

# Particle Physics

FOURTH  
EDITION

B.R. Martin and G. Shaw



WILEY



# **Particle Physics**

Fourth Edition

---

# The Manchester Physics Series

*General Editors*

**J.R. FORSHAW, H.F. GLEESON, F.K. LOEBINGER**

*School of Physics and Astronomy,  
University of Manchester*

---

<b>Properties of Matter</b>	B.H. Flowers and E. Mendoza
<b>Statistical Physics</b> <i>Second Edition</i>	F. Mandl
<b>Electromagnetism</b> <i>Second Edition</i>	I.S. Grant and W.R. Phillips
<b>Statistics</b>	R.J. Barlow
<b>Solid State Physics</b> <i>Second Edition</i>	J.R. Hook and H.E. Hall
<b>Quantum Mechanics</b>	F. Mandl
<b>Computing for Scientists</b>	R.J. Barlow and A.R. Barnett
<b>The Physics of Stars</b> <i>Second Edition</i>	A.C. Phillips
<b>Nuclear Physics</b>	J.S. Lilley
<b>Introduction to Quantum Mechanics</b>	A.C. Phillips
<b>Dynamics and Relativity</b>	J.R. Forshaw and A.G. Smith
<b>Vibrations and Waves</b>	G.C. King
<b>Mathematics for Physicists</b>	B.R. Martin and G. Shaw
<b>Particle Physics</b> <i>Fourth Edition</i>	B.R. Martin and G. Shaw

# **Particle Physics**

**Fourth Edition**

**B.R. MARTIN**

Department of Physics and Astronomy  
University College London

**G. SHAW**

School of Physics and Astronomy  
Manchester University

**WILEY**

This edition first published 2017

© 2017 John Wiley & Sons, Ltd

*Registered office*

John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at [www.wiley.com](http://www.wiley.com).

The right of the authors to be identified as the authors of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

*Library of Congress Cataloging-in-Publication Data applied for.*

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

ISBN: 9781118911907 (hardback), 9781118912164 (paperback)

Set in 11/13pt Computer Modern by Aptara Inc., New Delhi, India.

10 9 8 7 6 5 4 3 2 1

The pictures on the front cover show two Higgs boson production events obtained at the LHC accelerator at CERN. The larger circle shows the decay of the Higgs to four electrons (indicated by the blue and red lines) obtained by the ATLAS collaboration and the smaller circle shows the decay of the Higgs to two gamma rays (indicated by the green lines) obtained by the CMS collaboration.

'But why are such terrific efforts made just to find new particles?' asked Mr Tompkins.

'Well, this is science,' replied the professor, 'the attempt of the human mind to understand everything around us, be it giant stellar galaxies, microscopic bacteria, or these elementary particles. It is interesting and exciting, and that is why we are doing it.'

From *Mr Tompkins Tastes a Japanese Meal*, by George Gamow (*Mr Tompkins in Paperback*, Cambridge University Press (1965), p.186).



---

# Contents

---

<i>Editors' preface to the Manchester Physics Series</i>	xiii
<i>Authors' preface</i>	xv
<i>Suggested Short Course</i>	xvii
<i>Notes</i>	xix
<i>Physical Constants, Conversion Factors and Natural Units</i>	xxi
<b>1 Some basic concepts</b>	<b>1</b>
1.1 Introduction	1
1.2 Antiparticles	3
1.2.1 Relativistic wave equations	3
1.2.2 Hole theory and the positron	6
1.3 Interactions and Feynman diagrams	9
1.3.1 Basic electromagnetic processes	10
1.3.2 Real processes	11
1.3.3 Electron–positron pair production and annihilation	13
1.3.4 Other processes	15
1.4 Particle exchange	15
1.4.1 Range of forces	15
1.4.2 The Yukawa potential	17
1.4.3 The zero-range approximation	18
1.5 Units and dimensions	19
Problems 1	22
<b>2 Leptons and the weak interaction</b>	<b>24</b>
2.1 Lepton multiplets and lepton numbers	24
2.1.1 Electron neutrinos	25
2.1.2 Further generations	28
2.2 Leptonic weak interactions	31
2.2.1 $W^\pm$ and $Z^0$ exchange	31
2.2.2 Lepton decays and universality	33
2.3 Neutrino masses and neutrino mixing	35
2.3.1 Neutrino mixing	35
2.3.2 Neutrino oscillations	38
2.3.3 Neutrino masses	46
2.3.4 Lepton numbers revisited	48
Problems 2	50

