster is shown in the lower panel of [Figure 2.3](#). The individual galaxies seen in the visible light image in the upper panel are almost all invisible in X-rays, with the bright diffuse X-ray emission from the hot gas dominating the image.

Radio Waves: A powerful way of gaining high-resolution maps of distant galaxies is by mapping in the radio part of the spectrum. An important probe of cosmology is observations of the “21 cm” emission line, due to spin-flip of the electron in a hydrogen atom, enabling maps of the distribution of neutral hydrogen gas in the distant Universe. This can probe both the early development of structures in the Universe, and the transition of the Universe to its present state of high ionization of the intergalactic medium. The Square Kilometre Array, nearing completion as we write, will take these studies to the next level.

2.3 Homogeneity and isotropy

The evidence that the Universe becomes smooth on a large scale supports the use of the cosmological principle. It is therefore believed that our large-scale Universe possesses two important properties: **homogeneity** and **isotropy**. Homogeneity is the statement that the Universe looks the same at each point, while isotropy states that the Universe looks the same in all directions.

These do not automatically imply one another. For example, a universe with a uniform magnetic field is homogeneous, as all points are the same, but it fails to be isotropic because directions along the field lines can be distinguished from those perpendicular to them. Alternatively, a spherically symmetric distribution, viewed from its central point, is isotropic but not necessarily homogeneous. However, if we require that a distribution is isotropic at *every* point, then that does enforce homogeneity as well.

As mentioned earlier, the cosmological principle is not exact, and so our Universe does not respect exact homogeneity and isotropy. Indeed, the study of departures from homogeneity is currently the most prominent research topic in cosmology. We'll introduce this in [Advanced Topic 5](#), but in the main body of this book we are concerned only with the behavior of the Universe as a whole, and so will be assuming large-scale homogeneity and isotropy.

2.4 The expansion of the Universe

A key piece of observational evidence in cosmology is that almost everything in the Universe appears to be moving away from us, and the further away something is, the more rapid its recession appears to be. These velocities are measured via the **redshift**, which is basically the Doppler effect applied to light waves. Galaxies have a set of absorption and emission lines identifiable in their spectra, whose characteristic frequencies are well known. However, if a galaxy is moving toward us, the light waves get crowded together, raising the frequency. Because blue light is at the high-frequency end of the visible spectrum, this is known as a blueshift. If the galaxy is receding, the characteristic lines move toward the red end of the spectrum and the effect is known as a redshift. This technique was first used to measure a galaxy's velocity by Vesto Slipher around 1912, and was applied systematically by one of the most famous cosmologists, Edwin Hubble, in the following decades.

It turns out that almost all galaxies are receding from us, so the standard terminology is redshift z , defined by

$$z = \frac{\lambda_{\text{obs}} - \lambda_{\text{em}}}{\lambda_{\text{em}}} \quad (2.1)$$

where λ_{em} and λ_{obs} are the wavelengths of light at the points of emission (the galaxy) and observation (us). If a nearby object is receding at a speed v , then its redshift is

$$z = \frac{v}{c} \quad (2.2)$$

where c is the speed of light. ² [Figure 2.7](#) shows velocity against distance, a plot known as the Hubble diagram, for a sample of 1355 galaxies.

 Scatter plot of galaxy velocity versus estimated distance, plotted in Mpc, details a scattered and linear trend implying Hubble's law.

Figure 2.7 A plot of velocity versus estimated distance for a set of 1355 galaxies. A straight-line relation implies Hubble's law. The considerable scatter is due to observational uncertainties and random galaxy motions, but the best-fit line accurately gives Hubble's law. (The ***x***-axis scale assumes a particular value of ***H***.)

Hubble realized that his observations, which were of course much less extensive than those available to us now, showed that the velocity of recession was proportional to the distance of an object from us:

$$v = H_0 r \quad (2.3)$$

This is known as **Hubble's law**, and the constant of proportionality H_0 is known as **Hubble's constant**. Hubble's law isn't exact, as the cosmological principle doesn't hold perfectly for nearby galaxies, which typically possess some random motions known as peculiar velocities. But it does describe the average behavior of galaxies extremely well. Hubble's law gives the picture of our Universe illustrated in [Figure 2.8](#), where the nearby galaxies have the smallest velocity relative to ours. Over the years, many attempts have been made to determine accurate values for the proportionality constant, which is now known to within a few percent accuracy. Still, as we will see in [Chapter 6](#) and [Advanced Topic 6.5](#), there remains some controversy to this day.

 Diagram of a central point and six surrounding points, each with an arrow details speed proportional to distance from the center.

Figure 2.8 According to Hubble's law, the further away from us a galaxy is, the faster it is receding.

At first sight, it seems that the cosmological principle must be violated if we observe everything to be moving away from us, since that apparently places us at the center of the Universe. However, nothing could be further from the truth. In fact, *every* observer sees all objects rushing away from them with velocity proportional to distance. It is perhaps easiest to convince yourself of this by setting up a square grid with recession velocity proportional to distance from the central grid point. Then, transform the frame of reference to a nearby grid-point, and you'll find that the Hubble law still holds about the new "centre." This only works because of the linear relationship between velocity and distance; any other law and it wouldn't work.

So, although expanding, the Universe looks just the same whichever galaxy we choose to imagine ourselves within. A common analogy is to imagine baking a cake with raisins in it, or blowing up a balloon with dots on its surface. As the cake rises (or the balloon is inflated), the raisins (or dots) move apart. From each one, it seems that all the others are receding, and the further away they are, the faster that recession is.

Because everything is flying away from everything else, we conclude that in the distant past everything in the Universe was much closer together. Indeed, trace the history back far enough and everything comes together. The initial explosion is known as the **Big Bang**, and a model of the evolution of the Universe from such a beginning is known as the **Big Bang Cosmology**. Later on, we will find out why it is commonly called the **Hot Big Bang**.

2.5 Particles in the Universe

2.5.1 What particles are there?

Everything in the Universe is made up of fundamental particles, and the behavior of the Universe as a whole depends on the properties of these particles.

One crucial question is whether a particle is moving relativistically or not. Any particle has two contributions to its energy, one being kinetic energy and the other being mass–energy, which combine to give

$$E_{\text{total}} = mc^2 + p^2 c^2 \quad (2.4)$$

where m is the particle rest mass and p the particle momentum. If the mass–energy dominates, the particle will be moving at much less than the speed of light, and we say it is nonrelativistic. In that limit, we can carry out an expansion

$$E_{\text{total}} = mc^2 \left(1 + \frac{p^2}{2m^2 c^2} \right) \approx mc^2 + \frac{1}{2} p^2 \quad (2.5)$$

We recognize the first term as Einstein’s famous $E = mc^2$, known as the rest mass–energy as it is the energy of the particle when it is stationary. The second term is the usual kinetic energy ($\frac{1}{2} p^2$ equals $\frac{1}{2} m v^2$ in the nonrelativistic limit). If the mass–energy does not dominate, the particle will be moving at a substantial fraction of the speed of light and so is relativistic. In particular, any particle with zero rest mass is always relativistic and moves at the speed of light, the simplest example being light itself.

Let’s review the nature of the particles which are believed to exist in our Universe.

Baryons

We ourselves are built from atoms, the bulk of whose mass is attributable to the protons and neutrons in the atomic nuclei. Protons and neutrons are believed to be made up of more fundamental particles known as quarks, a proton being made of two up quarks and a down quark, while a neutron is an up and two downs. A general term for particles made up of three quarks is **baryons**. Of all the possible baryons, only the proton and neutron can be stable,³ and so these are thought to be the only types of baryonic particle significantly represented in the Universe. Yet another piece of terminology, nucleon, refers to just protons and neutrons, but I’ll follow the standard practice of using the term baryon. In particle physics units, the mass–energies of a proton and a neutron are 938.3 and 939.6 MeV, respectively, where “MeV” is a Mega-electron volt, a unit of energy equal to a million electron volts (eV) and rather more convenient than a Joule.

Although electrons are certainly not made from quarks, they are traditionally also included under the title baryon by cosmologists (to the annoyance of particle physicists). A crucial property of the Universe is that it is charge neutral, so there must be one electron for every proton. Weighing in at a puny 0.511 MeV, well under a thousandth of a proton mass, the contribution of electrons to the total mass is a tiny fraction, not meriting separate discussion.

In the present Universe, baryons are typically moving nonrelativistically, meaning that their kinetic energy is much less than their mass–energy.

Radiation

Our visual perception of the Universe comes from electromagnetic radiation, and such radiation, at a large variety of frequencies, pervades the Universe. In the quantum mechanical view of light, it can be thought of as made up of individual particles – like packets of energy – known as **photons** and usually indicated by the symbol γ . Photons propagate, naturally enough, at the speed of light; since they have zero rest mass, their total energy is always given by their kinetic energy, and is related to their frequency f by

$$E = hf \quad (2.6)$$

where h is Planck’s constant.

Photons can interact with the baryons and electrons; for example, a high-energy photon can knock an electron out of an atom (a process known as ionization), or can scatter off a free electron (known as Thomson scattering in the non-relativistic case $hf \ll mc^2$, otherwise Compton scattering). The more energetic the photons are, the more devastating their effects on other particles.

Neutrinos

Neutrinos are extremely weakly interacting particles, produced, for example, in radioactive decay. There is now significant experimental evidence that they possess a non-zero rest mass, but so far there has been no clear evidence that this mass is large enough to have observable cosmological effects. Hence, for now, it remains a working assumption in cosmology to treat them as effectively massless. We will adopt that assumption for the main body of this book, and in that case they, like photons, are always relativistic. The combination of photons and neutrinos makes up the relativistic material in our Universe. Confusingly, sometimes the term “radiation” is used to refer to all the relativistic material.

There are three types of neutrino, the electron neutrino, muon neutrino, and tau neutrino, and if they are indeed all massless, they should all exist in our Universe. Unfortunately, their interactions are so weak that for now there is no hope of detecting cosmological neutrinos directly. Originally, their presence was inferred on purely theoretical grounds, though we will see that the existence of the cosmic neutrino background may be inferred indirectly by some cosmological observations.

Because they are so weakly interacting, the experimental limits on the neutrino masses, especially of the latter two types, are quite weak, and it is in fact perfectly possible that they are massive enough to be nonrelativistic. The possible effects of neutrino masses are explored in [Advanced Topic 3](#). Cosmological effects of neutrino masses might well have been spotted by the time you read this book.

Dark matter

In this book, we'll encounter one further kind of particle that may exist in our Universe, which is not part of the Standard Model of particle theory. It is known as dark matter, and its properties are highly uncertain and a matter of constant debate among cosmologists. We'll return to it in [Chapter 9](#).

2.5.2 Thermal distributions and the blackbody spectrum

We end this section with some discussion of the physics of radiation. If this is unfamiliar to you, the details aren't all that crucial, though some of the final results will be used later in the book.

If particles are frequently interacting with one another, then the distribution of their energies can be described by equilibrium thermodynamics. In a thermal distribution, interactions are frequent, but a balance has been reached so that all interactions proceed equally frequently in both the forward and backward directions, so that the overall distribution of particle numbers and energies remains fixed. The number of particles of a given energy then depends only on the temperature.

The precise distribution depends on whether the particles considered are fermions, which obey the Pauli exclusion principle, or bosons, which do not. In this book, the most interesting case is that of photons, which are bosons, and their characteristic distribution at temperature T is the **Planck** or **blackbody** spectrum. Photons have two possible polarizations, and each has an occupation number per mode N given by the Planck function

$$N = \frac{1}{\exp\left(\frac{hf}{k_B T}\right) - 1} \quad (2.7)$$

where h is Planck's constant and k_B is one of the fundamental constants of nature, the **Boltzmann constant**, whose value is $1.381 \times 10^{-23} \text{ J K}^{-1} = 8.619 \times 10^{-5} \text{ eV K}^{-1}$.

To interpret this equation, remember that hf is the photon energy. The purpose of the Boltzmann constant is to convert temperature into a characteristic energy $k_B T$. Below this characteristic energy, $hf \ll k_B T$, it is easy to make photons and the occupation number is large (as photons are bosons, the Pauli exclusion principle doesn't apply and there may be arbitrarily many photons in a given mode). Above the characteristic energy, $hf \gg k_B T$, it is energetically unfavorable to make photons, and the number is exponentially suppressed, as shown in [Figure 2.9](#).

 Graph of Planck function N versus hf/k_B details a sharp decay, reflecting far more low-energy photons than high-energy.

Figure 2.9 The Planck function of [Equation \(2.7\)](#). There are far more photons with very low energy than very high energy.

More interesting than the number of photons in a mode is the distribution of energy among the modes. We focus on the energy per unit volume, known as the **energy density** ϵ . Because there are very few photons with

$hf \gg k_B T$ there isn't much energy at high frequencies. But, despite their large numbers, there also isn't much total energy at low frequencies

$hf \ll k_B T$, both because those photons have less energy each $\epsilon = hf N$, and because their wavelength is longer and so each photon occupies a greater volume. With quite a bit of work, the energy density in a frequency interval df about f can be shown to be

$$\epsilon(f) df = \frac{8\pi h}{c^3} \frac{f^3 df}{\exp\left(\frac{hf}{k_B T}\right) - 1} \quad (2.8)$$

which tells us how energy is distributed among the different frequencies. We see in [Figure 2.10](#) that the peak of the distribution is at $f_{\text{peak}} \approx 2.8 k_B T / h$,

corresponding to an energy

$\epsilon = hf$ asymptotically equals $2.8 k_B T$. That is to say, the total energy in the radiation is dominated by photons with energies of order

$k_B T$. Indeed, the mean energy of a photon in this distribution is

$\epsilon_{\text{mean}} \approx 3 k_B T$.



Graph of the normalized density versus $hf/k_B T$ details the energy density distribution of a spectrum.

Figure 2.10 The energy density distribution of a blackbody spectrum, given by Equation (2.8). Most of the energy is contributed by photons of energy

$h f$ **$k_B T$** .

When we study the early history of the Universe, an important question will be how this typical energy compares to atomic and nuclear binding energies.

A further quantity of interest will be the total energy density of the blackbody radiation, obtained by integrating Equation (2.8) over all frequencies. Setting $y = hf/k_B T$ quickly leads to

$$\epsilon_{\text{rad}} = \frac{8\pi k_B^4}{15} T^4 \int_0^\infty \frac{y^3 dy}{e^y - 1} \quad (2.9)$$

The integral is not particularly easy to compute, but you might like to try it as a challenge. The answer is

$\pi^4/15$, giving a radiation energy density

$$\epsilon_{\text{rad}} = \alpha T^4 \quad (2.10)$$

where the radiation constant α is defined as

$$\alpha = \frac{\pi^5 k_B^4}{15 h^3} = 7.565 \times 10^{-16} \text{ J m}^{-3} \text{ K}^{-4} \quad (2.11)$$

Here, h is the reduced Planck constant.

Problems

2.1. Suppose that the Milky Way galaxy is a typical size, containing say 10^{11} stars, and that galaxies are typically separated by a distance of one megaparsec. Estimate the density of the Universe in SI units. How does this compare with the density of the Earth?

$1 \text{ Mpc} \approx 3.1 \times 10^{22} \text{ m}$

2.2. In the real Universe, the expansion is not completely uniform. Rather, galaxies exhibit some random motion relative to the overall Hubble expansion, known as their *peculiar velocity* and caused by the gravitational pull of their near neighbors. Supposing that a typical (e.g. root mean square) galaxy's peculiar velocity is 600 km s^{-1} , how far away would a galaxy have to be before it could be used to determine the Hubble constant to 10% accuracy, supposing

a. the true value of the Hubble constant is $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$?

b. the true value of the Hubble constant is $50 \text{ km s}^{-1} \text{ Mpc}^{-1}$?

Assume in your calculation that the galaxy distance and redshift could be measured exactly. Unfortunately, that is not true of real observations.

2.3. What evidence can you think of to support the assertion that the Universe is charge neutral, and hence contains an equal number of protons and electrons?

2.4. The binding energy of the electron in a hydrogen atom is 13.6 eV. What is the frequency of a photon with this energy? At what temperature does the mean photon energy equal this energy?

2.5. The peak of the energy density distribution of a blackbody at

$\epsilon_{\text{peak}} \approx 2.8 k_B T$ implies that $hf/k_B T$ is a constant. Evaluate this constant in SI units (see page xiv for

useful numbers). The Sun radiates approximately like a blackbody with

$T_{\text{Sun}} \approx 5800 \text{ K}$. Compute λ_{peak} for solar radiation. Where in the electromagnetic spectrum does the peak emission lie?

2.6. The cosmic microwave background has a blackbody spectrum at a temperature of 2.725 K. Repeat [Problem 2.5](#) to find the peak frequency of its emission, and also find the corresponding wavelength and compare to [Figure 2.5](#). Confirm that the peak emission lies in the microwave part of the electromagnetic spectrum. Finally, compute the total energy density of the microwave background.

Notes

¹ A parsec is defined as the distance at which the mean distance between the Earth and Sun subtends a second of arc. The mean Earth–Sun distance (called an Astronomical Unit) is $1.496 \times 10^{11} \text{ m}$, and dividing that by $\tan(1 \text{ arcsec})$ gives $1 \text{ pc} = 3.086 \times 10^{16} \text{ m}$.

² This formula ignores special relativity and so is valid only for speeds v much less than c . If you're interested, the special relativity result, of which this is an expansion for small v divided by c , is

$$1 + z = \sqrt{\frac{1 + v/c}{1 - v/c}}$$

However, for distant objects in cosmology, there are further considerations concerning the propagation time of the light and how the relative velocity might change during it, and so this expression should not be used.

³ The proton's lifetime is known to be either infinite, corresponding to the proton being absolutely stable, or much longer than the age of the Universe so that they are effectively stable. Isolated neutrons are unstable (decaying into a proton, an electron, and an anti-neutrino), but those bound in nuclei may be stable; this will prove crucial in [Chapter 12](#).

Chapter 3

Newtonian Gravity

It is perfectly possible to discuss cosmology without having already learned general relativity. In fact, the most crucial equation, the Friedmann equation, which describes the expansion of the Universe, turns out to be the same when derived from Newton's theory of gravity as it is when derived from the equations of general relativity. The Newtonian derivation is, however, some way from being completely rigorous, and general relativity is required to fully patch it up, a detail that need not concern us at this stage. Readers with some familiarity with general relativity may wish to study the relativistic derivation given in [Advanced Topic 1](#).

In Newtonian gravity, all matter attracts, with the force exerted by an object of mass M on one of mass m given by the famous relationship

$$F = \frac{GMm}{r^2}, \quad (3.1)$$

where r is the distance between the objects and G is Newton's gravitational constant. That is, gravity obeys an inverse square law. Because a force on an object induces an acceleration which is also proportional to its mass, via $F = ma$, the acceleration an object feels under gravity is *independent* of its mass.

The force exerted means there is a gravitational potential energy

$$V = -\frac{GMm}{r}, \quad (3.2)$$

with the force being in the direction which decreases the potential energy the fastest. Like the electric potential of two opposite charges, the gravitational potential is negative, favoring the two objects being close together. But with gravity there is no analog of the repulsion of like charges. Gravity always attracts.

The derivation of the Friedmann equation requires a famous result due originally to Newton, which we won't attempt to prove here. This result states that in a spherically symmetric distribution of matter, a particle feels no force at all from the material at greater radii, and the material at smaller radii gives exactly the force which one would get if all the material was concentrated at the central point. This property arises from the inverse square law; the same results exist for electromagnetism. One example of its use is that the gravitational (or electromagnetic) force outside a spherical object of unknown density profile depends only on the total mass (charge). Another is that an "astronaut" inside a spherical shell feels no gravitational force, not only if they are at the center but if they are at any position inside the shell.

3.1 The Friedmann equation

The Friedmann equation describes the expansion of the Universe, and is therefore the most important equation in cosmology. One of the routine tasks of a working cosmologist is solving this equation under different assumptions concerning the material content of the Universe. To derive the Friedmann equation, we need to compute the gravitational potential energy and the kinetic energy of a test particle (it doesn't matter which one, since everywhere in the Universe is the same according to the cosmological principle), and then use energy conservation.

Let's consider an observer in a uniform expanding medium, with mass density ρ , the mass density being the mass per unit volume. We begin by realizing that because the Universe looks the same from anywhere, we can consider any point to be its center. Now, consider a particle a distance r away with mass m , as shown in [Figure 3.1](#). (By "particle," we really mean a small volume containing the mass m .) Due to Newton's theorem, this particle only feels a force from the material at smaller radii, shown as the shaded region. This material has total mass given by $M = 4\pi\rho r^3/3$, contributing a force

$$F = \frac{GMm}{r^2} = \frac{4\pi G\rho r m}{3}, \quad (3.3)$$

and our particle has a gravitational potential energy

$$V = -\frac{GMm}{r} = -\frac{4\pi G\rho r^2 m}{3}. \quad (3.4)$$

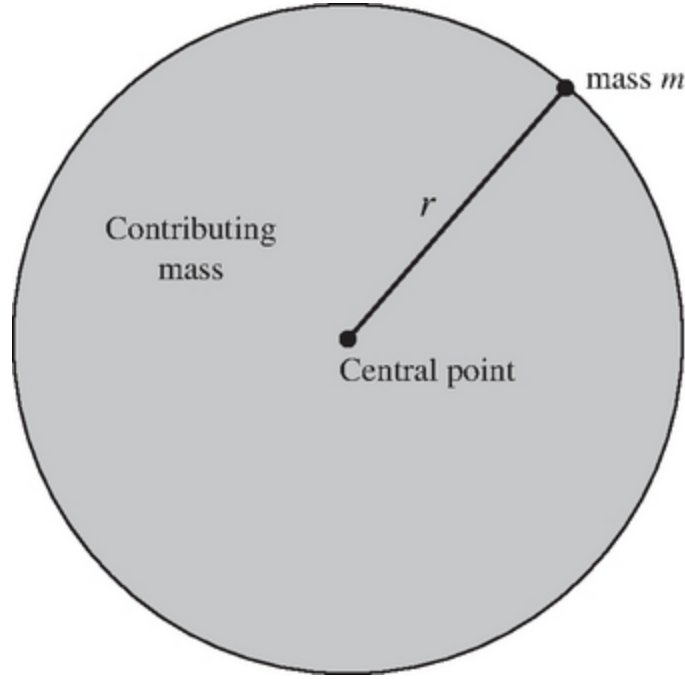


Figure 3.1 The particle at radius r only feels gravitational attraction from the shaded region. Any gravitational attraction from the material outside cancels out, according to Newton's theorem.

The kinetic energy is easy; the velocity of the particle is $\dot{r}$ (we'll always use dots to mean time derivatives) giving

$$T = \frac{1}{2} m \dot{r}^2. \quad (3.5)$$

The equation describing how the separation r changes can now be derived from energy conservation for that particle, namely

$$U = T + V, \quad (3.6)$$

where U is a constant. Note that U need not be the same constant for particles separated by different distances. Substituting gives

$$U = \frac{1}{2} m \dot{r}^2 - \frac{4\pi}{3} G\rho r^2 m. \quad (3.7)$$

This equation gives the evolution of the separation r between the two particles.

We now take a crucial step in this derivation, which is to realize that this argument applies to any two particles, because the Universe is homogeneous. This allows us to change to a different coordinate system, known as **comoving coordinates**. These are coordinates which are carried along with the expansion. Because the expansion is uniform, the relationship between real distance $\vec{r}$ and the comoving distance, which we can call $\vec{x}$, can be written

$$\vec{r} = a(t) \vec{x}, \quad (3.8)$$

where the homogeneity property has been used to ensure that a is a function of time alone. Note that these distances have been written as vector distances. What you should think of when studying this equation is a coordinate grid which expands with time, as shown in [Figure 3.2](#). The galaxies remain at fixed locations in the $\vec{x}$ coordinate system. The original $\vec{r}$ coordinate system, which does not expand, is usually known as **physical coordinates**.

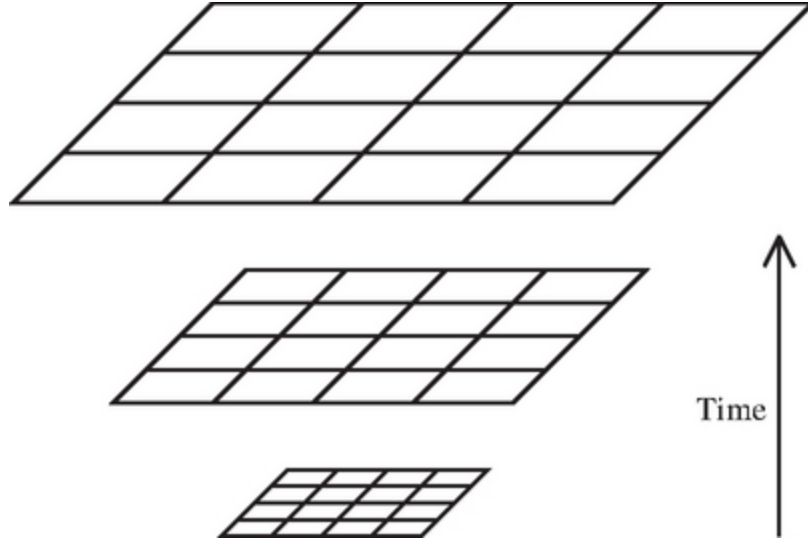


Figure 3.2 The comoving coordinate system is carried along with the expansion, so that any objects remain at fixed coordinate values.

The quantity $a(t)$ is a crucial one, and is known as the **scale factor of the Universe**. It measures the universal expansion rate. It is a function of time alone, and it tells us how physical separations are growing with time, since the coordinate distances $\vec{x}$ are by definition fixed. For example, if, between times t_1 and t_2 , the scale factor doubles in value, that tells us that the Universe has expanded in size by a factor two, and it will take us twice as long to get from one galaxy to another.

We can use the scale factor to rewrite [Equation \(3.7\)](#) for the expansion. Substituting [Equation \(3.8\)](#) into it, remembering $\dot{x} = 0$ by definition as objects are fixed in comoving coordinates, gives

$$U = \frac{1}{2}m\dot{a}^2x^2 - \frac{4\pi}{3}G\rho a^2x^2m. \quad (3.9)$$

Multiplying each side by $2/ma^2x^2$ and rearranging the terms then gives

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2}, \quad (3.10)$$

where $kc^2 = -2U/mx^2$. This is the standard form of the **Friedmann equation**, and it will appear frequently throughout this book. In this expression, k must be independent of x since all the other terms in the equation are, otherwise homogeneity will not be maintained. So, in fact, we learn that homogeneity requires that the quantity U , while constant for a given particle, does indeed change if we look at different separations x , with $U \propto x^2$.

Finally, since $k = -2U/mc^2x^2$ which is time independent (as the total energy U is conserved, and the comoving separation x is fixed), we learn that k is just a constant, unchanging with either space or time. It has the units of $[\text{length}]^{-2}$. An expanding Universe has a unique value of k , which it retains throughout its evolution. In [Chapter 4](#), we will see that k tells us about the geometry of the Universe, and it is often called the curvature.

3.2 On the meaning of the expansion

So what does the expansion of the Universe mean? Well, let's start with what it does *not* mean. It does not mean that your body is gradually going to get bigger with time (and certainly isn't an excuse if it does). That does not mean that the Earth's orbit is going to get further from the Sun. It doesn't even mean that the stars within our galaxy are going to become more widely separated with time.

But it does mean that distant galaxies are getting further apart.

The distinction lies in whether the motion of objects is governed by the cumulative gravitational effect of a homogeneous distribution of matter between them, as shown in [Figure 3.1](#). The atoms in your body are not: their separation is dictated by the strength of chemical bonds, with gravity playing no significant role. So, molecular structures will not be affected by the expansion. Likewise, the Earth's motion in its orbit is completely dominated by the gravitational attraction of the Sun (with a minor contribution from the other planets). And even the stars in our galaxy are orbiting in the common gravitational potential well that they themselves create, and are not moving apart relative to one another. The common feature of these environments

is that they are the ones of considerable excess density, very different from the smooth distribution of matter we used to derive the Friedmann equation.

But if we go to large enough scales, in practice tens of megaparsecs, the Universe does become effectively homogeneous and isotropic, with the galaxies flying apart from one another in accordance with the Friedmann equation. It is on these large scales that the expansion of the Universe is felt, and on which the cosmological principle applies.

3.3 Things that go faster than light

A common question that concerns people is whether faraway galaxies are receding from us faster than the speed of light. That is to say, if velocity is proportional to distance, then if we consider galaxies far enough away, can we not make the velocity as large as we like, in violation of special relativity?

The answer is that, indeed, in our theoretical predictions, distant objects can *appear* to move away faster than the speed of light. However, it is space itself which is expanding. There is no violation of causality, because no signal can be sent between such galaxies. Further, special relativity is not violated, because it refers to the relative speeds of objects passing each other, and cannot be used to compare the relative speeds of distant objects.

One way to think of this is to imagine a colony of ants on a balloon. Suppose that the fastest the ants can move is a centimeter per second. If any two ants happen to pass each other, their fastest relative speed would be two centimeters per second, if they happened to be moving in opposite directions. Now, start to blow the balloon up. Although the ants wandering around the surface still cannot exceed one centimeter per second, the balloon is now expanding under them, and ants that are far apart on the balloon could easily be moving apart at faster than two centimeters per second if the balloon is blown up fast enough. But if they are, they will never get to tell each other about it, because the balloon is pulling them apart faster than they can move together, even at full speed. Any ants that start close enough to be able to pass one another must do so at no more than two centimeters per second, even if the Universe is expanding.

The expansion of space is just like that of the balloon, and pulls the galaxies along with it.

3.4 The fluid equation

Fundamental though it is, the Friedmann equation is of no use without an equation to describe how the density ρ of material in the Universe is evolving with time. This involves the pressure p of the material, and is called the fluid equation. (Unfortunately, the standard symbol p for pressure is the same as for momentum, which we've already used. Almost always in this book, p will be pressure.) As we'll shortly see, the different types of material that might exist in our Universe have different pressures, and lead to different evolution of the density ρ .

We can derive the fluid equation by considering the first law of thermodynamics

$$dE + p dV = T dS, \quad (3.11)$$

applied to an expanding volume V of unit comoving radius.¹ This is exactly the same as applying thermodynamics to a gas in a cylinder. The volume has physical radius a , so the energy is given, using $E = mc^2$, by

$$E = \frac{4\pi}{3} a^3 \rho c^2. \quad (3.12)$$

The change of energy in a time dt , using the product rule, is

$$\frac{dE}{dt} = 4\pi a^2 \rho c^2 \frac{da}{dt} + \frac{4\pi}{3} a^3 \frac{d\rho}{dt} c^2, \quad (3.13)$$

while the rate of change in volume is

$$\frac{dV}{dt} = 4\pi a^2 \frac{da}{dt}. \quad (3.14)$$

Assuming a reversible adiabatic expansion $dS = 0$, putting these into Equation (3.11) and rearranging gives

$$\dot{\rho} + 3 \frac{\dot{a}}{a} \left(\rho + \frac{p}{c^2} \right) = 0, \quad (3.15)$$

where, as always, dots are shorthand for time derivatives. This is the **fluid equation**. As we see, there are two terms contributing to the change in the density. The first term in the brackets corresponds to the dilution in the density because the volume has increased, while the second corresponds to the loss of energy because the pressure of the material has done work

as the Universe's volume increased. This energy has not disappeared entirely, of course; energy is always conserved. The energy lost from the fluid via the work done has gone into gravitational potential energy.

Let me stress that there are no pressure forces in a homogeneous universe, because the density and pressure are everywhere the same. A pressure *gradient* is required to supply a force. So, pressure does not contribute a force helping the expansion along; its effect is solely through the work done as the Universe expands.

We are still not in a position to solve the equations, because now we only know what ρ is doing if we know what the pressure p is. It is in specifying the pressure that we are saying what kind of material our model universe is filled with. The usual assumption in cosmology is that there is a unique pressure associated with each density, so that $p \equiv p(\rho)$. Such a relationship is known as the **equation of state**, and we'll see two different examples in [Chapter 5](#). The simplest possibility is that there is no pressure at all, and that particular case is known as (nonrelativistic) matter.

Once the equation of state is specified, the Friedmann and fluid equations are all we need to describe the evolution of the Universe. However, before discussing this evolution, we are going to spend some time exploring some general properties of the equations, as well as devoting [Chapter 4](#) to consideration of the meaning of the constant k . If you prefer to immediately see how to solve these equations, feel free to jump straight away to [Sections 5.3 – 5.5](#), and come back to the intervening material later. On the way, you might want to glance at [Section 3.6](#) to find out why a factor of c^2 mysteriously vanishes from the Friedmann equation between here and there.

3.5 The acceleration equation

The Friedmann and fluid equations can be used to derive a third equation, not independent of the first two, of course, which describes the acceleration of the scale factor. By differentiating [Equation \(3.10\)](#) with respect to time, we obtain

$$2 \frac{\dot{a}}{a} \frac{a\ddot{a} - \dot{a}^2}{a^2} = \frac{8\pi G}{3} \dot{\rho} + 2 \frac{kc^2 \dot{a}}{a^3}. \quad (3.16)$$

Substituting in for $\dot{\rho}$ from [Equation \(3.15\)](#) and canceling the factor $2\dot{a}/a$ in each term gives

$$\frac{\ddot{a}}{a} - \left(\frac{\dot{a}}{a}\right)^2 = -4\pi G \left(\rho + \frac{p}{c^2}\right) + \frac{kc^2}{a^2}, \quad (3.17)$$

and finally, using [Equation \(3.10\)](#) again, we arrive at an important equation known as the **acceleration equation**

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3} \left(\rho + \frac{3p}{c^2}\right). \quad (3.18)$$

Notice that if the material has any pressure, this *increases* the gravitational force, and so further decelerates the expansion. We remind you that there are no forces associated with pressure in an isotropic universe, as there are no pressure gradients.

The acceleration equation does not feature the constant k that appears in the Friedmann equation; it cancels out during the derivation.

3.6 On mass, energy, and vanishing factors of c^2

You should be aware that cosmologists have a habit of using mass density ρ and energy density ϵ interchangeably. They are related via Einstein's most famous equation as $\epsilon = \rho c^2$, and if one chooses so-called "natural units" in which c is set equal to one, the two become the same. For clarity, however, we will be careful to maintain the distinction. Note that the phrase "mass density" is used in Einstein's sense – it includes the contributions to the mass from the energy of the various particles, as well as any rest mass they might have.

The habit of setting $c = 1$ means that the Friedmann equation is normally written without the c^2 in the final term, so that it reads

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3} \rho - \frac{k}{a^2}. \quad (3.19)$$

The constant k then appears to have units $[\text{time}]^{-2}$ – setting $c = 1$ makes time and length units interchangeable. Since the practice of omitting the c^2 in the Friedmann equation is widely adopted in other cosmology textbooks, we will drop it for the remainder of this book too. In practice, it is a rare situation indeed where one has to be careful about this.

Problems

- 3.1. Starting with the Friedmann and acceleration equations, derive the fluid equation.
- 3.2. Derive an equation for the third derivative of the scale factor in terms of the density (not its derivatives), in the case of a pressureless fluid.

Note

- ¹ Don't confuse V for volume with $\mathcal{V}$ for gravitational potential energy.

Chapter 4

The Geometry of the Universe

We now consider the real meaning of the constant k , which appears in the Friedmann equation

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}. \quad (4.1)$$

While the Newtonian derivation in [Chapter 3](#) introduced this as a measure of the energy per particle, the true interpretation, apparent in the context of general relativity, is that it measures the curvature of space. General relativity tells us that gravity is due to the curvature of four-dimensional (4D) space-time, and a full analysis can be found in any general relativity textbook. Here, we will be purely descriptive, and focus on the interpretation of k as measuring the curvature of the three spatial dimensions. Further details of general relativistic cosmology can be found in [Advanced Topic 1](#).

We have demanded that our model universes be both homogeneous and isotropic. The simplest type of geometry which can have this property is what is called a **flat geometry**, in which the normal rules of Euclidean geometry apply. However, it turns out that the assumption of isotropy is not enough to demand that as the only choice. Instead, there are three possible geometries for the Universe, and they correspond to k being zero, positive, or negative.

4.1 Flat geometry

Euclidean geometry is based on a set of simple axioms (e.g. a straight line is the shortest distance between two points), plus one more complex axiom which says that parallel straight lines remain a fixed distance apart. These are the basis for the standard laws of geometry, and lead to the following conclusions:

- The angles of a triangle add up to 180° .
- The circumference of a circle of radius r is $2\pi r$.

Such a geometry might well apply to our own Universe. If that is the case, then the Universe must be infinite in extent, because if it came to a definite edge, then that would clearly violate the principle that the Universe should look the same from all points.

A universe with this geometry is usually called a **flat universe**.

4.2 Spherical geometry

Euclid always hoped that the more artificial final axiom could be proven by others. It wasn't until the nineteenth century that Riemann demonstrated that Euclid's final axiom was an arbitrary choice, and that one could make other assumptions. In doing so, he founded the subject of non-Euclidean geometry, which forms the mathematical foundation for Einstein's theory of general relativity.

The simplest kind of non-Euclidean geometry is actually very familiar to us; it is the spherical geometry which we use, for instance, to navigate around the Earth. Before worrying about the Universe having three dimensions, let's examine the properties of the two-dimensional (2D) surface of the Earth, shown in [Figure 4.1](#).

 Diagram of a spherical surface depicting a triangle with three right angles detailed at the vertices, indicating positive curvature k .

Figure 4.1 A sketch of a spherical surface, representing positive ***k*** . A triangle is shown, which has three right angles!

First, we know that a perfect sphere looks the same from all points on its surface, so the condition of isotropy is satisfied (e.g. if someone hands you a snooker ball and asks which way up it is, you're not going to be able to tell them). But, unlike the case of a flat geometry, the spherical surface is perfectly finite in extent, its area being given by $4\pi r^2$. Yet, there is no boundary, no "edge" to the surface of the Earth. So, it is perfectly possible to have a finite surface which nevertheless has no boundary.

If we draw parallel lines on the surface of the Earth, then they violate Euclid's final axiom. The definition of a straight line is the shortest distance between two points, which means that the straight lines in a spherical geometry are segments of great circles, such as the equator or the lines of longitude.² The lines of longitude are an excellent example of the failure of Euclid's axiom; as

they cross the equator they are all parallel to one another, but rather than remaining a constant distance apart they meet at both poles.

If we draw a triangle on a sphere, we find that the angles do not add up to 180° either. The easiest example to think about is to start at the North Pole. Draw two straight lines down to the equator, 90° apart, and then join them with a line on the equator. You have drawn a triangle in which all three angles are 90° , as shown in [Figure 4.1](#).

The circumference of a circle also fails to obey the normal law. Suppose we draw a circle at a radius r from the North Pole, and we'll choose r so that our circle is the equator. That radius, measured on the surface of the sphere, corresponds to a quarter of a complete circle around the Earth, so r equals πR divided by 2 where R is the radius of the Earth. However, the circumference c is given by $2\pi R$, so instead of the usual relation one has c equals $4r$ for a circle drawn at the equator. The circumference is less than $2\pi r$. [Problem 4.1](#) looks at the general case; you may find it helpful to glance at the figure on page 38 now.

Although we have only considered specific cases where the algebra is easy, it is true that whatever triangle or circle is drawn, you'll always find

- the angles of a triangle add up to more than 180° ;
- the circumference of a circle is less than $2\pi r$.

If you make triangles or circles much smaller than the size of the Earth, then the Euclidean laws start to become a good approximation; certainly we don't have to worry about Euclidean laws being broken in our everyday existence (though the appreciation that the Earth is spherical is vital for the planning of long distance journeys). So, a small triangle drawn on a sphere will have the sum of its angles only marginally larger than 180° . This property makes it rather hard to measure the geometry of our own Universe, because the neighboring region that we can measure accurately is only a small fraction of the size of the Universe and so will obey nearly Euclidean laws whatever the overall geometry.

One of the most important conceptual points that you need to grasp is that our three-dimensional (3D) Universe can have properties just like the 2D surface of the sphere. Unfortunately, our brains are not conditioned to think of three dimensions as being curved, so we must work by analogy with the 2D situation we have just described. When we think of the surface of a sphere being curved, we naturally imagine the sphere as an object in our own 3D Universe and think of it as curved in that sense. The important point is that the curvature is a property of the 2D surface of the sphere itself; the triangles and circles whose properties we've just described are drawn on the surface. A classic application of this was the ancient Greeks' use of these laws to deduce that the Earth is spherical, and even to obtain a good estimate of its diameter. When we discuss spherical geometry, there is actually no need to think of it as existing within our 3D space at all, and we must remember when discussing the geometry that we are assumed to be restricted to the surface of the sphere, and not allowed to move off the surface either toward or away from the notional center.

All this is an analogy to what might happen with our 3D Universe. The analog of the 2D sphere is called a three-sphere. In the rather unlikely event of 4D creatures existing, they would be able to visualize the curvature of 3D space in just the same way we can visualize that of 2D space. However, like the 2D sphere, the possible curvature of three-space is an intrinsic property, and there is no actual need for a higher-dimensional space for it to live in. Obtaining a correct mental picture of this is one of the big challenges in understanding our Universe!

A universe with a spherical geometry, like the surface of the Earth, has a finite size but no boundary. All points are equivalent. If we lived in a spherical geometry and traveled in a straight line, we would not go on forever. Rather, we would eventually come back to where we had started from, from the opposite direction, exactly in the manner that someone traveling outward from the North Pole on the Earth eventually returns there from the opposite direction. Such a universe corresponds to choosing a positive value for the quantity k appearing in the Friedmann equation. Because the special properties of the spherical geometry are due to its curvature, k is often therefore called the curvature term.

A universe with k greater than 0 is normally referred to as a **closed universe**, because of its finite size.

4.3 Hyperbolic geometry

The final choice is k negative. The corresponding geometry is known as hyperbolic, and is much less familiar than spherical geometry. It is normally represented by a saddle-like surface as in [Figure 4.2](#); it is hard to see that this is consistent with isotropy, but in fact it is. In a hyperbolic geometry, parallel lines never meet – in fact they break Euclid's axiom by diverging away from one another.

 Diagram of a saddle surface featuring a central dashed curvilinear triangle with angles sum to less than 180 degrees.

Figure 4.2 A sketch of a saddle surface, representing the hyperbolic geometry obtained when ***k*** is negative. A rather exaggerated triangle is shown with its sum of angles well below 180° .

The behavior of the hyperbolic geometry can be guessed from what has gone before; it is the opposite of spherical geometry. One finds that

- the angles of a triangle add up to less than 180° ;
- the circumference of a circle is greater than $2\pi r$.

Because parallel lines never meet, such a universe must be infinite in extent, just as in the flat case. The situation $k < 0$ is known as an **open universe**.

In [Chapter 9](#), we'll explore which of the three possible geometries, summarized in [Table 4.1](#), seems best suited to describe the real Universe.

TABLE 4.1

A summary of possible geometries.

Curvature	Geometry	Angles of triangle	Circumference of circle	Type of universe
$k > 0$	Spherical	$> 180^\circ$	$c < 2\pi r$	Closed
$k = 0$	Flat	180°	$c = 2\pi r$	Flat
$k < 0$	Hyperbolic	$< 180^\circ$	$c > 2\pi r$	Open

4.4 Infinite and observable universes

What does it actually mean for the Universe to be infinite, as in the flat and open cases? This property has nothing to do with the Universe lasting forever; what it means is that the Universe is already infinite in size even at a finite time. It genuinely goes on forever in all directions. Even despite this, it is still able to expand – the distances between objects can still increase regardless of whether the total Universe is infinite in extent or not. (You might consider the integers; these form an infinite set, but you can still multiply each by two to get a new infinite set where the separation between the numbers is twice as large.)

Nevertheless, this description is only a model, and we have no way of discovering whether the actual Universe does indeed go on forever. Cosmologists often talk about a different concept to the entire, possibly infinite, Universe, namely the **observable Universe**. This corresponds to the portion of the Universe we can actually see, and is limited by the finite speed of light. As the Universe gets older, the observable Universe becomes larger and larger, by a combination of two effects. First, the Universe is expanding, and second, light has had longer to travel across the Universe. In practice, our knowledge of the Universe is restricted to this portion, and we have no way of telling whether it does indeed continue into the infinite distance as required by the cosmological principle. It may be, for example, that the Universe becomes highly irregular on an extremely large scale, and there are some theoretical models that predict that this might happen. Another possibility is discussed in [Advanced Topic 1.3](#).

4.5 Where did the Big Bang happen?

A common question is “where did the Big Bang happen?”, suggesting perhaps that one could point in a specific direction and say “That way!” After all, in a conventional explosion that is a perfectly reasonable question to ask, as all the material flies outward from the ignition point. Unfortunately, for the Big Bang, things aren't so simple, and in a sense the answer is “everywhere and nowhere.”

First of all, remember that our entire foundation is the cosmological principle, telling us that no point in the Universe is special. If there were a particular point where the “Bang” happened, that would clearly be a special point and violate the cosmological principle. Rather, space and time themselves were created at the instant of the Big Bang (unlike a conventional explosion where the material flies through preexisting space). If we take any point in the present Universe and trace back its history, it would start out at the explosion point, and in that sense the Big Bang happened everywhere in space.

In another sense, the location of the Big Bang is nowhere, because space itself is evolving and expanding, and it has changed since the Big Bang took place. Imagine the Universe as an expanding sphere. At any instant, “space” is the surface of the sphere, which is becoming bigger with time (again we are thinking of a 2D analogy to our real 3D space). The place where the “Bang” happened is at the center of the sphere, but that's no longer part of the space, the surface of the sphere in which we live. In

particular, being constrained to the surface of the sphere means we are unable to “point” to the place where the explosion is supposed to have happened. However, all the points in our current space were once at the center of the sphere, when the Big Bang took place.

4.6 Three values of k

Because there are only three distinct possibilities for the geometry, many people indicate this explicitly by scaling their variables so that k takes on one of only three possible values, namely k equals negative 1, 0, or +1, corresponding to the open, flat, and closed cases, respectively. This can be achieved by rescaling the scale factor by multiplying it by a fixed constant, namely $\frac{1}{\sqrt{|k|}}$ if k is non-zero. This leaves H equal to $\frac{\dot{a}}{a}$ and removes the k from the final term in Equation (4.1), which now reads

$$\left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3} \rho + \frac{k}{a^2} \quad (4.2)$$

with “−” for positive k , “+” for negative k , and the last term absent if k equals 0.

If this rescaling is used, then (except in the k equals 0 case) one loses the freedom to set

$\frac{\dot{a}}{a} = 1$ at the present time, as will be used in Section 5.3. In effect, what one is doing is choosing to measure comoving distances in units of the so-called curvature scale, which is the scale on which the effects of curved space must be included, rather than in astronomical units such as megaparsecs.

In this book, we will *not* rescale k to one of these three discrete values, though most of the time we will be specializing to k equals 0 anyway.

Problems

4.1. Consider the surface of a 2D sphere of radius R , as illustrated.

Diagram of a sphere details radius R extending from the center, angle θ from the vertical axis, and arc length r along the surface.

Circles are drawn on the sphere that have radius r as measured on the surface of the sphere, centered on the North Pole. Show that the general formula for the circumference c of such a circle, as a function of r , is given by

$$c = 2\pi \frac{\sin \theta}{\theta} r = 2\pi R \frac{\sin \frac{r}{R}}{\frac{r}{R}}$$

where θ is the angle between a line drawn from the center of the sphere to the North Pole and one drawn to the circle. (Remember that by definition of angles in radians, we have $r = \theta R$.)

Demonstrate that for small θ (i.e. r much less than R) this gives the normal flat geometry relation. Evaluate the relation for the case when the circle is at the equator.

4.2. Consider the spherical geometry of the previous problem, staying with the 2D analogy to the real Universe. Suppose that galaxies are distributed evenly in such a universe, with a number density n per unit area. Show that the total number N of galaxies inside a radius r is given by

$$N = 2\pi n R^2 \left[1 - \cos \frac{r}{R} \right]$$

Expand this for r much less than R to show that the flat space result that the number is $\pi n r^2$ is recovered (remember we are working in only two dimensions). Do you see more or fewer galaxies out to the same radius, if the Universe is spherical rather than flat?

4.3. Is it possible for a closed universe to evolve to become an open universe? Give a reason for your answer.

Notes

¹ Well, that’s almost true. See Advanced Topic 1.3 for a way to bypass that conclusion.

² Note that, apart from the equator, lines of latitude are *not* straight lines; this is why aeroplanes do not follow lines of latitude when flying, because they are not the shortest way to go!

Chapter 5

Simple Cosmological Models

In [Chapter 3](#), we derived the equations satisfied by an expanding isotropic gas. They are the Friedmann equation ¹

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}, \quad (5.1)$$

which governs the time evolution of the scale factor $a(t)$, and the fluid equation

$$\dot{\rho} + 3\frac{\dot{a}}{a}\rho = 0, \quad (5.2)$$

which gives us the evolution of mass density $\rho(t)$. We stress that, despite their Newtonian derivation, these are the real equations used by cosmologists, more traditionally derived via the equations of general relativity as described in [Advanced Topic 1](#).

This chapter finds and discusses some simple solutions to these equations, which will be used extensively during the book. However, while they have wide applicability during the early stages of the Universe, it will turn out that these are not sufficient to describe the present state of the Universe, for which an extra ingredient, the cosmological constant, will be needed. It is introduced in [Chapter 7](#).

Before finding solutions to these equations, we can study two of their implications.

5.1 Hubble's law

The Friedmann equation allows us to explain Hubble's discovery that recession velocity is proportional to the distance. The velocity of recession is given by

$\vec{v} = d\vec{r}/dt$ and is in the same direction as $\vec{r}$, allowing us to write

$$\vec{v} = \frac{d\vec{r}}{dt} = \frac{d}{dt}\left(\frac{a}{a_0}\right)\vec{r} = \frac{\dot{a}}{a}\vec{r} \quad (5.3)$$

The last step used $\vec{r} = a\vec{x}$, remembering that the comoving position $\vec{x}$ is a constant by definition. Consequently, Hubble's law

$\vec{v} = H\vec{r}$ tells us that the proportionality constant, the Hubble parameter, should be identified as

$$H = \frac{\dot{a}}{a} \quad (5.4)$$

and the value as measured today can be denoted with a subscript "0" as H_0 . Because we measure Hubble's constant to be positive rather than negative, we know that the Universe is expanding rather than contracting.

We notice from this that the phrase Hubble's *constant* is a bit misleading. Although it is certainly constant in space due to the cosmological principle, there is no reason for it to be constant in time. In fact, using it as a more compact notation, we can write the Friedmann equation as an evolution equation for $H(t)$, as

$$H^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \quad (5.5)$$

It is best to use the phrase "Hubble parameter" for this quantity as a function of time, reserving "Hubble constant" for its present value. Normally, the Hubble parameter decreases with time, for instance, as the expansion is slowed by the gravitational attraction of the matter in the Universe.

5.2 Expansion and redshift

The redshift of spectral lines that we used to justify the assumption of an expanding universe can also be related to the scale factor. In this derivation, we will make the simplifying assumption that light is passed between two objects that are very close together, separated by a small distance d_r , as shown in [Figure 5.1](#). We have drawn the objects as galaxies, but we really mean two nearby points. According to Hubble's law, their relative velocity v will be

$$v \text{ equals } H d_r \text{ equals } \frac{1}{a} \frac{da}{dt} d_r \quad (5.6)$$

As the points are nearby, we can directly apply the Doppler law to say that the change in wavelength between emission and reception, $d\lambda$ identical to λ Subscript normal r Baseline minus λ Subscript normal e , is

$$\frac{d\lambda}{\lambda} \text{ equals } \frac{v}{c} \quad (5.7)$$

where $d\lambda$ is going to be positive since the wavelength is increased. The time between emission and reception is given by the light travel time dt equals d_r divided by c , and putting all that together gives

$$\frac{d\lambda}{\lambda} \text{ equals } \frac{1}{c} \frac{da}{dt} d_r \text{ equals } \frac{1}{a} \frac{da}{dt} d_r \quad (5.8)$$

Integrate and we find that $\ln \lambda$ equals $\ln a$ plus constant, i.e.

$$\lambda \text{ proportional to } a \quad (5.9)$$

where λ is now the instantaneous wavelength measured at any given time.

 Diagram of a photon wave traveling from spiral Galaxy A toward spiral Galaxy B across a distance measured as d_r .

Figure 5.1 A photon travels a distance d_r between two galaxies A and B.

Although as we have derived it, this result only applies to objects that are very close to each other, it turns out that it is completely general (a rigorous treatment is given in [Advanced Topic 2](#)). It tells us that as space expands, wavelengths become longer in direct proportion. One can think of the wavelength as being stretched by the expansion of the Universe, and its change therefore tells us how much the Universe has expanded since light began its travels. For example, if the wavelength had doubled, the Universe must have been half its present size when the light was emitted.

The redshift as defined in [Equation \(2.1\)](#) is related to the scale factor by

$$1 + z \text{ equals } \frac{\lambda_r}{\lambda_e} \text{ equals } \frac{a(t_r)}{a(t_e)} \quad (5.10)$$

and is normally only used to refer to light received by us at the present epoch.

5.3 Solving the equations

In order to discover how the Universe might evolve, we need some idea of what is in it. In a cosmological context, this is done by specifying the relationship between the mass density ρ and the pressure p . This relationship is known as the **equation of state**. At this point, we shall only consider two possibilities.

Matter: In this context, the term “matter” is used by cosmologists as shorthand for “non-relativistic matter,” and refers to any type of material which exerts negligible pressure, p equals 0. Occasionally, care is needed to avoid confusion between “matter” used in this sense, and used to indicate all types of matter, whether nonrelativistic or not. A pressureless universe is the simplest assumption that can be made. It is a good approximation to use for the atoms in the Universe once they have cooled down, as they are quite well separated and seldom interact, and it is also a good description of a collection of galaxies in the Universe, as they have no interactions other than gravitational ones. Occasionally the term “dust” is used instead of “matter.”

Radiation: Particles of light move, naturally enough, at the speed of light. Their kinetic energy leads to a pressure force, the radiation pressure, which, using the standard theory of radiation, can be shown to be

p equals ρc^2 divided by 3. [Problem 5.2](#) gives a rather hand-waving derivation of this result. More generally, any particles moving at highly relativistic speeds have this equation of state, neutrinos being an obvious example.

We will concentrate on the case where the constant k in the Friedmann equation is set equal to zero, corresponding to a flat geometry.

5.3.1 Matter

We start by solving the fluid equation, having set p equals 0 for matter. One way to solve it is to notice a clever way of rewriting it, as follows

$$\frac{d}{dt} \left(\rho a^3 \right) = 0 \quad (5.11)$$

though one could also solve it more formally by noting that it is a separable equation. Integrating tells us that ρa^3 equals a constant, i.e.

$$\rho \propto \frac{1}{a^3} \quad (5.12)$$

This is not a surprising result. It says that the density falls off in proportion to the volume of the Universe. It is very natural that if the volume of the Universe increases by a factor of say two, then the density of the matter must fall by the same factor. After all, material cannot come from nowhere, and there is no pressure to do any work.

The equations we are solving (with k equals 0) have one very useful symmetry; their form is unchanged if we multiply the scale factor a by a constant, since only the combination $\frac{da}{dt} a^3$ appears. This means that we are free to rescale a as we choose, and the normal convention is to choose a equals 1 at the present time. With this choice, physical and comoving coordinate systems coincide at the present, since $\frac{dr}{dt} = \frac{dx}{dt}$. Throughout this book we will use the subscript "0" to indicate the present value of quantities. Denoting the present density by ρ_0 fixes the proportionality constant

$$\rho = \frac{\rho_0}{a^3} \quad (5.13)$$

Having solved for the evolution of the density in terms of a , we must now find how a varies with time by using the Friedmann equation. Substituting in for ρ , and remembering we are assuming k equals 0, gives

$$\left(\frac{da}{dt} \right)^2 = \frac{8\pi G \rho_0}{3} \frac{1}{a} \quad (5.14)$$

Faced with an equation like this, one can use formal techniques to solve it (this equation is separable, allowing it to be integrated), or alternatively make an educated guess as to the solution and confirm it by substitution. In cosmology, a good educated guess is normally a power-law a proportional to t^q . Substituting this in, the left-hand side has time dependence t^{2q-2} and the right-hand side t^{-q} . This can only be a solution if these match, which requires q equals 2 divided by 3, and so the solution is a proportional to $t^{2/3}$. As we have fixed a equals 1 at the present time t equals t_0 , the full solution is therefore

$$a(t) = \left(\frac{3}{2} H_0 t \right)^{2/3} \quad (5.15)$$

In this solution, the Universe expands forever, but the rate of expansion H decreases with time

$$H = \frac{2}{3t} \quad (5.16)$$

becoming infinitely slow as the Universe becomes infinitely old. Notice that, despite the pull of gravity, the material in the Universe does not recollapse but rather expands forever.

This is one of the classic cosmological solutions, and will be much used throughout this book.

5.3.2 Radiation

Radiation obeys $p = \frac{1}{3} \rho c^2$. Consequently, the fluid equation is changed from the matter-dominated case, now reading

$$\frac{d}{dt}(\rho a^4) = 0 \quad (5.17)$$

This is amenable to the same trick as before, with the a^3 replaced by a^4 in Equation (5.11), giving

$$\rho \propto \frac{1}{a^4} \quad (5.18)$$

Carrying out the same analysis we did in the matter-dominated case gives

$$a(t) = \left(\frac{3}{4} \right)^{1/4} \left(\frac{t}{t_0} \right)^{1/2} \quad (5.19)$$

This is the second classic cosmological solution.

Notice that the Universe expands more slowly if radiation dominated than if matter dominated, a consequence of the extra deceleration that the pressure supplies – see Equation (3.18). So, it is definitely wrong to think of pressure as somehow “blowing” the Universe apart. However, in each case, the density of material falls off as a^{-3} .

We’d better examine the fall off of the radiation density with volume more carefully. It drops as the fourth power of the scale factor. Three of those powers we have already identified as the increase in volume, leading naturally to a drop in the density. The final power arises from a different effect, the stretching of the wavelength of the light. Since the stretching is proportional to a , and the energy of radiation proportional to its frequency via $E = hf$, this results in a further drop in energy by the remaining power of a . This lowering of energy is exactly the redshifting effect we use to measure distances.

The rate of decrease of the radiation density also has an explanation in terms of thermodynamics, which is macroscopic rather than microscopic. Since the Universe in this case has a pressure, when it expands, there is work done which is given by $p dV$, in exactly the same way as work is done on a piston when the gas is allowed to expand and cool. This work done corresponds to the extra diminution of the radiation density by the final factor of a .

5.3.3 Mixtures

A more general situation is when one has a mixture of both matter and radiation. Then there are two separate fluid equations, one for each of the two components. The trick which allows us to write ρ as a function of a still works, so we still have

$$\rho_m a^3 \propto 1, \quad \rho_r a^4 \propto 1 \quad (5.20)$$

However, there is still only a single Friedmann equation (after all, there is only one Universe!), which now has

$$\rho = \rho_m + \rho_r \quad (5.21)$$

This means that the scale factor will have a more complicated behavior, and so to convert

$\rho(a)$ into $\rho(t)$ is much harder. It is possible to obtain exact solutions for this situation, but they are quite messy, so we won’t include them here. Instead, we will consider a simpler situation where one or other of the densities is by far the larger.

In that case, we can say that the Friedmann equation is accurately solved by just including the dominant component. That is, we can use the expansion rates we have already found. For example, suppose radiation is much more important. Then, one would have

$$\rho \propto t^{-1} \quad ; \quad \rho_m \propto t^{-3/2} \quad ; \quad \rho_r \propto t^{-4} \quad (5.22)$$

Notice that the density in matter falls off more slowly than that in radiation. So, the situation of radiation dominating cannot last forever; however, small the matter component might be originally, it will eventually come to dominate. We can say that domination of the Universe by radiation is an unstable situation.

In the opposite situation, where it is the matter which is dominant, we obtain the solution

$$\rho \propto t^{-2/3} \quad ; \quad \rho_m \propto t^{-2} \quad ; \quad \rho_r \propto t^{-4} \quad (5.23)$$

Matter domination is a stable situation, the matter becoming increasingly dominant over the radiation as time goes by.

[Figure 5.2](#) shows the evolution of a universe containing matter and radiation, with the radiation initially dominating. Eventually, the matter comes to dominate, and as it does so, the expansion rate speeds up from

$\propto t^{-1/2}$ to $\propto t^{-2/3}$. It is likely that this is the situation that applied in our Universe until fairly recently, as we'll see in [Chapter 11](#).

 Graph of log density versus log time with two downward sloping lines, Radiation and Matter, converging at a central point.

Figure 5.2 A schematic illustration of the evolution of a universe containing radiation and matter. Once matter comes to dominate the expansion rate speeds up, so the densities fall more quickly with time.

5.4 Particle number densities

An important alternative view of the evolution of particles, which will be much used later in the book, is that of the **number density** n of particles rather than of their mass or energy density.

The number density is simply the number of particles in a given volume. If the mean energy per particle (including any mass-energy) is ϵ , then the number density is related to the energy density by

$$\epsilon = n \epsilon \quad (5.24)$$

The number density is useful because, in most circumstances, particle number is conserved. For example, if particle interactions are negligible, you wouldn't expect an electron to suddenly vanish into oblivion, and the same is true of a photon of light. The particle number can change through interactions. For example, an electron and positron could annihilate and create two photons. However, if the interaction rate is high, we expect the Universe to be in a state of thermal equilibrium. If so, then particle numbers are conserved even in a highly interacting state, since, by definition, thermal equilibrium means that any interaction, which may change the number density of a particular type of particle, must proceed at the same rate in both forward and backward directions so that any change cancels out.

So, barring brief periods where thermal equilibrium does not hold, we expect the number of particles to be conserved. The only thing that changes the number *density*, therefore, is that the volume is getting bigger, so that these particles are spread out in a larger volume. This implies

$$n \propto t^{-3} \quad (5.25)$$

This looks encouragingly like the behavior we have already seen for matter, but it's also true for radiation as well!

How does this relate to our earlier results? The energy of nonrelativistic particles is dominated by their rest mass-energy, which is constant, so

$$\rho_m \propto n m c^2 \propto t^{-3} \quad (5.26)$$

But photons lose energy as the Universe expands and their wavelength is stretched, so their energy is

$\epsilon \propto 1/a$ as we have already seen. So

$$\rho \propto a^{-3} \quad \text{Baseline proportional to } \epsilon \propto a^{-4} \quad \text{Baseline proportional to } n \propto a^{-3} \quad \text{Baseline times}$$

$$\frac{1}{a^3} \propto \frac{1}{a^3} \quad \text{Baseline proportional to } \frac{1}{a^3} \quad \text{Baseline proportional to } \frac{1}{a^4} \quad \text{Baseline times } \frac{1}{a^3}$$

$$\frac{1}{a^3} \propto \frac{1}{a^3} \quad \text{Baseline proportional to } \frac{1}{a^3} \quad \text{Baseline proportional to } \frac{1}{a^4} \quad \text{Baseline times } \frac{1}{a^3}$$

These are exactly the results we saw before: [Equations \(5.12\)](#) and [\(5.18\)](#).

Although the energy densities of matter and radiation evolve in different ways, their particle numbers evolve in the same way. So, apart from epochs during which the assumption of thermal equilibrium fails, the relative number densities of the different particles (e.g. electrons and photons) do not change as the Universe expands.

5.5 Evolution including curvature

We can now re-introduce the possibility that the constant k is nonzero, corresponding to spherical or hyperbolic geometry. Rather than seeking precise solutions, we will concentrate on the qualitative properties of the solutions. These are actually of rather limited use in describing our own Universe, because, as we will see, the cosmological models discussed so far are not general enough, and we will need to consider a cosmological constant (see [Chapter 7](#)). Nevertheless, studying the possible behaviors of these simple models is a useful exercise, even if one should be cautious about drawing general conclusions.

In analyzing the possible dynamics, we will assume that the Universe is always dominated by nonrelativistic matter, which in practice is not a restrictive assumption. We have already seen that if we assume that the constant k in the Friedmann equation is zero, then the Universe expands for ever, $a \propto t^{2/3}$, but slows down arbitrarily at late times. So, we know the fate of the Universe in that case. But what happens if $k \neq 0$?

The principal question to ask is whether it is possible for the expansion of the Universe to stop, which since $H = \dot{a}/a$ corresponds to $\dot{a} = 0$. Looking at the Friedmann equation

$$H^2 = \frac{8\pi G}{3} \rho - \frac{k}{a^2} \quad (5.28)$$

it is immediately apparent that this is not possible if k is negative, for then both the terms on the right-hand side of the Friedmann equation are positive. Consequently, such a universe must expand forever. That enables us to study the late-time behavior, because we can see that the term k/a^2 falls off more slowly with the expansion than it does $\rho \propto a^{-3}$. Since a becomes arbitrarily large for the matter-dominated solution for negligible k , the k/a^2 term must eventually come to dominate. When it does, the Friedmann equation becomes

$$-\frac{k}{a^2} = -\frac{k}{a^2} \quad (5.29)$$

Cancel off the a^2 terms and you'll find the solution is $a \propto t$. So, when the last term comes to dominate, the expansion of the Universe becomes yet faster. In this case, the velocity does not tend to zero at late times, but instead becomes constant. This is sometimes known as free expansion.

Things are very different if k is positive. It then becomes possible for H to be zero, by the two terms on the right-hand side of the Friedmann equation canceling each other out. Indeed, this is inevitable, because the negative influence of the k/a^2 term will become more and more important relative to the $\rho \propto a^{-3}$ term as time goes by. In such a universe, therefore, the expansion must come to an end after a finite amount of time. As gravitational attraction persists, the recollapse of the Universe becomes inevitable.

The collapse of the Universe is fairly easy to describe, because the equations governing evolution are time reversible. That is, if one substitutes negative t for t , they remain the same. The collapse phase is therefore just like the expansion in reverse, and so, after a finite time, the Universe will come to an end in a Big Crunch. [Problem 5.5](#) investigates this in more detail.

These three behaviors, illustrated in [Figure 5.3](#), can be related to the particle energy U in our Newtonian derivation of the Friedmann equation. If the particle's energy is positive, then it can escape to infinity, with a final kinetic energy given by U . If the total energy is zero, then the particle can just escape but with zero velocity. Finally, if the energy is negative, it cannot escape the gravitational attraction and is destined to recollapse inward.

 Graph of scale factor versus time with three curves labeled $k < 0$, $k = 0$, and $k > 0$ details universe evolution.

Figure 5.3 Three possible evolutions for the Universe, corresponding to the different signs of k . The middle line corresponds to the $k = 0$ case, where the expansion rate approaches zero in the infinite future. During the early phases of the expansion the lines are very close and so observationally it can be difficult to distinguish which path the actual Universe will follow.

There is a fairly precise analogy with escape velocity from the Earth (or the Moon, if you want to worry about the atmosphere). If you throw a rock up in the air hard enough, gravity will be unable to stop it, and eventually it will sail off into space at a constant velocity. If your throw is too puny, it will rapidly fall back. And in between is the escape velocity, where the rock is just able to escape the gravitational field and no more.

Problems

5.1. Is the energy of the Universe conserved as it expands?

5.2. This problem indicates the origin of the equation of state $p = \frac{1}{3} \rho c^2$ for radiation. An ideal gas has pressure

$$p = \frac{1}{3} n \langle v^2 \rangle$$

where $\langle \rangle$ indicates an average over the direction of particle motions. Here, n is the number density, and be careful not to confuse the unfortunate notation p for pressure and $\mathbf{p}$ for momentum. Using Equation (2.4) to relate the photon energy and momentum show that this gives

$$p = \frac{1}{3} n \langle E \rangle$$

where $\langle E \rangle$ is the mean photon energy. Hence, it demonstrates the equation of state for radiation.

5.3. In this chapter, we examined solutions for the expansion when the Universe contained either matter

or radiation

or a more general equation of state, $p = w \rho c^2$, where w is a constant in the range $-1 < w < 1$. Find solutions for $\rho(t)$ and hence

for universes containing such matter. Assume $k = 0$ in the Friedmann equation. What is the solution if $p = -\rho c^2$?

5.4. Using your answer to Problem 5.3, what value of w would be needed so that ρ has the same time dependence as the curvature term k/a^2 ? Find the solution $a(t)$ to the full Friedmann equation for a fluid with this w , assuming negative k .

5.5. The full Friedmann equation is

$$\left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3} \rho - \frac{k}{a^2}$$

Consider the case $k > 0$, with a universe containing only matter

so that $\rho = \rho_0 / a^3$. Demonstrate that the parametric solution

$$a(1 - \cos \theta) = \frac{4\pi G \rho_0}{3k} \quad t(1 - \cos \theta) = \frac{4\pi G \rho_0}{3k} \frac{1 - \cos^3 \theta}{2}$$

solves this equation, where θ is a variable which runs from 0 to 2π .

Sketch a and t as functions of θ . Describe qualitatively the behavior of the Universe. Attempt to sketch a as a function of t .

5.6. Now consider the case $k < 0$, with a universe containing only matter

so that $p = 0$ so that $\rho = \rho_0 / a^3$. What is the solution

in a situation where the final term of the Friedmann equation dominates over the density term? How does the density of matter vary with time? Is domination by the curvature term a stable situation that will continue forever?

Note

¹In accord with the discussion of [Section 3.6](#), the c^2 squared on the final term has been dropped, so that its appearance matches that of other cosmology textbooks.

Chapter 6

Observational Parameters

The Big Bang model does not give a unique description of our present Universe, but rather leaves quantities such as the present expansion rate, or the present composition of the Universe, to be fixed by observation.

It is a standard practice to specify cosmological models via a few parameters, which one then tries to determine observationally to decide which version of the model best describes our Universe. In this chapter and [Chapter 7](#), we will discuss the most commonly considered parameters, including ones we have already seen and new ones.

6.1 The expansion rate H_0

The Hubble constant H_0 tells us the present expansion rate of the Universe, making it a vital cosmological parameter as measurements of other properties of the Universe often depend on it. It also ought to be the easiest to measure, since all galaxies are supposed to obey $v = H_0 r$. So, all we have to do is measure the velocities and distances of as many galaxies as we can and get an answer. However, each measurement has its problems.

Velocities are given by the redshift of spectral lines, a measurement which is now easy enough that the velocity of an individual galaxy can be measured to high accuracy. However, remember that the cosmological principle isn't perfect, and so, as well as the uniform expansion we are trying to measure, galaxies also have motions relative to one another, the so-called "peculiar velocity." The peculiar velocities are randomly oriented, and for a given galaxy we cannot split its measured velocity into the Hubble expansion and the peculiar velocity. However, the cosmological principle does tell us that the typical size of the peculiar velocity should not depend on where in the Universe the galaxy is. It is therefore independent of distance, whereas the Hubble velocity is proportional to distance. If we look far enough away (in practice many tens of megaparsecs) then the Hubble velocity dominates and the (unknown) peculiar velocity can be ignored.

Given that the expansion velocity can only be accurately distinguished from the peculiar velocity at large distances, we need to be able to estimate these large distances accurately in order to carry out the calculation $H_0 = v/r$. These distances are much harder to obtain, because galaxies are far too distant to be located by parallax. Remember that an object one parsec away has a parallax (i.e. an apparent motion when viewed from different parts of Earth's orbit) of one arcsecond, by definition. A galaxy many megaparsecs away will have an immeasurably small parallax of less than a micro-arcsecond. The usual method is known as "standard candles," where some type of object is assumed to have exactly the same properties in all parts of the Universe. This is the cosmological equivalent of saying that if one light bulb looks a quarter as bright as another, then from the inverse square law it must be twice as far away – fine as long as you believe that all light bulbs have precisely the same brightness. A classic example is the period–luminosity relation in cepheid variable stars. The period of variability of those stars is readily measured, and there is reasonable empirical evidence of a relation between the period and the luminosity of the stars, which lets us convert the measured period into a brightness. Other standard candles which have been used include the brightness of certain types of supernovae, and the brightest galaxies in galaxy clusters. All these methods have good success at determining relative distances between two galaxies, which requires that the objects be good standard candles, and relative distances are all that is needed to confirm the Hubble law.

However, to give an absolute distance, by actually measuring the proportionality constant H_0 , we also need a calibration against an object of known distance, which proves much harder. In the light bulb analogy, to get relative distances we need only the inverse square law and the belief that all bulbs have the same brightness; we don't need to know how bright the bulbs are. But to be able to say how far away a bulb of a given observed brightness is, we need to know its absolute brightness.

It is only relatively recently that the calibration problem has begun to come under control, principally through Hubble Space Telescope observations. The Hubble constant is usually parametrized as

$$H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}. \quad (6.1)$$

In the 1990s, the Hubble Space Telescope Key Project, led by Wendy Freedman, gave

$$h = 0.72 \pm 0.08, \quad (6.2)$$

where the uncertainty is a one-sigma error (meaning it should be doubled to indicate 95% confidence, at least if the uncertainty is approximately Gaussian-distributed). This was considered the first precision direct measurement of the Hubble constant, and has subsequently been improved to $h = 0.73 \pm 0.01$ by the SH0ES (Supernovae and H_0 for the Equation of State) project.

On the other hand, indirect measurements, which assess the impact of the expansion rate on other cosmological observables, such as the cosmic microwave background, have suggested a lower value. The *Planck* satellite observations described in [Advanced Topic 6](#) give a very precise determination $h = 0.674 \pm 0.005$. There is a clear discrepancy between this measurement and the direct one, which has persisted for several years and become known as the **Hubble tension**. Its

implications will be discussed further in [Advanced Topic A6.5](#). Notwithstanding this controversy, it is safe to conclude that h is now quite accurately measured to be 0.70 within at most a few per cent. An object with a recession velocity of 7000 km s^{-1} would be expected to be at a distance $v/H_0 = 100 \text{ Mpc}$. The overall scale of the Universe is therefore well determined.

6.2 The density parameter Ω_0

The density parameter is a very useful way of specifying the density of the Universe. Let's start with the Friedmann equation again. Recalling that $H = \dot{a}/a$, it reads

$$H^2 = \frac{8\pi G}{3} \rho - \frac{k}{a^2}. \quad (6.3)$$

For a given value of H , there is a special value of the density which would be required in order to make the geometry of the Universe flat, $k = 0$. This is known as the **critical density** ρ_c , which we see is given by

$$\rho_c(t) = \frac{3H^2}{8\pi G}. \quad (6.4)$$

Note that the critical density changes with time, since H does. As we know the present value of the Hubble constant (at least in terms of h defined in [Equation \(6.1\)](#)), we can compute the present critical density. Since $G = 6.67 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ sec}^{-2}$, and converting megaparsecs to meters using conversion factors quoted on page xiii, it is

$$\rho_c(t_0) = 1.88 h^2 \times 10^{-26} \text{ kg m}^{-3}. \quad (6.5)$$

This is a startlingly small number; compare, for example, the density of water which is 10^3 kg m^{-3} . If there is any more matter than this apparently tiny amount, it is enough to tip the balance beyond a flat universe to a closed one with $k > 0$. So only a very tiny density of matter is needed in order to provide enough gravitational attraction to halt and reverse the expansion of the Universe.

However, let us write that another way, since kilograms and meters are rather small and inconvenient units for dealing with something as big as the Universe. Let's try measuring masses in solar masses and distances in megaparsecs. It becomes

$$\rho_c(t_0) = 2.78 h^{-1} \times 10^{11} M_\odot / (h^{-1} \text{ Mpc})^3. \quad (6.6)$$

Suddenly this doesn't look so small. In fact, $10^{11} - 10^{12}$ solar mass is about the mass of a typical galaxy, and a megaparsec more or less the typical galaxy separation, so the Universe cannot be far away (within an order of magnitude or so) from the critical density. Its density really must be around $10^{-26} \text{ kg m}^{-3}$.

Be sure to understand that the critical density is not necessarily the true density of the Universe, since the Universe need not be flat. However, it sets a natural scale for the density of the Universe. Consequently, rather than quote the density of the Universe directly, it is often useful to quote its value relative to the critical density. This dimensionless quantity is known as the **density parameter** Ω , defined by

$$\Omega(t) \equiv \frac{\rho}{\rho_c}. \quad (6.7)$$

Again, in general, Ω is a function of time, since both ρ and ρ_c depend on time. The present value of the density parameter is denoted Ω_0 .

With this new notation, we can rewrite the Friedmann equation in a very useful form. Substituting in for ρ in [Equation \(6.3\)](#) using the definitions we have made, [Equations \(6.4\)](#) and [\(6.7\)](#), leads to

$$H^2 = \frac{8\pi G}{3} \rho_c \Omega - \frac{k}{a^2} = H^2 \Omega - \frac{k}{a^2}, \quad (6.8)$$

and rearranging gives

$$\Omega - 1 = \frac{k}{a^2 H^2}. \quad (6.9)$$

We see that the case $\Omega = 1$ is very special, because in that case k must equal zero and since k is a fixed constant, this equation becomes $\Omega = 1$ for all time. That is true independent of the type of matter we have in the Universe, and this is often called a **critical-density universe**. When $\Omega \neq 1$, this form of the Friedmann equation is very useful for analyzing the evolution of the density, as we will see later in the chapter on inflationary cosmology.

Our Universe contains several different types of matter, and this notation can be used not just for the total density but also for each individual component of the density, so one talks of Ω_{mat} , Ω_{rad} , etc. We can even define a “density parameter” associated with the curvature term, by writing

$$\Omega_k \equiv -\frac{k}{a^2 H^2}. \quad (6.10)$$

This can be positive or negative, and using it the Friedmann equation can be written as

$$\Omega + \Omega_k = 1. \quad (6.11)$$

We'll return to the observational status of Ω_0 in [Chapter 9](#).

6.3 The deceleration parameter q_0

As we've discovered, not only is the Universe expanding, but also the rate at which it is expanding, given by the Hubble parameter, is changing with time. The deceleration parameter is a way of quantifying this.

Consider a Taylor expansion of the scale factor about the present time. The general form of this (with dots as always indicating time derivatives) is

$$a(t) = a(t_0) + \dot{a}(t_0)[t - t_0] + \frac{1}{2}\ddot{a}(t_0)[t - t_0]^2 + \dots \quad (6.12)$$

Let's divide through by $a(t_0)$. Then the coefficient of the $[t - t_0]$ term will just be the present Hubble parameter, and we can write

$$\frac{a(t)}{a(t_0)} = 1 + H_0[t - t_0] - \frac{q_0}{2}H_0^2[t - t_0]^2 + \dots, \quad (6.13)$$

which defines the deceleration parameter q_0 as

$$q_0 = -\frac{\ddot{a}(t_0)}{a(t_0)} \frac{1}{H_0^2} = -\frac{a(t_0)\ddot{a}(t_0)}{\dot{a}^2(t_0)}. \quad (6.14)$$

The larger the value of q_0 , the more rapid the deceleration.

The simplest situation is if the Universe is matter dominated, $p = 0$. Remember that by “matter” we mean any pressureless material; it could be a collection of elementary particles, or, equally well, a collection of galaxies. Then from the acceleration [Equation \(3.18\)](#) and the definition of critical density, [Equation \(6.4\)](#), we find

$$q_0 = \frac{4\pi G}{3}\rho \frac{3}{8\pi G\rho_c} = \frac{\Omega_0}{2}. \quad (6.15)$$

So, in this case, a measurement of q_0 would immediately tell us Ω_0 .

If we know the properties of the matter in the Universe, then q_0 is not independent of the first two parameters we have discussed, H_0 and Ω_0 . Those two are sufficient to describe all the possibilities. However, we don't know everything about the material in the Universe, so q_0 can provide a new way of looking at the Universe. It can in principle be measured directly by making observations of objects at very large distances, such as the numbers of distant galaxies, because the deceleration governs how large the Universe would be at an earlier time.

In the late 1990s, the first convincing measurements of q_0 were made by two research groups studying distant supernovae of a class known as type Ia, which are believed to be good standard candles. To widespread surprise at the time, the result is that the Universe is *accelerating* at present, $q_0 < 0$.¹ None of the cosmological models that we have discussed so far is capable of satisfying this condition, as can be seen directly from the acceleration [Equation \(3.18\)](#). This result is now firmly established, and among the most dramatic observational results in modern cosmology. The following chapter discusses how to extend our simple cosmological models to account for it.

Problems

- 6.1. Verify the numerical value of the critical density in terms of solar masses and megaparsecs, [Equation \(6.6\)](#).

6.2. The deceleration parameter is defined by [Equation \(6.14\)](#). Use the acceleration [Equation \(3.18\)](#) and the definition of critical density to show that a radiation-dominated universe has $q_0 = \Omega_0$.

6.3. Identify a sufficient and necessary condition that must be satisfied by the equation of state if q_0 is to be negative.

Note

¹The mathematical tools required to analyze such data are beyond the scope of the main body of this book, but are described in [Advanced Topic 2](#), where the supernova observations are discussed in greater detail.

Chapter 7

The Cosmological Constant

7.1 Introducing Λ

When formulating general relativity, Einstein believed that the Universe was static, but found that his theory of general relativity did not permit it. This is simply because all matter attracts gravitationally; none of the solutions we have found correspond to a static universe with constant a . In order to arrange a static universe, he proposed a change to the equations, something he would later famously call his “greatest blunder.” That was the introduction of a cosmological constant.

The introduction of such a term is permitted by general relativity, and although Einstein’s original motivation has long since faded, it is currently seen as one of the most important and enigmatic objects in cosmology. The cosmological constant Λ appears in the Friedmann equation as an extra term, giving

$$H^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} + \frac{\Lambda}{3}. \quad (7.1)$$

Here, Λ has units $[\text{time}]^{-2}$, though some people include an explicit factor of c^2 in this equation to instead measure it as $[\text{length}]^{-2}$.

In principle, Λ can be positive or negative, though the positive case is much more commonly considered. Einstein’s original idea was to balance curvature, Λ , and ρ to get $H(t) = 0$, and hence, a static universe (see [Problem 7.2](#)). In fact, this idea was rather misguided, since such a balance proves to be unstable to small perturbations, and hence presumably couldn’t arise in practice. Nowadays, the cosmological constant is most often discussed in the context of universes with the flat Euclidean geometry, $k = 0$.

The effect of Λ can be seen more directly from the acceleration equation. Following the derivation of [Section 3.5](#), but now using the Friedmann equation as given above, gives

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda}{3}. \quad (7.2)$$

A positive cosmological constant gives a positive contribution to $\ddot{a}$, and so acts effectively as a repulsive force. In particular, if the cosmological constant is sufficiently large, it can overcome the gravitational attraction represented by the first term and lead to an accelerating universe. It can therefore explain the observed acceleration of the Universe described in [Section 6.3](#).

In the same way that it is useful to express the density as a fraction of the critical density, it is convenient to define a density parameter for the cosmological constant as

$$\Omega_\Lambda = \frac{\Lambda}{3H^2}. \quad (7.3)$$

Although Λ is a constant, Ω_Λ is not since H varies with time. Repeating the steps used to write the Friedmann equation in the form of [Equation \(6.9\)](#), we then find

$$\Omega + \Omega_\Lambda - 1 = \frac{k}{a^2 H^2}. \quad (7.4)$$

The condition to have a flat universe, $k = 0$, generalizes to

$$\Omega + \Omega_\Lambda = 1. \quad (7.5)$$

The usual convention among astronomers, which we will follow in this book, is that the cosmological constant term is not considered to be part of the matter density Ω . (Particle physicists, on the other hand, often include the cosmological constant as one of the components of the total density.) The relation between the density parameters and the geometry now becomes

Open universe:	$0 < \Omega + \Omega_\Lambda < 1.$
Flat universe:	$\Omega + \Omega_\Lambda = 1.$
Closed universe:	$\Omega + \Omega_\Lambda > 1.$

7.2 Fluid description of Λ

It is often helpful to describe Λ as if it were a fluid with energy density ρ_Λ and pressure p_Λ . From [Equation \(7.1\)](#), we see that the definition

$$\rho_\Lambda \equiv \frac{\Lambda}{8\pi G} \quad (7.6)$$

brings the Friedmann equation into the form

$$H^2 = \frac{8\pi G}{3}(\rho + \rho_\Lambda) - \frac{k}{a^2}. \quad (7.7)$$

This definition also ensures that $\Omega_\Lambda \equiv \rho_\Lambda / \rho_c$, where ρ_c is the critical density.

In order to determine the effective pressure corresponding to Λ , one can seek a definition so that the acceleration equation with Λ reduces to its standard form, [Equation \(3.18\)](#), with $\rho \rightarrow \rho + \rho_\Lambda$ and $p \rightarrow p + p_\Lambda$. More directly, we can consider the fluid equation for Λ

$$\dot{\rho}_\Lambda + 3\frac{\dot{a}}{a}\left(\rho_\Lambda + \frac{p_\Lambda}{c^2}\right) = 0. \quad (7.8)$$

Since ρ_Λ is constant by definition, we must have

$$p_\Lambda = -\rho_\Lambda c^2. \quad (7.9)$$

The cosmological constant has a *negative* effective pressure. This means that as the Universe expands, work is done on the cosmological constant fluid. This permits its energy density to remain constant even though the volume of the Universe is increasing.

Concerning its physical interpretation, Λ is sometimes thought of as the energy density of “empty” space. In particular, in quantum physics, one possible origin is as a type of “zero-point energy,” which remains even if no particles are present, though unfortunately, particle physics theories tend to predict that the cosmological constant is far larger than observations allow. This discrepancy is known as the **cosmological constant problem**, and is one of the key unsolved problems in elementary particle physics.

It may be that the cosmological constant is only a transient phenomenon, which will disappear in the future. Another possibility, often called **dark energy** or **quintessence**, is that the cosmological constant is not actually perfectly constant but exhibits slow variation. For instance, one could assume the dark energy “fluid” to have equation of state

$$p_{de} = w\rho_{de}c^2, \quad (7.10)$$

where w could be a constant or, more generally, also itself evolve with redshift. The case $w = -1$ corresponds to a cosmological constant, while more generally accelerated expansion is possible provided $w < -1/3$ (you explored some solutions of this type in [Problem 5.3](#)). However, in this book, we will only consider the case of a perfect cosmological constant, except for a short comment in the very last Advanced Topic.

7.3 Cosmological models with Λ

The introduction of Λ has forced cosmologists to rethink some of the standard lore of cosmology, as it greatly increases the range of possible behaviors of the Universe. For instance, it is no longer necessarily true that a closed universe ($k > 0$) recollapses, nor that an open universe expands forever. In fact, if the cosmological constant is powerful enough, there need not even be a Big Bang, with the Universe instead beginning in a collapsing phase, followed by a bounce at finite size under the influence of the cosmological constant (though such models are ruled out by observations). It is also possible to have a prolonged phase where the Universe remains almost static, known as “loitering,” by arranging parameters so that the Universe closely approaches the unstable Einstein static universe.

As the Hubble parameter only provides an overall scaling factor, a useful way to parametrize possible models is to focus on the two other parameters, the present densities of matter and the cosmological constant. An excellent assumption is to assume the matter in the present Universe is pressureless. Different models can then be identified by their location in the plane of Ω_0 and

Ω_Λ , as shown in [Figure 7.1](#).¹ We have already seen that the line $\Omega_0 + \Omega_\Lambda = 1$ gives a flat universe, and divides the plane into open and closed cosmologies.

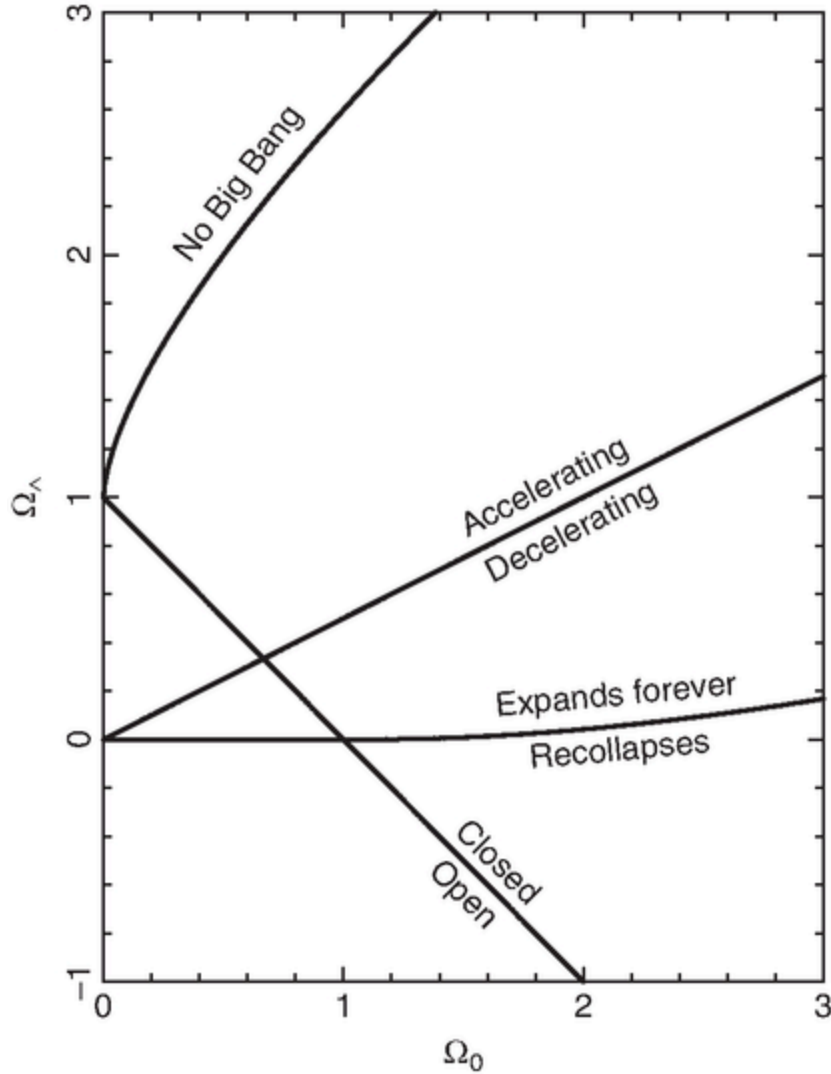


Figure 7.1 Different models for the Universe can be identified by their location in the plane showing the densities of matter and Λ . This figure indicates the main properties in different regions, with the labels indicating the behavior on each side of the dividing lines.

To identify where in the plane we have an accelerating universe, we need an expression for the deceleration parameter q_0 . A pressureless universe with a cosmological constant has

$$q_0 = \frac{\Omega_0}{2} - \Omega_\Lambda, \quad (7.11)$$

which you are asked to derive in [Problem 7.3](#), and so we have acceleration provided $\Omega_\Lambda > \Omega_0/2$. If we additionally assume that the geometry is flat, this relation simplifies further to $q_0 = 3\Omega_0/2 - 1$, and we have acceleration if $\Omega_\Lambda > 1/3$.

The other two main properties are whether there is a Big Bang, and whether the Universe will eventually collapse or not. There are analytic expressions for these curves, shown in [Figure 7.1](#), but they are too complicated to give here. For $\Omega_0 \leq 1$, whether there is recollapse or not depends simply on the sign of Λ , but for $\Omega_0 > 1$, the gravitational attraction of matter can overcome a small positive cosmological constant and cause recollapse.

While most cosmologists would have preferred the cosmological constant to equal zero, the Universe itself appears to have other ideas, with the observations of distant type Ia supernovae mentioned at the end of the previous chapter arguing strongly in favor of a presently accelerating universe. The observations leading to that conclusion are explored in more detail in

[Advanced Topic 2.3](#), and even if you are not planning to read that section, we suggest you have a look at [Figure A2.4](#) on page 142, which shows the observational constraints superimposed on the $\Omega_0 - \Omega_\Lambda$ plane. These observations demand inclusion of the cosmological constant; it is now regarded as an essential part of cosmological models aiming to explain observational data, and understanding its value is one of the mysteries of fundamental physics.

Problems

7.1. Suppose that the Universe contains four different contributions to the Friedmann equation, namely radiation, nonrelativistic matter, a cosmological constant, and a negative (hyperbolic) curvature. Write down the way in which each of these terms behaves as a function of the scale factor $a(t)$. Which of them would you expect to dominate the Friedmann equation at early times, and which at late times?

7.2. By considering both Friedmann and acceleration equations, and assuming a pressureless universe, demonstrate that in order to have a static universe, we must have a closed universe with a positive vacuum energy. Using either physical arguments or mathematics, demonstrate that this solution must be unstable.

7.3. Confirm the result, quoted in the main text, that a pressureless universe with a cosmological constant has a deceleration parameter given by

$$q_0 = \frac{\Omega_0}{2} - \Omega_\Lambda(t_0).$$

7.4. The most likely cosmology describing our own Universe has a flat geometry with a matter density of $\Omega_0 \simeq 0.3$ and a cosmological constant with $\Omega_\Lambda(t_0) \simeq 0.7$. What will the values of Ω and Ω_Λ be when the Universe has expanded to be five times its present size? Use an approximation suggested by this result to find the late-time solution to the Friedmann equation for our Universe. What is the late-time value of the deceleration parameter q ?

7.5. Show that in a spatially flat matter-dominated cosmology the density parameter evolves as

$$\Omega(z) = \Omega_0 \frac{(1+z)^3}{1 - \Omega_0 + (1+z)^3 \Omega_0}.$$

If our Universe has $\Omega_0 \simeq 0.3$, at what redshift did it begin accelerating?

Note

¹Beware the somewhat sloppy notation of sometimes using Ω_Λ to indicate the present value of this quantity.

Chapter 8

The Age of the Universe

One of the quantities that we can predict from a cosmological model, from the solution $a(t)$ for the expansion, is the age of the Universe t_0 . This offers an opportunity to connect the age of the Universe itself with the ages of objects within it. Historically, there was much concern as to whether the predicted age of the Universe was large enough to accommodate the ages of its contents, but in recent years such fears have largely disappeared.

Let's start with an approximate estimation. The characteristic rate of expansion is given by the Hubble parameter, so a first guess at the age of the Universe is that it is the timescale associated with the Hubble parameter, namely H_0^{-1} , since $v = H_0 r$ is basically distance equals velocity multiplied by time. This estimate is approximate, because it ignores the fact that v changes with time under the effect of gravity. We've been writing the Hubble constant as

$$H_0 = 100 \, h \, \text{km} \, \text{normal s}^{-1} \text{Baseline} \, \text{upper M p c}^{-1} \text{Baseline} \, \text{comma} \quad (8.1)$$

which isn't quite what we need, because the unusual units conceal the fact that H_0 simply has the units of $[\text{time}]^{-1}$. We have to convert the kilometers into megaparsecs (or vice versa) to cancel them out and then convert seconds to years. What we get, using values quoted on page xiii, is

$$H_0^{-1} \text{Baseline} = 9.77 \, h^{-1} \text{Baseline} \times 10^9 \text{Baseline} \, \text{yr} \quad (8.2)$$

s period

This is known as the **Hubble time**. The uncertainty in the Hubble constant, indicated by h , means we have to tolerate an uncertainty in this crude estimate of the age of the Universe. But the message is that 10 billion years is a good first guess at the age of the Universe.

Before progressing to a better calculation, let's find out what observations tell us of the age of the Universe. The geological timescale gives us a good estimate of the age of the Earth, which is about four and a half billion years. But it is not thought that the Earth is nearly as old as the Universe. There are various ways of dating other objects in the Universe. The relative amounts in the Galactic disk of Uranium isotopes, which have lifetimes comparable to the age of the Universe, suggests an age around ten billion years – see [Problem 8.1](#). Estimates consistent with this are also found from studying the cooling of white dwarf stars after they form. However, the best method is thought to be to use the chemical evolution of stars in old globular clusters, which are believed to be among the oldest objects in the Universe. The best estimate for the oldest ones is around 13 billion years, give or take a billion, with perhaps an extra billion years to be added in before they manage to form in the first place.

It should be stressed that it is remarkable that the cosmological theory and the age measurements are in the same ballpark at all. That in itself is a strong vindication that the ideas behind the Big Bang are along the right track. Our crude estimate above seems to adequately account for the observed ages of objects in the Universe.

What happens if we try to do better with our theoretical estimates? Although the precise cosmological model describing our Universe is uncertain, we are pretty sure that it has been matter dominated (i.e. dominated by some form of pressureless material) for some considerable time, and so we can use the matter-dominated evolution to calculate the age. Let's first suppose that the Universe has the critical density. Then we already have the solution, [Equation \(5.15\)](#), which is

$$t \text{ right parenthesis} = \text{left parenthesis} \frac{t}{t_0} \text{EndFraction} \text{ right parenthesis} \quad (8.3)$$

$$\text{Superscript 2 divided by 3 Baseline} \text{ right double arrow} \text{upper H identical to} \frac{2}{3} t \text{EndFraction} \text{ comma}$$

and so, the present Hubble parameter is

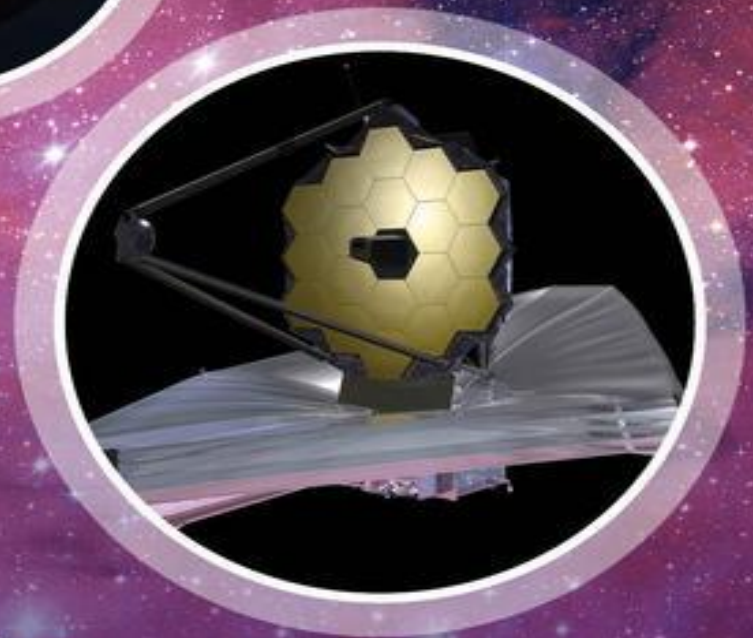
$$H_0 = \frac{2}{3} t_0^{-1} \text{period} \quad (8.4)$$

So, in such a universe, the age is actually shorter than our naïve estimate – it's only

$$t_0 = \frac{3}{2} H_0^{-1} = 6.51 \, h^{-1} \text{Baseline} \times 10^9 \text{Baseline} \, \text{yr} \text{ s period} \quad (8.5)$$

The extra factor of $2/3$ has put us in an uncomfortable position; with the measured value h asymptotically equals 0.7, we predict an age of only 9.3 billion years.

**Marina Cortês
Andrew Liddle**



An Introduction to **MODERN COSMOLOGY**

Fourth Edition

WILEY

Table of Contents

[Cover](#)

[Table of Contents](#)

[Title Page](#)

[Copyright](#)

[Dedication](#)

[Preface](#)

[About the Companion Website](#)

[Chapter 1: A Brief History of Cosmological Ideas](#)

[Chapter 2: Observational Overview](#)

[2.1 In visible light](#)

[2.2 In other wavebands](#)

[2.3 Homogeneity and isotropy](#)

[2.4 The expansion of the Universe](#)

[2.5 Particles in the Universe](#)

[Problems](#)

[Notes](#)

[Chapter 3: Newtonian Gravity](#)

[3.1 The Friedmann equation](#)

[3.2 On the meaning of the expansion](#)

[3.3 Things that go faster than light](#)

[3.4 The fluid equation](#)

[3.5 The acceleration equation](#)

[3.6 On mass, energy, and vanishing factors of \$c^2\$](#)

[Problems](#)

[Note](#)

[Chapter 4: The Geometry of the Universe](#)

[4.1 Flat geometry](#)

[4.2 Spherical geometry](#)

[4.3 Hyperbolic geometry](#)

[4.4 Infinite and observable universes](#)

[4.5 Where did the Big Bang happen?](#)

[4.6 Three values of \$k\$](#)

[Problems](#)

[Notes](#)

[Chapter 5: Simple Cosmological Models](#)

[5.1 Hubble's law](#)

[5.2 Expansion and redshift](#)

[5.3 Solving the equations](#)

[5.4 Particle number densities](#)

[5.5 Evolution including curvature](#)

[Problems](#)

[Note](#)

[Chapter 6: Observational Parameters](#)

[6.1 The expansion rate \$H_0\$](#)

[6.2 The density parameter \$\Omega_0\$](#)

[6.3 The deceleration parameter \$q_0\$](#)

[Problems](#)

[Note](#)

[Chapter 7: The Cosmological Constant](#)

[7.1 Introducing \$\Lambda\$](#)

[7.2 Fluid description of \$\Lambda\$](#)

[7.3 Cosmological models with \$\Lambda\$](#)

[Problems](#)

[Note](#)

[Chapter 8: The Age of the Universe](#)

[Problems](#)

[Chapter 9: The Density of the Universe and Dark Matter](#)

[9.1 Weighing the Universe](#)

[9.2 What might the dark matter be?](#)

[9.3 Dark matter searches](#)

[Problems](#)

[Notes](#)

[Chapter 10: The Cosmic Microwave Background](#)

[10.1 Properties of the microwave background](#)

[10.2 The photon to baryon ratio](#)

[10.3 The origin of the microwave background](#)

[10.4 The origin of the microwave background \(optional advanced treatment\)](#)

[Problems](#)

[Chapter 11: The Early Universe](#)

[Problems](#)

[Chapter 12: Nucleosynthesis: The Origin of the Light Elements](#)

[12.1 Hydrogen and helium](#)

[12.2 Comparing with observations](#)

[12.3 Contrasting decoupling and nucleosynthesis](#)

[Problems](#)

[Note](#)

[Chapter 13: The Inflationary Universe](#)

[13.1 Problems with the Hot Big Bang](#)

[13.2 Inflationary expansion](#)

[13.3 Solving the Big Bang problems](#)

[13.4 How much inflation?](#)

[13.5 Inflation and particle physics](#)

[Problems](#)

[Chapter 14: The Initial Singularity](#)

[Problem](#)

[Chapter 15: Overview: The Standard Cosmological Model](#)

[15.1 Expansion](#)

[15.2 Geometry](#)

[15.3 Age](#)

[15.4 Fate](#)

[15.5 Contents](#)

[15.6 Early history](#)

[15.7 Outlook](#)

[Advanced Topic 1: General Relativistic Cosmology](#)

[A1.1 The metric of space-time](#)

[A1.2 The Einstein equations](#)

[A1.3 Aside: Topology of the Universe](#)

[Problems](#)

[Advanced Topic 2: Classic Cosmology: Distances and Luminosities](#)

[A2.1 Light propagation and redshift](#)

[A2.2 The observable Universe](#)

[A2.3 Luminosity distance](#)

[A2.4 Angular diameter distance](#)

[A2.5 Source counts](#)

[Problems](#)

[Advanced Topic 3: Neutrino Cosmology](#)

[A3.1 The massless case](#)

[A3.2 Massive neutrinos](#)

[A3.3 Neutrinos and structure formation](#)

[Problems](#)

[Advanced Topic 4: Baryogenesis](#)

[Problem](#)

[Advanced Topic 5: Structures in the Universe](#)

[A5.1 The observed structures](#)

[A5.2 Gravitational instability](#)

[A5.3 The clustering of galaxies](#)

[A5.4 Cosmic microwave background anisotropies](#)

[A5.5 The origin of structure](#)

[Problems](#)

[Advanced Topic 6 Constraining Cosmological Models](#)

[A6.1 Cosmological models and parameters](#)

[A6.2 Key cosmological observations](#)

[A6.3 Cosmological data analysis](#)

[A6.4 The Standard Cosmological Model: 2025 edition](#)

[A6.5 Tensions and hints of new physics](#)

[A6.6 The future](#)

[Problems](#)

[Bibliography](#)

[Numerical Answers and Hints to Problems](#)

[Index](#)

[End User License Agreement](#)

List of Illustrations

Chapter 2

[Figure 2.1 If viewed from above the disk, our own Milky Way galaxy would probably resemble ...](#)

[Figure 2.2 A map of galaxy positions in a narrow slice of the Universe, as measured by the...](#)

[Figure 2.3 Images of the Coma cluster of galaxies in visible/infrared light \(top\) and in X-...](#)

[Figure 2.4 A computer simulation showing the predicted distribution of matter in the Univer...](#)

[Figure 2.5 The cosmic microwave background spectrum as measured by the FIRAS experiment...](#)

[Figure 2.6 The cosmic microwave background temperature variations on the sky as measured b...](#)

[Figure 2.7 A plot of velocity versus estimated distance for a set of 1355 galaxies. A straight-line relation...](#)

[Figure 2.8 According to Hubble's law, the further away from us a galaxy is, the faster it i...](#)

[Figure 2.9 The Planck function of Equation \(2.7\). There are far more photons with very low...](#)

[Figure 2.10 The energy density distribution of a blackbody spectrum, given by Equation \(2.8\). Most of...](#)

Chapter 3

[Figure 3.1 The particle at radius \$r\$ only feels gravitational attraction from the shaded region. Any...](#)

[Figure 3.2 The comoving coordinate system is carried along with the expansion, so that any...](#)

Chapter 4

[Figure 4.1 A sketch of a spherical surface, representing positive \$k\$. A triangle is shown, which has...](#)

[Figure 4.2 A sketch of a saddle surface, representing the hyperbolic geometry obtained when \$k\$ is...](#)

Chapter 5

[Figure 5.1 A photon travels a distance \$dr\$ between two galaxies A and B.](#)

[Figure 5.2 A schematic illustration of the evolution of a universe containing radiation and...](#)

[Figure 5.3 Three possible evolutions for the Universe, corresponding to the different signs of \$k\$. The...](#)

Chapter 7

[Figure 7.1 Different models for the Universe can be identified by their location in the plane showing the...](#)

Chapter 8

[Figure 8.1 Predicted ages as fractions of the Hubble time \$H_0^{-1}\$ for open universes and for universes with...](#)

Chapter 9

[Figure 9.1 The rotation curve of the spiral galaxy NGC3198. We see that it remains roughly...](#)

[Figure 9.2 A schematic illustration of a galactic disk, with a few globular clusters, embed...](#)

Chapter 10

[Figure 10.1 The evolution of the blackbody spectrum as the Universe expands. The expansion ...](#)

[Figure 10.2 Because they have traveled toward us uninterrupted since the Universe was at 3000 K, the...](#)

Chapter 11

[Figure 11.1 A schematic illustration of the temperature–time relation, assuming \$\Omega_0 = 0.3\$ and \$h = 0.7\$...](#)

Chapter 12

[Figure 12.1 The predicted abundances of light nuclei, as a function of \$\Omega_B h^2\$ along the top and of the...](#)

Chapter 13

[Figure 13.1 An illustration of the horizon problem. We receive microwave radiation from poi...](#)

[Figure 13.2 Possible evolution of the density parameter \$\Omega_{\text{tot}}\$. There might or might not be a period before inflation,...](#)

[Figure 13.3 A schematic illustration of the inflationary solution to the horizon problem, w...](#)

Chapter 14

[Figure 14.1 The solid line shows the true \(decelerating\) scale factor. The dotted line extr...](#)

List of Tables

Chapter 4

[Table 4.1 A summary of possible geometries.](#)

Chapter 11

[Table 11.1 Different stages of the Universe's evolution \(taking \$\Omega_0 = 0.3\$ and \$h = 0.7\$ \). Some nu...](#)

Chapter 12

[Table 12.1 A comparison of nucleosynthesis and decoupling.](#)

An Introduction to Modern Cosmology

Marina Cortês and Andrew Liddle

Institute for Astrophysics and Space Sciences

Faculty of Sciences

University of Lisbon

Campo Grande

Lisbon, Portugal

Fourth Edition

WILEY

This fourth edition first published 2026

© 2026 John Wiley & Sons Ltd

Edition History

John Wiley & Sons Ltd (1e, 1998); John Wiley & Sons Ltd (2e, 2003); John Wiley & Sons Ltd (3e, 2015)

All rights reserved, including rights for text and data mining and training of artificial intelligence technologies or similar technologies. 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, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

The right of Marina Cortês and Andrew Liddle to be identified as the author(s) of this work has been asserted in accordance with law.

Registered Office(s)

John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA

John Wiley & Sons Ltd, New Era House, 8 Oldlands Way, Bognor Regis, West Sussex, PO22 0NO, UK

For details of our global editorial offices, customer services, and more information about Wiley products, visit us at www.wiley.com.

The manufacturer's authorized representative according to the EU General Product Safety Regulation is Wiley-VCH GmbH, Boschstr. 12, 69469 Weinheim, Germany, e-mail: Product_Safety@wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Trademarks: Wiley and the Wiley logo are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates in the United States and other countries and may not be used without written permission. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc. is not associated with any product or vendor mentioned in this book.

Limit of Liability/Disclaimer of Warranty

While the publisher and the authors have used their best efforts in preparing this work, including a review of the content of the work, neither the publisher nor the authors make any representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

Library of Congress Cataloging-in-Publication Data:

Names: Liddle, Andrew R. author | Cortês, Marina V. author

Title: An introduction to modern cosmology / Marina V. Cortês, Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal, Andrew Liddle, Institute for Astrophysics and Space Sciences, Faculty of Sciences, University of Lisbon, Campo Grande, Lisbon, Portugal.

Description: Fourth edition. | Hoboken, NJ, USA : John Wiley & Sons Ltd, 2026. | Includes index.

Identifiers: LCCN 2025045575 | ISBN 9781394278435 print | ISBN 9781394278503 ePDF | ISBN 9781394278442 ePUB

Subjects: LCSH: Cosmology

Classification: LCC QB981 .L567 2026

LC record available at <https://lcn.loc.gov/2025045575>

Cover Design: Wiley

Cover Images: Courtesy of NASA, © Nihit RPG/Shutterstock, © Outer Space/Shutterstock, © Vadim Sadovski/Shutterstock

To Dawa and Kylo

Preface

The development of cosmology will no doubt be seen as one of the scientific triumphs of the twentieth century. At its beginning, cosmology hardly existed as a scientific discipline. By its end, the Hot Big Bang cosmology stood secure as the accepted description of the universe as a whole. Telescopes such as the Hubble and James Webb Space Telescopes are capable of seeing light from galaxies so distant that the light has been traveling toward us for most of the lifetime of the universe. The cosmic microwave background, a fossil relic of a time when the universe was both denser and hotter, is routinely detected and its properties examined. That our universe is presently expanding is established without doubt, and its material composition is accurately determined.

We are now in an era where understanding of cosmology has shifted from the qualitative to the quantitative. The turn of the millennium saw the establishment of what has come to be known as the Standard Cosmological Model, representing an almost universal consensus among cosmologists as to the best description of our universe. Nevertheless, it is a model with a major surprise – the belief that our universe is presently experiencing accelerated expansion. Add to that ongoing mysteries such as the properties of the so-called dark matter, which is believed to be the dominant form of matter in the universe, and it is clear that we have some way to go before we can say that a full picture of the physics of the universe is in our grasp. The first decades of the twenty-first century have seen dramatic leaps in the precision of cosmological probes, but this accuracy has yet to yield any substantial improvements to our overall understanding.

Such a bold endeavor as cosmology easily captures the imagination, and there has been increasing demand for cosmology to be taught at university in an accessible manner. Traditionally, cosmology was taught as the tail end of a general relativity course, with a derivation of the metric for an expanding universe and a few solutions. Such a course fails to capture the flavor of modern cosmology, which takes classic physical sciences like thermodynamics, atomic physics, and gravitation and apply them on a grand scale.

In fact, introductory modern cosmology can be tackled in a different way, by avoiding general relativity altogether. By a lucky chance and a subtle bit of cheating, the correct equations describing an expanding universe can be obtained from Newtonian gravity. From this basis, one can study all the triumphs of the Hot Big Bang cosmology – the expansion of the Universe, the prediction of its age, the existence of the cosmic microwave background, and the abundances of light elements such as helium and deuterium – and even go on to discuss more speculative ideas such as the inflationary cosmology.

The origin of this book, first published in 1998, was a short lecture course at the University of Sussex, around 20 lectures, taught by Andrew to students in the final year of a bachelor's degree or the penultimate year of a master's degree. The prerequisites are all very standard physics, and the emphasis is aimed at physical intuition rather than mathematical rigor. It has been further developed through his experiences of teaching cosmology at Imperial College London and at the University of Edinburgh, plus much useful feedback from readers.

The structure of the book is a central “spine,” the main chapters from 1 to 15, which provide a self-contained introduction to modern cosmology. In addition, there are six Advanced Topic chapters, each with prerequisites, which can be added to extend the course as desired. Ordinarily, the best time to tackle those Advanced Topics is immediately after their prerequisites have been attained, though they could also be included at any later stage.

Cosmology is an interesting course to teach, as it is not like most of the other subjects taught in undergraduate physics courses. There is no perceived wisdom, built up over a century or more, which provides an unquestionable foundation, as in thermodynamics, electromagnetism, and even quantum mechanics and general relativity. Each successive edition of this book has introduced new discoveries, and lecturers can expect opportunities to crop up during their course to discuss new results which impact on cosmologists' views of the universe.

You can follow our own evolving views by checking out this book's WWW Home Page. There you may find some updates on observations, a detailed solution set for instructors, and also a list of any errors in the book that we are aware of. If you are confident you've found one yourself, and it's not on the list, we'd be very pleased to hear of it.

Marina Cortês and Andrew Liddle
Lisbon
July 2025

Some fundamental constants

Newton's constant	G	$6.672 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ sec}^{-2}$
Speed of light	c	$2.998 \times 10^8 \text{ m sec}^{-1}$ or $3.076 \times 10^{-7} \text{ Mpc yr}^{-1}$
Reduced Planck constant	$\hbar = h/2\pi$	$1.055 \times 10^{-34} \text{ m}^2 \text{ kg sec}^{-1}$
Boltzmann constant	k_B	$1.381 \times 10^{-23} \text{ J K}^{-1}$ or $8.619 \times 10^{-5} \text{ eV K}^{-1}$
Radiation constant	$\alpha = \pi^2 k_B^4 / 15 \hbar^3 c^3$	$7.565 \times 10^{-16} \text{ J m}^{-3} \text{ K}^{-4}$
Electron mass-energy	$m_e c^2$	0.511 MeV
Proton mass-energy	$m_p c^2$	938.3 MeV
Neutron mass-energy	$m_n c^2$	939.6 MeV
Thomson cross-section	σ_e	$6.652 \times 10^{-29} \text{ m}^2$
Free neutron half-life	t_{half}	609 sec

Some conversion factors

$$\begin{aligned}
 1 \text{ pc} &= 3.261 \text{ light years} = 3.086 \times 10^{16} \text{ m} \\
 1 \text{ yr} &= 3.156 \times 10^7 \text{ sec} \\
 1 \text{ eV} &= 1.602 \times 10^{-19} \text{ J} \\
 1 M_{\odot} &= 1.989 \times 10^{30} \text{ kg} \\
 1 \text{ J} &= 1 \text{ kg m}^2 \text{ sec}^{-2} \\
 1 \text{ Hz} &= 1 \text{ sec}^{-1}
 \end{aligned}$$

Commonly used symbols

z	redshift	<i>defined on page</i> 12 , 48
H_0	Hubble constant	13 , 59
r	physical distance	13
v	velocity	13
f	frequency	16
T	temperature	17
k_B	Boltzmann constant	17
ϵ	energy density	19
α	radiation constant	19
G	Newton's gravitational constant	23
ρ	mass density	24
a	scale factor	26
x	comoving distance	26
k	curvature	27
p	pressure	28
	(or occasionally momentum)	15
n	number density	53
H	Hubble parameter	60
h	Hubble constant	60
	(or Planck's constant)	16
Ω_0	present density parameter	61
ρ_c	critical density	61
Ω	density parameter	62
Ω_k	curvature "density parameter"	63
q_0	deceleration parameter	63
Λ	cosmological constant	65
Ω_Λ	cosmological constant density parameter	66
t	time	73
t_0	present age	73
Ω_B	baryon density parameter	80
Y_4	helium abundance	114
d_{lum}	luminosity distance	151
d_{diam}	angular diameter distance	156
$\Delta T/T$, C_ℓ	cosmic microwave background anisotropies	181 , 182

About the Companion Website

This book is accompanied by a companion website for instructors:



<http://www.wiley.com/go/AnIntroductiontoModernCosmology4e>

This website includes:

- Full model solutions

Chapter 1

A Brief History of Cosmological Ideas

The cornerstone of modern cosmology is the belief that the place that we occupy in the Universe is in no way special. This is known as the **cosmological principle**, and it is an idea that is both powerful and simple. It is intriguing, then, that for the bulk of the history of civilization it was believed that we occupy a very special location, usually the center, in the scheme of things.

The ancient Greeks, in a model further developed by the Alexandrian Ptolemy, believed that the Earth must lie at the center of the cosmos. It would be circled by the Moon, the Sun, and the planets, and then the “fixed” stars would be yet further away. A complex combination of circular motions, Ptolemy’s epicycles, was devised in order to explain the motions of the planets, especially the phenomenon of retrograde motion where planets appear to temporarily reverse their direction of motion. It was not until the early 1500s that Copernicus stated forcefully the view, initiated nearly 2000 years before by Aristarchus, that one should regard the Earth, and the other planets, as going around the Sun. By ensuring that the planets moved at different speeds, retrograde motion could easily be explained by this theory. However, although Copernicus is credited with removing the anthropocentric view of the Universe, which placed the Earth at its center, he still believed that the Sun was at the center.

Newton’s theory of gravity put what had been an empirical science (Kepler’s discovery that the planets moved on elliptical orbits) on a solid footing, and it appears that Newton believed that the stars were also suns pretty much like our own, distributed evenly throughout infinite space, in a static configuration. However, it seems that Newton was aware that such a static configuration is unstable.

Over the next 200 years, it became increasingly understood that the nearby stars are not evenly distributed, but rather are located in a disk-shaped assembly, which we now know as the Milky Way galaxy. The Herschels were able to identify the disk structure in the late 1700s, but their observations were not perfect, and they wrongly concluded that the Solar System lay at its center. Only in the early 1900s was this convincingly overturned, by Shapley, who realized that we are some two-thirds of the radius away from the center of the Galaxy. Even then, he apparently still believed our galaxy to be at the center of the Universe. It was only in 1952 that Baade finally demonstrated that the Milky Way is a fairly typical galaxy, leading to the modern view, known as the **cosmological principle** (or sometimes the Copernican principle), that the Universe looks the same wherever you are.

It is important to stress that the cosmological principle isn’t exact. For example, no one thinks that sitting in a lecture theater is exactly the same as relaxing at the beach, and the interior of the Sun is a very different environment from the interstellar regions. Rather, it is an approximation that we believe holds better and better the larger the length scales we consider. Even on the scale of individual galaxies, it is not very good, but once we take very large regions (though still much smaller than the Universe itself), containing say a million galaxies, we expect every such region to look more or less like every other one. The cosmological principle is therefore a property of the global Universe, breaking down if one looks at local phenomena.

The cosmological principle is the basis of Big Bang cosmology. The Big Bang is the best description we have of our Universe, and the aim of this book is to explain why. The Big Bang is a picture of our Universe as an evolving entity, which was very different in the past as compared to the present. Originally, it was forced to compete with a rival idea, the Steady State Universe, which holds that the Universe does not evolve but rather has looked the same forever, with new material being created to fill the gaps as the Universe expands. However, the observations we will describe now support the Big Bang so strongly that the Steady State theory is almost never considered.

Chapter 2

Observational Overview

For most of history, astronomers have had to rely on light in the visible part of the spectrum in order to study the Universe. One of the great astronomical achievements of the twentieth century was the exploitation of the full electromagnetic spectrum for astronomical measurements. We now have instruments capable of making observations of radio waves, microwaves, infrared light, visible light, ultraviolet light, X-rays, and gamma rays, which all correspond to light waves of different (in this case increasing) frequency. We have also entered an epoch where we can go beyond the electromagnetic spectrum and receive information of other types. A remarkable feature of observations of a nearby supernova in 1987 was that it was also seen through the detection of neutrinos, an extraordinarily weakly interacting type of particle normally associated with radioactive decay. Very high-energy cosmic rays, consisting of highly relativistic elementary particles, are now routinely detected, though as yet there is no clear understanding of their astronomical origin. And, most excitingly of all, in 2015, the Advanced Laser Interferometer Gravitational-wave Observatory (LIGO) made the first direct detection of gravitational waves, ripples in space-time itself, which has become a technique to observe astronomical events such as colliding stars and black hole mergers.

The advent of large ground- and satellite-based telescopes operating in all parts of the electromagnetic spectrum has revolutionized our picture of the Universe. While there are probably gaps in our knowledge, some of which may be important for all we know, we do seem to have a consistent picture, based on the cosmological principle, of how material is distributed in the Universe. Our discussion here is brief; for a more detailed discussion of the observed Universe, see some of the more advanced textbooks listed in the Bibliography. A huge array of images can be found online; the ones shown here are just a taster for what you can find for yourself.

2.1 In visible light

Historically, our picture of the Universe was built up through ever more careful observations using visible light.

Stars: The main source of visible light in the Universe is nuclear fusion within stars. The Sun is a fairly typical star, with a mass of about 2×10^{30} kg. This is known as a solar mass, indicated $M_{\odot}$, and is a convenient unit for measuring masses.

The nearest stars to us are a few light years away, a light year being the distance (about 10^{16} m) that light can travel in a year. For historical reasons, an alternative unit, known as the **parsec** and denoted “pc,” ¹ is more commonly used in cosmology. A parsec equals 3.261 light years. In cosmology, one seldom considers individual stars, instead preferring to adopt as the smallest considered unit the conglomerations of stars known as ...

Galaxies: Our Solar System lies some way off-center in a giant disk structure known as the Milky Way galaxy. It contains a staggering hundred thousand million (10^{11}) or so stars, with masses ranging from about a tenth that of our Sun to tens of times larger. It consists of a central bulge, plus a disk of radius 12.5 kiloparsecs (kpc, equal to 10^3 pc), and a thickness of only about 0.3 kpc. We are located in the disk about 8 kpc from the center. The disk rotates slowly (and also differentially, with the outer edges moving more slowly, just as more distant planets in the Solar System orbit more slowly). At our radius, the galaxy rotates over a period of 200 million years. Because we are within it, we can’t get an image of our own galaxy, but it probably looks not unlike the M100 galaxy shown in [Figure 2.1](#).

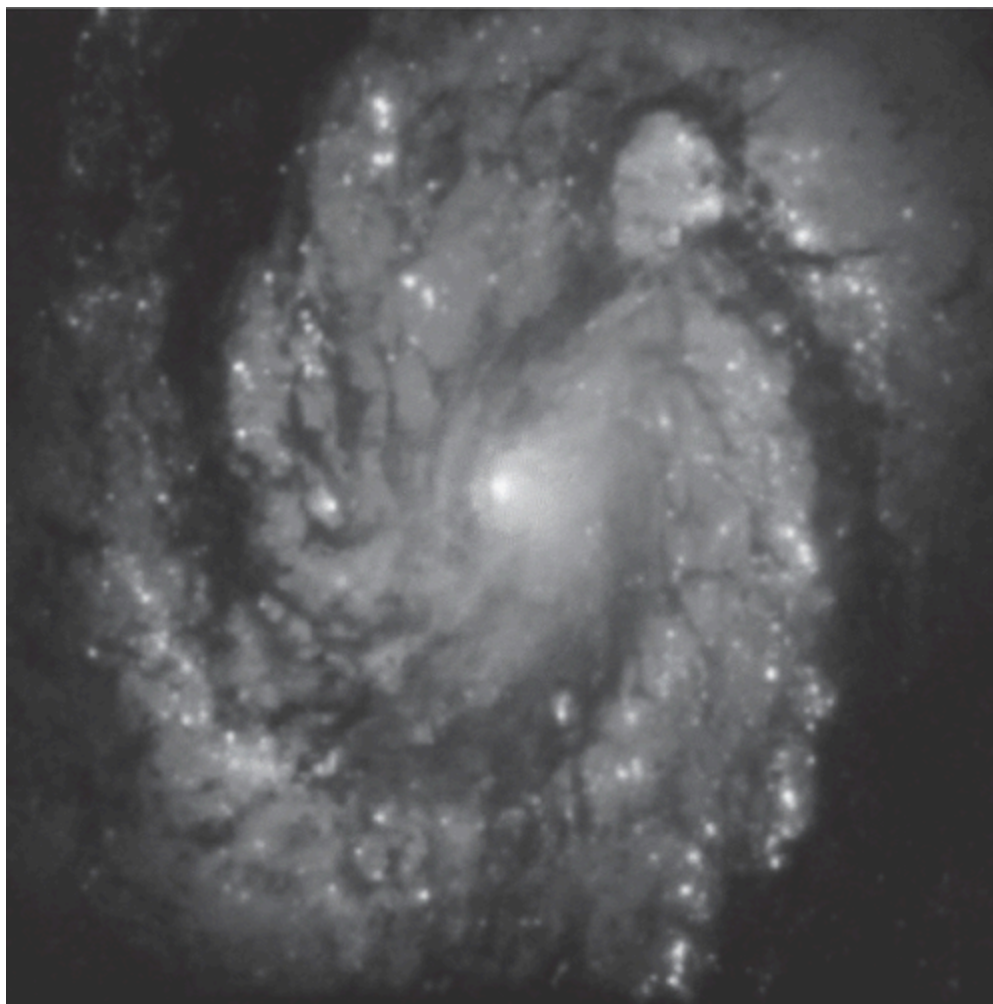


Figure 2.1 If viewed from above the disk, our own Milky Way galaxy would probably resemble the M100 galaxy, imaged here by the Hubble Space telescope. (Courtesy of NASA.)

Our galaxy is surrounded by smaller collections of stars, known as globular clusters. These are distributed more or less symmetrically about the bulge, at distances of 5–30 kpc. Typically, they each contain a million stars, and are thought to be remnants of the formation of the galaxy. As we shall discuss later, it is believed that the entire disk and globular cluster system are embedded in a larger spherical structure known as the Galactic halo.

Galaxies are the most visually striking and beautiful astronomical objects in the Universe, exhibiting a wide range of properties. However, in cosmology, the detailed structure of a galaxy is usually irrelevant, and galaxies are normally thought of as point-like objects emitting light, often broken into subclasses according to colors, luminosities, and morphologies.

The Local Group: Our galaxy resides within a small concentrated group of galaxies known as the local group. The nearest galaxy is a small irregular galaxy known as the Large Magellanic Cloud (LMC), which is 50 kpc away from the Sun. The nearest galaxy of similar size to our own is the Andromeda Galaxy, at a distance of 770 kpc. The Milky Way is one of the largest galaxies in the local group. A typical galaxy group occupies a volume of a few cubic **megaparsecs**. The megaparsec, denoted Mpc and equal to a million parsecs, is the cosmologist's favorite unit for measuring distances, because it is roughly the separation between neighboring galaxies. It equals 3.086×10^{22} m.

Clusters of Galaxies, Superclusters, and Voids: Surveying larger regions of the Universe, say on a scale of hundreds of Mpc, one sees a variety of large-scale structures, as shown in [Figure 2.2](#). This figure is not a photograph, but rather a carefully constructed map of the nearby region of our Universe, on a scale of about $1:10^{27}$! In some places, galaxies are clearly grouped into clusters of galaxies.

 Circular map of the Universe details galaxy positions forming filamentary structure across angular slices, labeled with redshift and hours.

Figure 2.2 A map of galaxy positions in a narrow slice of the Universe, as measured by the Sloan Digital Sky Survey. Our galaxy is located at the center, and the survey radius is around 600 Mpc. The galaxy positions were obtained by measurement of the shift of spectral lines, as described in [Section 2.4](#). (Figure courtesy M. Blanton and the Sloan Digital Sky Survey, www.sdss3.org.)

A famous example of a cluster of galaxies is the Coma cluster, which is about 100 Mpc away from our own Galaxy. The upper panel of [Figure 2.3](#) shows a combined optical/infrared image of Coma; although the image resembles a star field, almost every source is a distinct galaxy (the main exception being two bright stars in the upper right quadrant). Coma contains perhaps 10 000 galaxies, mostly too faint to show in this image, orbiting in their common gravitational field.

 Two panels detail infrared light and a labeled extended X-ray of the Coma cluster, with a scale of 10 arcmin.

Figure 2.3 Images of the Coma cluster of galaxies in visible/infrared light (top) and in X-rays (bottom), the latter being on a larger angular scale. (Courtesy of NASA/Spitzer satellite and ESA/U. Briel/MPE Garching/XMM-Newton satellite.)

However, most galaxies, sometimes called field galaxies, are not part of a cluster. Galaxy clusters are the largest gravitationally collapsed objects in the Universe, and they themselves are grouped into superclusters, joined by filaments and walls of galaxies. In between this “foamlike” structure lie large voids, some as large as 50 Mpc across. Structures in the Universe will be further described in [Advanced Topic 5](#). [Figure 2.4](#) shows an example of a computer simulation aiming to model the distribution of material within the Universe.

 Simulated large scale structure of the universe details a dense web of filaments and nodes with a scale bar for 125 Mpc per hour.

Figure 2.4 A computer simulation showing the predicted distribution of matter in the Universe on large scales. (Courtesy of V. Springel and the Virgo Consortium.)

Large-scale Smoothness: Only once we get to scales of hundreds of megaparsecs or more does the Universe begin to appear smooth, as revealed by extremely large galaxy surveys such as the 2dF galaxy redshift survey and the Sloan Digital Sky Survey. Such surveys do not find any huge structures on scales greater than those described above; the galaxy superclusters and voids are believed to be the biggest structures in the present Universe. The observation that the Universe does indeed become smooth on the largest scales, the cosmological principle, is the underpinning of modern cosmology. It is interesting that, while the smoothness of the matter distribution on large scales has been a key assumption of cosmology for decades, it is only fairly recently that a convincing observational demonstration has become possible.

2.2 In other wavebands

Observations using visible light provide us with a good picture of what’s going on in the present-day Universe. However, many other wavebands make vital contributions to our understanding, and in particular our best knowledge of cosmology comes not from visible light but from microwaves.

Microwaves: For cosmology, this is by far the most important waveband. Penzias & Wilson’s accidental discovery in 1965 that the Earth is bathed in microwave radiation, with a blackbody spectrum at a temperature of around 3 K, was and is one of the most powerful pieces of information in support of the Big Bang theory, around which cosmology is based. This is now known as the cosmic microwave background or simply CMB. Observations by the Far InfraRed Absolute Spectrometer (FIRAS) experiment on board the COsmic Background Explorer (COBE) satellite have confirmed that the radiation is extremely close to the blackbody form at a temperature 2.725 ± 0.001 K. This data is shown in [Figure 2.5](#). Furthermore, the temperature coming from different parts of the sky is astonishingly uniform, and this presents the best evidence that we can use the cosmological principle as the foundation of cosmology. Even more importantly, it is possible to identify tiny variations, only one part in a hundred thousand, between the intensities of the microwaves coming from different directions. These were first discovered by the COBE satellite in 1992, and [Figure 2.6](#) shows precision measurements published by the *Planck* Satellite collaboration in 2018. It is believed that these patterns, known as cosmic microwave background anisotropies, are intimately related to the origin of structure in the Universe, and at present such measurements give us by far the most accurate understanding of cosmology. This topic is explored further in [Advanced Topics 5](#) and [6](#).



Graph of Intensity versus Waves per centimeter, plotting cosmic microwave background curve data points with error bars multiplied by 400.

Figure 2.5 The cosmic microwave background spectrum as measured by the FIRAS experiment on the COBE satellite. The error bars are so small that they have been multiplied by 400 to make them visible on this plot, and the best-fit blackbody spectrum at **2.725 K**, shown by the line, is an excellent fit.



Map of cosmic microwave temperature variations across an oval projection scaled from negative 300 to 300 microkelvin.

Figure 2.6 The cosmic microwave background temperature variations on the sky as measured by the *Planck* Satellite, with scale in microKelvin. (Courtesy of ESA/Planck collaboration.)

Infrared: Carrying out surveys in the infrared part of the spectrum, as was first done by the highly successful InfraRed Astronomical Satellite (IRAS) in the 1980s, is an excellent way of spotting young galaxies, in which star formation is at an early stage. Infrared surveys pick up a somewhat different population of galaxies to surveys carried out in optical light, though obviously the brightest galaxies are seen in both. Infrared is particularly good for looking through the dust in our own Galaxy to see distant objects, as it is absorbed and scattered much less strongly than visible radiation. Accordingly, it is best for studying the region close to our Galactic plane, where obscuration by dust is strongest. Infrared is also vital for studying the most distant galaxies, whose emission has been shifted predominantly into the infrared part of the spectrum by the expansion of the Universe after the light was emitted; NASA's James Webb Space Telescope is the current state-of-the-art for such studies. The far infrared can probe emission from interstellar dust grains to learn about dust-enshrouded environments that are opaque at visible and near-infrared wavelengths.

X-rays: These are a vital probe of clusters of galaxies; in between the galaxies lies gas so hot that it emits in the X-ray part of the spectrum, corresponding to temperatures of tens of millions of Kelvin. This gas is thought to be a remnant material from the formation of the galaxies, which failed to collapse to form stars. X-ray emission from the Coma galaxy cluster is shown in the lower panel of [Figure 2.3](#). The individual galaxies seen in the visible light image in the upper panel are almost all invisible in X-rays, with the bright diffuse X-ray emission from the hot gas dominating the image.

Radio Waves: A powerful way of gaining high-resolution maps of distant galaxies is by mapping in the radio part of the spectrum. An important probe of cosmology is observations of the "21 cm" emission line, due to spin-flip of the electron in a hydrogen atom, enabling maps of the distribution of neutral hydrogen gas in the distant Universe. This can probe both the early development of structures in the Universe, and the transition of the Universe to its present state of high ionization of the intergalactic medium. The Square Kilometre Array, nearing completion as we write, will take these studies to the next level.

2.3 Homogeneity and isotropy

The evidence that the Universe becomes smooth on a large scale supports the use of the cosmological principle. It is therefore believed that our large-scale Universe possesses two important properties: **homogeneity** and **isotropy**. Homogeneity is the statement that the Universe looks the same at each point, while isotropy states that the Universe looks the same in all directions.

These do not automatically imply one another. For example, a universe with a uniform magnetic field is homogeneous, as all points are the same, but it fails to be isotropic because directions along the field lines can be distinguished from those perpendicular to them. Alternatively, a spherically symmetric distribution, viewed from its central point, is isotropic but not necessarily homogeneous. However, if we require that a distribution is isotropic at *every* point, then that does enforce homogeneity as well.

As mentioned earlier, the cosmological principle is not exact, and so our Universe does not respect exact homogeneity and isotropy. Indeed, the study of departures from homogeneity is currently the most prominent research topic in cosmology. We'll introduce this in [Advanced Topic 5](#), but in the main body of this book we are concerned only with the behavior of the Universe as a whole, and so will be assuming large-scale homogeneity and isotropy.

2.4 The expansion of the Universe

A key piece of observational evidence in cosmology is that almost everything in the Universe appears to be moving away from us, and the further away something is, the more rapid its recession appears to be. These velocities are measured via the **redshift**, which is basically the Doppler effect applied to light waves. Galaxies have a set of absorption and emission lines identifiable in their spectra, whose characteristic frequencies are well known. However, if a galaxy is moving toward us, the light waves get crowded together, raising the frequency. Because blue light is at the high-frequency end of the visible spectrum, this is known as a blueshift. If the galaxy is receding, the characteristic lines move toward the red end of the spectrum and the effect is known as a redshift. This technique was first used to measure a galaxy's velocity by Vesto Slipher around 1912, and was applied systematically by one of the most famous cosmologists, Edwin Hubble, in the following decades.

It turns out that almost all galaxies are receding from us, so the standard terminology is redshift z , defined by

$$z = \frac{\lambda_{\text{obs}} - \lambda_{\text{em}}}{\lambda_{\text{em}}} \quad (2.1)$$

where λ_{em} and λ_{obs} are the wavelengths of light at the points of emission (the galaxy) and observation (us). If a nearby object is receding at a speed v , then its redshift is

$$z = \frac{v}{c} \quad (2.2)$$

where c is the speed of light. ² Figure 2.7 shows velocity against distance, a plot known as the Hubble diagram, for a sample of 1355 galaxies.

Scatter plot of galaxy velocity versus estimated distance, plotted in Mpc, details a scattered and linear trend implying Hubble's law.

Figure 2.7 A plot of velocity versus estimated distance for a set of 1355 galaxies. A straight-line relation implies Hubble's law. The considerable scatter is due to observational uncertainties and random galaxy motions, but the best-fit line accurately gives Hubble's law. (The bold italic x -axis scale assumes a particular value of bold italic upper H_0 .)

Hubble realized that his observations, which were of course much less extensive than those available to us now, showed that the velocity of recession was proportional to the distance of an object from us:

$$v = H_0 r \quad (2.3)$$

This is known as **Hubble's law**, and the constant of proportionality H_0 is known as **Hubble's constant**. Hubble's law isn't exact, as the cosmological principle doesn't hold perfectly for nearby galaxies, which typically possess some random motions known as peculiar velocities. But it does describe the average behavior of galaxies extremely well. Hubble's law gives the picture of our Universe illustrated in Figure 2.8, where the nearby galaxies have the smallest velocity relative to ours. Over the years, many attempts have been made to determine accurate values for the proportionality constant, which is now known to within a few percent accuracy. Still, as we will see in Chapter 6 and Advanced Topic 6.5, there remains some controversy to this day.

Diagram of a central point and six surrounding points, each with an arrow details speed proportional to distance from the center.

Figure 2.8 According to Hubble's law, the further away from us a galaxy is, the faster it is receding.

At first sight, it seems that the cosmological principle must be violated if we observe everything to be moving away from us, since that apparently places us at the center of the Universe. However, nothing could be further from the truth. In fact, *every* observer sees all objects rushing away from them with velocity proportional to distance. It is perhaps easiest to convince yourself of this by setting up a square grid with recession velocity proportional to distance from the central grid point. Then, transform the frame of reference to a nearby grid-point, and you'll find that the Hubble law still holds about the new "centre." This only works because of the linear relationship between velocity and distance; any other law and it wouldn't work.

So, although expanding, the Universe looks just the same whichever galaxy we choose to imagine ourselves within. A common analogy is to imagine baking a cake with raisins in it, or blowing up a balloon with dots on its surface. As the cake rises (or the balloon is inflated), the raisins (or dots) move apart. From each one, it seems that all the others are receding, and the further away they are, the faster that recession is.

Because everything is flying away from everything else, we conclude that in the distant past everything in the Universe was much closer together. Indeed, trace the history back far enough and everything comes together. The initial explosion is known as the **Big Bang**, and a model of the evolution of the Universe from such a beginning is known as the **Big Bang Cosmology**. Later on, we will find out why it is commonly called the **Hot Big Bang**.

2.5 Particles in the Universe

2.5.1 What particles are there?

Everything in the Universe is made up of fundamental particles, and the behavior of the Universe as a whole depends on the properties of these particles.

One crucial question is whether a particle is moving relativistically or not. Any particle has two contributions to its energy, one being kinetic energy and the other being mass-energy, which combine to give

$$E_{\text{total}} = mc^2 + p^2 c^2 \quad (2.4)$$

where m is the particle rest mass and p the particle momentum. If the mass-energy dominates, the particle will be moving at much less than the speed of light, and we say it is nonrelativistic. In that limit, we can carry out an expansion

$$E_{\text{total}} = mc^2 \left(1 + \frac{p^2}{2m^2 c^2} \right) \approx mc^2 + \frac{p^2}{2m} \quad (2.5)$$

We recognize the first term as Einstein's famous $E = mc^2$, known as the rest mass-energy as it is the energy of the particle when it is stationary. The second term is the usual kinetic energy ($p = mv$ in the nonrelativistic limit). If the mass-energy does not dominate, the particle will be moving at a substantial fraction of the speed of light and so is relativistic. In particular, any particle with zero rest mass is always relativistic and moves at the speed of light, the simplest example being light itself.

Let's review the nature of the particles which are believed to exist in our Universe.

Baryons

We ourselves are built from atoms, the bulk of whose mass is attributable to the protons and neutrons in the atomic nuclei. Protons and neutrons are believed to be made up of more fundamental particles known as quarks, a proton being made of two up quarks and a down quark, while a neutron is an up and two downs. A general term for particles made up of three quarks is **baryons**. Of all the possible baryons, only the proton and neutron can be stable,³ and so these are thought to be the only types of baryonic particle significantly represented in the Universe. Yet another piece of terminology, nucleon, refers to just protons and neutrons, but I'll follow the standard practice of using the term baryon. In particle physics units, the mass-energies of a proton and a neutron are 938.3 and 939.6 MeV, respectively, where "MeV" is a Mega-electron volt, a unit of energy equal to a million electron volts (eV) and rather more convenient than a Joule.

Although electrons are certainly not made from quarks, they are traditionally also included under the title baryon by cosmologists (to the annoyance of particle physicists). A crucial property of the Universe is that it is charge neutral, so there must be one electron for every proton. Weighing in at a puny 0.511 MeV, well under a thousandth of a proton mass, the contribution of electrons to the total mass is a tiny fraction, not meriting separate discussion.

In the present Universe, baryons are typically moving nonrelativistically, meaning that their kinetic energy is much less than their mass-energy.

Radiation

Our visual perception of the Universe comes from electromagnetic radiation, and such radiation, at a large variety of frequencies, pervades the Universe. In the quantum mechanical view of light, it can be thought of as made up of individual particles – like packets of energy – known as **photons** and usually indicated by the symbol γ . Photons propagate, naturally enough, at the speed of light; since they have zero rest mass, their total energy is always given by their kinetic energy, and is related to their frequency f by

$$E = hf \quad (2.6)$$

where h is Planck's constant.

Photons can interact with the baryons and electrons; for example, a high-energy photon can knock an electron out of an atom (a process known as ionization), or can scatter off a free electron (known as Thomson scattering in the non-relativistic case $hf \ll mc^2$, otherwise Compton scattering). The more energetic the photons are, the more devastating their effects on other particles.

Neutrinos

Neutrinos are extremely weakly interacting particles, produced, for example, in radioactive decay. There is now significant experimental evidence that they possess a non-zero rest mass, but so far there has been no clear evidence that this mass is large enough to have observable cosmological effects. Hence, for now, it remains a working assumption in cosmology to treat them as effectively massless. We will adopt that assumption for the main body of this book, and in that case they, like photons, are always relativistic. The combination of photons and neutrinos makes up the relativistic material in our Universe. Confusingly, sometimes the term "radiation" is used to refer to all the relativistic material.

There are three types of neutrino, the electron neutrino, muon neutrino, and tau neutrino, and if they are indeed all massless, they should all exist in our Universe. Unfortunately, their interactions are so weak that for now there is no hope of detecting cosmological neutrinos directly. Originally, their presence was inferred on purely theoretical grounds, though we will see that the existence of the cosmic neutrino background may be inferred indirectly by some cosmological observations.

Because they are so weakly interacting, the experimental limits on the neutrino masses, especially of the latter two types, are quite weak, and it is in fact perfectly possible that they are massive enough to be nonrelativistic. The possible effects of neutrino masses are explored in [Advanced Topic 3](#). Cosmological effects of neutrino masses might well have been spotted by the time you read this book.

Dark matter

In this book, we'll encounter one further kind of particle that may exist in our Universe, which is not part of the Standard Model of particle theory. It is known as dark matter, and its properties are highly uncertain and a matter of constant debate among cosmologists. We'll return to it in [Chapter 9](#).

2.5.2 Thermal distributions and the blackbody spectrum

We end this section with some discussion of the physics of radiation. If this is unfamiliar to you, the details aren't all that crucial, though some of the final results will be used later in the book.

If particles are frequently interacting with one another, then the distribution of their energies can be described by equilibrium thermodynamics. In a thermal distribution, interactions are frequent, but a balance has been reached so that all interactions proceed equally frequently in both the forward and backward directions, so that the overall distribution of particle numbers and energies remains fixed. The number of particles of a given energy then depends only on the temperature.

The precise distribution depends on whether the particles considered are fermions, which obey the Pauli exclusion principle, or bosons, which do not. In this book, the most interesting case is that of photons, which are bosons, and their characteristic distribution at temperature T is the **Planck** or **blackbody** spectrum. Photons have two possible polarizations, and each has an occupation number per mode N given by the Planck function

$$N = \frac{1}{\exp\left(\frac{hf}{k_B T}\right) - 1} \quad (2.7)$$

where h is Planck's constant and k_B is one of the fundamental constants of nature, the **Boltzmann constant**, whose value is $1.381 \times 10^{-23} \text{ J K}^{-1} = 8.619 \times 10^{-5} \text{ eV K}^{-1}$.

To interpret this equation, remember that hf is the photon energy. The purpose of the Boltzmann constant is to convert temperature into a characteristic energy $k_B T$. Below this characteristic energy, $hf \ll k_B T$, it is easy to make photons and the occupation number is large (as photons are bosons, the Pauli exclusion principle doesn't apply and there may be arbitrarily many photons in a given mode). Above the characteristic energy, $hf \gg k_B T$, it is energetically unfavorable to make photons, and the number is exponentially suppressed, as shown in [Figure 2.9](#).

 Graph of Planck function N versus hf/k_B details a sharp decay, reflecting far more low-energy photons than high-energy.

Figure 2.9 The Planck function of [Equation \(2.7\)](#). There are far more photons with very low energy than very high energy.

More interesting than the number of photons in a mode is the distribution of energy among the modes. We focus on the energy per unit volume, known as the **energy density** ϵ . Because there are very few photons with

$hf \gg k_B T$ there isn't much energy at high frequencies. But, despite their large numbers, there also isn't much total energy at low frequencies

$hf \ll k_B T$, both because those photons have less energy each $\epsilon = hf N$, and because their wavelength is longer and so each photon occupies a greater volume. With quite a bit of work, the energy density in a frequency interval df about f can be shown to be

$$\epsilon(f) df = \frac{8\pi h}{c^3} \frac{f^3 df}{\exp\left(\frac{hf}{k_B T}\right) - 1} \quad (2.8)$$

which tells us how energy is distributed among the different frequencies. We see in [Figure 2.10](#) that the peak of the distribution is at $f_{\text{peak}} \approx 2.8 k_B T / h$,

corresponding to an energy

$\epsilon_{\text{peak}} = hf_{\text{peak}}$ asymptotically equals $2.8 k_B T$. That is to say, the total energy in the radiation is dominated by photons with energies of order

$k_B T$. Indeed, the mean energy of a photon in this distribution is

$\epsilon_{\text{mean}} = 3 k_B T$.



Graph of the normalized density versus $hf/k_B T$ details the energy density distribution of a spectrum.

Figure 2.10 The energy density distribution of a blackbody spectrum, given by Equation (2.8). Most of the energy is contributed by photons of energy

hf_{peak} .

When we study the early history of the Universe, an important question will be how this typical energy compares to atomic and nuclear binding energies.

A further quantity of interest will be the total energy density of the blackbody radiation, obtained by integrating Equation (2.8) over all frequencies. Setting $y = hf/k_B T$ quickly leads to

$$\epsilon_{\text{rad}} = \frac{8\pi k_B^4}{15} T^4 \int_0^\infty \frac{y^3 dy}{e^y - 1} \quad (2.9)$$

The integral is not particularly easy to compute, but you might like to try it as a challenge. The answer is

$\pi^4/15$, giving a radiation energy density

$$\epsilon_{\text{rad}} = \alpha T^4 \quad (2.10)$$

where the radiation constant α is defined as

$$\alpha = \frac{\pi^5 k_B^4}{15 h^3} = 7.565 \times 10^{-16} \text{ J m}^{-3} \text{ K}^{-4} \quad (2.11)$$

Here, h with stroke equals $h/2\pi$ is the reduced Planck constant.

Problems

2.1. Suppose that the Milky Way galaxy is a typical size, containing say 10^{11} stars, and that galaxies are typically separated by a distance of one megaparsec. Estimate the density of the Universe in SI units. How does this compare with the density of the Earth?

$1 \text{ Mpc} \approx 3.1 \times 10^{22} \text{ m}$

2.2. In the real Universe, the expansion is not completely uniform. Rather, galaxies exhibit some random motion relative to the overall Hubble expansion, known as their *peculiar velocity* and caused by the gravitational pull of their near neighbors. Supposing that a typical (e.g. root mean square) galaxy's peculiar velocity is 600 km s^{-1} , how far away would a galaxy have to be before it could be used to determine the Hubble constant to 10% accuracy, supposing

a. the true value of the Hubble constant is $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$?

b. the true value of the Hubble constant is $50 \text{ km s}^{-1} \text{ Mpc}^{-1}$?

Assume in your calculation that the galaxy distance and redshift could be measured exactly. Unfortunately, that is not true of real observations.

2.3. What evidence can you think of to support the assertion that the Universe is charge neutral, and hence contains an equal number of protons and electrons?

2.4. The binding energy of the electron in a hydrogen atom is 13.6 eV. What is the frequency of a photon with this energy? At what temperature does the mean photon energy equal this energy?

2.5. The peak of the energy density distribution of a blackbody at

$\epsilon_{\text{peak}} = 2.8 k_B T$ divided by h implies that $hf_{\text{peak}}/k_B T$ is a constant. Evaluate this constant in SI units (see page xiv for

useful numbers). The Sun radiates approximately like a blackbody with $T_{\text{Sun}} \approx 5800 \text{ K}$. Compute λ_{peak} for solar radiation. Where in the electromagnetic spectrum does the peak emission lie?

2.6. The cosmic microwave background has a blackbody spectrum at a temperature of 2.725 K . Repeat [Problem 2.5](#) to find the peak frequency of its emission, and also find the corresponding wavelength and compare to [Figure 2.5](#). Confirm that the peak emission lies in the microwave part of the electromagnetic spectrum. Finally, compute the total energy density of the microwave background.

Notes

¹ A parsec is defined as the distance at which the mean distance between the Earth and Sun subtends a second of arc. The mean Earth–Sun distance (called an Astronomical Unit) is $1.496 \times 10^{11} \text{ m}$, and dividing that by $\tan(1 \text{ arcsec})$ gives $1 \text{ pc} = 3.086 \times 10^{16} \text{ m}$.

² This formula ignores special relativity and so is valid only for speeds v much less than c . If you're interested, the special relativity result, of which this is an expansion for small v/c , is

$$1 + z = \sqrt{\frac{1 + v/c}{1 - v/c}}$$

However, for distant objects in cosmology, there are further considerations concerning the propagation time of the light and how the relative velocity might change during it, and so this expression should not be used.

³ The proton's lifetime is known to be either infinite, corresponding to the proton being absolutely stable, or much longer than the age of the Universe so that they are effectively stable. Isolated neutrons are unstable (decaying into a proton, an electron, and an anti-neutrino), but those bound in nuclei may be stable; this will prove crucial in [Chapter 12](#).

Chapter 3

Newtonian Gravity

It is perfectly possible to discuss cosmology without having already learned general relativity. In fact, the most crucial equation, the Friedmann equation, which describes the expansion of the Universe, turns out to be the same when derived from Newton's theory of gravity as it is when derived from the equations of general relativity. The Newtonian derivation is, however, some way from being completely rigorous, and general relativity is required to fully patch it up, a detail that need not concern us at this stage. Readers with some familiarity with general relativity may wish to study the relativistic derivation given in [Advanced Topic 1](#).

In Newtonian gravity, all matter attracts, with the force exerted by an object of mass M on one of mass m given by the famous relationship

$$F = \frac{GMm}{r^2}, \quad (3.1)$$

where r is the distance between the objects and G is Newton's gravitational constant. That is, gravity obeys an inverse square law. Because a force on an object induces an acceleration which is also proportional to its mass, via $F = ma$, the acceleration an object feels under gravity is *independent* of its mass.

The force exerted means there is a gravitational potential energy

$$V = -\frac{GMm}{r}, \quad (3.2)$$

with the force being in the direction which decreases the potential energy the fastest. Like the electric potential of two opposite charges, the gravitational potential is negative, favoring the two objects being close together. But with gravity there is no analog of the repulsion of like charges. Gravity always attracts.

The derivation of the Friedmann equation requires a famous result due originally to Newton, which we won't attempt to prove here. This result states that in a spherically symmetric distribution of matter, a particle feels no force at all from the material at greater radii, and the material at smaller radii gives exactly the force which one would get if all the material was concentrated at the central point. This property arises from the inverse square law; the same results exist for electromagnetism. One example of its use is that the gravitational (or electromagnetic) force outside a spherical object of unknown density profile depends only on the total mass (charge). Another is that an "astronaut" inside a spherical shell feels no gravitational force, not only if they are at the center but if they are at any position inside the shell.

3.1 The Friedmann equation

The Friedmann equation describes the expansion of the Universe, and is therefore the most important equation in cosmology. One of the routine tasks of a working cosmologist is solving this equation under different assumptions concerning the material content of the Universe. To derive the Friedmann equation, we need to compute the gravitational potential energy and the kinetic energy of a test particle (it doesn't matter which one, since everywhere in the Universe is the same according to the cosmological principle), and then use energy conservation.

Let's consider an observer in a uniform expanding medium, with mass density ρ , the mass density being the mass per unit volume. We begin by realizing that because the Universe looks the same from anywhere, we can consider any point to be its center. Now, consider a particle a distance r away with mass m , as shown in [Figure 3.1](#). (By "particle," we really mean a small volume containing the mass m .) Due to Newton's theorem, this particle only feels a force from the material at smaller radii, shown as the shaded region. This material has total mass given by $M = 4\pi\rho r^3/3$, contributing a force

$$F = \frac{GMm}{r^2} = \frac{4\pi G\rho r m}{3}, \quad (3.3)$$

and our particle has a gravitational potential energy

$$V = -\frac{GMm}{r} = -\frac{4\pi G\rho r^2 m}{3}. \quad (3.4)$$

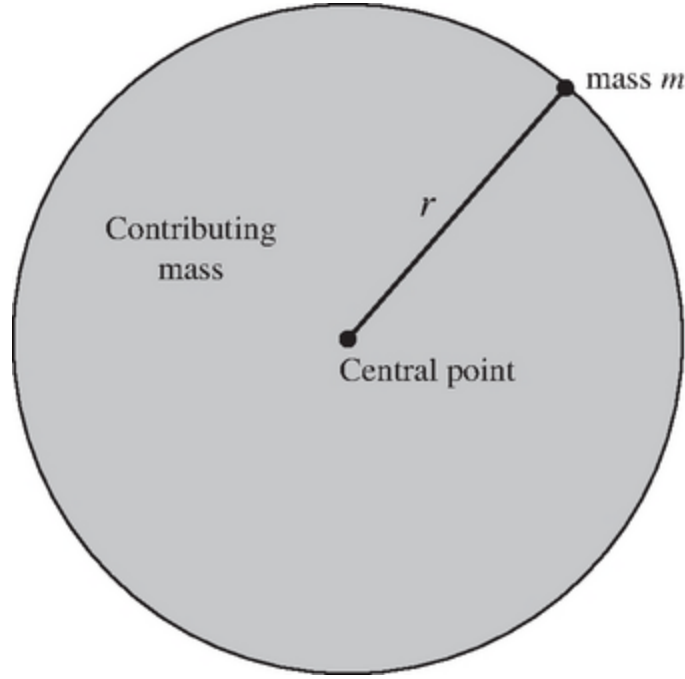


Figure 3.1 The particle at radius r only feels gravitational attraction from the shaded region. Any gravitational attraction from the material outside cancels out, according to Newton's theorem.

The kinetic energy is easy; the velocity of the particle is $\dot{r}$ (we'll always use dots to mean time derivatives) giving

$$T = \frac{1}{2} m \dot{r}^2. \quad (3.5)$$

The equation describing how the separation r changes can now be derived from energy conservation for that particle, namely

$$U = T + V, \quad (3.6)$$

where U is a constant. Note that U need not be the same constant for particles separated by different distances. Substituting gives

$$U = \frac{1}{2} m \dot{r}^2 - \frac{4\pi}{3} G\rho r^2 m. \quad (3.7)$$

This equation gives the evolution of the separation r between the two particles.

We now take a crucial step in this derivation, which is to realize that this argument applies to any two particles, because the Universe is homogeneous. This allows us to change to a different coordinate system, known as **comoving coordinates**. These are coordinates which are carried along with the expansion. Because the expansion is uniform, the relationship between real distance $\vec{r}$ and the comoving distance, which we can call $\vec{x}$, can be written

$$\vec{r} = a(t) \vec{x}, \quad (3.8)$$

where the homogeneity property has been used to ensure that a is a function of time alone. Note that these distances have been written as vector distances. What you should think of when studying this equation is a coordinate grid which expands with time, as shown in [Figure 3.2](#). The galaxies remain at fixed locations in the $\vec{x}$ coordinate system. The original $\vec{r}$ coordinate system, which does not expand, is usually known as **physical coordinates**.

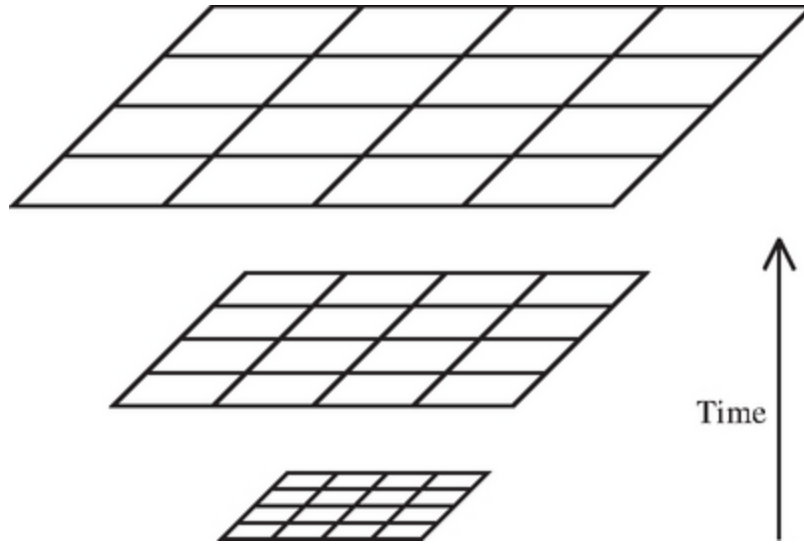


Figure 3.2 The comoving coordinate system is carried along with the expansion, so that any objects remain at fixed coordinate values.

The quantity $a(t)$ is a crucial one, and is known as the **scale factor of the Universe**. It measures the universal expansion rate. It is a function of time alone, and it tells us how physical separations are growing with time, since the coordinate distances $\vec{x}$ are by definition fixed. For example, if, between times t_1 and t_2 , the scale factor doubles in value, that tells us that the Universe has expanded in size by a factor two, and it will take us twice as long to get from one galaxy to another.

We can use the scale factor to rewrite [Equation \(3.7\)](#) for the expansion. Substituting [Equation \(3.8\)](#) into it, remembering $\dot{x} = 0$ by definition as objects are fixed in comoving coordinates, gives

$$U = \frac{1}{2} m \dot{a}^2 x^2 - \frac{4\pi}{3} G \rho a^2 x^2 m. \quad (3.9)$$

Multiplying each side by $2/ma^2 x^2$ and rearranging the terms then gives

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3} \rho - \frac{kc^2}{a^2}, \quad (3.10)$$

where $kc^2 = -2U/mx^2$. This is the standard form of the **Friedmann equation**, and it will appear frequently throughout this book. In this expression, k must be independent of x since all the other terms in the equation are, otherwise homogeneity will not be maintained. So, in fact, we learn that homogeneity requires that the quantity U , while constant for a given particle, does indeed change if we look at different separations x , with $U \propto x^2$.

Finally, since $k = -2U/mc^2 x^2$ which is time independent (as the total energy U is conserved, and the comoving separation x is fixed), we learn that k is just a constant, unchanging with either space or time. It has the units of $[\text{length}]^{-2}$. An expanding Universe has a unique value of k , which it retains throughout its evolution. In [Chapter 4](#), we will see that k tells us about the geometry of the Universe, and it is often called the curvature.

3.2 On the meaning of the expansion

So what does the expansion of the Universe mean? Well, let's start with what it does *not* mean. It does not mean that your body is gradually going to get bigger with time (and certainly isn't an excuse if it does). That does not mean that the Earth's orbit is going to get further from the Sun. It doesn't even mean that the stars within our galaxy are going to become more widely separated with time.

But it does mean that distant galaxies are getting further apart.

The distinction lies in whether the motion of objects is governed by the cumulative gravitational effect of a homogeneous distribution of matter between them, as shown in [Figure 3.1](#). The atoms in your body are not: their separation is dictated by the strength of chemical bonds, with gravity playing no significant role. So, molecular structures will not be affected by the expansion. Likewise, the Earth's motion in its orbit is completely dominated by the gravitational attraction of the Sun (with a minor contribution from the other planets). And even the stars in our galaxy are orbiting in the common gravitational potential well that they themselves create, and are not moving apart relative to one another. The common feature of these environments

is that they are the ones of considerable excess density, very different from the smooth distribution of matter we used to derive the Friedmann equation.

But if we go to large enough scales, in practice tens of megaparsecs, the Universe does become effectively homogeneous and isotropic, with the galaxies flying apart from one another in accordance with the Friedmann equation. It is on these large scales that the expansion of the Universe is felt, and on which the cosmological principle applies.

3.3 Things that go faster than light

A common question that concerns people is whether faraway galaxies are receding from us faster than the speed of light. That is to say, if velocity is proportional to distance, then if we consider galaxies far enough away, can we not make the velocity as large as we like, in violation of special relativity?

The answer is that, indeed, in our theoretical predictions, distant objects can *appear* to move away faster than the speed of light. However, it is space itself which is expanding. There is no violation of causality, because no signal can be sent between such galaxies. Further, special relativity is not violated, because it refers to the relative speeds of objects passing each other, and cannot be used to compare the relative speeds of distant objects.

One way to think of this is to imagine a colony of ants on a balloon. Suppose that the fastest the ants can move is a centimeter per second. If any two ants happen to pass each other, their fastest relative speed would be two centimeters per second, if they happened to be moving in opposite directions. Now, start to blow the balloon up. Although the ants wandering around the surface still cannot exceed one centimeter per second, the balloon is now expanding under them, and ants that are far apart on the balloon could easily be moving apart at faster than two centimeters per second if the balloon is blown up fast enough. But if they are, they will never get to tell each other about it, because the balloon is pulling them apart faster than they can move together, even at full speed. Any ants that start close enough to be able to pass one another must do so at no more than two centimeters per second, even if the Universe is expanding.

The expansion of space is just like that of the balloon, and pulls the galaxies along with it.

3.4 The fluid equation

Fundamental though it is, the Friedmann equation is of no use without an equation to describe how the density ρ of material in the Universe is evolving with time. This involves the pressure p of the material, and is called the fluid equation. (Unfortunately, the standard symbol p for pressure is the same as for momentum, which we've already used. Almost always in this book, p will be pressure.) As we'll shortly see, the different types of material that might exist in our Universe have different pressures, and lead to different evolution of the density ρ .

We can derive the fluid equation by considering the first law of thermodynamics

$$dE + p dV = T dS, \quad (3.11)$$

applied to an expanding volume V of unit comoving radius.¹ This is exactly the same as applying thermodynamics to a gas in a cylinder. The volume has physical radius a , so the energy is given, using $E = mc^2$, by

$$E = \frac{4\pi}{3} a^3 \rho c^2. \quad (3.12)$$

The change of energy in a time dt , using the product rule, is

$$\frac{dE}{dt} = 4\pi a^2 \rho c^2 \frac{da}{dt} + \frac{4\pi}{3} a^3 \frac{d\rho}{dt} c^2, \quad (3.13)$$

while the rate of change in volume is

$$\frac{dV}{dt} = 4\pi a^2 \frac{da}{dt}. \quad (3.14)$$

Assuming a reversible adiabatic expansion $dS = 0$, putting these into Equation (3.11) and rearranging gives

$$\dot{\rho} + 3 \frac{\dot{a}}{a} \left(\rho + \frac{p}{c^2} \right) = 0, \quad (3.15)$$

where, as always, dots are shorthand for time derivatives. This is the **fluid equation**. As we see, there are two terms contributing to the change in the density. The first term in the brackets corresponds to the dilution in the density because the volume has increased, while the second corresponds to the loss of energy because the pressure of the material has done work

as the Universe's volume increased. This energy has not disappeared entirely, of course; energy is always conserved. The energy lost from the fluid via the work done has gone into gravitational potential energy.

Let me stress that there are no pressure forces in a homogeneous universe, because the density and pressure are everywhere the same. A pressure *gradient* is required to supply a force. So, pressure does not contribute a force helping the expansion along; its effect is solely through the work done as the Universe expands.

We are still not in a position to solve the equations, because now we only know what ρ is doing if we know what the pressure p is. It is in specifying the pressure that we are saying what kind of material our model universe is filled with. The usual assumption in cosmology is that there is a unique pressure associated with each density, so that $p \equiv p(\rho)$. Such a relationship is known as the **equation of state**, and we'll see two different examples in [Chapter 5](#). The simplest possibility is that there is no pressure at all, and that particular case is known as (nonrelativistic) matter.

Once the equation of state is specified, the Friedmann and fluid equations are all we need to describe the evolution of the Universe. However, before discussing this evolution, we are going to spend some time exploring some general properties of the equations, as well as devoting [Chapter 4](#) to consideration of the meaning of the constant k . If you prefer to immediately see how to solve these equations, feel free to jump straight away to [Sections 5.3 – 5.5](#), and come back to the intervening material later. On the way, you might want to glance at [Section 3.6](#) to find out why a factor of c^2 mysteriously vanishes from the Friedmann equation between here and there.

3.5 The acceleration equation

The Friedmann and fluid equations can be used to derive a third equation, not independent of the first two, of course, which describes the acceleration of the scale factor. By differentiating [Equation \(3.10\)](#) with respect to time, we obtain

$$2 \frac{\dot{a}}{a} \frac{a\ddot{a} - \dot{a}^2}{a^2} = \frac{8\pi G}{3} \dot{\rho} + 2 \frac{kc^2 \dot{a}}{a^3}. \quad (3.16)$$

Substituting in for $\dot{\rho}$ from [Equation \(3.15\)](#) and canceling the factor $2\dot{a}/a$ in each term gives

$$\frac{\ddot{a}}{a} - \left(\frac{\dot{a}}{a}\right)^2 = -4\pi G \left(\rho + \frac{p}{c^2}\right) + \frac{kc^2}{a^2}, \quad (3.17)$$

and finally, using [Equation \(3.10\)](#) again, we arrive at an important equation known as the **acceleration equation**

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3} \left(\rho + \frac{3p}{c^2}\right). \quad (3.18)$$

Notice that if the material has any pressure, this *increases* the gravitational force, and so further decelerates the expansion. We remind you that there are no forces associated with pressure in an isotropic universe, as there are no pressure gradients.

The acceleration equation does not feature the constant k that appears in the Friedmann equation; it cancels out during the derivation.

3.6 On mass, energy, and vanishing factors of c^2

You should be aware that cosmologists have a habit of using mass density ρ and energy density ϵ interchangeably. They are related via Einstein's most famous equation as $\epsilon = \rho c^2$, and if one chooses so-called "natural units" in which c is set equal to one, the two become the same. For clarity, however, we will be careful to maintain the distinction. Note that the phrase "mass density" is used in Einstein's sense – it includes the contributions to the mass from the energy of the various particles, as well as any rest mass they might have.

The habit of setting $c = 1$ means that the Friedmann equation is normally written without the c^2 in the final term, so that it reads

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3} \rho - \frac{k}{a^2}. \quad (3.19)$$

The constant k then appears to have units $[\text{time}]^{-2}$ – setting $c = 1$ makes time and length units interchangeable. Since the practice of omitting the c^2 in the Friedmann equation is widely adopted in other cosmology textbooks, we will drop it for the remainder of this book too. In practice, it is a rare situation indeed where one has to be careful about this.

Problems

- 3.1. Starting with the Friedmann and acceleration equations, derive the fluid equation.
- 3.2. Derive an equation for the third derivative of the scale factor in terms of the density (not its derivatives), in the case of a pressureless fluid.

Note

- ¹ Don't confuse V for volume with $\mathcal{V}$ for gravitational potential energy.

Chapter 4

The Geometry of the Universe

We now consider the real meaning of the constant k , which appears in the Friedmann equation

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}. \quad (4.1)$$

While the Newtonian derivation in [Chapter 3](#) introduced this as a measure of the energy per particle, the true interpretation, apparent in the context of general relativity, is that it measures the curvature of space. General relativity tells us that gravity is due to the curvature of four-dimensional (4D) space-time, and a full analysis can be found in any general relativity textbook. Here, we will be purely descriptive, and focus on the interpretation of k as measuring the curvature of the three spatial dimensions. Further details of general relativistic cosmology can be found in [Advanced Topic 1](#).

We have demanded that our model universes be both homogeneous and isotropic. The simplest type of geometry which can have this property is what is called a **flat geometry**, in which the normal rules of Euclidean geometry apply. However, it turns out that the assumption of isotropy is not enough to demand that as the only choice. Instead, there are three possible geometries for the Universe, and they correspond to k being zero, positive, or negative.

4.1 Flat geometry

Euclidean geometry is based on a set of simple axioms (e.g. a straight line is the shortest distance between two points), plus one more complex axiom which says that parallel straight lines remain a fixed distance apart. These are the basis for the standard laws of geometry, and lead to the following conclusions:

- The angles of a triangle add up to 180° .
- The circumference of a circle of radius r is $2\pi r$.

Such a geometry might well apply to our own Universe. If that is the case, then the Universe must be infinite in extent, because if it came to a definite edge, then that would clearly violate the principle that the Universe should look the same from all points.

A universe with this geometry is usually called a **flat universe**.

4.2 Spherical geometry

Euclid always hoped that the more artificial final axiom could be proven by others. It wasn't until the nineteenth century that Riemann demonstrated that Euclid's final axiom was an arbitrary choice, and that one could make other assumptions. In doing so, he founded the subject of non-Euclidean geometry, which forms the mathematical foundation for Einstein's theory of general relativity.

The simplest kind of non-Euclidean geometry is actually very familiar to us; it is the spherical geometry which we use, for instance, to navigate around the Earth. Before worrying about the Universe having three dimensions, let's examine the properties of the two-dimensional (2D) surface of the Earth, shown in [Figure 4.1](#).

 Diagram of a spherical surface depicting a triangle with three right angles detailed at the vertices, indicating positive curvature k .

Figure 4.1 A sketch of a spherical surface, representing positive k . A triangle is shown, which has three right angles!

First, we know that a perfect sphere looks the same from all points on its surface, so the condition of isotropy is satisfied (e.g. if someone hands you a snooker ball and asks which way up it is, you're not going to be able to tell them). But, unlike the case of a flat geometry, the spherical surface is perfectly finite in extent, its area being given by $4\pi r^2$. Yet, there is no boundary, no "edge" to the surface of the Earth. So, it is perfectly possible to have a finite surface which nevertheless has no boundary.

If we draw parallel lines on the surface of the Earth, then they violate Euclid's final axiom. The definition of a straight line is the shortest distance between two points, which means that the straight lines in a spherical geometry are segments of great circles, such as the equator or the lines of longitude.² The lines of longitude are an excellent example of the failure of Euclid's axiom; as

they cross the equator they are all parallel to one another, but rather than remaining a constant distance apart they meet at both poles.

If we draw a triangle on a sphere, we find that the angles do not add up to 180° either. The easiest example to think about is to start at the North Pole. Draw two straight lines down to the equator, 90° apart, and then join them with a line on the equator. You have drawn a triangle in which all three angles are 90° , as shown in [Figure 4.1](#).

The circumference of a circle also fails to obey the normal law. Suppose we draw a circle at a radius r from the North Pole, and we'll choose r so that our circle is the equator. That radius, measured on the surface of the sphere, corresponds to a quarter of a complete circle around the Earth, so r equals πR divided by 2 where R is the radius of the Earth. However, the circumference c is given by $2\pi r$, so instead of the usual relation one has c equals $4r$ for a circle drawn at the equator. The circumference is less than $2\pi r$. [Problem 4.1](#) looks at the general case; you may find it helpful to glance at the figure on page 38 now.

Although we have only considered specific cases where the algebra is easy, it is true that whatever triangle or circle is drawn, you'll always find

- the angles of a triangle add up to more than 180° ;
- the circumference of a circle is less than $2\pi r$.

If you make triangles or circles much smaller than the size of the Earth, then the Euclidean laws start to become a good approximation; certainly we don't have to worry about Euclidean laws being broken in our everyday existence (though the appreciation that the Earth is spherical is vital for the planning of long distance journeys). So, a small triangle drawn on a sphere will have the sum of its angles only marginally larger than 180° . This property makes it rather hard to measure the geometry of our own Universe, because the neighboring region that we can measure accurately is only a small fraction of the size of the Universe and so will obey nearly Euclidean laws whatever the overall geometry.

One of the most important conceptual points that you need to grasp is that our three-dimensional (3D) Universe can have properties just like the 2D surface of the sphere. Unfortunately, our brains are not conditioned to think of three dimensions as being curved, so we must work by analogy with the 2D situation we have just described. When we think of the surface of a sphere being curved, we naturally imagine the sphere as an object in our own 3D Universe and think of it as curved in that sense. The important point is that the curvature is a property of the 2D surface of the sphere itself; the triangles and circles whose properties we've just described are drawn on the surface. A classic application of this was the ancient Greeks' use of these laws to deduce that the Earth is spherical, and even to obtain a good estimate of its diameter. When we discuss spherical geometry, there is actually no need to think of it as existing within our 3D space at all, and we must remember when discussing the geometry that we are assumed to be restricted to the surface of the sphere, and not allowed to move off the surface either toward or away from the notional center.

All this is an analogy to what might happen with our 3D Universe. The analog of the 2D sphere is called a three-sphere. In the rather unlikely event of 4D creatures existing, they would be able to visualize the curvature of 3D space in just the same way we can visualize that of 2D space. However, like the 2D sphere, the possible curvature of three-space is an intrinsic property, and there is no actual need for a higher-dimensional space for it to live in. Obtaining a correct mental picture of this is one of the big challenges in understanding our Universe!

A universe with a spherical geometry, like the surface of the Earth, has a finite size but no boundary. All points are equivalent. If we lived in a spherical geometry and traveled in a straight line, we would not go on forever. Rather, we would eventually come back to where we had started from, from the opposite direction, exactly in the manner that someone traveling outward from the North Pole on the Earth eventually returns there from the opposite direction. Such a universe corresponds to choosing a positive value for the quantity k appearing in the Friedmann equation. Because the special properties of the spherical geometry are due to its curvature, k is often therefore called the curvature term.

A universe with k greater than 0 is normally referred to as a **closed universe**, because of its finite size.

4.3 Hyperbolic geometry

The final choice is k negative. The corresponding geometry is known as hyperbolic, and is much less familiar than spherical geometry. It is normally represented by a saddle-like surface as in [Figure 4.2](#); it is hard to see that this is consistent with isotropy, but in fact it is. In a hyperbolic geometry, parallel lines never meet – in fact they break Euclid's axiom by diverging away from one another.

 Diagram of a saddle surface featuring a central dashed curvilinear triangle with angles sum to less than 180 degrees.

Figure 4.2 A sketch of a saddle surface, representing the hyperbolic geometry obtained when ***k is negative***. A rather exaggerated triangle is shown with its sum of angles well below 180° .

The behavior of the hyperbolic geometry can be guessed from what has gone before; it is the opposite of spherical geometry. One finds that

- the angles of a triangle add up to less than 180° ;
- the circumference of a circle is greater than $2\pi r$.

Because parallel lines never meet, such a universe must be infinite in extent, just as in the flat case. The situation

 k less than 0 is known as an **open universe**.

In [Chapter 9](#), we'll explore which of the three possible geometries, summarized in [Table 4.1](#), seems best suited to describe the real Universe.

TABLE 4.1

A summary of possible geometries.

Curvature	Geometry	Angles of triangle	Circumference of circle	Type of universe
 k greater than 0	Spherical	$>180^\circ$	 c less than $2\pi r$	Closed
 k equals 0	Flat	180°	 c equals $2\pi r$	Flat
 k less than 0	Hyperbolic	$<180^\circ$	 c greater than $2\pi r$	Open

4.4 Infinite and observable universes

What does it actually mean for the Universe to be infinite, as in the flat and open cases? This property has nothing to do with the Universe lasting forever; what it means is that the Universe is already infinite in size even at a finite time. It genuinely goes on forever in all directions. Even despite this, it is still able to expand – the distances between objects can still increase regardless of whether the total Universe is infinite in extent or not. (You might consider the integers; these form an infinite set, but you can still multiply each by two to get a new infinite set where the separation between the numbers is twice as large.)

Nevertheless, this description is only a model, and we have no way of discovering whether the actual Universe does indeed go on forever. Cosmologists often talk about a different concept to the entire, possibly infinite, Universe, namely the **observable Universe**. This corresponds to the portion of the Universe we can actually see, and is limited by the finite speed of light. As the Universe gets older, the observable Universe becomes larger and larger, by a combination of two effects. First, the Universe is expanding, and second, light has had longer to travel across the Universe. In practice, our knowledge of the Universe is restricted to this portion, and we have no way of telling whether it does indeed continue into the infinite distance as required by the cosmological principle. It may be, for example, that the Universe becomes highly irregular on an extremely large scale, and there are some theoretical models that predict that this might happen. Another possibility is discussed in [Advanced Topic 1.3](#).

4.5 Where did the Big Bang happen?

A common question is “where did the Big Bang happen?”, suggesting perhaps that one could point in a specific direction and say “That way!” After all, in a conventional explosion that is a perfectly reasonable question to ask, as all the material flies outward from the ignition point. Unfortunately, for the Big Bang, things aren't so simple, and in a sense the answer is “everywhere and nowhere.”

First of all, remember that our entire foundation is the cosmological principle, telling us that no point in the Universe is special. If there were a particular point where the “Bang” happened, that would clearly be a special point and violate the cosmological principle. Rather, space and time themselves were created at the instant of the Big Bang (unlike a conventional explosion where the material flies through preexisting space). If we take any point in the present Universe and trace back its history, it would start out at the explosion point, and in that sense the Big Bang happened everywhere in space.

In another sense, the location of the Big Bang is nowhere, because space itself is evolving and expanding, and it has changed since the Big Bang took place. Imagine the Universe as an expanding sphere. At any instant, “space” is the surface of the sphere, which is becoming bigger with time (again we are thinking of a 2D analogy to our real 3D space). The place where the “Bang” happened is at the center of the sphere, but that's no longer part of the space, the surface of the sphere in which we live. In

particular, being constrained to the surface of the sphere means we are unable to “point” to the place where the explosion is supposed to have happened. However, all the points in our current space were once at the center of the sphere, when the Big Bang took place.

4.6 Three values of k

Because there are only three distinct possibilities for the geometry, many people indicate this explicitly by scaling their variables so that k takes on one of only three possible values, namely k equals negative 1, 0, or +1, corresponding to the open, flat, and closed cases, respectively. This can be achieved by rescaling the scale factor by multiplying it by a fixed constant, namely $\frac{1}{\sqrt{|k|}}$ if k is non-zero. This leaves $\dot{H}$ equals $\frac{\dot{a}}{a}$ unchanged and removes the k from the final term in Equation (4.1), which now reads

$$\left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3} \rho \pm \frac{1}{a^2} \quad (4.2)$$

with “−” for positive k , “+” for negative k , and the last term absent if k equals 0.

If this rescaling is used, then (except in the k equals 0 case) one loses the freedom to set

$\frac{\dot{a}}{a} = 1$ at the present time, as will be used in Section 5.3. In effect, what one is doing is choosing to measure comoving distances in units of the so-called curvature scale, which is the scale on which the effects of curved space must be included, rather than in astronomical units such as megaparsecs.

In this book, we will *not* rescale k to one of these three discrete values, though most of the time we will be specializing to k equals 0 anyway.

Problems

4.1. Consider the surface of a 2D sphere of radius R , as illustrated.

Diagram of a sphere details radius R extending from the center, angle θ from the vertical axis, and arc length r along the surface.

Circles are drawn on the sphere that have radius r as measured on the surface of the sphere, centered on the North Pole. Show that the general formula for the circumference c of such a circle, as a function of r , is given by

$$c = 2\pi \sin\left(\frac{r}{R}\right) r$$

where θ is the angle between a line drawn from the center of the sphere to the North Pole and one drawn to the circle. (Remember that by definition of angles in radians, we have $r = \theta R$.)

Demonstrate that for small θ (i.e. r much less than R) this gives the normal flat geometry relation. Evaluate the relation for the case when the circle is at the equator.

4.2. Consider the spherical geometry of the previous problem, staying with the 2D analogy to the real Universe. Suppose that galaxies are distributed evenly in such a universe, with a number density n per unit area. Show that the total number N of galaxies inside a radius r is given by

$$N = 2\pi n R^2 \left[1 - \cos\left(\frac{r}{R}\right) \right]$$

Expand this for r much less than R to show that the flat space result that the number is $\pi n r^2$ is recovered (remember we are working in only two dimensions). Do you see more or fewer galaxies out to the same radius, if the Universe is spherical rather than flat?

4.3. Is it possible for a closed universe to evolve to become an open universe? Give a reason for your answer.

Notes

¹ Well, that’s almost true. See Advanced Topic 1.3 for a way to bypass that conclusion.

² Note that, apart from the equator, lines of latitude are *not* straight lines; this is why aeroplanes do not follow lines of latitude when flying, because they are not the shortest way to go!

Chapter 5

Simple Cosmological Models

In [Chapter 3](#), we derived the equations satisfied by an expanding isotropic gas. They are the Friedmann equation ¹

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}, \quad (5.1)$$

which governs the time evolution of the scale factor $a(t)$, and the fluid equation

$$\frac{d\rho}{dt} + 3\frac{\dot{a}}{a}\rho = 0, \quad (5.2)$$

which gives us the evolution of mass density $\rho(t)$. We stress that, despite their Newtonian derivation, these are the real equations used by cosmologists, more traditionally derived via the equations of general relativity as described in [Advanced Topic 1](#).

This chapter finds and discusses some simple solutions to these equations, which will be used extensively during the book. However, while they have wide applicability during the early stages of the Universe, it will turn out that these are not sufficient to describe the present state of the Universe, for which an extra ingredient, the cosmological constant, will be needed. It is introduced in [Chapter 7](#).

Before finding solutions to these equations, we can study two of their implications.

5.1 Hubble's law

The Friedmann equation allows us to explain Hubble's discovery that recession velocity is proportional to the distance. The velocity of recession is given by

$\vec{v} = d\vec{r}/dt$ and is in the same direction as $\vec{r}$, allowing us to write

$$\vec{v} = \frac{d\vec{r}}{dt} = \frac{d}{dt}(\vec{r}) = \frac{d}{dt}(a\vec{r}_0) = \dot{a}\vec{r}_0 = \frac{\dot{a}}{a}\vec{r}, \quad (5.3)$$

The last step used $\vec{r}_0$ is a constant by definition. Consequently, Hubble's law

$\vec{v} = H\vec{r}$ tells us that the proportionality constant, the Hubble parameter, should be identified as

$$H = \frac{\dot{a}}{a} \quad (5.4)$$

and the value as measured today can be denoted with a subscript "0" as H_0 . Because we measure Hubble's constant to be positive rather than negative, we know that the Universe is expanding rather than contracting.

We notice from this that the phrase Hubble's *constant* is a bit misleading. Although it is certainly constant in space due to the cosmological principle, there is no reason for it to be constant in time. In fact, using it as a more compact notation, we can write the Friedmann equation as an evolution equation for $H(t)$, as

$$H^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \quad (5.5)$$

It is best to use the phrase "Hubble parameter" for this quantity as a function of time, reserving "Hubble constant" for its present value. Normally, the Hubble parameter decreases with time, for instance, as the expansion is slowed by the gravitational attraction of the matter in the Universe.

5.2 Expansion and redshift

The redshift of spectral lines that we used to justify the assumption of an expanding universe can also be related to the scale factor. In this derivation, we will make the simplifying assumption that light is passed between two objects that are very close together, separated by a small distance d_r , as shown in [Figure 5.1](#). We have drawn the objects as galaxies, but we really mean two nearby points. According to Hubble's law, their relative velocity $d v$ will be

$$d v \text{ equals } H d_r \text{ equals } \frac{H}{c} \frac{d r}{dt} \quad (5.6)$$

As the points are nearby, we can directly apply the Doppler law to say that the change in wavelength between emission and reception, $d \lambda$ identical to λ Subscript normal r Baseline minus λ Subscript normal e , is

$$\frac{d \lambda}{\lambda} \text{ equals } \frac{d v}{c} \quad (5.7)$$

where $d \lambda$ is going to be positive since the wavelength is increased. The time between emission and reception is given by the light travel time $d t$ equals d_r divided by c , and putting all that together gives

$$\frac{d \lambda}{\lambda} \text{ equals } \frac{d r}{c d t} \text{ equals } \frac{H}{c} \frac{d r}{d t} \quad (5.8)$$

Integrate and we find that $\ln \lambda$ equals $\ln$ plus c on s t a n t, i.e.

$$\lambda \text{ proportional to } a \quad (5.9)$$

where λ is now the instantaneous wavelength measured at any given time.

 Diagram of a photon wave traveling from spiral Galaxy A toward spiral Galaxy B across a distance measured as d_r .

Figure 5.1 A photon travels a distance d_r between two galaxies A and B.

Although as we have derived it, this result only applies to objects that are very close to each other, it turns out that it is completely general (a rigorous treatment is given in [Advanced Topic 2](#)). It tells us that as space expands, wavelengths become longer in direct proportion. One can think of the wavelength as being stretched by the expansion of the Universe, and its change therefore tells us how much the Universe has expanded since light began its travels. For example, if the wavelength had doubled, the Universe must have been half its present size when the light was emitted.

The redshift as defined in [Equation \(2.1\)](#) is related to the scale factor by

$$1 + z \text{ equals } \frac{\lambda_r}{\lambda_e} \quad (5.10)$$

and is normally only used to refer to light received by us at the present epoch.

5.3 Solving the equations

In order to discover how the Universe might evolve, we need some idea of what is in it. In a cosmological context, this is done by specifying the relationship between the mass density ρ and the pressure p . This relationship is known as the **equation of state**. At this point, we shall only consider two possibilities.

Matter: In this context, the term “matter” is used by cosmologists as shorthand for “non-relativistic matter,” and refers to any type of material which exerts negligible pressure, p equals 0. Occasionally, care is needed to avoid confusion between “matter” used in this sense, and used to indicate all types of matter, whether nonrelativistic or not. A pressureless universe is the simplest assumption that can be made. It is a good approximation to use for the atoms in the Universe once they have cooled down, as they are quite well separated and seldom interact, and it is also a good description of a collection of galaxies in the Universe, as they have no interactions other than gravitational ones. Occasionally the term “dust” is used instead of “matter.”

Radiation: Particles of light move, naturally enough, at the speed of light. Their kinetic energy leads to a pressure force, the radiation pressure, which, using the standard theory of radiation, can be shown to be

ρ equals ρ_0 divided by 3. [Problem 5.2](#) gives a rather hand-waving derivation of this result. More generally, any particles moving at highly relativistic speeds have this equation of state, neutrinos being an obvious example.

We will concentrate on the case where the constant k in the Friedmann equation is set equal to zero, corresponding to a flat geometry.

5.3.1 Matter

We start by solving the fluid equation, having set $p = 0$ for matter. One way to solve it is to notice a clever way of rewriting it, as follows

$$\frac{d}{dt} \left(\rho a^3 \right) = 0 \quad (5.11)$$

though one could also solve it more formally by noting that it is a separable equation. Integrating tells us that ρa^3 equals a constant, i.e.

$$\rho a^3 \propto \frac{1}{a^3} \quad (5.12)$$

This is not a surprising result. It says that the density falls off in proportion to the volume of the Universe. It is very natural that if the volume of the Universe increases by a factor of say two, then the density of the matter must fall by the same factor. After all, material cannot come from nowhere, and there is no pressure to do any work.

The equations we are solving (with $k = 0$) have one very useful symmetry; their form is unchanged if we multiply the scale factor a by a constant, since only the combination $\frac{da}{dt} a^3$ appears. This means that we are free to rescale a as we choose, and the normal convention is to choose $a = 1$ at the present time. With this choice, physical and comoving coordinate systems coincide at the present, since $\frac{dr}{dt} = \frac{dx}{dt}$. Throughout this book we will use the subscript “0” to indicate the present value of quantities. Denoting the present density by ρ_0 fixes the proportionality constant

$$\rho a^3 = \rho_0 \quad (5.13)$$

Having solved for the evolution of the density in terms of a , we must now find how a varies with time by using the Friedmann equation. Substituting in for ρ , and remembering we are assuming $k = 0$, gives

$$\left(\frac{da}{dt} \right)^2 = \frac{8\pi G}{3} \rho_0 \frac{1}{a^3} \quad (5.14)$$

Faced with an equation like this, one can use formal techniques to solve it (this equation is separable, allowing it to be integrated), or alternatively make an educated guess as to the solution and confirm it by substitution. In cosmology, a good educated guess is normally a power-law $a \propto t^q$. Substituting this in, the left-hand side has time dependence t^{2q-2} and the right-hand side t^{-3q} . This can only be a solution if these match, which requires $q = 2/3$, and so the solution is $a \propto t^{2/3}$. As we have fixed $a = 1$ at the present time $t = t_0$, the full solution is therefore

$$a(t) = \left(\frac{t}{t_0} \right)^{2/3} \quad (5.15)$$

In this solution, the Universe expands forever, but the rate of expansion $H(t)$ decreases with time

$$H(t) = \frac{2}{3t} \quad (5.16)$$

becoming infinitely slow as the Universe becomes infinitely old. Notice that, despite the pull of gravity, the material in the Universe does not recollapse but rather expands forever.

This is one of the classic cosmological solutions, and will be much used throughout this book.

5.3.2 Radiation

Radiation obeys $p = \frac{1}{3} \rho c^2$. Consequently, the fluid equation is changed from the matter-dominated case, now reading

$$\frac{d\rho}{dt} + 4\frac{\dot{a}}{a}\rho = 0 \quad (5.17)$$

This is amenable to the same trick as before, with the a^3 replaced by a^4 in Equation (5.11), giving

$$\rho \propto \frac{1}{a^4} \quad (5.18)$$

Carrying out the same analysis we did in the matter-dominated case gives

$$\left(\frac{dt}{da} \right)^2 = \frac{1}{2\rho} \quad (5.19)$$

$$\frac{dt}{da} = \frac{1}{\sqrt{2\rho}} = \frac{1}{\sqrt{2}} \frac{a^2}{a^2} = \frac{1}{\sqrt{2}} \frac{a^2}{a^2}$$

This is the second classic cosmological solution.

Notice that the Universe expands more slowly if radiation dominated than if matter dominated, a consequence of the extra deceleration that the pressure supplies – see Equation (3.18). So, it is definitely wrong to think of pressure as somehow “blowing” the Universe apart. However, in each case, the density of material falls off as a^{-3} .

We’d better examine the fall off of the radiation density with volume more carefully. It drops as the fourth power of the scale factor. Three of those powers we have already identified as the increase in volume, leading naturally to a drop in the density. The final power arises from a different effect, the stretching of the wavelength of the light. Since the stretching is proportional to a , and the energy of radiation proportional to its frequency via $E = hf$, this results in a further drop in energy by the remaining power of a . This lowering of energy is exactly the redshifting effect we use to measure distances.

The rate of decrease of the radiation density also has an explanation in terms of thermodynamics, which is macroscopic rather than microscopic. Since the Universe in this case has a pressure, when it expands, there is work done which is given by $p dV$, in exactly the same way as work is done on a piston when the gas is allowed to expand and cool. This work done corresponds to the extra diminution of the radiation density by the final factor of a .

5.3.3 Mixtures

A more general situation is when one has a mixture of both matter and radiation. Then there are two separate fluid equations, one for each of the two components. The trick which allows us to write ρ as a function of a still works, so we still have

$$\rho_m \propto \frac{1}{a^3} \quad (5.20)$$

$$\rho_r \propto \frac{1}{a^4}$$

However, there is still only a single Friedmann equation (after all, there is only one Universe!), which now has

$$\dot{a}^2 = \frac{8\pi G}{3} (\rho_m + \rho_r) \quad (5.21)$$

This means that the scale factor will have a more complicated behavior, and so to convert

ρ into $\rho(a)$ is much harder. It is possible to obtain exact solutions for this situation, but they are quite messy, so we won’t include them here. Instead, we will consider a simpler situation where one or other of the densities is by far the larger.

In that case, we can say that the Friedmann equation is accurately solved by just including the dominant component. That is, we can use the expansion rates we have already found. For example, suppose radiation is much more important. Then, one would have

$$\rho_m \propto \frac{1}{t^2} \quad \rho_r \propto \frac{1}{t^4} \quad (5.22)$$

Notice that the density in matter falls off more slowly than that in radiation. So, the situation of radiation dominating cannot last forever; however, small the matter component might be originally, it will eventually come to dominate. We can say that domination of the Universe by radiation is an unstable situation.

In the opposite situation, where it is the matter which is dominant, we obtain the solution

$$\rho_m \propto \frac{1}{t^3} \quad \rho_r \propto \frac{1}{t^4} \quad (5.23)$$

Matter domination is a stable situation, the matter becoming increasingly dominant over the radiation as time goes by.

[Figure 5.2](#) shows the evolution of a universe containing matter and radiation, with the radiation initially dominating. Eventually, the matter comes to dominate, and as it does so, the expansion rate speeds up from

$\rho_m \propto \frac{1}{t^3}$ to $\rho_m \propto \frac{1}{t^2}$ law. It is likely that this is the situation that applied in our Universe until fairly recently, as we'll see in [Chapter 11](#).

 Graph of log density versus log time with two downward sloping lines, Radiation and Matter, converging at a central point.

Figure 5.2 A schematic illustration of the evolution of a universe containing radiation and matter. Once matter comes to dominate the expansion rate speeds up, so the densities fall more quickly with time.

5.4 Particle number densities

An important alternative view of the evolution of particles, which will be much used later in the book, is that of the **number density** n of particles rather than of their mass or energy density.

The number density is simply the number of particles in a given volume. If the mean energy per particle (including any mass-energy) is ϵ , then the number density is related to the energy density by

$$\epsilon = n \epsilon \quad (5.24)$$

The number density is useful because, in most circumstances, particle number is conserved. For example, if particle interactions are negligible, you wouldn't expect an electron to suddenly vanish into oblivion, and the same is true of a photon of light. The particle number can change through interactions. For example, an electron and positron could annihilate and create two photons. However, if the interaction rate is high, we expect the Universe to be in a state of thermal equilibrium. If so, then particle numbers are conserved even in a highly interacting state, since, by definition, thermal equilibrium means that any interaction, which may change the number density of a particular type of particle, must proceed at the same rate in both forward and backward directions so that any change cancels out.

So, barring brief periods where thermal equilibrium does not hold, we expect the number of particles to be conserved. The only thing that changes the number *density*, therefore, is that the volume is getting bigger, so that these particles are spread out in a larger volume. This implies

$$n \propto \frac{1}{t^3} \quad (5.25)$$

This looks encouragingly like the behavior we have already seen for matter, but it's also true for radiation as well!

How does this relate to our earlier results? The energy of nonrelativistic particles is dominated by their rest mass-energy, which is constant, so

$$\rho_m \propto n \epsilon \propto \frac{1}{t^3} \quad (5.26)$$

But photons lose energy as the Universe expands and their wavelength is stretched, so their energy is

$\epsilon \propto \frac{1}{t}$ as we have already seen. So

$$\rho \propto a^{-3} \quad \text{and} \quad \epsilon \propto a^{-4} \quad (5.27)$$

These are exactly the results we saw before: [Equations \(5.12\)](#) and [\(5.18\)](#).

Although the energy densities of matter and radiation evolve in different ways, their particle numbers evolve in the same way. So, apart from epochs during which the assumption of thermal equilibrium fails, the relative number densities of the different particles (e.g. electrons and photons) do not change as the Universe expands.

5.5 Evolution including curvature

We can now re-introduce the possibility that the constant k is nonzero, corresponding to spherical or hyperbolic geometry. Rather than seeking precise solutions, we will concentrate on the qualitative properties of the solutions. These are actually of rather limited use in describing our own Universe, because, as we will see, the cosmological models discussed so far are not general enough, and we will need to consider a cosmological constant (see [Chapter 7](#)). Nevertheless, studying the possible behaviors of these simple models is a useful exercise, even if one should be cautious about drawing general conclusions.

In analyzing the possible dynamics, we will assume that the Universe is always dominated by nonrelativistic matter, which in practice is not a restrictive assumption. We have already seen that if we assume that the constant k in the Friedmann equation is zero, then the Universe expands for ever, $a \propto t^{2/3}$, but slows down arbitrarily at late times. So, we know the fate of the Universe in that case. But what happens if $k \neq 0$?

The principal question to ask is whether it is possible for the expansion of the Universe to stop, which since $\dot{H} = -H^2 - \frac{1}{6}\Lambda - \frac{1}{2}k/a^2$ corresponds to $\dot{H} = 0$. Looking at the Friedmann equation

$$\dot{H}^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \quad (5.28)$$

it is immediately apparent that this is not possible if k is negative, for then both the terms on the right-hand side of the Friedmann equation are positive. Consequently, such a universe must expand forever. That enables us to study the late-time behavior, because we can see that the term k/a^2 falls off more slowly with the expansion than it does $\rho \propto a^{-3}$. Since a becomes arbitrarily large for the matter-dominated solution for negligible k , the k/a^2 term must eventually come to dominate. When it does, the Friedmann equation becomes

$$-\frac{k}{a^2} = -\frac{1}{2}\Lambda \quad (5.29)$$

Cancel off the a^2 terms and you'll find the solution is $a \propto t$. So, when the last term comes to dominate, the expansion of the Universe becomes yet faster. In this case, the velocity does not tend to zero at late times, but instead becomes constant. This is sometimes known as free expansion.

Things are very different if k is positive. It then becomes possible for $\dot{H}$ to be zero, by the two terms on the right-hand side of the Friedmann equation canceling each other out. Indeed, this is inevitable, because the negative influence of the k/a^2 term will become more and more important relative to the ρ term as time goes by. In such a universe, therefore, the expansion must come to an end after a finite amount of time. As gravitational attraction persists, the recollapse of the Universe becomes inevitable.

The collapse of the Universe is fairly easy to describe, because the equations governing evolution are time reversible. That is, if one substitutes negative t for t , they remain the same. The collapse phase is therefore just like the expansion in reverse, and so, after a finite time, the Universe will come to an end in a Big Crunch. [Problem 5.5](#) investigates this in more detail.

These three behaviors, illustrated in [Figure 5.3](#), can be related to the particle energy U in our Newtonian derivation of the Friedmann equation. If the particle's energy is positive, then it can escape to infinity, with a final kinetic energy given by U . If the total energy is zero, then the particle can just escape but with zero velocity. Finally, if the energy is negative, it cannot escape the gravitational attraction and is destined to recollapse inward.

 Graph of scale factor versus time with three curves labeled $k < 0$, $k = 0$, and $k > 0$ details universe evolution.

Figure 5.3 Three possible evolutions for the Universe, corresponding to the different signs of k . The middle line corresponds to the $k = 0$ case, where the expansion rate approaches zero in the infinite future. During the early phases of the expansion the lines are very close and so observationally it can be difficult to distinguish which path the actual Universe will follow.

There is a fairly precise analogy with escape velocity from the Earth (or the Moon, if you want to worry about the atmosphere). If you throw a rock up in the air hard enough, gravity will be unable to stop it, and eventually it will sail off into space at a constant velocity. If your throw is too puny, it will rapidly fall back. And in between is the escape velocity, where the rock is just able to escape the gravitational field and no more.

Problems

5.1. Is the energy of the Universe conserved as it expands?

5.2. This problem indicates the origin of the equation of state $p = \frac{1}{3} \rho c^2$ for radiation. An ideal gas has pressure

$$p = \frac{1}{3} n \langle v^2 \rangle$$

where $\langle \rangle$ indicates an average over the direction of particle motions. Here, n is the number density, and be careful not to confuse the unfortunate notation p for pressure and $\mathbf{p}$ for momentum. Using Equation (2.4) to relate the photon energy and momentum show that this gives

$$p = \frac{1}{3} n \langle E \rangle$$

where $\langle E \rangle$ is the mean photon energy. Hence, it demonstrates the equation of state for radiation.

5.3. In this chapter, we examined solutions for the expansion when the Universe contained either matter

or radiation

Suppose we have a more general equation of state, $p = w \rho c^2$, where w is a constant in the range $-1 < w < 1$. Find solutions for $\rho(t)$ and hence $a(t)$ for universes containing such matter. Assume $k = 0$ in the Friedmann equation. What is the solution if $p = -\rho c^2$?

5.4. Using your answer to Problem 5.3, what value of w would be needed so that ρ has the same time dependence as the curvature term k/a^2 ? Find the solution $a(t)$ to the full Friedmann equation for a fluid with this w , assuming negative k .

5.5. The full Friedmann equation is

$$\left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3} \rho - \frac{k}{a^2}$$

Consider the case $k > 0$, with a universe containing only matter

so that $\rho = \rho_0 / a^3$. Demonstrate that the parametric solution

$$a = \frac{4\pi G \rho_0}{3k} (1 - \cos \theta) \quad t = \frac{4\pi G \rho_0}{3k} \left(\theta - \sin \theta \right)$$

solves this equation, where θ is a variable which runs from 0 to 2π .

Sketch a and t as functions of θ . Describe qualitatively the behavior of the Universe. Attempt to sketch a as a function of t .

5.6. Now consider the case $k < 0$, with a universe containing only matter

left parenthesis p equals 0 right parenthesis so that ρ equals ρ_0 divided by a^3 . What is the solution

left parenthesis t right parenthesis in a situation where the final term of the Friedmann equation dominates over the density term? How does the density of matter vary with time? Is domination by the curvature term a stable situation that will continue forever?

Note

¹In accord with the discussion of [Section 3.6](#), the c^2 squared on the final term has been dropped, so that its appearance matches that of other cosmology textbooks.

Chapter 6

Observational Parameters

The Big Bang model does not give a unique description of our present Universe, but rather leaves quantities such as the present expansion rate, or the present composition of the Universe, to be fixed by observation.

It is a standard practice to specify cosmological models via a few parameters, which one then tries to determine observationally to decide which version of the model best describes our Universe. In this chapter and [Chapter 7](#), we will discuss the most commonly considered parameters, including ones we have already seen and new ones.

6.1 The expansion rate H_0

The Hubble constant H_0 tells us the present expansion rate of the Universe, making it a vital cosmological parameter as measurements of other properties of the Universe often depend on it. It also ought to be the easiest to measure, since all galaxies are supposed to obey $v = H_0 r$. So, all we have to do is measure the velocities and distances of as many galaxies as we can and get an answer. However, each measurement has its problems.

Velocities are given by the redshift of spectral lines, a measurement which is now easy enough that the velocity of an individual galaxy can be measured to high accuracy. However, remember that the cosmological principle isn't perfect, and so, as well as the uniform expansion we are trying to measure, galaxies also have motions relative to one another, the so-called "peculiar velocity." The peculiar velocities are randomly oriented, and for a given galaxy we cannot split its measured velocity into the Hubble expansion and the peculiar velocity. However, the cosmological principle does tell us that the typical size of the peculiar velocity should not depend on where in the Universe the galaxy is. It is therefore independent of distance, whereas the Hubble velocity is proportional to distance. If we look far enough away (in practice many tens of megaparsecs) then the Hubble velocity dominates and the (unknown) peculiar velocity can be ignored.

Given that the expansion velocity can only be accurately distinguished from the peculiar velocity at large distances, we need to be able to estimate these large distances accurately in order to carry out the calculation $H_0 = v/r$. These distances are much harder to obtain, because galaxies are far too distant to be located by parallax. Remember that an object one parsec away has a parallax (i.e. an apparent motion when viewed from different parts of Earth's orbit) of one arcsecond, by definition. A galaxy many megaparsecs away will have an immeasurably small parallax of less than a micro-arcsecond. The usual method is known as "standard candles," where some type of object is assumed to have exactly the same properties in all parts of the Universe. This is the cosmological equivalent of saying that if one light bulb looks a quarter as bright as another, then from the inverse square law it must be twice as far away – fine as long as you believe that all light bulbs have precisely the same brightness. A classic example is the period–luminosity relation in cepheid variable stars. The period of variability of those stars is readily measured, and there is reasonable empirical evidence of a relation between the period and the luminosity of the stars, which lets us convert the measured period into a brightness. Other standard candles which have been used include the brightness of certain types of supernovae, and the brightest galaxies in galaxy clusters. All these methods have good success at determining relative distances between two galaxies, which requires that the objects be good standard candles, and relative distances are all that is needed to confirm the Hubble law.

However, to give an absolute distance, by actually measuring the proportionality constant H_0 , we also need a calibration against an object of known distance, which proves much harder. In the light bulb analogy, to get relative distances we need only the inverse square law and the belief that all bulbs have the same brightness; we don't need to know how bright the bulbs are. But to be able to say how far away a bulb of a given observed brightness is, we need to know its absolute brightness.

It is only relatively recently that the calibration problem has begun to come under control, principally through Hubble Space Telescope observations. The Hubble constant is usually parametrized as

$$H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}. \quad (6.1)$$

In the 1990s, the Hubble Space Telescope Key Project, led by Wendy Freedman, gave

$$h = 0.72 \pm 0.08, \quad (6.2)$$

where the uncertainty is a one-sigma error (meaning it should be doubled to indicate 95% confidence, at least if the uncertainty is approximately Gaussian-distributed). This was considered the first precision direct measurement of the Hubble constant, and has subsequently been improved to $h = 0.73 \pm 0.01$ by the SH0ES (Supernovae and H_0 for the Equation of State) project.

On the other hand, indirect measurements, which assess the impact of the expansion rate on other cosmological observables, such as the cosmic microwave background, have suggested a lower value. The *Planck* satellite observations described in [Advanced Topic 6](#) give a very precise determination $h = 0.674 \pm 0.005$. There is a clear discrepancy between this measurement and the direct one, which has persisted for several years and become known as the **Hubble tension**. Its

implications will be discussed further in [Advanced Topic A6.5](#). Notwithstanding this controversy, it is safe to conclude that h is now quite accurately measured to be 0.70 within at most a few per cent. An object with a recession velocity of 7000 km s^{-1} would be expected to be at a distance $v/H_0 = 100 \text{ Mpc}$. The overall scale of the Universe is therefore well determined.

6.2 The density parameter Ω_0

The density parameter is a very useful way of specifying the density of the Universe. Let's start with the Friedmann equation again. Recalling that $H = \dot{a}/a$, it reads

$$H^2 = \frac{8\pi G}{3} \rho - \frac{k}{a^2}. \quad (6.3)$$

For a given value of H , there is a special value of the density which would be required in order to make the geometry of the Universe flat, $k = 0$. This is known as the **critical density** ρ_c , which we see is given by

$$\rho_c(t) = \frac{3H^2}{8\pi G}. \quad (6.4)$$

Note that the critical density changes with time, since H does. As we know the present value of the Hubble constant (at least in terms of h defined in [Equation \(6.1\)](#)), we can compute the present critical density. Since $G = 6.67 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ sec}^{-2}$, and converting megaparsecs to meters using conversion factors quoted on page xiii, it is

$$\rho_c(t_0) = 1.88 h^2 \times 10^{-26} \text{ kg m}^{-3}. \quad (6.5)$$

This is a startlingly small number; compare, for example, the density of water which is 10^3 kg m^{-3} . If there is any more matter than this apparently tiny amount, it is enough to tip the balance beyond a flat universe to a closed one with $k > 0$. So only a very tiny density of matter is needed in order to provide enough gravitational attraction to halt and reverse the expansion of the Universe.

However, let us write that another way, since kilograms and meters are rather small and inconvenient units for dealing with something as big as the Universe. Let's try measuring masses in solar masses and distances in megaparsecs. It becomes

$$\rho_c(t_0) = 2.78 h^{-1} \times 10^{11} M_\odot / (h^{-1} \text{ Mpc})^3. \quad (6.6)$$

Suddenly this doesn't look so small. In fact, $10^{11} - 10^{12}$ solar mass is about the mass of a typical galaxy, and a megaparsec more or less the typical galaxy separation, so the Universe cannot be far away (within an order of magnitude or so) from the critical density. Its density really must be around $10^{-26} \text{ kg m}^{-3}$.

Be sure to understand that the critical density is not necessarily the true density of the Universe, since the Universe need not be flat. However, it sets a natural scale for the density of the Universe. Consequently, rather than quote the density of the Universe directly, it is often useful to quote its value relative to the critical density. This dimensionless quantity is known as the **density parameter** Ω , defined by

$$\Omega(t) \equiv \frac{\rho}{\rho_c}. \quad (6.7)$$

Again, in general, Ω is a function of time, since both ρ and ρ_c depend on time. The present value of the density parameter is denoted Ω_0 .

With this new notation, we can rewrite the Friedmann equation in a very useful form. Substituting in for ρ in [Equation \(6.3\)](#) using the definitions we have made, [Equations \(6.4\)](#) and [\(6.7\)](#), leads to

$$H^2 = \frac{8\pi G}{3} \rho_c \Omega - \frac{k}{a^2} = H^2 \Omega - \frac{k}{a^2}, \quad (6.8)$$

and rearranging gives

$$\Omega - 1 = \frac{k}{a^2 H^2}. \quad (6.9)$$

We see that the case $\Omega = 1$ is very special, because in that case k must equal zero and since k is a fixed constant, this equation becomes $\Omega = 1$ for all time. That is true independent of the type of matter we have in the Universe, and this is often called a **critical-density universe**. When $\Omega \neq 1$, this form of the Friedmann equation is very useful for analyzing the evolution of the density, as we will see later in the chapter on inflationary cosmology.

Our Universe contains several different types of matter, and this notation can be used not just for the total density but also for each individual component of the density, so one talks of Ω_{mat} , Ω_{rad} , etc. We can even define a “density parameter” associated with the curvature term, by writing

$$\Omega_k \equiv -\frac{k}{a^2 H^2}. \quad (6.10)$$

This can be positive or negative, and using it the Friedmann equation can be written as

$$\Omega + \Omega_k = 1. \quad (6.11)$$

We'll return to the observational status of Ω_0 in [Chapter 9](#).

6.3 The deceleration parameter q_0

As we've discovered, not only is the Universe expanding, but also the rate at which it is expanding, given by the Hubble parameter, is changing with time. The deceleration parameter is a way of quantifying this.

Consider a Taylor expansion of the scale factor about the present time. The general form of this (with dots as always indicating time derivatives) is

$$a(t) = a(t_0) + \dot{a}(t_0)[t - t_0] + \frac{1}{2}\ddot{a}(t_0)[t - t_0]^2 + \dots \quad (6.12)$$

Let's divide through by $a(t_0)$. Then the coefficient of the $[t - t_0]$ term will just be the present Hubble parameter, and we can write

$$\frac{a(t)}{a(t_0)} = 1 + H_0[t - t_0] - \frac{q_0}{2}H_0^2[t - t_0]^2 + \dots, \quad (6.13)$$

which defines the deceleration parameter q_0 as

$$q_0 = -\frac{\ddot{a}(t_0)}{a(t_0)} \frac{1}{H_0^2} = -\frac{a(t_0)\ddot{a}(t_0)}{\dot{a}^2(t_0)}. \quad (6.14)$$

The larger the value of q_0 , the more rapid the deceleration.

The simplest situation is if the Universe is matter dominated, $p = 0$. Remember that by “matter” we mean any pressureless material; it could be a collection of elementary particles, or, equally well, a collection of galaxies. Then from the acceleration [Equation \(3.18\)](#) and the definition of critical density, [Equation \(6.4\)](#), we find

$$q_0 = \frac{4\pi G}{3}\rho \frac{3}{8\pi G\rho_c} = \frac{\Omega_0}{2}. \quad (6.15)$$

So, in this case, a measurement of q_0 would immediately tell us Ω_0 .

If we know the properties of the matter in the Universe, then q_0 is not independent of the first two parameters we have discussed, H_0 and Ω_0 . Those two are sufficient to describe all the possibilities. However, we don't know everything about the material in the Universe, so q_0 can provide a new way of looking at the Universe. It can in principle be measured directly by making observations of objects at very large distances, such as the numbers of distant galaxies, because the deceleration governs how large the Universe would be at an earlier time.

In the late 1990s, the first convincing measurements of q_0 were made by two research groups studying distant supernovae of a class known as type Ia, which are believed to be good standard candles. To widespread surprise at the time, the result is that the Universe is *accelerating* at present, $q_0 < 0$.¹ None of the cosmological models that we have discussed so far is capable of satisfying this condition, as can be seen directly from the acceleration [Equation \(3.18\)](#). This result is now firmly established, and among the most dramatic observational results in modern cosmology. The following chapter discusses how to extend our simple cosmological models to account for it.

Problems

- 6.1. Verify the numerical value of the critical density in terms of solar masses and megaparsecs, [Equation \(6.6\)](#).

6.2. The deceleration parameter is defined by [Equation \(6.14\)](#). Use the acceleration [Equation \(3.18\)](#) and the definition of critical density to show that a radiation-dominated universe has $q_0 = \Omega_0$.

6.3. Identify a sufficient and necessary condition that must be satisfied by the equation of state if q_0 is to be negative.

Note

¹The mathematical tools required to analyze such data are beyond the scope of the main body of this book, but are described in [Advanced Topic 2](#), where the supernova observations are discussed in greater detail.

Chapter 7

The Cosmological Constant

7.1 Introducing Λ

When formulating general relativity, Einstein believed that the Universe was static, but found that his theory of general relativity did not permit it. This is simply because all matter attracts gravitationally; none of the solutions we have found correspond to a static universe with constant a . In order to arrange a static universe, he proposed a change to the equations, something he would later famously call his “greatest blunder.” That was the introduction of a cosmological constant.

The introduction of such a term is permitted by general relativity, and although Einstein’s original motivation has long since faded, it is currently seen as one of the most important and enigmatic objects in cosmology. The cosmological constant Λ appears in the Friedmann equation as an extra term, giving

$$H^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} + \frac{\Lambda}{3}. \quad (7.1)$$

Here, Λ has units $[\text{time}]^{-2}$, though some people include an explicit factor of c^2 in this equation to instead measure it as $[\text{length}]^{-2}$.

In principle, Λ can be positive or negative, though the positive case is much more commonly considered. Einstein’s original idea was to balance curvature, Λ , and ρ to get $H(t) = 0$, and hence, a static universe (see [Problem 7.2](#)). In fact, this idea was rather misguided, since such a balance proves to be unstable to small perturbations, and hence presumably couldn’t arise in practice. Nowadays, the cosmological constant is most often discussed in the context of universes with the flat Euclidean geometry, $k = 0$.

The effect of Λ can be seen more directly from the acceleration equation. Following the derivation of [Section 3.5](#), but now using the Friedmann equation as given above, gives

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda}{3}. \quad (7.2)$$

A positive cosmological constant gives a positive contribution to $\ddot{a}$, and so acts effectively as a repulsive force. In particular, if the cosmological constant is sufficiently large, it can overcome the gravitational attraction represented by the first term and lead to an accelerating universe. It can therefore explain the observed acceleration of the Universe described in [Section 6.3](#).

In the same way that it is useful to express the density as a fraction of the critical density, it is convenient to define a density parameter for the cosmological constant as

$$\Omega_\Lambda = \frac{\Lambda}{3H^2}. \quad (7.3)$$

Although Λ is a constant, Ω_Λ is not since H varies with time. Repeating the steps used to write the Friedmann equation in the form of [Equation \(6.9\)](#), we then find

$$\Omega + \Omega_\Lambda - 1 = \frac{k}{a^2 H^2}. \quad (7.4)$$

The condition to have a flat universe, $k = 0$, generalizes to

$$\Omega + \Omega_\Lambda = 1. \quad (7.5)$$

The usual convention among astronomers, which we will follow in this book, is that the cosmological constant term is not considered to be part of the matter density Ω . (Particle physicists, on the other hand, often include the cosmological constant as one of the components of the total density.) The relation between the density parameters and the geometry now becomes

$$\begin{aligned}\text{Open universe:} & \quad 0 < \Omega + \Omega_\Lambda < 1. \\ \text{Flat universe:} & \quad \Omega + \Omega_\Lambda = 1. \\ \text{Closed universe:} & \quad \Omega + \Omega_\Lambda > 1.\end{aligned}$$

7.2 Fluid description of Λ

It is often helpful to describe Λ as if it were a fluid with energy density ρ_Λ and pressure p_Λ . From [Equation \(7.1\)](#), we see that the definition

$$\rho_\Lambda \equiv \frac{\Lambda}{8\pi G} \quad (7.6)$$

brings the Friedmann equation into the form

$$H^2 = \frac{8\pi G}{3}(\rho + \rho_\Lambda) - \frac{k}{a^2}. \quad (7.7)$$

This definition also ensures that $\Omega_\Lambda \equiv \rho_\Lambda / \rho_c$, where ρ_c is the critical density.

In order to determine the effective pressure corresponding to Λ , one can seek a definition so that the acceleration equation with Λ reduces to its standard form, [Equation \(3.18\)](#), with $\rho \rightarrow \rho + \rho_\Lambda$ and $p \rightarrow p + p_\Lambda$. More directly, we can consider the fluid equation for Λ

$$\dot{\rho}_\Lambda + 3\frac{\dot{a}}{a}\left(\rho_\Lambda + \frac{p_\Lambda}{c^2}\right) = 0. \quad (7.8)$$

Since ρ_Λ is constant by definition, we must have

$$p_\Lambda = -\rho_\Lambda c^2. \quad (7.9)$$

The cosmological constant has a *negative* effective pressure. This means that as the Universe expands, work is done on the cosmological constant fluid. This permits its energy density to remain constant even though the volume of the Universe is increasing.

Concerning its physical interpretation, Λ is sometimes thought of as the energy density of “empty” space. In particular, in quantum physics, one possible origin is as a type of “zero-point energy,” which remains even if no particles are present, though unfortunately, particle physics theories tend to predict that the cosmological constant is far larger than observations allow. This discrepancy is known as the **cosmological constant problem**, and is one of the key unsolved problems in elementary particle physics.

It may be that the cosmological constant is only a transient phenomenon, which will disappear in the future. Another possibility, often called **dark energy** or **quintessence**, is that the cosmological constant is not actually perfectly constant but exhibits slow variation. For instance, one could assume the dark energy “fluid” to have equation of state

$$p_{de} = w\rho_{de}c^2, \quad (7.10)$$

where w could be a constant or, more generally, also itself evolve with redshift. The case $w = -1$ corresponds to a cosmological constant, while more generally accelerated expansion is possible provided $w < -1/3$ (you explored some solutions of this type in [Problem 5.3](#)). However, in this book, we will only consider the case of a perfect cosmological constant, except for a short comment in the very last Advanced Topic.

7.3 Cosmological models with Λ

The introduction of Λ has forced cosmologists to rethink some of the standard lore of cosmology, as it greatly increases the range of possible behaviors of the Universe. For instance, it is no longer necessarily true that a closed universe ($k > 0$) recollapses, nor that an open universe expands forever. In fact, if the cosmological constant is powerful enough, there need not even be a Big Bang, with the Universe instead beginning in a collapsing phase, followed by a bounce at finite size under the influence of the cosmological constant (though such models are ruled out by observations). It is also possible to have a prolonged phase where the Universe remains almost static, known as “loitering,” by arranging parameters so that the Universe closely approaches the unstable Einstein static universe.

As the Hubble parameter only provides an overall scaling factor, a useful way to parametrize possible models is to focus on the two other parameters, the present densities of matter and the cosmological constant. An excellent assumption is to assume the matter in the present Universe is pressureless. Different models can then be identified by their location in the plane of Ω_0 and

Ω_Λ , as shown in Figure 7.1.¹ We have already seen that the line $\Omega_0 + \Omega_\Lambda = 1$ gives a flat universe, and divides the plane into open and closed cosmologies.

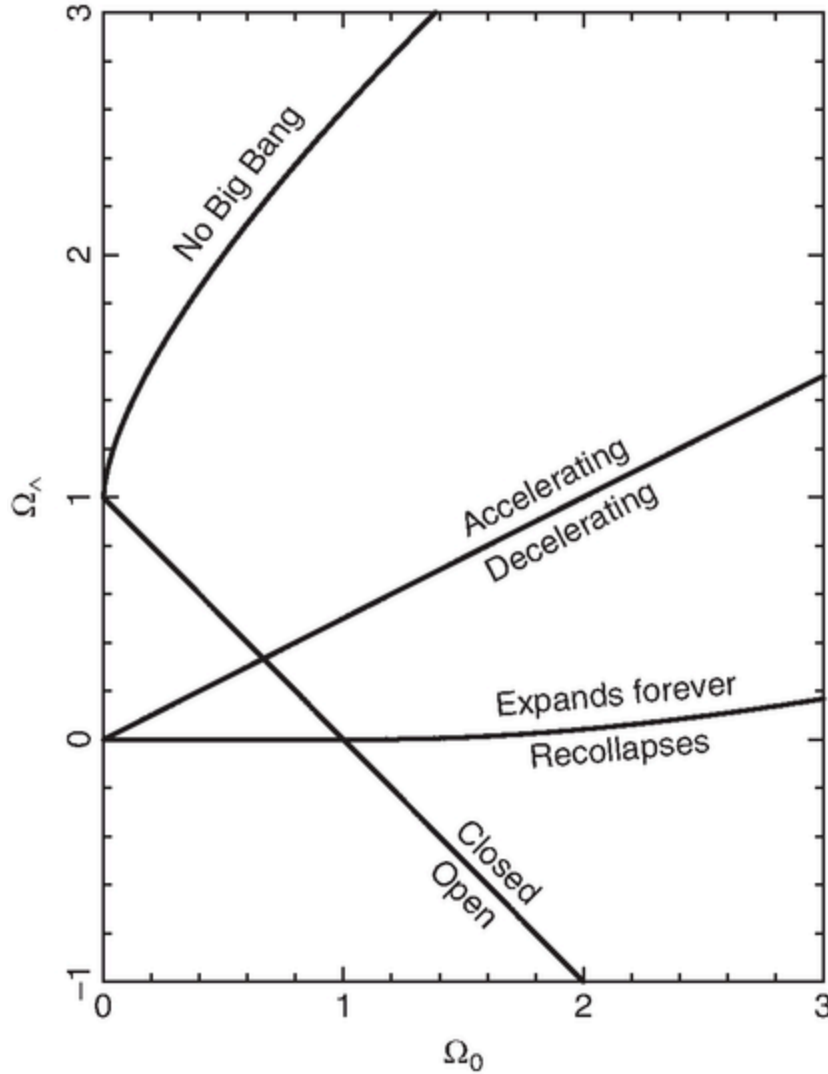


Figure 7.1 Different models for the Universe can be identified by their location in the plane showing the densities of matter and Λ . This figure indicates the main properties in different regions, with the labels indicating the behavior on each side of the dividing lines.

To identify where in the plane we have an accelerating universe, we need an expression for the deceleration parameter q_0 . A pressureless universe with a cosmological constant has

$$q_0 = \frac{\Omega_0}{2} - \Omega_\Lambda, \quad (7.11)$$

which you are asked to derive in Problem 7.3, and so we have acceleration provided $\Omega_\Lambda > \Omega_0/2$. If we additionally assume that the geometry is flat, this relation simplifies further to $q_0 = 3\Omega_0/2 - 1$, and we have acceleration if $\Omega_\Lambda > 1/3$.

The other two main properties are whether there is a Big Bang, and whether the Universe will eventually collapse or not. There are analytic expressions for these curves, shown in Figure 7.1, but they are too complicated to give here. For $\Omega_0 \leq 1$, whether there is recollapse or not depends simply on the sign of Λ , but for $\Omega_0 > 1$, the gravitational attraction of matter can overcome a small positive cosmological constant and cause recollapse.

While most cosmologists would have preferred the cosmological constant to equal zero, the Universe itself appears to have other ideas, with the observations of distant type Ia supernovae mentioned at the end of the previous chapter arguing strongly in favor of a presently accelerating universe. The observations leading to that conclusion are explored in more detail in

[Advanced Topic 2.3](#), and even if you are not planning to read that section, we suggest you have a look at [Figure A2.4](#) on page 142, which shows the observational constraints superimposed on the $\Omega_0 - \Omega_\Lambda$ plane. These observations demand inclusion of the cosmological constant; it is now regarded as an essential part of cosmological models aiming to explain observational data, and understanding its value is one of the mysteries of fundamental physics.

Problems

7.1. Suppose that the Universe contains four different contributions to the Friedmann equation, namely radiation, nonrelativistic matter, a cosmological constant, and a negative (hyperbolic) curvature. Write down the way in which each of these terms behaves as a function of the scale factor $a(t)$. Which of them would you expect to dominate the Friedmann equation at early times, and which at late times?

7.2. By considering both Friedmann and acceleration equations, and assuming a pressureless universe, demonstrate that in order to have a static universe, we must have a closed universe with a positive vacuum energy. Using either physical arguments or mathematics, demonstrate that this solution must be unstable.

7.3. Confirm the result, quoted in the main text, that a pressureless universe with a cosmological constant has a deceleration parameter given by

$$q_0 = \frac{\Omega_0}{2} - \Omega_\Lambda(t_0).$$

7.4. The most likely cosmology describing our own Universe has a flat geometry with a matter density of $\Omega_0 \simeq 0.3$ and a cosmological constant with $\Omega_\Lambda(t_0) \simeq 0.7$. What will the values of Ω and Ω_Λ be when the Universe has expanded to be five times its present size? Use an approximation suggested by this result to find the late-time solution to the Friedmann equation for our Universe. What is the late-time value of the deceleration parameter q ?

7.5. Show that in a spatially flat matter-dominated cosmology the density parameter evolves as

$$\Omega(z) = \Omega_0 \frac{(1+z)^3}{1 - \Omega_0 + (1+z)^3 \Omega_0}.$$

If our Universe has $\Omega_0 \simeq 0.3$, at what redshift did it begin accelerating?

Note

¹Beware the somewhat sloppy notation of sometimes using Ω_Λ to indicate the present value of this quantity.

Chapter 8

The Age of the Universe

One of the quantities that we can predict from a cosmological model, from the solution $a(t)$ for the expansion, is the age of the Universe t_0 . This offers an opportunity to connect the age of the Universe itself with the ages of objects within it. Historically, there was much concern as to whether the predicted age of the Universe was large enough to accommodate the ages of its contents, but in recent years such fears have largely disappeared.

Let's start with an approximate estimation. The characteristic rate of expansion is given by the Hubble parameter, so a first guess at the age of the Universe is that it is the timescale associated with the Hubble parameter, namely H_0^{-1} , since $v = H_0 r$ is basically distance equals velocity multiplied by time. This estimate is approximate, because it ignores the fact that v changes with time under the effect of gravity. We've been writing the Hubble constant as

$$H_0 = 100 \, h \, \text{km} \, \text{normal s}^{-1} \text{Baseline} \, \text{upper M p c}^{-1} \text{Baseline} \, \text{comma} \quad (8.1)$$

which isn't quite what we need, because the unusual units conceal the fact that H_0 simply has the units of $[\text{time}]^{-1}$. We have to convert the kilometers into megaparsecs (or vice versa) to cancel them out and then convert seconds to years. What we get, using values quoted on page xiii, is

$$H_0^{-1} \text{Baseline} = 9.77 \, h^{-1} \text{Baseline} \times 10^9 \text{Baseline y r s period} \quad (8.2)$$

This is known as the **Hubble time**. The uncertainty in the Hubble constant, indicated by h , means we have to tolerate an uncertainty in this crude estimate of the age of the Universe. But the message is that 10 billion years is a good first guess at the age of the Universe.

Before progressing to a better calculation, let's find out what observations tell us of the age of the Universe. The geological timescale gives us a good estimate of the age of the Earth, which is about four and a half billion years. But it is not thought that the Earth is nearly as old as the Universe. There are various ways of dating other objects in the Universe. The relative amounts in the Galactic disk of Uranium isotopes, which have lifetimes comparable to the age of the Universe, suggests an age around ten billion years – see [Problem 8.1](#). Estimates consistent with this are also found from studying the cooling of white dwarf stars after they form. However, the best method is thought to be to use the chemical evolution of stars in old globular clusters, which are believed to be among the oldest objects in the Universe. The best estimate for the oldest ones is around 13 billion years, give or take a billion, with perhaps an extra billion years to be added in before they manage to form in the first place.

It should be stressed that it is remarkable that the cosmological theory and the age measurements are in the same ballpark at all. That in itself is a strong vindication that the ideas behind the Big Bang are along the right track. Our crude estimate above seems to adequately account for the observed ages of objects in the Universe.

What happens if we try to do better with our theoretical estimates? Although the precise cosmological model describing our Universe is uncertain, we are pretty sure that it has been matter dominated (i.e. dominated by some form of pressureless material) for some considerable time, and so we can use the matter-dominated evolution to calculate the age. Let's first suppose that the Universe has the critical density. Then we already have the solution, [Equation \(5.15\)](#), which is

$$a(t) = \left(\frac{t}{t_0} \right)^{2/3} \quad \text{where } t_0 = \frac{2}{3H_0} \quad (8.3)$$

and so, the present Hubble parameter is

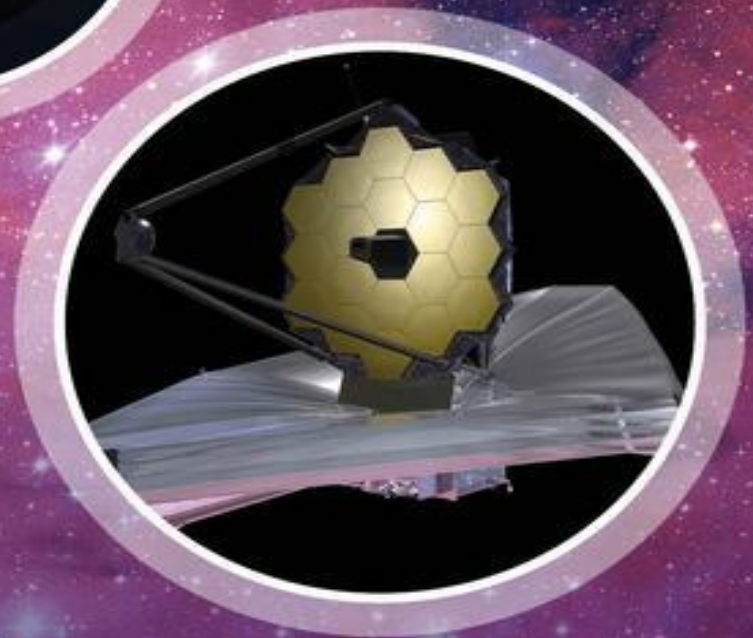
$$H_0 = \frac{2}{3t_0} \quad (8.4)$$

So, in such a universe, the age is actually shorter than our naïve estimate – it's only

$$t_0 = \frac{2}{3} H_0^{-1} = 6.51 \, h^{-1} \text{Baseline} \times 10^9 \text{Baseline y r s period} \quad (8.5)$$

The extra factor of $2/3$ has put us in an uncomfortable position; with the measured value h asymptotically equals 0.7, we predict an age of only 9.3 billion years.

**Marina Cortês
Andrew Liddle**



An Introduction to
MODERN
COSMOLOGY

Fourth Edition

WILEY

Table of Contents

[Cover](#)

[Table of Contents](#)

[Title Page](#)

[Copyright](#)

[Dedication](#)

[Preface](#)

[About the Companion Website](#)

[Chapter 1: A Brief History of Cosmological Ideas](#)

[Chapter 2: Observational Overview](#)

[2.1 In visible light](#)

[2.2 In other wavebands](#)

[2.3 Homogeneity and isotropy](#)

[2.4 The expansion of the Universe](#)

[2.5 Particles in the Universe](#)

[Problems](#)

[Notes](#)

[Chapter 3: Newtonian Gravity](#)

[3.1 The Friedmann equation](#)

[3.2 On the meaning of the expansion](#)

[3.3 Things that go faster than light](#)

[3.4 The fluid equation](#)

[3.5 The acceleration equation](#)

[3.6 On mass, energy, and vanishing factors of \$c^2\$](#)

[Problems](#)

[Note](#)

[Chapter 4: The Geometry of the Universe](#)

[4.1 Flat geometry](#)

[4.2 Spherical geometry](#)

[4.3 Hyperbolic geometry](#)

[4.4 Infinite and observable universes](#)

[4.5 Where did the Big Bang happen?](#)

[4.6 Three values of \$k\$](#)

[Problems](#)

[Notes](#)

[Chapter 5: Simple Cosmological Models](#)

[5.1 Hubble's law](#)

[5.2 Expansion and redshift](#)

[5.3 Solving the equations](#)

[5.4 Particle number densities](#)

[5.5 Evolution including curvature](#)

[Problems](#)

[Note](#)

[Chapter 6: Observational Parameters](#)

[6.1 The expansion rate \$H_0\$](#)

[6.2 The density parameter \$\Omega_0\$](#)

[6.3 The deceleration parameter \$q_0\$](#)

[Problems](#)

[Note](#)

[Chapter 7: The Cosmological Constant](#)

[7.1 Introducing \$\Lambda\$](#)

[7.2 Fluid description of \$\Lambda\$](#)

[7.3 Cosmological models with \$\Lambda\$](#)

[Problems](#)

[Note](#)

[Chapter 8: The Age of the Universe](#)

[Problems](#)

[Chapter 9: The Density of the Universe and Dark Matter](#)

[9.1 Weighing the Universe](#)

[9.2 What might the dark matter be?](#)

[9.3 Dark matter searches](#)

[Problems](#)

[Notes](#)

[Chapter 10: The Cosmic Microwave Background](#)

[10.1 Properties of the microwave background](#)

[10.2 The photon to baryon ratio](#)

[10.3 The origin of the microwave background](#)

[10.4 The origin of the microwave background \(optional advanced treatment\)](#)

[Problems](#)

[Chapter 11: The Early Universe](#)

[Problems](#)

[Chapter 12: Nucleosynthesis: The Origin of the Light Elements](#)

[12.1 Hydrogen and helium](#)

[12.2 Comparing with observations](#)

[12.3 Contrasting decoupling and nucleosynthesis](#)

[Problems](#)

[Note](#)

[Chapter 13: The Inflationary Universe](#)

[13.1 Problems with the Hot Big Bang](#)

[13.2 Inflationary expansion](#)

[13.3 Solving the Big Bang problems](#)

[13.4 How much inflation?](#)

[13.5 Inflation and particle physics](#)

[Problems](#)

[Chapter 14: The Initial Singularity](#)

[Problem](#)

[Chapter 15: Overview: The Standard Cosmological Model](#)

[15.1 Expansion](#)

[15.2 Geometry](#)

[15.3 Age](#)

[15.4 Fate](#)

[15.5 Contents](#)

[15.6 Early history](#)

[15.7 Outlook](#)

[Advanced Topic 1: General Relativistic Cosmology](#)

[A1.1 The metric of space-time](#)

[A1.2 The Einstein equations](#)

[A1.3 Aside: Topology of the Universe](#)

[Problems](#)

[Advanced Topic 2: Classic Cosmology: Distances and Luminosities](#)

[A2.1 Light propagation and redshift](#)

[A2.2 The observable Universe](#)

[A2.3 Luminosity distance](#)

[A2.4 Angular diameter distance](#)

[A2.5 Source counts](#)

[Problems](#)

[Advanced Topic 3: Neutrino Cosmology](#)

[A3.1 The massless case](#)

[A3.2 Massive neutrinos](#)

[A3.3 Neutrinos and structure formation](#)

[Problems](#)

[Advanced Topic 4: Baryogenesis](#)

[Problem](#)

[Advanced Topic 5: Structures in the Universe](#)

[A5.1 The observed structures](#)

[A5.2 Gravitational instability](#)

[A5.3 The clustering of galaxies](#)

[A5.4 Cosmic microwave background anisotropies](#)

[A5.5 The origin of structure](#)

[Problems](#)

[Advanced Topic 6 Constraining Cosmological Models](#)

[A6.1 Cosmological models and parameters](#)

[A6.2 Key cosmological observations](#)

[A6.3 Cosmological data analysis](#)

[A6.4 The Standard Cosmological Model: 2025 edition](#)

[A6.5 Tensions and hints of new physics](#)

[A6.6 The future](#)

[Problems](#)

[Bibliography](#)

[Numerical Answers and Hints to Problems](#)

[Index](#)

[End User License Agreement](#)

List of Illustrations

Chapter 2

[Figure 2.1 If viewed from above the disk, our own Milky Way galaxy would probably resemble ...](#)

[Figure 2.2 A map of galaxy positions in a narrow slice of the Universe, as measured by the...](#)

[Figure 2.3 Images of the Coma cluster of galaxies in visible/infrared light \(top\) and in X-...](#)

[Figure 2.4 A computer simulation showing the predicted distribution of matter in the Univer...](#)

[Figure 2.5 The cosmic microwave background spectrum as measured by the FIRAS experiment...](#)

[Figure 2.6 The cosmic microwave background temperature variations on the sky as measured b...](#)

[Figure 2.7 A plot of velocity versus estimated distance for a set of 1355 galaxies. A straight-line relation...](#)

[Figure 2.8 According to Hubble's law, the further away from us a galaxy is, the faster it i...](#)

[Figure 2.9 The Planck function of Equation \(2.7\). There are far more photons with very low...](#)

[Figure 2.10 The energy density distribution of a blackbody spectrum, given by Equation \(2.8\). Most of...](#)

Chapter 3

[Figure 3.1 The particle at radius \$r\$ only feels gravitational attraction from the shaded region. Any...](#)

[Figure 3.2 The comoving coordinate system is carried along with the expansion, so that any...](#)

Chapter 4

[Figure 4.1 A sketch of a spherical surface, representing positive \$k\$. A triangle is shown, which has...](#)

[Figure 4.2 A sketch of a saddle surface, representing the hyperbolic geometry obtained when \$k\$ is...](#)

Chapter 5

[Figure 5.1 A photon travels a distance \$dr\$ between two galaxies A and B.](#)

[Figure 5.2 A schematic illustration of the evolution of a universe containing radiation and...](#)

[Figure 5.3 Three possible evolutions for the Universe, corresponding to the different signs of \$k\$. The...](#)

Chapter 7

[Figure 7.1 Different models for the Universe can be identified by their location in the plane showing the...](#)

Chapter 8

[Figure 8.1 Predicted ages as fractions of the Hubble time \$H_0^{-1}\$ for open universes and for universes with...](#)

Chapter 9

[Figure 9.1 The rotation curve of the spiral galaxy NGC3198. We see that it remains roughly...](#)

[Figure 9.2 A schematic illustration of a galactic disk, with a few globular clusters, embed...](#)

Chapter 10

[Figure 10.1 The evolution of the blackbody spectrum as the Universe expands. The expansion ...](#)

[Figure 10.2 Because they have traveled toward us uninterrupted since the Universe was at 3000 K, the...](#)

Chapter 11

[Figure 11.1 A schematic illustration of the temperature–time relation, assuming \$\Omega_0 = 0.3\$ and \$h = 0.7\$...](#)

Chapter 12

[Figure 12.1 The predicted abundances of light nuclei, as a function of \$\Omega_B h^2\$ along the top and of the...](#)

Chapter 13

[Figure 13.1 An illustration of the horizon problem. We receive microwave radiation from poi...](#)

[Figure 13.2 Possible evolution of the density parameter \$\Omega_{\text{tot}}\$. There might or might not be a period before inflation,...](#)

[Figure 13.3 A schematic illustration of the inflationary solution to the horizon problem, w...](#)

Chapter 14

[Figure 14.1 The solid line shows the true \(decelerating\) scale factor. The dotted line extr...](#)

List of Tables

Chapter 4

[Table 4.1 A summary of possible geometries.](#)

Chapter 11

[Table 11.1 Different stages of the Universe's evolution \(taking \$\Omega_0 = 0.3\$ and \$h = 0.7\$ \). Some nu...](#)

Chapter 12

[Table 12.1 A comparison of nucleosynthesis and decoupling.](#)

An Introduction to Modern Cosmology

Marina Cortês and Andrew Liddle
Institute for Astrophysics and Space Sciences
Faculty of Sciences
University of Lisbon
Campo Grande
Lisbon, Portugal

Fourth Edition

WILEY

This fourth edition first published 2026

© 2026 John Wiley & Sons Ltd

Edition History

John Wiley & Sons Ltd (1e, 1998); John Wiley & Sons Ltd (2e, 2003); John Wiley & Sons Ltd (3e, 2015)

All rights reserved, including rights for text and data mining and training of artificial intelligence technologies or similar technologies. 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, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

The right of Marina Cortês and Andrew Liddle to be identified as the author(s) of this work has been asserted in accordance with law.

Registered Office(s)

John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA

John Wiley & Sons Ltd, New Era House, 8 Oldlands Way, Bognor Regis, West Sussex, PO22 0NO, UK

For details of our global editorial offices, customer services, and more information about Wiley products, visit us at www.wiley.com.

The manufacturer's authorized representative according to the EU General Product Safety Regulation is Wiley-VCH GmbH, Boschstr. 12, 69469 Weinheim, Germany, e-mail: Product.Safety@wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Trademarks: Wiley and the Wiley logo are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates in the United States and other countries and may not be used without written permission. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc. is not associated with any product or vendor mentioned in this book.

Limit of Liability/Disclaimer of Warranty

While the publisher and the authors have used their best efforts in preparing this work, including a review of the content of the work, neither the publisher nor the authors make any representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

Library of Congress Cataloging-in-Publication Data:

Names: Liddle, Andrew R. author | Cortês, Marina V. author

Title: An introduction to modern cosmology / Marina V. Cortês, Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal, Andrew Liddle, Institute for Astrophysics and Space Sciences, Faculty of Sciences, University of Lisbon, Campo Grande, Lisbon, Portugal.

Description: Fourth edition. | Hoboken, NJ, USA : John Wiley & Sons Ltd, 2026. | Includes index.

Identifiers: LCCN 2025045575 | ISBN 9781394278435 print | ISBN 9781394278503 ePDF | ISBN 9781394278442 ePUB

Subjects: LCSH: Cosmology

Classification: LCC QB981 .L567 2026

LC record available at <https://lcn.loc.gov/2025045575>

Cover Design: Wiley

Cover Images: Courtesy of NASA, © Nihit RPG/Shutterstock, © Outer Space/Shutterstock, © Vadim Sadovskii/Shutterstock

To Dawa and Kylo

Preface

The development of cosmology will no doubt be seen as one of the scientific triumphs of the twentieth century. At its beginning, cosmology hardly existed as a scientific discipline. By its end, the Hot Big Bang cosmology stood secure as the accepted description of the universe as a whole. Telescopes such as the Hubble and James Webb Space Telescopes are capable of seeing light from galaxies so distant that the light has been traveling toward us for most of the lifetime of the universe. The cosmic microwave background, a fossil relic of a time when the universe was both denser and hotter, is routinely detected and its properties examined. That our universe is presently expanding is established without doubt, and its material composition is accurately determined.

We are now in an era where understanding of cosmology has shifted from the qualitative to the quantitative. The turn of the millennium saw the establishment of what has come to be known as the Standard Cosmological Model, representing an almost universal consensus among cosmologists as to the best description of our universe. Nevertheless, it is a model with a major surprise – the belief that our universe is presently experiencing accelerated expansion. Add to that ongoing mysteries such as the properties of the so-called dark matter, which is believed to be the dominant form of matter in the universe, and it is clear that we have some way to go before we can say that a full picture of the physics of the universe is in our grasp. The first decades of the twenty-first century have seen dramatic leaps in the precision of cosmological probes, but this accuracy has yet to yield any substantial improvements to our overall understanding.

Such a bold endeavor as cosmology easily captures the imagination, and there has been increasing demand for cosmology to be taught at university in an accessible manner. Traditionally, cosmology was taught as the tail end of a general relativity course, with a derivation of the metric for an expanding universe and a few solutions. Such a course fails to capture the flavor of modern cosmology, which takes classic physical sciences like thermodynamics, atomic physics, and gravitation and apply them on a grand scale.

In fact, introductory modern cosmology can be tackled in a different way, by avoiding general relativity altogether. By a lucky chance and a subtle bit of cheating, the correct equations describing an expanding universe can be obtained from Newtonian gravity. From this basis, one can study all the triumphs of the Hot Big Bang cosmology – the expansion of the Universe, the prediction of its age, the existence of the cosmic microwave background, and the abundances of light elements such as helium and deuterium – and even go on to discuss more speculative ideas such as the inflationary cosmology.

The origin of this book, first published in 1998, was a short lecture course at the University of Sussex, around 20 lectures, taught by Andrew to students in the final year of a bachelor's degree or the penultimate year of a master's degree. The prerequisites are all very standard physics, and the emphasis is aimed at physical intuition rather than mathematical rigor. It has been further developed through his experiences of teaching cosmology at Imperial College London and at the University of Edinburgh, plus much useful feedback from readers.

The structure of the book is a central “spine,” the main chapters from 1 to 15, which provide a self-contained introduction to modern cosmology. In addition, there are six Advanced Topic chapters, each with prerequisites, which can be added to extend the course as desired. Ordinarily, the best time to tackle those Advanced Topics is immediately after their prerequisites have been attained, though they could also be included at any later stage.

Cosmology is an interesting course to teach, as it is not like most of the other subjects taught in undergraduate physics courses. There is no perceived wisdom, built up over a century or more, which provides an unquestionable foundation, as in thermodynamics, electromagnetism, and even quantum mechanics and general relativity. Each successive edition of this book has introduced new discoveries, and lecturers can expect opportunities to crop up during their course to discuss new results which impact on cosmologists' views of the universe.

You can follow our own evolving views by checking out this book's WWW Home Page. There you may find some updates on observations, a detailed solution set for instructors, and also a list of any errors in the book that we are aware of. If you are confident you've found one yourself, and it's not on the list, we'd be very pleased to hear of it.

Marina Cortês and Andrew Liddle
Lisbon
July 2025

Some fundamental constants

Newton's constant	G	$6.672 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ sec}^{-2}$
Speed of light	c	$2.998 \times 10^8 \text{ m sec}^{-1}$ or $3.076 \times 10^{-7} \text{ Mpc yr}^{-1}$
Reduced Planck constant	$\hbar = h/2\pi$	$1.055 \times 10^{-34} \text{ m}^2 \text{ kg sec}^{-1}$
Boltzmann constant	k_B	$1.381 \times 10^{-23} \text{ J K}^{-1}$ or $8.619 \times 10^{-5} \text{ eV K}^{-1}$
Radiation constant	$\alpha = \pi^2 k_B^4 / 15 \hbar^3 c^3$	$7.565 \times 10^{-16} \text{ J m}^{-3} \text{ K}^{-4}$
Electron mass-energy	$m_e c^2$	0.511 MeV
Proton mass-energy	$m_p c^2$	938.3 MeV
Neutron mass-energy	$m_n c^2$	939.6 MeV
Thomson cross-section	σ_e	$6.652 \times 10^{-29} \text{ m}^2$
Free neutron half-life	t_{half}	609 sec

Some conversion factors

$$\begin{aligned}
 1 \text{ pc} &= 3.261 \text{ light years} = 3.086 \times 10^{16} \text{ m} \\
 1 \text{ yr} &= 3.156 \times 10^7 \text{ sec} \\
 1 \text{ eV} &= 1.602 \times 10^{-19} \text{ J} \\
 1 M_{\odot} &= 1.989 \times 10^{30} \text{ kg} \\
 1 \text{ J} &= 1 \text{ kg m}^2 \text{ sec}^{-2} \\
 1 \text{ Hz} &= 1 \text{ sec}^{-1}
 \end{aligned}$$

Commonly used symbols

z	redshift	<i>defined on page</i> 12 , 48
H_0	Hubble constant	13 , 59
r	physical distance	13
v	velocity	13
f	frequency	16
T	temperature	17
k_B	Boltzmann constant	17
ϵ	energy density	19
α	radiation constant	19
G	Newton's gravitational constant	23
ρ	mass density	24
a	scale factor	26
x	comoving distance	26
k	curvature	27
p	pressure	28
	(or occasionally momentum)	15
n	number density	53
H	Hubble parameter	60
h	Hubble constant	60
	(or Planck's constant)	16
Ω_0	present density parameter	61
ρ_c	critical density	61
Ω	density parameter	62
Ω_k	curvature "density parameter"	63
q_0	deceleration parameter	63
Λ	cosmological constant	65
Ω_Λ	cosmological constant density parameter	66
t	time	73
t_0	present age	73
Ω_B	baryon density parameter	80
Y_4	helium abundance	114
d_{lum}	luminosity distance	151
d_{diam}	angular diameter distance	156
$\Delta T/T$, C_ℓ	cosmic microwave background anisotropies	181 , 182

About the Companion Website

This book is accompanied by a companion website for instructors:



<http://www.wiley.com/go/AnIntroductiontoModernCosmology4e>

This website includes:

- Full model solutions

Chapter 1

A Brief History of Cosmological Ideas

The cornerstone of modern cosmology is the belief that the place that we occupy in the Universe is in no way special. This is known as the **cosmological principle**, and it is an idea that is both powerful and simple. It is intriguing, then, that for the bulk of the history of civilization it was believed that we occupy a very special location, usually the center, in the scheme of things.

The ancient Greeks, in a model further developed by the Alexandrian Ptolemy, believed that the Earth must lie at the center of the cosmos. It would be circled by the Moon, the Sun, and the planets, and then the “fixed” stars would be yet further away. A complex combination of circular motions, Ptolemy’s epicycles, was devised in order to explain the motions of the planets, especially the phenomenon of retrograde motion where planets appear to temporarily reverse their direction of motion. It was not until the early 1500s that Copernicus stated forcefully the view, initiated nearly 2000 years before by Aristarchus, that one should regard the Earth, and the other planets, as going around the Sun. By ensuring that the planets moved at different speeds, retrograde motion could easily be explained by this theory. However, although Copernicus is credited with removing the anthropocentric view of the Universe, which placed the Earth at its center, he still believed that the Sun was at the center.

Newton’s theory of gravity put what had been an empirical science (Kepler’s discovery that the planets moved on elliptical orbits) on a solid footing, and it appears that Newton believed that the stars were also suns pretty much like our own, distributed evenly throughout infinite space, in a static configuration. However, it seems that Newton was aware that such a static configuration is unstable.

Over the next 200 years, it became increasingly understood that the nearby stars are not evenly distributed, but rather are located in a disk-shaped assembly, which we now know as the Milky Way galaxy. The Herschels were able to identify the disk structure in the late 1700s, but their observations were not perfect, and they wrongly concluded that the Solar System lay at its center. Only in the early 1900s was this convincingly overturned, by Shapley, who realized that we are some two-thirds of the radius away from the center of the Galaxy. Even then, he apparently still believed our galaxy to be at the center of the Universe. It was only in 1952 that Baade finally demonstrated that the Milky Way is a fairly typical galaxy, leading to the modern view, known as the **cosmological principle** (or sometimes the Copernican principle), that the Universe looks the same wherever you are.

It is important to stress that the cosmological principle isn’t exact. For example, no one thinks that sitting in a lecture theater is exactly the same as relaxing at the beach, and the interior of the Sun is a very different environment from the interstellar regions. Rather, it is an approximation that we believe holds better and better the larger the length scales we consider. Even on the scale of individual galaxies, it is not very good, but once we take very large regions (though still much smaller than the Universe itself), containing say a million galaxies, we expect every such region to look more or less like every other one. The cosmological principle is therefore a property of the global Universe, breaking down if one looks at local phenomena.

The cosmological principle is the basis of Big Bang cosmology. The Big Bang is the best description we have of our Universe, and the aim of this book is to explain why. The Big Bang is a picture of our Universe as an evolving entity, which was very different in the past as compared to the present. Originally, it was forced to compete with a rival idea, the Steady State Universe, which holds that the Universe does not evolve but rather has looked the same forever, with new material being created to fill the gaps as the Universe expands. However, the observations we will describe now support the Big Bang so strongly that the Steady State theory is almost never considered.

Chapter 2

Observational Overview

For most of history, astronomers have had to rely on light in the visible part of the spectrum in order to study the Universe. One of the great astronomical achievements of the twentieth century was the exploitation of the full electromagnetic spectrum for astronomical measurements. We now have instruments capable of making observations of radio waves, microwaves, infrared light, visible light, ultraviolet light, X-rays, and gamma rays, which all correspond to light waves of different (in this case increasing) frequency. We have also entered an epoch where we can go beyond the electromagnetic spectrum and receive information of other types. A remarkable feature of observations of a nearby supernova in 1987 was that it was also seen through the detection of neutrinos, an extraordinarily weakly interacting type of particle normally associated with radioactive decay. Very high-energy cosmic rays, consisting of highly relativistic elementary particles, are now routinely detected, though as yet there is no clear understanding of their astronomical origin. And, most excitingly of all, in 2015, the Advanced Laser Interferometer Gravitational-wave Observatory (LIGO) made the first direct detection of gravitational waves, ripples in space-time itself, which has become a technique to observe astronomical events such as colliding stars and black hole mergers.

The advent of large ground- and satellite-based telescopes operating in all parts of the electromagnetic spectrum has revolutionized our picture of the Universe. While there are probably gaps in our knowledge, some of which may be important for all we know, we do seem to have a consistent picture, based on the cosmological principle, of how material is distributed in the Universe. Our discussion here is brief; for a more detailed discussion of the observed Universe, see some of the more advanced textbooks listed in the Bibliography. A huge array of images can be found online; the ones shown here are just a taster for what you can find for yourself.

2.1 In visible light

Historically, our picture of the Universe was built up through ever more careful observations using visible light.

Stars: The main source of visible light in the Universe is nuclear fusion within stars. The Sun is a fairly typical star, with a mass of about 2×10^{30} kg. This is known as a solar mass, indicated $M_{\odot}$, and is a convenient unit for measuring masses.

The nearest stars to us are a few light years away, a light year being the distance (about 10^{16} m) that light can travel in a year. For historical reasons, an alternative unit, known as the **parsec** and denoted “pc,” ¹ is more commonly used in cosmology. A parsec equals 3.261 light years. In cosmology, one seldom considers individual stars, instead preferring to adopt as the smallest considered unit the conglomerations of stars known as ...

Galaxies: Our Solar System lies some way off-center in a giant disk structure known as the Milky Way galaxy. It contains a staggering hundred thousand million (10^{11}) or so stars, with masses ranging from about a tenth that of our Sun to tens of times larger. It consists of a central bulge, plus a disk of radius 12.5 kiloparsecs (kpc, equal to 10^3 pc), and a thickness of only about 0.3 kpc. We are located in the disk about 8 kpc from the center. The disk rotates slowly (and also differentially, with the outer edges moving more slowly, just as more distant planets in the Solar System orbit more slowly). At our radius, the galaxy rotates over a period of 200 million years. Because we are within it, we can’t get an image of our own galaxy, but it probably looks not unlike the M100 galaxy shown in [Figure 2.1](#).

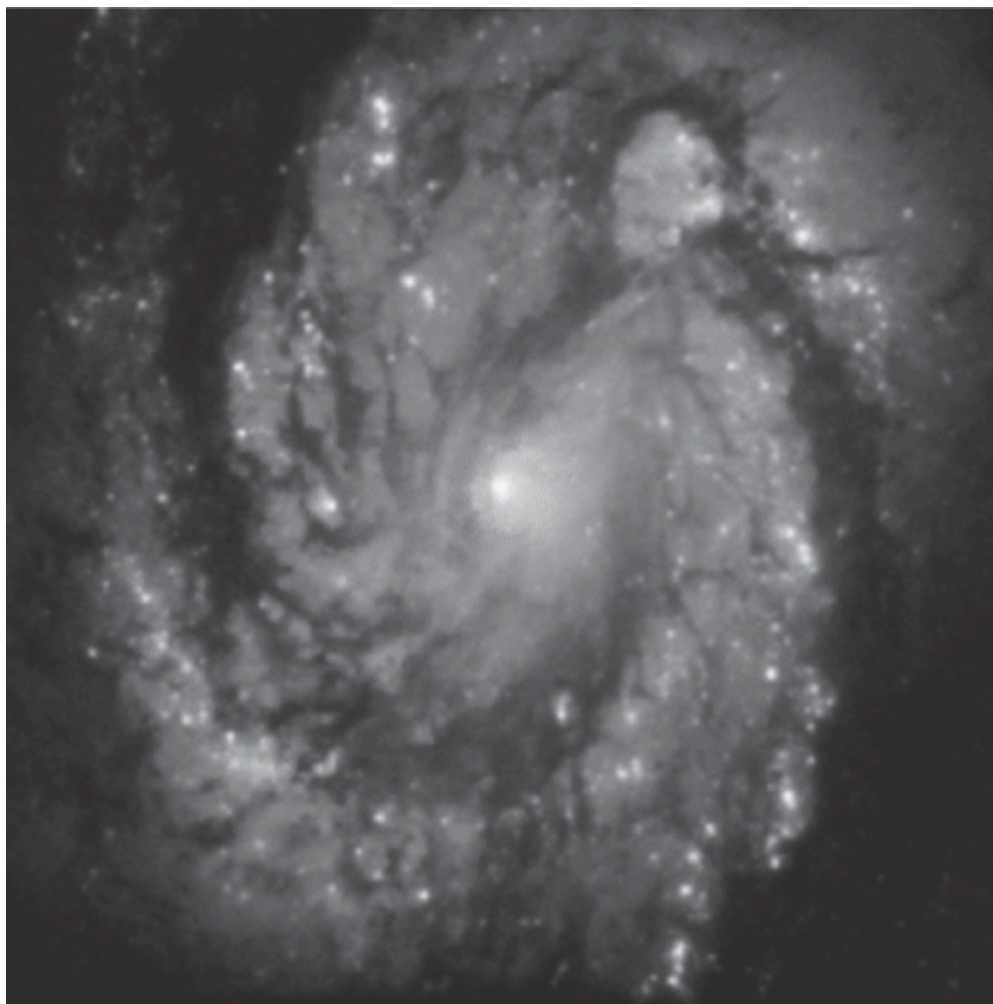


Figure 2.1 If viewed from above the disk, our own Milky Way galaxy would probably resemble the M100 galaxy, imaged here by the Hubble Space telescope. (Courtesy of NASA.)

Our galaxy is surrounded by smaller collections of stars, known as globular clusters. These are distributed more or less symmetrically about the bulge, at distances of 5–30 kpc. Typically, they each contain a million stars, and are thought to be remnants of the formation of the galaxy. As we shall discuss later, it is believed that the entire disk and globular cluster system are embedded in a larger spherical structure known as the Galactic halo.

Galaxies are the most visually striking and beautiful astronomical objects in the Universe, exhibiting a wide range of properties. However, in cosmology, the detailed structure of a galaxy is usually irrelevant, and galaxies are normally thought of as point-like objects emitting light, often broken into subclasses according to colors, luminosities, and morphologies.

The Local Group: Our galaxy resides within a small concentrated group of galaxies known as the local group. The nearest galaxy is a small irregular galaxy known as the Large Magellanic Cloud (LMC), which is 50 kpc away from the Sun. The nearest galaxy of similar size to our own is the Andromeda Galaxy, at a distance of 770 kpc. The Milky Way is one of the largest galaxies in the local group. A typical galaxy group occupies a volume of a few cubic **megaparsecs**. The megaparsec, denoted Mpc and equal to a million parsecs, is the cosmologist's favorite unit for measuring distances, because it is roughly the separation between neighboring galaxies. It equals 3.086×10^{22} m.

Clusters of Galaxies, Superclusters, and Voids: Surveying larger regions of the Universe, say on a scale of hundreds of Mpc, one sees a variety of large-scale structures, as shown in [Figure 2.2](#). This figure is not a photograph, but rather a carefully constructed map of the nearby region of our Universe, on a scale of about $1:10^{27}$! In some places, galaxies are clearly grouped into clusters of galaxies.



Circular map of the Universe details galaxy positions forming filamentary structure across angular slices, labeled with redshift and hours.

Figure 2.2 A map of galaxy positions in a narrow slice of the Universe, as measured by the Sloan Digital Sky Survey. Our galaxy is located at the center, and the survey radius is around 600 Mpc. The galaxy positions were obtained by measurement of the shift of spectral lines, as described in [Section 2.4](#). (Figure courtesy M. Blanton and the Sloan Digital Sky Survey, www.sdss3.org.)

A famous example of a cluster of galaxies is the Coma cluster, which is about 100 Mpc away from our own Galaxy. The upper panel of [Figure 2.3](#) shows a combined optical/infrared image of Coma; although the image resembles a star field, almost every source is a distinct galaxy (the main exception being two bright stars in the upper right quadrant). Coma contains perhaps 10 000 galaxies, mostly too faint to show in this image, orbiting in their common gravitational field.



Two panels detail infrared light and a labeled extended X-ray of the Coma cluster, with a scale of 10 arcmin.

Figure 2.3 Images of the Coma cluster of galaxies in visible/infrared light (top) and in X-rays (bottom), the latter being on a larger angular scale. (Courtesy of NASA/Spitzer satellite and ESA/U. Briel/MPE Garching/XMM-Newton satellite.)

However, most galaxies, sometimes called field galaxies, are not part of a cluster. Galaxy clusters are the largest gravitationally collapsed objects in the Universe, and they themselves are grouped into superclusters, joined by filaments and walls of galaxies. In between this “foamlike” structure lie large voids, some as large as 50 Mpc across. Structures in the Universe will be further described in [Advanced Topic 5](#). [Figure 2.4](#) shows an example of a computer simulation aiming to model the distribution of material within the Universe.



Simulated large scale structure of the universe details a dense web of filaments and nodes with a scale bar for 125 Mpc per hour.

Figure 2.4 A computer simulation showing the predicted distribution of matter in the Universe on large scales. (Courtesy of V. Springel and the Virgo Consortium.)

Large-scale Smoothness: Only once we get to scales of hundreds of megaparsecs or more does the Universe begin to appear smooth, as revealed by extremely large galaxy surveys such as the 2dF galaxy redshift survey and the Sloan Digital Sky Survey. Such surveys do not find any huge structures on scales greater than those described above; the galaxy superclusters and voids are believed to be the biggest structures in the present Universe. The observation that the Universe does indeed become smooth on the largest scales, the cosmological principle, is the underpinning of modern cosmology. It is interesting that, while the smoothness of the matter distribution on large scales has been a key assumption of cosmology for decades, it is only fairly recently that a convincing observational demonstration has become possible.

2.2 In other wavebands

Observations using visible light provide us with a good picture of what’s going on in the present-day Universe. However, many other wavebands make vital contributions to our understanding, and in particular our best knowledge of cosmology comes not from visible light but from microwaves.

Microwaves: For cosmology, this is by far the most important waveband. Penzias & Wilson’s accidental discovery in 1965 that the Earth is bathed in microwave radiation, with a blackbody spectrum at a temperature of around 3 K, was and is one of the most powerful pieces of information in support of the Big Bang theory, around which cosmology is based. This is now known as the cosmic microwave background or simply CMB. Observations by the Far InfraRed Absolute Spectrometer (FIRAS) experiment on board the COsmic Background Explorer (COBE) satellite have confirmed that the radiation is extremely close to the blackbody form at a temperature 2.725 ± 0.001 K. This data is shown in [Figure 2.5](#). Furthermore, the temperature coming from different parts of the sky is astonishingly uniform, and this presents the best evidence that we can use the cosmological principle as the foundation of cosmology. Even more importantly, it is possible to identify tiny variations, only one part in a hundred thousand, between the intensities of the microwaves coming from different directions. These were first discovered by the COBE satellite in 1992, and [Figure 2.6](#) shows precision measurements published by the *Planck* Satellite collaboration in 2018. It is believed that these patterns, known as cosmic microwave background anisotropies, are intimately related to the origin of structure in the Universe, and at present such measurements give us by far the most accurate understanding of cosmology. This topic is explored further in [Advanced Topics 5](#) and [6](#).



Graph of Intensity versus Waves per centimeter, plotting cosmic microwave background curve data points with error bars multiplied by 400.

Figure 2.5 The cosmic microwave background spectrum as measured by the FIRAS experiment on the COBE satellite. The error bars are so small that they have been multiplied by 400 to make them visible on this plot, and the best-fit blackbody spectrum at **2.725 K**, shown by the line, is an excellent fit.



Map of cosmic microwave temperature variations across an oval projection scaled from negative 300 to 300 microkelvin.

Figure 2.6 The cosmic microwave background temperature variations on the sky as measured by the *Planck* Satellite, with scale in microKelvin. (Courtesy of ESA/Planck collaboration.)

Infrared: Carrying out surveys in the infrared part of the spectrum, as was first done by the highly successful InfraRed Astronomical Satellite (IRAS) in the 1980s, is an excellent way of spotting young galaxies, in which star formation is at an early stage. Infrared surveys pick up a somewhat different population of galaxies to surveys carried out in optical light, though obviously the brightest galaxies are seen in both. Infrared is particularly good for looking through the dust in our own Galaxy to see distant objects, as it is absorbed and scattered much less strongly than visible radiation. Accordingly, it is best for studying the region close to our Galactic plane, where obscuration by dust is strongest. Infrared is also vital for studying the most distant galaxies, whose emission has been shifted predominantly into the infrared part of the spectrum by the expansion of the Universe after the light was emitted; NASA's James Webb Space Telescope is the current state-of-the-art for such studies. The far infrared can probe emission from interstellar dust grains to learn about dust-enshrouded environments that are opaque at visible and near-infrared wavelengths.

X-rays: These are a vital probe of clusters of galaxies; in between the galaxies lies gas so hot that it emits in the X-ray part of the spectrum, corresponding to temperatures of tens of millions of Kelvin. This gas is thought to be a remnant material from the formation of the galaxies, which failed to collapse to form stars. X-ray emission from the Coma galaxy cluster is shown in the lower panel of [Figure 2.3](#). The individual galaxies seen in the visible light image in the upper panel are almost all invisible in X-rays, with the bright diffuse X-ray emission from the hot gas dominating the image.

Radio Waves: A powerful way of gaining high-resolution maps of distant galaxies is by mapping in the radio part of the spectrum. An important probe of cosmology is observations of the “21 cm” emission line, due to spin-flip of the electron in a hydrogen atom, enabling maps of the distribution of neutral hydrogen gas in the distant Universe. This can probe both the early development of structures in the Universe, and the transition of the Universe to its present state of high ionization of the intergalactic medium. The Square Kilometre Array, nearing completion as we write, will take these studies to the next level.

2.3 Homogeneity and isotropy

The evidence that the Universe becomes smooth on a large scale supports the use of the cosmological principle. It is therefore believed that our large-scale Universe possesses two important properties: **homogeneity** and **isotropy**. Homogeneity is the statement that the Universe looks the same at each point, while isotropy states that the Universe looks the same in all directions.

These do not automatically imply one another. For example, a universe with a uniform magnetic field is homogeneous, as all points are the same, but it fails to be isotropic because directions along the field lines can be distinguished from those perpendicular to them. Alternatively, a spherically symmetric distribution, viewed from its central point, is isotropic but not necessarily homogeneous. However, if we require that a distribution is isotropic at *every* point, then that does enforce homogeneity as well.

As mentioned earlier, the cosmological principle is not exact, and so our Universe does not respect exact homogeneity and isotropy. Indeed, the study of departures from homogeneity is currently the most prominent research topic in cosmology. We'll introduce this in [Advanced Topic 5](#), but in the main body of this book we are concerned only with the behavior of the Universe as a whole, and so will be assuming large-scale homogeneity and isotropy.

2.4 The expansion of the Universe

A key piece of observational evidence in cosmology is that almost everything in the Universe appears to be moving away from us, and the further away something is, the more rapid its recession appears to be. These velocities are measured via the **redshift**, which is basically the Doppler effect applied to light waves. Galaxies have a set of absorption and emission lines identifiable in their spectra, whose characteristic frequencies are well known. However, if a galaxy is moving toward us, the light waves get crowded together, raising the frequency. Because blue light is at the high-frequency end of the visible spectrum, this is known as a blueshift. If the galaxy is receding, the characteristic lines move toward the red end of the spectrum and the effect is known as a redshift. This technique was first used to measure a galaxy's velocity by Vesto Slipher around 1912, and was applied systematically by one of the most famous cosmologists, Edwin Hubble, in the following decades.

It turns out that almost all galaxies are receding from us, so the standard terminology is redshift z , defined by

$$z = \frac{\lambda_{\text{obs}} - \lambda_{\text{em}}}{\lambda_{\text{em}}} \quad (2.1)$$

where λ_{em} and λ_{obs} are the wavelengths of light at the points of emission (the galaxy) and observation (us). If a nearby object is receding at a speed v , then its redshift is

$$z = \frac{v}{c} \quad (2.2)$$

where c is the speed of light. ² Figure 2.7 shows velocity against distance, a plot known as the Hubble diagram, for a sample of 1355 galaxies.

Scatter plot of galaxy velocity versus estimated distance, plotted in Mpc, details a scattered and linear trend implying Hubble's law.

Figure 2.7 A plot of velocity versus estimated distance for a set of 1355 galaxies. A straight-line relation implies Hubble's law. The considerable scatter is due to observational uncertainties and random galaxy motions, but the best-fit line accurately gives Hubble's law. (The bold italic x -axis scale assumes a particular value of bold italic upper H_0 .)

Hubble realized that his observations, which were of course much less extensive than those available to us now, showed that the velocity of recession was proportional to the distance of an object from us:

$$v = H_0 r \quad (2.3)$$

This is known as **Hubble's law**, and the constant of proportionality H_0 is known as **Hubble's constant**. Hubble's law isn't exact, as the cosmological principle doesn't hold perfectly for nearby galaxies, which typically possess some random motions known as peculiar velocities. But it does describe the average behavior of galaxies extremely well. Hubble's law gives the picture of our Universe illustrated in Figure 2.8, where the nearby galaxies have the smallest velocity relative to ours. Over the years, many attempts have been made to determine accurate values for the proportionality constant, which is now known to within a few percent accuracy. Still, as we will see in Chapter 6 and Advanced Topic 6.5, there remains some controversy to this day.

Diagram of a central point and six surrounding points, each with an arrow details speed proportional to distance from the center.

Figure 2.8 According to Hubble's law, the further away from us a galaxy is, the faster it is receding.

At first sight, it seems that the cosmological principle must be violated if we observe everything to be moving away from us, since that apparently places us at the center of the Universe. However, nothing could be further from the truth. In fact, *every* observer sees all objects rushing away from them with velocity proportional to distance. It is perhaps easiest to convince yourself of this by setting up a square grid with recession velocity proportional to distance from the central grid point. Then, transform the frame of reference to a nearby grid-point, and you'll find that the Hubble law still holds about the new "centre." This only works because of the linear relationship between velocity and distance; any other law and it wouldn't work.

So, although expanding, the Universe looks just the same whichever galaxy we choose to imagine ourselves within. A common analogy is to imagine baking a cake with raisins in it, or blowing up a balloon with dots on its surface. As the cake rises (or the balloon is inflated), the raisins (or dots) move apart. From each one, it seems that all the others are receding, and the further away they are, the faster that recession is.

Because everything is flying away from everything else, we conclude that in the distant past everything in the Universe was much closer together. Indeed, trace the history back far enough and everything comes together. The initial explosion is known as the **Big Bang**, and a model of the evolution of the Universe from such a beginning is known as the **Big Bang Cosmology**. Later on, we will find out why it is commonly called the **Hot Big Bang**.

2.5 Particles in the Universe

2.5.1 What particles are there?

Everything in the Universe is made up of fundamental particles, and the behavior of the Universe as a whole depends on the properties of these particles.

One crucial question is whether a particle is moving relativistically or not. Any particle has two contributions to its energy, one being kinetic energy and the other being mass–energy, which combine to give

$$E_{\text{total}} = mc^2 + p^2 c^2 \quad (2.4)$$

where m is the particle rest mass and p the particle momentum. If the mass–energy dominates, the particle will be moving at much less than the speed of light, and we say it is nonrelativistic. In that limit, we can carry out an expansion

$$E_{\text{total}} = mc^2 \left(1 + \frac{p^2}{2m^2 c^2} \right) \approx mc^2 + \frac{p^2}{2m} \quad (2.5)$$

We recognize the first term as Einstein’s famous $E = mc^2$, known as the rest mass–energy as it is the energy of the particle when it is stationary. The second term is the usual kinetic energy ($p = mv$ in the nonrelativistic limit). If the mass–energy does not dominate, the particle will be moving at a substantial fraction of the speed of light and so is relativistic. In particular, any particle with zero rest mass is always relativistic and moves at the speed of light, the simplest example being light itself.

Let’s review the nature of the particles which are believed to exist in our Universe.

Baryons

We ourselves are built from atoms, the bulk of whose mass is attributable to the protons and neutrons in the atomic nuclei. Protons and neutrons are believed to be made up of more fundamental particles known as quarks, a proton being made of two up quarks and a down quark, while a neutron is an up and two downs. A general term for particles made up of three quarks is **baryons**. Of all the possible baryons, only the proton and neutron can be stable,³ and so these are thought to be the only types of baryonic particle significantly represented in the Universe. Yet another piece of terminology, nucleon, refers to just protons and neutrons, but I’ll follow the standard practice of using the term baryon. In particle physics units, the mass–energies of a proton and a neutron are 938.3 and 939.6 MeV, respectively, where “MeV” is a Mega-electron volt, a unit of energy equal to a million electron volts (eV) and rather more convenient than a Joule.

Although electrons are certainly not made from quarks, they are traditionally also included under the title baryon by cosmologists (to the annoyance of particle physicists). A crucial property of the Universe is that it is charge neutral, so there must be one electron for every proton. Weighing in at a puny 0.511 MeV, well under a thousandth of a proton mass, the contribution of electrons to the total mass is a tiny fraction, not meriting separate discussion.

In the present Universe, baryons are typically moving nonrelativistically, meaning that their kinetic energy is much less than their mass–energy.

Radiation

Our visual perception of the Universe comes from electromagnetic radiation, and such radiation, at a large variety of frequencies, pervades the Universe. In the quantum mechanical view of light, it can be thought of as made up of individual particles – like packets of energy – known as **photons** and usually indicated by the symbol γ . Photons propagate, naturally enough, at the speed of light; since they have zero rest mass, their total energy is always given by their kinetic energy, and is related to their frequency f by

$$E = hf \quad (2.6)$$

where h is Planck’s constant.

Photons can interact with the baryons and electrons; for example, a high-energy photon can knock an electron out of an atom (a process known as ionization), or can scatter off a free electron (known as Thomson scattering in the non-relativistic case $hf \ll mc^2$, otherwise Compton scattering). The more energetic the photons are, the more devastating their effects on other particles.

Neutrinos

Neutrinos are extremely weakly interacting particles, produced, for example, in radioactive decay. There is now significant experimental evidence that they possess a non-zero rest mass, but so far there has been no clear evidence that this mass is large enough to have observable cosmological effects. Hence, for now, it remains a working assumption in cosmology to treat them as effectively massless. We will adopt that assumption for the main body of this book, and in that case they, like photons, are always relativistic. The combination of photons and neutrinos makes up the relativistic material in our Universe. Confusingly, sometimes the term “radiation” is used to refer to all the relativistic material.

There are three types of neutrino, the electron neutrino, muon neutrino, and tau neutrino, and if they are indeed all massless, they should all exist in our Universe. Unfortunately, their interactions are so weak that for now there is no hope of detecting cosmological neutrinos directly. Originally, their presence was inferred on purely theoretical grounds, though we will see that the existence of the cosmic neutrino background may be inferred indirectly by some cosmological observations.

Because they are so weakly interacting, the experimental limits on the neutrino masses, especially of the latter two types, are quite weak, and it is in fact perfectly possible that they are massive enough to be nonrelativistic. The possible effects of neutrino masses are explored in [Advanced Topic 3](#). Cosmological effects of neutrino masses might well have been spotted by the time you read this book.

Dark matter

In this book, we'll encounter one further kind of particle that may exist in our Universe, which is not part of the Standard Model of particle theory. It is known as dark matter, and its properties are highly uncertain and a matter of constant debate among cosmologists. We'll return to it in [Chapter 9](#).

2.5.2 Thermal distributions and the blackbody spectrum

We end this section with some discussion of the physics of radiation. If this is unfamiliar to you, the details aren't all that crucial, though some of the final results will be used later in the book.

If particles are frequently interacting with one another, then the distribution of their energies can be described by equilibrium thermodynamics. In a thermal distribution, interactions are frequent, but a balance has been reached so that all interactions proceed equally frequently in both the forward and backward directions, so that the overall distribution of particle numbers and energies remains fixed. The number of particles of a given energy then depends only on the temperature.

The precise distribution depends on whether the particles considered are fermions, which obey the Pauli exclusion principle, or bosons, which do not. In this book, the most interesting case is that of photons, which are bosons, and their characteristic distribution at temperature T is the **Planck** or **blackbody** spectrum. Photons have two possible polarizations, and each has an occupation number per mode N given by the Planck function

$$N = \frac{1}{\exp\left(\frac{hf}{k_B T}\right) - 1} \quad (2.7)$$

where h is Planck's constant and k_B is one of the fundamental constants of nature, the **Boltzmann constant**, whose value is $1.381 \times 10^{-23} \text{ J K}^{-1} = 8.619 \times 10^{-5} \text{ eV K}^{-1}$.

To interpret this equation, remember that hf is the photon energy. The purpose of the Boltzmann constant is to convert temperature into a characteristic energy $k_B T$. Below this characteristic energy, $hf \ll k_B T$, it is easy to make photons and the occupation number is large (as photons are bosons, the Pauli exclusion principle doesn't apply and there may be arbitrarily many photons in a given mode). Above the characteristic energy, $hf \gg k_B T$, it is energetically unfavorable to make photons, and the number is exponentially suppressed, as shown in [Figure 2.9](#).

 Graph of Planck function N versus hf over $k_B T$ details a sharp decay, reflecting far more low-energy photons than high-energy.

Figure 2.9 The Planck function of [Equation \(2.7\)](#). There are far more photons with very low energy than very high energy.

More interesting than the number of photons in a mode is the distribution of energy among the modes. We focus on the energy per unit volume, known as the **energy density** ϵ . Because there are very few photons with

$hf \gg k_B T$ there isn't much energy at high frequencies. But, despite their large numbers, there also isn't much total energy at low frequencies

$hf \ll k_B T$, both because those photons have less energy each $\epsilon = hf$, and because their wavelength is longer and so each photon occupies a greater volume. With quite a bit of work, the energy density in a frequency interval df about f can be shown to be

$$\epsilon(f) df = \frac{8\pi h}{c^3} \frac{f^3 df}{\exp\left(\frac{hf}{k_B T}\right) - 1} \quad (2.8)$$

which tells us how energy is distributed among the different frequencies. We see in [Figure 2.10](#) that the peak of the distribution is at $f_{\text{peak}} \approx 2.8 k_B T$.

corresponding to an energy

$\epsilon_{\text{peak}} = hf_{\text{peak}}$ asymptotically equals $2.8 k_B T$. That is to say, the total energy in the radiation is dominated by photons with energies of order

$k_B T$. Indeed, the mean energy of a photon in this distribution is

$\epsilon_{\text{mean}} = 3 k_B T$.



Graph of the normalized density versus $hf/k_B T$ details the energy density distribution of a spectrum.

Figure 2.10 The energy density distribution of a blackbody spectrum, given by Equation (2.8). Most of the energy is contributed by photons of energy

hf_{peak} .

When we study the early history of the Universe, an important question will be how this typical energy compares to atomic and nuclear binding energies.

A further quantity of interest will be the total energy density of the blackbody radiation, obtained by integrating Equation (2.8) over all frequencies. Setting $y = hf/k_B T$ quickly leads to

$$\epsilon_{\text{rad}} = \frac{8\pi k_B^4}{15} T^4 \int_0^\infty \frac{y^3 dy}{e^y - 1} \quad (2.9)$$

The integral is not particularly easy to compute, but you might like to try it as a challenge. The answer is

$\pi^4/15$, giving a radiation energy density

$$\epsilon_{\text{rad}} = \alpha T^4 \quad (2.10)$$

where the radiation constant α is defined as

$$\alpha = \frac{\pi^5 k_B^4}{15 h^3} = 7.565 \times 10^{-16} \text{ J m}^{-3} \text{ K}^{-4} \quad (2.11)$$

Here, h is the reduced Planck constant.

Problems

2.1. Suppose that the Milky Way galaxy is a typical size, containing say 10^{11} stars, and that galaxies are typically separated by a distance of one megaparsec. Estimate the density of the Universe in SI units. How does this compare with the density of the Earth?

$1 \text{ Mpc} \approx 3.1 \times 10^{22} \text{ m}$

2.2. In the real Universe, the expansion is not completely uniform. Rather, galaxies exhibit some random motion relative to the overall Hubble expansion, known as their *peculiar velocity* and caused by the gravitational pull of their near neighbors. Supposing that a typical (e.g. root mean square) galaxy's peculiar velocity is 600 km s^{-1} , how far away would a galaxy have to be before it could be used to determine the Hubble constant to 10% accuracy, supposing

a. the true value of the Hubble constant is $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$?

b. the true value of the Hubble constant is $50 \text{ km s}^{-1} \text{ Mpc}^{-1}$?

Assume in your calculation that the galaxy distance and redshift could be measured exactly. Unfortunately, that is not true of real observations.

2.3. What evidence can you think of to support the assertion that the Universe is charge neutral, and hence contains an equal number of protons and electrons?

2.4. The binding energy of the electron in a hydrogen atom is 13.6 eV. What is the frequency of a photon with this energy? At what temperature does the mean photon energy equal this energy?

2.5. The peak of the energy density distribution of a blackbody at

$\epsilon_{\text{peak}} = 2.8 k_B T$ divided by h implies that $hf_{\text{peak}}/k_B T$ is a constant. Evaluate this constant in SI units (see page xiv for

useful numbers). The Sun radiates approximately like a blackbody with $T_{\text{Sun}} \approx 5800 \text{ K}$. Compute λ_{peak} for solar radiation. Where in the electromagnetic spectrum does the peak emission lie?

2.6. The cosmic microwave background has a blackbody spectrum at a temperature of 2.725 K . Repeat [Problem 2.5](#) to find the peak frequency of its emission, and also find the corresponding wavelength and compare to [Figure 2.5](#). Confirm that the peak emission lies in the microwave part of the electromagnetic spectrum. Finally, compute the total energy density of the microwave background.

Notes

¹ A parsec is defined as the distance at which the mean distance between the Earth and Sun subtends a second of arc. The mean Earth–Sun distance (called an Astronomical Unit) is $1.496 \times 10^{11} \text{ m}$, and dividing that by $\tan(1 \text{ arcsec})$ gives $1 \text{ pc} = 3.086 \times 10^{16} \text{ m}$.

² This formula ignores special relativity and so is valid only for speeds v much less than c . If you're interested, the special relativity result, of which this is an expansion for small v/c , is

$$1 + z = \sqrt{\frac{1 + v/c}{1 - v/c}}$$

However, for distant objects in cosmology, there are further considerations concerning the propagation time of the light and how the relative velocity might change during it, and so this expression should not be used.

³ The proton's lifetime is known to be either infinite, corresponding to the proton being absolutely stable, or much longer than the age of the Universe so that they are effectively stable. Isolated neutrons are unstable (decaying into a proton, an electron, and an anti-neutrino), but those bound in nuclei may be stable; this will prove crucial in [Chapter 12](#).

Chapter 3

Newtonian Gravity

It is perfectly possible to discuss cosmology without having already learned general relativity. In fact, the most crucial equation, the Friedmann equation, which describes the expansion of the Universe, turns out to be the same when derived from Newton's theory of gravity as it is when derived from the equations of general relativity. The Newtonian derivation is, however, some way from being completely rigorous, and general relativity is required to fully patch it up, a detail that need not concern us at this stage. Readers with some familiarity with general relativity may wish to study the relativistic derivation given in [Advanced Topic 1](#).

In Newtonian gravity, all matter attracts, with the force exerted by an object of mass M on one of mass m given by the famous relationship

$$F = \frac{GMm}{r^2}, \quad (3.1)$$

where r is the distance between the objects and G is Newton's gravitational constant. That is, gravity obeys an inverse square law. Because a force on an object induces an acceleration which is also proportional to its mass, via $F = ma$, the acceleration an object feels under gravity is *independent* of its mass.

The force exerted means there is a gravitational potential energy

$$V = -\frac{GMm}{r}, \quad (3.2)$$

with the force being in the direction which decreases the potential energy the fastest. Like the electric potential of two opposite charges, the gravitational potential is negative, favoring the two objects being close together. But with gravity there is no analog of the repulsion of like charges. Gravity always attracts.

The derivation of the Friedmann equation requires a famous result due originally to Newton, which we won't attempt to prove here. This result states that in a spherically symmetric distribution of matter, a particle feels no force at all from the material at greater radii, and the material at smaller radii gives exactly the force which one would get if all the material was concentrated at the central point. This property arises from the inverse square law; the same results exist for electromagnetism. One example of its use is that the gravitational (or electromagnetic) force outside a spherical object of unknown density profile depends only on the total mass (charge). Another is that an "astronaut" inside a spherical shell feels no gravitational force, not only if they are at the center but if they are at any position inside the shell.

3.1 The Friedmann equation

The Friedmann equation describes the expansion of the Universe, and is therefore the most important equation in cosmology. One of the routine tasks of a working cosmologist is solving this equation under different assumptions concerning the material content of the Universe. To derive the Friedmann equation, we need to compute the gravitational potential energy and the kinetic energy of a test particle (it doesn't matter which one, since everywhere in the Universe is the same according to the cosmological principle), and then use energy conservation.

Let's consider an observer in a uniform expanding medium, with mass density ρ , the mass density being the mass per unit volume. We begin by realizing that because the Universe looks the same from anywhere, we can consider any point to be its center. Now, consider a particle a distance r away with mass m , as shown in [Figure 3.1](#). (By "particle," we really mean a small volume containing the mass m .) Due to Newton's theorem, this particle only feels a force from the material at smaller radii, shown as the shaded region. This material has total mass given by $M = 4\pi\rho r^3/3$, contributing a force

$$F = \frac{GMm}{r^2} = \frac{4\pi G\rho r m}{3}, \quad (3.3)$$

and our particle has a gravitational potential energy

$$V = -\frac{GMm}{r} = -\frac{4\pi G\rho r^2 m}{3}. \quad (3.4)$$

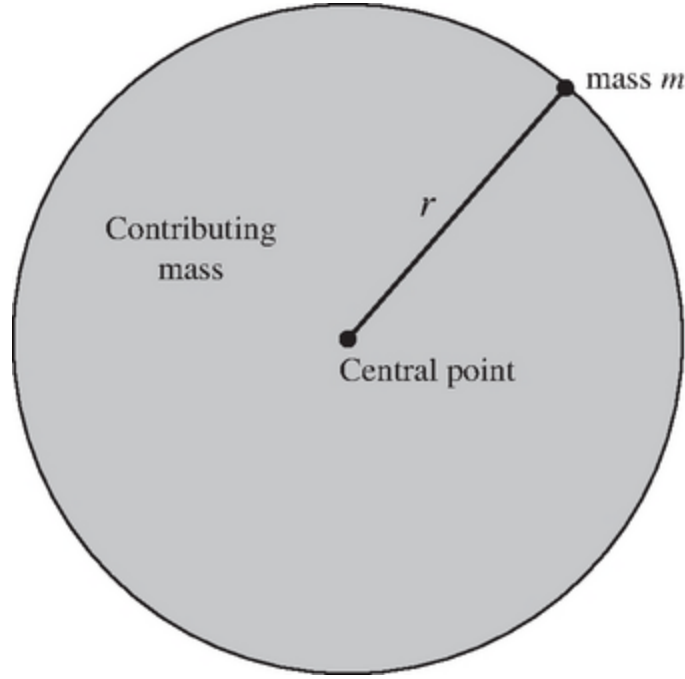


Figure 3.1 The particle at radius r only feels gravitational attraction from the shaded region. Any gravitational attraction from the material outside cancels out, according to Newton's theorem.

The kinetic energy is easy; the velocity of the particle is $\dot{r}$ (we'll always use dots to mean time derivatives) giving

$$T = \frac{1}{2} m \dot{r}^2. \quad (3.5)$$

The equation describing how the separation r changes can now be derived from energy conservation for that particle, namely

$$U = T + V, \quad (3.6)$$

where U is a constant. Note that U need not be the same constant for particles separated by different distances. Substituting gives

$$U = \frac{1}{2} m \dot{r}^2 - \frac{4\pi}{3} G\rho r^2 m. \quad (3.7)$$

This equation gives the evolution of the separation r between the two particles.

We now take a crucial step in this derivation, which is to realize that this argument applies to any two particles, because the Universe is homogeneous. This allows us to change to a different coordinate system, known as **comoving coordinates**. These are coordinates which are carried along with the expansion. Because the expansion is uniform, the relationship between real distance $\vec{r}$ and the comoving distance, which we can call $\vec{x}$, can be written

$$\vec{r} = a(t) \vec{x}, \quad (3.8)$$

where the homogeneity property has been used to ensure that a is a function of time alone. Note that these distances have been written as vector distances. What you should think of when studying this equation is a coordinate grid which expands with time, as shown in [Figure 3.2](#). The galaxies remain at fixed locations in the $\vec{x}$ coordinate system. The original $\vec{r}$ coordinate system, which does not expand, is usually known as **physical coordinates**.

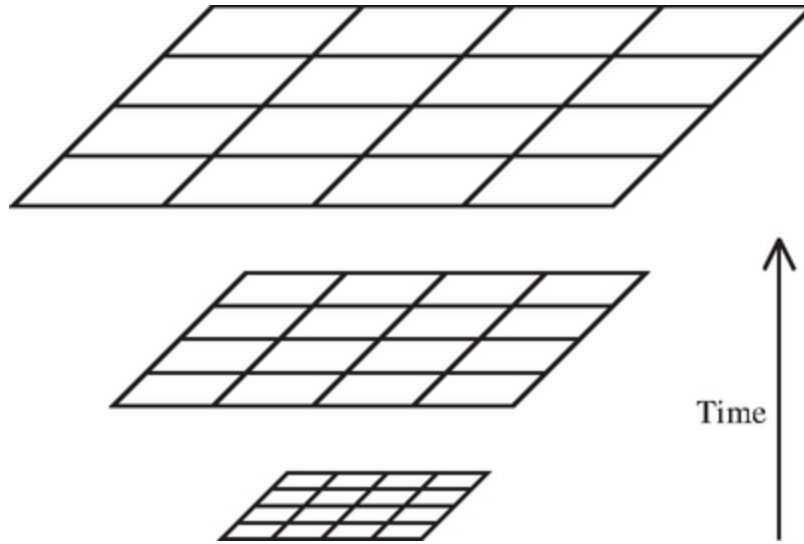


Figure 3.2 The comoving coordinate system is carried along with the expansion, so that any objects remain at fixed coordinate values.

The quantity $a(t)$ is a crucial one, and is known as the **scale factor of the Universe**. It measures the universal expansion rate. It is a function of time alone, and it tells us how physical separations are growing with time, since the coordinate distances $\vec{x}$ are by definition fixed. For example, if, between times t_1 and t_2 , the scale factor doubles in value, that tells us that the Universe has expanded in size by a factor two, and it will take us twice as long to get from one galaxy to another.

We can use the scale factor to rewrite [Equation \(3.7\)](#) for the expansion. Substituting [Equation \(3.8\)](#) into it, remembering $\dot{x} = 0$ by definition as objects are fixed in comoving coordinates, gives

$$U = \frac{1}{2}m\dot{a}^2x^2 - \frac{4\pi}{3}G\rho a^2x^2m. \quad (3.9)$$

Multiplying each side by $2/ma^2x^2$ and rearranging the terms then gives

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2}, \quad (3.10)$$

where $kc^2 = -2U/mx^2$. This is the standard form of the **Friedmann equation**, and it will appear frequently throughout this book. In this expression, k must be independent of x since all the other terms in the equation are, otherwise homogeneity will not be maintained. So, in fact, we learn that homogeneity requires that the quantity U , while constant for a given particle, does indeed change if we look at different separations x , with $U \propto x^2$.

Finally, since $k = -2U/mc^2x^2$ which is time independent (as the total energy U is conserved, and the comoving separation x is fixed), we learn that k is just a constant, unchanging with either space or time. It has the units of $[\text{length}]^{-2}$. An expanding Universe has a unique value of k , which it retains throughout its evolution. In [Chapter 4](#), we will see that k tells us about the geometry of the Universe, and it is often called the curvature.

3.2 On the meaning of the expansion

So what does the expansion of the Universe mean? Well, let's start with what it does *not* mean. It does not mean that your body is gradually going to get bigger with time (and certainly isn't an excuse if it does). That does not mean that the Earth's orbit is going to get further from the Sun. It doesn't even mean that the stars within our galaxy are going to become more widely separated with time.

But it does mean that distant galaxies are getting further apart.

The distinction lies in whether the motion of objects is governed by the cumulative gravitational effect of a homogeneous distribution of matter between them, as shown in [Figure 3.1](#). The atoms in your body are not: their separation is dictated by the strength of chemical bonds, with gravity playing no significant role. So, molecular structures will not be affected by the expansion. Likewise, the Earth's motion in its orbit is completely dominated by the gravitational attraction of the Sun (with a minor contribution from the other planets). And even the stars in our galaxy are orbiting in the common gravitational potential well that they themselves create, and are not moving apart relative to one another. The common feature of these environments

is that they are the ones of considerable excess density, very different from the smooth distribution of matter we used to derive the Friedmann equation.

But if we go to large enough scales, in practice tens of megaparsecs, the Universe does become effectively homogeneous and isotropic, with the galaxies flying apart from one another in accordance with the Friedmann equation. It is on these large scales that the expansion of the Universe is felt, and on which the cosmological principle applies.

3.3 Things that go faster than light

A common question that concerns people is whether faraway galaxies are receding from us faster than the speed of light. That is to say, if velocity is proportional to distance, then if we consider galaxies far enough away, can we not make the velocity as large as we like, in violation of special relativity?

The answer is that, indeed, in our theoretical predictions, distant objects can *appear* to move away faster than the speed of light. However, it is space itself which is expanding. There is no violation of causality, because no signal can be sent between such galaxies. Further, special relativity is not violated, because it refers to the relative speeds of objects passing each other, and cannot be used to compare the relative speeds of distant objects.

One way to think of this is to imagine a colony of ants on a balloon. Suppose that the fastest the ants can move is a centimeter per second. If any two ants happen to pass each other, their fastest relative speed would be two centimeters per second, if they happened to be moving in opposite directions. Now, start to blow the balloon up. Although the ants wandering around the surface still cannot exceed one centimeter per second, the balloon is now expanding under them, and ants that are far apart on the balloon could easily be moving apart at faster than two centimeters per second if the balloon is blown up fast enough. But if they are, they will never get to tell each other about it, because the balloon is pulling them apart faster than they can move together, even at full speed. Any ants that start close enough to be able to pass one another must do so at no more than two centimeters per second, even if the Universe is expanding.

The expansion of space is just like that of the balloon, and pulls the galaxies along with it.

3.4 The fluid equation

Fundamental though it is, the Friedmann equation is of no use without an equation to describe how the density ρ of material in the Universe is evolving with time. This involves the pressure p of the material, and is called the fluid equation. (Unfortunately, the standard symbol p for pressure is the same as for momentum, which we've already used. Almost always in this book, p will be pressure.) As we'll shortly see, the different types of material that might exist in our Universe have different pressures, and lead to different evolution of the density ρ .

We can derive the fluid equation by considering the first law of thermodynamics

$$dE + p dV = T dS, \quad (3.11)$$

applied to an expanding volume V of unit comoving radius.¹ This is exactly the same as applying thermodynamics to a gas in a cylinder. The volume has physical radius a , so the energy is given, using $E = mc^2$, by

$$E = \frac{4\pi}{3} a^3 \rho c^2. \quad (3.12)$$

The change of energy in a time dt , using the product rule, is

$$\frac{dE}{dt} = 4\pi a^2 \rho c^2 \frac{da}{dt} + \frac{4\pi}{3} a^3 \frac{d\rho}{dt} c^2, \quad (3.13)$$

while the rate of change in volume is

$$\frac{dV}{dt} = 4\pi a^2 \frac{da}{dt}. \quad (3.14)$$

Assuming a reversible adiabatic expansion $dS = 0$, putting these into Equation (3.11) and rearranging gives

$$\dot{\rho} + 3 \frac{\dot{a}}{a} \left(\rho + \frac{p}{c^2} \right) = 0, \quad (3.15)$$

where, as always, dots are shorthand for time derivatives. This is the **fluid equation**. As we see, there are two terms contributing to the change in the density. The first term in the brackets corresponds to the dilution in the density because the volume has increased, while the second corresponds to the loss of energy because the pressure of the material has done work

as the Universe's volume increased. This energy has not disappeared entirely, of course; energy is always conserved. The energy lost from the fluid via the work done has gone into gravitational potential energy.

Let me stress that there are no pressure forces in a homogeneous universe, because the density and pressure are everywhere the same. A pressure *gradient* is required to supply a force. So, pressure does not contribute a force helping the expansion along; its effect is solely through the work done as the Universe expands.

We are still not in a position to solve the equations, because now we only know what ρ is doing if we know what the pressure p is. It is in specifying the pressure that we are saying what kind of material our model universe is filled with. The usual assumption in cosmology is that there is a unique pressure associated with each density, so that $p \equiv p(\rho)$. Such a relationship is known as the **equation of state**, and we'll see two different examples in [Chapter 5](#). The simplest possibility is that there is no pressure at all, and that particular case is known as (nonrelativistic) matter.

Once the equation of state is specified, the Friedmann and fluid equations are all we need to describe the evolution of the Universe. However, before discussing this evolution, we are going to spend some time exploring some general properties of the equations, as well as devoting [Chapter 4](#) to consideration of the meaning of the constant k . If you prefer to immediately see how to solve these equations, feel free to jump straight away to [Sections 5.3 – 5.5](#), and come back to the intervening material later. On the way, you might want to glance at [Section 3.6](#) to find out why a factor of c^2 mysteriously vanishes from the Friedmann equation between here and there.

3.5 The acceleration equation

The Friedmann and fluid equations can be used to derive a third equation, not independent of the first two, of course, which describes the acceleration of the scale factor. By differentiating [Equation \(3.10\)](#) with respect to time, we obtain

$$2 \frac{\dot{a}}{a} \frac{a\ddot{a} - \dot{a}^2}{a^2} = \frac{8\pi G}{3} \dot{\rho} + 2 \frac{kc^2 \dot{a}}{a^3}. \quad (3.16)$$

Substituting in for $\dot{\rho}$ from [Equation \(3.15\)](#) and canceling the factor $2\dot{a}/a$ in each term gives

$$\frac{\ddot{a}}{a} - \left(\frac{\dot{a}}{a}\right)^2 = -4\pi G \left(\rho + \frac{p}{c^2}\right) + \frac{kc^2}{a^2}, \quad (3.17)$$

and finally, using [Equation \(3.10\)](#) again, we arrive at an important equation known as the **acceleration equation**

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3} \left(\rho + \frac{3p}{c^2}\right). \quad (3.18)$$

Notice that if the material has any pressure, this *increases* the gravitational force, and so further decelerates the expansion. We remind you that there are no forces associated with pressure in an isotropic universe, as there are no pressure gradients.

The acceleration equation does not feature the constant k that appears in the Friedmann equation; it cancels out during the derivation.

3.6 On mass, energy, and vanishing factors of c^2

You should be aware that cosmologists have a habit of using mass density ρ and energy density ϵ interchangeably. They are related via Einstein's most famous equation as $\epsilon = \rho c^2$, and if one chooses so-called "natural units" in which c is set equal to one, the two become the same. For clarity, however, we will be careful to maintain the distinction. Note that the phrase "mass density" is used in Einstein's sense – it includes the contributions to the mass from the energy of the various particles, as well as any rest mass they might have.

The habit of setting $c = 1$ means that the Friedmann equation is normally written without the c^2 in the final term, so that it reads

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3} \rho - \frac{k}{a^2}. \quad (3.19)$$

The constant k then appears to have units $[\text{time}]^{-2}$ – setting $c = 1$ makes time and length units interchangeable. Since the practice of omitting the c^2 in the Friedmann equation is widely adopted in other cosmology textbooks, we will drop it for the remainder of this book too. In practice, it is a rare situation indeed where one has to be careful about this.

Problems

- 3.1. Starting with the Friedmann and acceleration equations, derive the fluid equation.
- 3.2. Derive an equation for the third derivative of the scale factor in terms of the density (not its derivatives), in the case of a pressureless fluid.

Note

- ¹ Don't confuse V for volume with $\mathcal{V}$ for gravitational potential energy.

Chapter 4

The Geometry of the Universe

We now consider the real meaning of the constant k , which appears in the Friedmann equation

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}. \quad (4.1)$$

While the Newtonian derivation in [Chapter 3](#) introduced this as a measure of the energy per particle, the true interpretation, apparent in the context of general relativity, is that it measures the curvature of space. General relativity tells us that gravity is due to the curvature of four-dimensional (4D) space-time, and a full analysis can be found in any general relativity textbook. Here, we will be purely descriptive, and focus on the interpretation of k as measuring the curvature of the three spatial dimensions. Further details of general relativistic cosmology can be found in [Advanced Topic 1](#).

We have demanded that our model universes be both homogeneous and isotropic. The simplest type of geometry which can have this property is what is called a **flat geometry**, in which the normal rules of Euclidean geometry apply. However, it turns out that the assumption of isotropy is not enough to demand that as the only choice. Instead, there are three possible geometries for the Universe, and they correspond to k being zero, positive, or negative.

4.1 Flat geometry

Euclidean geometry is based on a set of simple axioms (e.g. a straight line is the shortest distance between two points), plus one more complex axiom which says that parallel straight lines remain a fixed distance apart. These are the basis for the standard laws of geometry, and lead to the following conclusions:

- The angles of a triangle add up to 180° .
- The circumference of a circle of radius r is $2\pi r$.

Such a geometry might well apply to our own Universe. If that is the case, then the Universe must be infinite in extent, because if it came to a definite edge, then that would clearly violate the principle that the Universe should look the same from all points.

A universe with this geometry is usually called a **flat universe**.

4.2 Spherical geometry

Euclid always hoped that the more artificial final axiom could be proven by others. It wasn't until the nineteenth century that Riemann demonstrated that Euclid's final axiom was an arbitrary choice, and that one could make other assumptions. In doing so, he founded the subject of non-Euclidean geometry, which forms the mathematical foundation for Einstein's theory of general relativity.

The simplest kind of non-Euclidean geometry is actually very familiar to us; it is the spherical geometry which we use, for instance, to navigate around the Earth. Before worrying about the Universe having three dimensions, let's examine the properties of the two-dimensional (2D) surface of the Earth, shown in [Figure 4.1](#).

 Diagram of a spherical surface depicting a triangle with three right angles detailed at the vertices, indicating positive curvature k .

Figure 4.1 A sketch of a spherical surface, representing positive k . A triangle is shown, which has three right angles!

First, we know that a perfect sphere looks the same from all points on its surface, so the condition of isotropy is satisfied (e.g. if someone hands you a snooker ball and asks which way up it is, you're not going to be able to tell them). But, unlike the case of a flat geometry, the spherical surface is perfectly finite in extent, its area being given by $4\pi r^2$. Yet, there is no boundary, no "edge" to the surface of the Earth. So, it is perfectly possible to have a finite surface which nevertheless has no boundary.

If we draw parallel lines on the surface of the Earth, then they violate Euclid's final axiom. The definition of a straight line is the shortest distance between two points, which means that the straight lines in a spherical geometry are segments of great circles, such as the equator or the lines of longitude.² The lines of longitude are an excellent example of the failure of Euclid's axiom; as

they cross the equator they are all parallel to one another, but rather than remaining a constant distance apart they meet at both poles.

If we draw a triangle on a sphere, we find that the angles do not add up to 180° either. The easiest example to think about is to start at the North Pole. Draw two straight lines down to the equator, 90° apart, and then join them with a line on the equator. You have drawn a triangle in which all three angles are 90° , as shown in [Figure 4.1](#).

The circumference of a circle also fails to obey the normal law. Suppose we draw a circle at a radius r from the North Pole, and we'll choose r so that our circle is the equator. That radius, measured on the surface of the sphere, corresponds to a quarter of a complete circle around the Earth, so r equals πR divided by 2 where R is the radius of the Earth. However, the circumference c is given by $2\pi r$, so instead of the usual relation one has c equals $4r$ for a circle drawn at the equator. The circumference is less than $2\pi r$. [Problem 4.1](#) looks at the general case; you may find it helpful to glance at the figure on page 38 now.

Although we have only considered specific cases where the algebra is easy, it is true that whatever triangle or circle is drawn, you'll always find

- the angles of a triangle add up to more than 180° ;
- the circumference of a circle is less than $2\pi r$.

If you make triangles or circles much smaller than the size of the Earth, then the Euclidean laws start to become a good approximation; certainly we don't have to worry about Euclidean laws being broken in our everyday existence (though the appreciation that the Earth is spherical is vital for the planning of long distance journeys). So, a small triangle drawn on a sphere will have the sum of its angles only marginally larger than 180° . This property makes it rather hard to measure the geometry of our own Universe, because the neighboring region that we can measure accurately is only a small fraction of the size of the Universe and so will obey nearly Euclidean laws whatever the overall geometry.

One of the most important conceptual points that you need to grasp is that our three-dimensional (3D) Universe can have properties just like the 2D surface of the sphere. Unfortunately, our brains are not conditioned to think of three dimensions as being curved, so we must work by analogy with the 2D situation we have just described. When we think of the surface of a sphere being curved, we naturally imagine the sphere as an object in our own 3D Universe and think of it as curved in that sense. The important point is that the curvature is a property of the 2D surface of the sphere itself; the triangles and circles whose properties we've just described are drawn on the surface. A classic application of this was the ancient Greeks' use of these laws to deduce that the Earth is spherical, and even to obtain a good estimate of its diameter. When we discuss spherical geometry, there is actually no need to think of it as existing within our 3D space at all, and we must remember when discussing the geometry that we are assumed to be restricted to the surface of the sphere, and not allowed to move off the surface either toward or away from the notional center.

All this is an analogy to what might happen with our 3D Universe. The analog of the 2D sphere is called a three-sphere. In the rather unlikely event of 4D creatures existing, they would be able to visualize the curvature of 3D space in just the same way we can visualize that of 2D space. However, like the 2D sphere, the possible curvature of three-space is an intrinsic property, and there is no actual need for a higher-dimensional space for it to live in. Obtaining a correct mental picture of this is one of the big challenges in understanding our Universe!

A universe with a spherical geometry, like the surface of the Earth, has a finite size but no boundary. All points are equivalent. If we lived in a spherical geometry and traveled in a straight line, we would not go on forever. Rather, we would eventually come back to where we had started from, from the opposite direction, exactly in the manner that someone traveling outward from the North Pole on the Earth eventually returns there from the opposite direction. Such a universe corresponds to choosing a positive value for the quantity k appearing in the Friedmann equation. Because the special properties of the spherical geometry are due to its curvature, k is often therefore called the curvature term.

A universe with k greater than 0 is normally referred to as a **closed universe**, because of its finite size.

4.3 Hyperbolic geometry

The final choice is k negative. The corresponding geometry is known as hyperbolic, and is much less familiar than spherical geometry. It is normally represented by a saddle-like surface as in [Figure 4.2](#); it is hard to see that this is consistent with isotropy, but in fact it is. In a hyperbolic geometry, parallel lines never meet – in fact they break Euclid's axiom by diverging away from one another.

 Diagram of a saddle surface featuring a central dashed curvilinear triangle with angles sum to less than 180 degrees.

Figure 4.2 A sketch of a saddle surface, representing the hyperbolic geometry obtained when ***k is negative***. A rather exaggerated triangle is shown with its sum of angles well below 180° .

The behavior of the hyperbolic geometry can be guessed from what has gone before; it is the opposite of spherical geometry. One finds that

- the angles of a triangle add up to less than 180° ;
- the circumference of a circle is greater than $2\pi r$.

Because parallel lines never meet, such a universe must be infinite in extent, just as in the flat case. The situation

 k less than 0 is known as an **open universe**.

In [Chapter 9](#), we'll explore which of the three possible geometries, summarized in [Table 4.1](#), seems best suited to describe the real Universe.

TABLE 4.1

A summary of possible geometries.

Curvature	Geometry	Angles of triangle	Circumference of circle	Type of universe
 k greater than 0	Spherical	$>180^\circ$	 c less than $2\pi r$	Closed
 k equals 0	Flat	180°	 c equals $2\pi r$	Flat
 k less than 0	Hyperbolic	$<180^\circ$	 c greater than $2\pi r$	Open

4.4 Infinite and observable universes

What does it actually mean for the Universe to be infinite, as in the flat and open cases? This property has nothing to do with the Universe lasting forever; what it means is that the Universe is already infinite in size even at a finite time. It genuinely goes on forever in all directions. Even despite this, it is still able to expand – the distances between objects can still increase regardless of whether the total Universe is infinite in extent or not. (You might consider the integers; these form an infinite set, but you can still multiply each by two to get a new infinite set where the separation between the numbers is twice as large.)

Nevertheless, this description is only a model, and we have no way of discovering whether the actual Universe does indeed go on forever. Cosmologists often talk about a different concept to the entire, possibly infinite, Universe, namely the **observable Universe**. This corresponds to the portion of the Universe we can actually see, and is limited by the finite speed of light. As the Universe gets older, the observable Universe becomes larger and larger, by a combination of two effects. First, the Universe is expanding, and second, light has had longer to travel across the Universe. In practice, our knowledge of the Universe is restricted to this portion, and we have no way of telling whether it does indeed continue into the infinite distance as required by the cosmological principle. It may be, for example, that the Universe becomes highly irregular on an extremely large scale, and there are some theoretical models that predict that this might happen. Another possibility is discussed in [Advanced Topic 1.3](#).

4.5 Where did the Big Bang happen?

A common question is “where did the Big Bang happen?”, suggesting perhaps that one could point in a specific direction and say “That way!” After all, in a conventional explosion that is a perfectly reasonable question to ask, as all the material flies outward from the ignition point. Unfortunately, for the Big Bang, things aren't so simple, and in a sense the answer is “everywhere and nowhere.”

First of all, remember that our entire foundation is the cosmological principle, telling us that no point in the Universe is special. If there were a particular point where the “Bang” happened, that would clearly be a special point and violate the cosmological principle. Rather, space and time themselves were created at the instant of the Big Bang (unlike a conventional explosion where the material flies through preexisting space). If we take any point in the present Universe and trace back its history, it would start out at the explosion point, and in that sense the Big Bang happened everywhere in space.

In another sense, the location of the Big Bang is nowhere, because space itself is evolving and expanding, and it has changed since the Big Bang took place. Imagine the Universe as an expanding sphere. At any instant, “space” is the surface of the sphere, which is becoming bigger with time (again we are thinking of a 2D analogy to our real 3D space). The place where the “Bang” happened is at the center of the sphere, but that's no longer part of the space, the surface of the sphere in which we live. In

particular, being constrained to the surface of the sphere means we are unable to “point” to the place where the explosion is supposed to have happened. However, all the points in our current space were once at the center of the sphere, when the Big Bang took place.

4.6 Three values of k

Because there are only three distinct possibilities for the geometry, many people indicate this explicitly by scaling their variables so that k takes on one of only three possible values, namely k equals negative 1, 0, or +1, corresponding to the open, flat, and closed cases, respectively. This can be achieved by rescaling the scale factor by multiplying it by a fixed constant, namely $\frac{1}{\sqrt{|k|}}$ if k is non-zero. This leaves $\dot{H}$ equals $\frac{\dot{a}}{a}$ unchanged and removes the k from the final term in Equation (4.1), which now reads

$$\left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3} \rho \pm \frac{1}{a^2} \quad (4.2)$$

with “−” for positive k , “+” for negative k , and the last term absent if k equals 0.

If this rescaling is used, then (except in the k equals 0 case) one loses the freedom to set

$\frac{\dot{a}}{a} = 1$ at the present time, as will be used in Section 5.3. In effect, what one is doing is choosing to measure comoving distances in units of the so-called curvature scale, which is the scale on which the effects of curved space must be included, rather than in astronomical units such as megaparsecs.

In this book, we will *not* rescale k to one of these three discrete values, though most of the time we will be specializing to k equals 0 anyway.

Problems

4.1. Consider the surface of a 2D sphere of radius R , as illustrated.

Diagram of a sphere details radius R extending from the center, angle θ from the vertical axis, and arc length r along the surface.

Circles are drawn on the sphere that have radius r as measured on the surface of the sphere, centered on the North Pole. Show that the general formula for the circumference c of such a circle, as a function of r , is given by

$$c = 2\pi \sin\left(\frac{r}{R}\right) r$$

where θ is the angle between a line drawn from the center of the sphere to the North Pole and one drawn to the circle. (Remember that by definition of angles in radians, we have $r = \theta R$.)

Demonstrate that for small θ (i.e. r much less than R) this gives the normal flat geometry relation. Evaluate the relation for the case when the circle is at the equator.

4.2. Consider the spherical geometry of the previous problem, staying with the 2D analogy to the real Universe. Suppose that galaxies are distributed evenly in such a universe, with a number density n per unit area. Show that the total number N of galaxies inside a radius r is given by

$$N = 2\pi n R^2 \left[1 - \cos\left(\frac{r}{R}\right) \right]$$

Expand this for r much less than R to show that the flat space result that the number is $\pi n r^2$ is recovered (remember we are working in only two dimensions). Do you see more or fewer galaxies out to the same radius, if the Universe is spherical rather than flat?

4.3. Is it possible for a closed universe to evolve to become an open universe? Give a reason for your answer.

Notes

¹ Well, that’s almost true. See Advanced Topic 1.3 for a way to bypass that conclusion.

² Note that, apart from the equator, lines of latitude are *not* straight lines; this is why aeroplanes do not follow lines of latitude when flying, because they are not the shortest way to go!

Chapter 5

Simple Cosmological Models

In [Chapter 3](#), we derived the equations satisfied by an expanding isotropic gas. They are the Friedmann equation ¹

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}, \quad (5.1)$$

which governs the time evolution of the scale factor $a(t)$, and the fluid equation

$$\frac{d\rho}{dt} + 3\frac{\dot{a}}{a}\rho = 0, \quad (5.2)$$

which gives us the evolution of mass density $\rho(t)$. We stress that, despite their Newtonian derivation, these are the real equations used by cosmologists, more traditionally derived via the equations of general relativity as described in [Advanced Topic 1](#).

This chapter finds and discusses some simple solutions to these equations, which will be used extensively during the book. However, while they have wide applicability during the early stages of the Universe, it will turn out that these are not sufficient to describe the present state of the Universe, for which an extra ingredient, the cosmological constant, will be needed. It is introduced in [Chapter 7](#).

Before finding solutions to these equations, we can study two of their implications.

5.1 Hubble's law

The Friedmann equation allows us to explain Hubble's discovery that recession velocity is proportional to the distance. The velocity of recession is given by

$\vec{v} = \frac{d\vec{r}}{dt}$ and is in the same direction as $\vec{r}$, allowing us to write

$$\vec{v} = \frac{d\vec{r}}{dt} = \frac{d}{dt}(\vec{r}) = \frac{d}{dt}(a\vec{r}_0) = \dot{a}\vec{r}_0 = \frac{\dot{a}}{a}\vec{r}, \quad (5.3)$$

The last step used $\vec{r}_0$ is a constant by definition. Consequently, Hubble's law

$\vec{v} = H\vec{r}$ tells us that the proportionality constant, the Hubble parameter, should be identified as

$$H = \frac{\dot{a}}{a} \quad (5.4)$$

and the value as measured today can be denoted with a subscript "0" as H_0 . Because we measure Hubble's constant to be positive rather than negative, we know that the Universe is expanding rather than contracting.

We notice from this that the phrase Hubble's *constant* is a bit misleading. Although it is certainly constant in space due to the cosmological principle, there is no reason for it to be constant in time. In fact, using it as a more compact notation, we can write the Friedmann equation as an evolution equation for $H(t)$, as

$$H^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \quad (5.5)$$

It is best to use the phrase "Hubble parameter" for this quantity as a function of time, reserving "Hubble constant" for its present value. Normally, the Hubble parameter decreases with time, for instance, as the expansion is slowed by the gravitational attraction of the matter in the Universe.

5.2 Expansion and redshift

The redshift of spectral lines that we used to justify the assumption of an expanding universe can also be related to the scale factor. In this derivation, we will make the simplifying assumption that light is passed between two objects that are very close together, separated by a small distance d_r , as shown in [Figure 5.1](#). We have drawn the objects as galaxies, but we really mean two nearby points. According to Hubble's law, their relative velocity $d v$ will be

$$d v \text{ equals } H d_r \text{ equals } \frac{H}{c} \frac{d r}{dt} \quad (5.6)$$

As the points are nearby, we can directly apply the Doppler law to say that the change in wavelength between emission and reception, $d \lambda$ identical to λ Subscript normal r Baseline minus λ Subscript normal e , is

$$\frac{d \lambda}{\lambda} \text{ equals } \frac{d v}{c} \quad (5.7)$$

where $d \lambda$ is going to be positive since the wavelength is increased. The time between emission and reception is given by the light travel time $d t$ equals d_r divided by c , and putting all that together gives

$$\frac{d \lambda}{\lambda} \text{ equals } \frac{d r}{c d t} \text{ equals } \frac{H}{c} \frac{d r}{d t} \quad (5.8)$$

Integrate and we find that $\ln \lambda$ equals $\ln$ plus c on s t a n t, i.e.

$$\lambda \text{ proportional to } a \quad (5.9)$$

where λ is now the instantaneous wavelength measured at any given time.

 Diagram of a photon wave traveling from spiral Galaxy A toward spiral Galaxy B across a distance measured as d_r .

Figure 5.1 A photon travels a distance d_r between two galaxies A and B.

Although as we have derived it, this result only applies to objects that are very close to each other, it turns out that it is completely general (a rigorous treatment is given in [Advanced Topic 2](#)). It tells us that as space expands, wavelengths become longer in direct proportion. One can think of the wavelength as being stretched by the expansion of the Universe, and its change therefore tells us how much the Universe has expanded since light began its travels. For example, if the wavelength had doubled, the Universe must have been half its present size when the light was emitted.

The redshift as defined in [Equation \(2.1\)](#) is related to the scale factor by

$$1 + z \text{ equals } \frac{\lambda_r}{\lambda_e} \quad (5.10)$$

and is normally only used to refer to light received by us at the present epoch.

5.3 Solving the equations

In order to discover how the Universe might evolve, we need some idea of what is in it. In a cosmological context, this is done by specifying the relationship between the mass density ρ and the pressure p . This relationship is known as the **equation of state**. At this point, we shall only consider two possibilities.

Matter: In this context, the term “matter” is used by cosmologists as shorthand for “non-relativistic matter,” and refers to any type of material which exerts negligible pressure, p equals 0. Occasionally, care is needed to avoid confusion between “matter” used in this sense, and used to indicate all types of matter, whether nonrelativistic or not. A pressureless universe is the simplest assumption that can be made. It is a good approximation to use for the atoms in the Universe once they have cooled down, as they are quite well separated and seldom interact, and it is also a good description of a collection of galaxies in the Universe, as they have no interactions other than gravitational ones. Occasionally the term “dust” is used instead of “matter.”

Radiation: Particles of light move, naturally enough, at the speed of light. Their kinetic energy leads to a pressure force, the radiation pressure, which, using the standard theory of radiation, can be shown to be

ρ equals ρ_0 divided by 3. [Problem 5.2](#) gives a rather hand-waving derivation of this result. More generally, any particles moving at highly relativistic speeds have this equation of state, neutrinos being an obvious example.

We will concentrate on the case where the constant k in the Friedmann equation is set equal to zero, corresponding to a flat geometry.

5.3.1 Matter

We start by solving the fluid equation, having set $p = 0$ for matter. One way to solve it is to notice a clever way of rewriting it, as follows

$$\frac{d}{dt} \left(\rho a^3 \right) = 0 \quad (5.11)$$

though one could also solve it more formally by noting that it is a separable equation. Integrating tells us that ρa^3 equals a constant, i.e.

$$\rho a^3 \propto \frac{1}{a^3} \quad (5.12)$$

This is not a surprising result. It says that the density falls off in proportion to the volume of the Universe. It is very natural that if the volume of the Universe increases by a factor of say two, then the density of the matter must fall by the same factor. After all, material cannot come from nowhere, and there is no pressure to do any work.

The equations we are solving (with $k = 0$) have one very useful symmetry; their form is unchanged if we multiply the scale factor a by a constant, since only the combination $\frac{da}{dt} a^3$ appears. This means that we are free to rescale a as we choose, and the normal convention is to choose $a = 1$ at the present time. With this choice, physical and comoving coordinate systems coincide at the present, since $\frac{dr}{dt} = \frac{dx}{dt}$. Throughout this book we will use the subscript “0” to indicate the present value of quantities. Denoting the present density by ρ_0 fixes the proportionality constant

$$\rho a^3 = \rho_0 \quad (5.13)$$

Having solved for the evolution of the density in terms of a , we must now find how a varies with time by using the Friedmann equation. Substituting in for ρ , and remembering we are assuming $k = 0$, gives

$$\left(\frac{da}{dt} \right)^2 = \frac{8\pi G}{3} \rho_0 \frac{1}{a^3} \quad (5.14)$$

Faced with an equation like this, one can use formal techniques to solve it (this equation is separable, allowing it to be integrated), or alternatively make an educated guess as to the solution and confirm it by substitution. In cosmology, a good educated guess is normally a power-law $a \propto t^q$. Substituting this in, the left-hand side has time dependence t^{2q-2} and the right-hand side t^{-3q} . This can only be a solution if these match, which requires $q = 2/3$, and so the solution is $a \propto t^{2/3}$. As we have fixed $a = 1$ at the present time $t = t_0$, the full solution is therefore

$$a(t) = \left(\frac{3}{2} H_0 t \right)^{2/3} \quad (5.15)$$

In this solution, the Universe expands forever, but the rate of expansion $H(t)$ decreases with time

$$H(t) = \frac{2}{3t} \quad (5.16)$$

becoming infinitely slow as the Universe becomes infinitely old. Notice that, despite the pull of gravity, the material in the Universe does not recollapse but rather expands forever.

This is one of the classic cosmological solutions, and will be much used throughout this book.

5.3.2 Radiation

Radiation obeys $p = \frac{1}{3} \rho c^2$. Consequently, the fluid equation is changed from the matter-dominated case, now reading

$$\frac{d\rho}{dt} + 4\frac{\dot{a}}{a}\rho = 0 \quad (5.17)$$

This is amenable to the same trick as before, with the a^3 replaced by a^4 in Equation (5.11), giving

$$\rho \propto \frac{1}{a^4} \quad (5.18)$$

Carrying out the same analysis we did in the matter-dominated case gives

$$\left(\frac{t}{t_0} \right)^2 \propto \frac{\rho(t_0)}{\rho(t)} \quad (5.19)$$

This is the second classic cosmological solution.

Notice that the Universe expands more slowly if radiation dominated than if matter dominated, a consequence of the extra deceleration that the pressure supplies – see Equation (3.18). So, it is definitely wrong to think of pressure as somehow “blowing” the Universe apart. However, in each case, the density of material falls off as t^{-2} .

We’d better examine the fall off of the radiation density with volume more carefully. It drops as the fourth power of the scale factor. Three of those powers we have already identified as the increase in volume, leading naturally to a drop in the density. The final power arises from a different effect, the stretching of the wavelength of the light. Since the stretching is proportional to a , and the energy of radiation proportional to its frequency via $E = hf$, this results in a further drop in energy by the remaining power of a . This lowering of energy is exactly the redshifting effect we use to measure distances.

The rate of decrease of the radiation density also has an explanation in terms of thermodynamics, which is macroscopic rather than microscopic. Since the Universe in this case has a pressure, when it expands, there is work done which is given by $p dV$, in exactly the same way as work is done on a piston when the gas is allowed to expand and cool. This work done corresponds to the extra diminution of the radiation density by the final factor of a .

5.3.3 Mixtures

A more general situation is when one has a mixture of both matter and radiation. Then there are two separate fluid equations, one for each of the two components. The trick which allows us to write ρ as a function of a still works, so we still have

$$\rho_m \propto \frac{1}{a^3}, \quad \rho_r \propto \frac{1}{a^4} \quad (5.20)$$

However, there is still only a single Friedmann equation (after all, there is only one Universe!), which now has

$$\rho = \rho_m + \rho_r \quad (5.21)$$

This means that the scale factor will have a more complicated behavior, and so to convert

$\rho(a)$ into $\rho(t)$ is much harder. It is possible to obtain exact solutions for this situation, but they are quite messy, so we won’t include them here. Instead, we will consider a simpler situation where one or other of the densities is by far the larger.

In that case, we can say that the Friedmann equation is accurately solved by just including the dominant component. That is, we can use the expansion rates we have already found. For example, suppose radiation is much more important. Then, one would have

$$\rho_m \propto \frac{1}{a^3} \propto \frac{1}{t^3} \quad (5.22)$$

Notice that the density in matter falls off more slowly than that in radiation. So, the situation of radiation dominating cannot last forever; however, small the matter component might be originally, it will eventually come to dominate. We can say that domination of the Universe by radiation is an unstable situation.

In the opposite situation, where it is the matter which is dominant, we obtain the solution

$$\rho_m \propto \frac{1}{a^3} \propto \frac{1}{t^3} \quad (5.23)$$

Matter domination is a stable situation, the matter becoming increasingly dominant over the radiation as time goes by.

[Figure 5.2](#) shows the evolution of a universe containing matter and radiation, with the radiation initially dominating. Eventually, the matter comes to dominate, and as it does so, the expansion rate speeds up from

$\rho_m \propto \frac{1}{a^3} \propto \frac{1}{t^3}$ to $\rho_m \propto \frac{1}{a^3} \propto \frac{1}{t^2}$. It is likely that this is the situation that applied in our Universe until fairly recently, as we'll see in [Chapter 11](#).

 Graph of log density versus log time with two downward sloping lines, Radiation and Matter, converging at a central point.

Figure 5.2 A schematic illustration of the evolution of a universe containing radiation and matter. Once matter comes to dominate the expansion rate speeds up, so the densities fall more quickly with time.

5.4 Particle number densities

An important alternative view of the evolution of particles, which will be much used later in the book, is that of the **number density** n of particles rather than of their mass or energy density.

The number density is simply the number of particles in a given volume. If the mean energy per particle (including any mass-energy) is ϵ , then the number density is related to the energy density by

$$\epsilon = n \epsilon \quad (5.24)$$

The number density is useful because, in most circumstances, particle number is conserved. For example, if particle interactions are negligible, you wouldn't expect an electron to suddenly vanish into oblivion, and the same is true of a photon of light. The particle number can change through interactions. For example, an electron and positron could annihilate and create two photons. However, if the interaction rate is high, we expect the Universe to be in a state of thermal equilibrium. If so, then particle numbers are conserved even in a highly interacting state, since, by definition, thermal equilibrium means that any interaction, which may change the number density of a particular type of particle, must proceed at the same rate in both forward and backward directions so that any change cancels out.

So, barring brief periods where thermal equilibrium does not hold, we expect the number of particles to be conserved. The only thing that changes the number *density*, therefore, is that the volume is getting bigger, so that these particles are spread out in a larger volume. This implies

$$n \propto \frac{1}{a^3} \propto \frac{1}{t^3} \quad (5.25)$$

This looks encouragingly like the behavior we have already seen for matter, but it's also true for radiation as well!

How does this relate to our earlier results? The energy of nonrelativistic particles is dominated by their rest mass-energy, which is constant, so

$$\rho_m \propto n m c^2 \propto \frac{1}{a^3} \propto \frac{1}{t^3} \quad (5.26)$$

But photons lose energy as the Universe expands and their wavelength is stretched, so their energy is

$\epsilon \propto \frac{1}{a} \propto \frac{1}{t^2}$ as we have already seen. So

$$\rho \propto a^{-3} \quad \text{Baseline proportional to } \epsilon \propto a^{-4} \quad \text{Baseline proportional to } n \propto a^{-3} \quad \text{Baseline times upper } E \propto a^{-4} \quad \text{Baseline proportional to } \frac{1}{a^3} \quad \text{EndFraction times } \frac{1}{a^4} \quad \text{EndFraction proportional to } \frac{1}{a^4} \quad \text{Superscript 4 Baseline EndFraction period} \quad (5.27)$$

These are exactly the results we saw before: [Equations \(5.12\)](#) and [\(5.18\)](#).

Although the energy densities of matter and radiation evolve in different ways, their particle numbers evolve in the same way. So, apart from epochs during which the assumption of thermal equilibrium fails, the relative number densities of the different particles (e.g. electrons and photons) do not change as the Universe expands.

5.5 Evolution including curvature

We can now re-introduce the possibility that the constant k is nonzero, corresponding to spherical or hyperbolic geometry. Rather than seeking precise solutions, we will concentrate on the qualitative properties of the solutions. These are actually of rather limited use in describing our own Universe, because, as we will see, the cosmological models discussed so far are not general enough, and we will need to consider a cosmological constant (see [Chapter 7](#)). Nevertheless, studying the possible behaviors of these simple models is a useful exercise, even if one should be cautious about drawing general conclusions.

In analyzing the possible dynamics, we will assume that the Universe is always dominated by nonrelativistic matter, which in practice is not a restrictive assumption. We have already seen that if we assume that the constant k in the Friedmann equation is zero, then the Universe expands for ever, $a \propto t^{2/3}$, but slows down arbitrarily at late times. So, we know the fate of the Universe in that case. But what happens if $k \neq 0$?

The principal question to ask is whether it is possible for the expansion of the Universe to stop, which since $\dot{H} = -\frac{1}{2}(\frac{\dot{a}}{a})^2 - \frac{1}{2}k$ corresponds to $\dot{H} = 0$. Looking at the Friedmann equation

$$\dot{H}^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \quad (5.28)$$

it is immediately apparent that this is not possible if k is negative, for then both the terms on the right-hand side of the Friedmann equation are positive. Consequently, such a universe must expand forever. That enables us to study the late-time behavior, because we can see that the term $\frac{k}{a^2}$ falls off more slowly with the expansion than it does $\rho \propto a^{-3}$. Since a becomes arbitrarily large for the matter-dominated solution for negligible k , the $\frac{k}{a^2}$ term must eventually come to dominate. When it does, the Friedmann equation becomes

$$-\frac{k}{a^2} = -\frac{1}{2}(\frac{\dot{a}}{a})^2 \quad (5.29)$$

Cancel off the a^2 terms and you'll find the solution is $a \propto t$. So, when the last term comes to dominate, the expansion of the Universe becomes yet faster. In this case, the velocity does not tend to zero at late times, but instead becomes constant. This is sometimes known as free expansion.

Things are very different if k is positive. It then becomes possible for $\dot{H}$ to be zero, by the two terms on the right-hand side of the Friedmann equation canceling each other out. Indeed, this is inevitable, because the negative influence of the $\frac{k}{a^2}$ term will become more and more important relative to the ρ term as time goes by. In such a universe, therefore, the expansion must come to an end after a finite amount of time. As gravitational attraction persists, the recollapse of the Universe becomes inevitable.

The collapse of the Universe is fairly easy to describe, because the equations governing evolution are time reversible. That is, if one substitutes negative t for t , they remain the same. The collapse phase is therefore just like the expansion in reverse, and so, after a finite time, the Universe will come to an end in a Big Crunch. [Problem 5.5](#) investigates this in more detail.

These three behaviors, illustrated in [Figure 5.3](#), can be related to the particle energy U in our Newtonian derivation of the Friedmann equation. If the particle's energy is positive, then it can escape to infinity, with a final kinetic energy given by U . If the total energy is zero, then the particle can just escape but with zero velocity. Finally, if the energy is negative, it cannot escape the gravitational attraction and is destined to recollapse inward.

 Graph of scale factor versus time with three curves labeled $k < 0$, $k = 0$, and $k > 0$ details universe evolution.

Figure 5.3 Three possible evolutions for the Universe, corresponding to the different signs of k . The middle line corresponds to the $k = 0$ case, where the expansion rate approaches zero in the infinite future. During the early phases of the expansion the lines are very close and so observationally it can be difficult to distinguish which path the actual Universe will follow.

There is a fairly precise analogy with escape velocity from the Earth (or the Moon, if you want to worry about the atmosphere). If you throw a rock up in the air hard enough, gravity will be unable to stop it, and eventually it will sail off into space at a constant velocity. If your throw is too puny, it will rapidly fall back. And in between is the escape velocity, where the rock is just able to escape the gravitational field and no more.

Problems

5.1. Is the energy of the Universe conserved as it expands?

5.2. This problem indicates the origin of the equation of state $p = \frac{1}{3} \rho c^2$ for radiation. An ideal gas has pressure

$$p = \frac{1}{3} n \langle v^2 \rangle$$

where $\langle \rangle$ indicates an average over the direction of particle motions. Here, n is the number density, and be careful not to confuse the unfortunate notation p for pressure and $\mathbf{p}$ for momentum. Using Equation (2.4) to relate the photon energy and momentum show that this gives

$$p = \frac{1}{3} n \langle E \rangle$$

where $\langle E \rangle$ is the mean photon energy. Hence, it demonstrates the equation of state for radiation.

5.3. In this chapter, we examined solutions for the expansion when the Universe contained either matter

or radiation

Suppose we have a more general equation of state, $p = w \rho c^2$, where w is a constant in the range $-1 < w < 1$. Find solutions for $\rho(t)$ and hence

for universes containing such matter. Assume $k = 0$ in the Friedmann equation. What is the solution if $p = -\rho c^2$?

5.4. Using your answer to Problem 5.3, what value of w would be needed so that ρ has the same time dependence as the curvature term k/a^2 ? Find the solution $a(t)$ to the full Friedmann equation for a fluid with this w , assuming negative k .

5.5. The full Friedmann equation is

$$\left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3} \rho - \frac{k}{a^2}$$

Consider the case $k > 0$, with a universe containing only matter

so that $\rho = \rho_0 / a^3$. Demonstrate that the parametric solution

$$a(1 - \cos \theta) = \frac{4\pi G \rho_0}{3k} \quad t(1 - \cos \theta) = \frac{4\pi G \rho_0}{3k} \frac{1 - \cos \theta}{\sin^3 \theta}$$

solves this equation, where θ is a variable which runs from 0 to 2π .

Sketch a and t as functions of θ . Describe qualitatively the behavior of the Universe. Attempt to sketch a as a function of t .

5.6. Now consider the case $k < 0$, with a universe containing only matter

left parenthesis p equals 0 right parenthesis so that ρ equals ρ_0 divided by a^3 . What is the solution

left parenthesis t right parenthesis in a situation where the final term of the Friedmann equation dominates over the density term? How does the density of matter vary with time? Is domination by the curvature term a stable situation that will continue forever?

Note

¹In accord with the discussion of [Section 3.6](#), the c^2 squared on the final term has been dropped, so that its appearance matches that of other cosmology textbooks.

Chapter 6

Observational Parameters

The Big Bang model does not give a unique description of our present Universe, but rather leaves quantities such as the present expansion rate, or the present composition of the Universe, to be fixed by observation.

It is a standard practice to specify cosmological models via a few parameters, which one then tries to determine observationally to decide which version of the model best describes our Universe. In this chapter and [Chapter 7](#), we will discuss the most commonly considered parameters, including ones we have already seen and new ones.

6.1 The expansion rate H_0

The Hubble constant H_0 tells us the present expansion rate of the Universe, making it a vital cosmological parameter as measurements of other properties of the Universe often depend on it. It also ought to be the easiest to measure, since all galaxies are supposed to obey $v = H_0 r$. So, all we have to do is measure the velocities and distances of as many galaxies as we can and get an answer. However, each measurement has its problems.

Velocities are given by the redshift of spectral lines, a measurement which is now easy enough that the velocity of an individual galaxy can be measured to high accuracy. However, remember that the cosmological principle isn't perfect, and so, as well as the uniform expansion we are trying to measure, galaxies also have motions relative to one another, the so-called "peculiar velocity." The peculiar velocities are randomly oriented, and for a given galaxy we cannot split its measured velocity into the Hubble expansion and the peculiar velocity. However, the cosmological principle does tell us that the typical size of the peculiar velocity should not depend on where in the Universe the galaxy is. It is therefore independent of distance, whereas the Hubble velocity is proportional to distance. If we look far enough away (in practice many tens of megaparsecs) then the Hubble velocity dominates and the (unknown) peculiar velocity can be ignored.

Given that the expansion velocity can only be accurately distinguished from the peculiar velocity at large distances, we need to be able to estimate these large distances accurately in order to carry out the calculation $H_0 = v/r$. These distances are much harder to obtain, because galaxies are far too distant to be located by parallax. Remember that an object one parsec away has a parallax (i.e. an apparent motion when viewed from different parts of Earth's orbit) of one arcsecond, by definition. A galaxy many megaparsecs away will have an immeasurably small parallax of less than a micro-arcsecond. The usual method is known as "standard candles," where some type of object is assumed to have exactly the same properties in all parts of the Universe. This is the cosmological equivalent of saying that if one light bulb looks a quarter as bright as another, then from the inverse square law it must be twice as far away – fine as long as you believe that all light bulbs have precisely the same brightness. A classic example is the period–luminosity relation in cepheid variable stars. The period of variability of those stars is readily measured, and there is reasonable empirical evidence of a relation between the period and the luminosity of the stars, which lets us convert the measured period into a brightness. Other standard candles which have been used include the brightness of certain types of supernovae, and the brightest galaxies in galaxy clusters. All these methods have good success at determining relative distances between two galaxies, which requires that the objects be good standard candles, and relative distances are all that is needed to confirm the Hubble law.

However, to give an absolute distance, by actually measuring the proportionality constant H_0 , we also need a calibration against an object of known distance, which proves much harder. In the light bulb analogy, to get relative distances we need only the inverse square law and the belief that all bulbs have the same brightness; we don't need to know how bright the bulbs are. But to be able to say how far away a bulb of a given observed brightness is, we need to know its absolute brightness.

It is only relatively recently that the calibration problem has begun to come under control, principally through Hubble Space Telescope observations. The Hubble constant is usually parametrized as

$$H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}. \quad (6.1)$$

In the 1990s, the Hubble Space Telescope Key Project, led by Wendy Freedman, gave

$$h = 0.72 \pm 0.08, \quad (6.2)$$

where the uncertainty is a one-sigma error (meaning it should be doubled to indicate 95% confidence, at least if the uncertainty is approximately Gaussian-distributed). This was considered the first precision direct measurement of the Hubble constant, and has subsequently been improved to $h = 0.73 \pm 0.01$ by the SH0ES (Supernovae and H_0 for the Equation of State) project.

On the other hand, indirect measurements, which assess the impact of the expansion rate on other cosmological observables, such as the cosmic microwave background, have suggested a lower value. The *Planck* satellite observations described in [Advanced Topic 6](#) give a very precise determination $h = 0.674 \pm 0.005$. There is a clear discrepancy between this measurement and the direct one, which has persisted for several years and become known as the **Hubble tension**. Its