<b>3 Quarks and hadrons</b>	<b>52</b>
3.1 Quarks	53
3.2 General properties of hadrons	55
3.3 Pions and nucleons	58
3.4 Strange particles, charm and bottom	61
3.5 Short-lived hadrons	66
3.6 Allowed and exotic quantum numbers	72
Problems 3	75
<b>4 Experimental methods</b>	<b>77</b>
4.1 Overview	77
4.2 Accelerators and beams	79
4.2.1 Linear accelerators	80
4.2.2 Cyclic accelerators	81
4.2.3 Fixed-target machines and colliders	83
4.2.4 Neutral and unstable particle beams	85
4.3 Particle interactions with matter	86
4.3.1 Short-range interactions with nuclei	86
4.3.2 Ionisation energy losses	89
4.3.3 Radiation energy losses	92
4.3.4 Interactions of photons in matter	93
4.3.5 Ranges and interaction lengths	94
4.4 Particle detectors	95
4.4.1 Introduction	96
4.4.2 Gaseous ionisation detectors	97
4.4.3 Semiconductor detectors	103
4.4.4 Scintillation counters	104
4.4.5 Čerenkov counters and transition radiation	105
4.4.6 Calorimeters	109
4.5 Detector systems and accelerator experiments	112
4.5.1 Discovery of the $W^\pm$ and $Z^0$ bosons	113
4.5.2 Some modern detector systems	117
4.6 Non-accelerator experiments	121
Problems 4	123
<b>5 Space-time symmetries</b>	<b>126</b>
5.1 Translational invariance	127
5.2 Rotational invariance	129
5.2.1 Angular momentum conservation	129
5.2.2 Classification of particles	132
5.2.3 Angular momentum in the quark model	134
5.3 Parity	135
5.3.1 Leptons and antileptons	137
5.3.2 Quarks and hadrons	139
5.3.3 Parity of the charged pion	140
5.3.4 Parity of the photon	141

---

5.4	Charge conjugation	142
5.4.1	$\pi^0$ and $\eta$ decays	144
5.5	Positronium	145
★5.5.1	Fine structure	147
★5.5.2	$C$ -parity and annihilations	148
*5.6	Time reversal	149
5.6.1	Principle of detailed balance	151
5.6.2	Spin of the charged pion	152
	Problems 5	153
<b>6</b>	<b>The quark model</b>	<b>155</b>
6.1	Isospin symmetry	156
6.1.1	Isospin quantum numbers	157
6.1.2	Allowed quantum numbers	158
6.1.3	An example: the sigma ( $\Sigma$ ) baryons	159
6.1.4	The $u$ , $d$ quark mass splitting	161
6.2	The lightest hadrons	162
6.2.1	The light mesons	162
6.2.2	The light baryons	164
★6.2.3	Baryon magnetic moments	167
★6.2.4	Hadron mass splittings	169
6.3	The $L = 0$ heavy quark states	174
6.4	Colour	177
6.4.1	Colour charges and confinement	178
★6.4.2	Colour wavefunctions and the Pauli principle	182
6.5	Charmonium and bottomonium	184
6.5.1	Charmonium	185
6.5.2	Bottomonium	189
★6.5.3	The quark–antiquark potential	189
	Problems 6	191
<b>7</b>	<b>QCD, jets and gluons</b>	<b>193</b>
7.1	Quantum chromodynamics	193
7.1.1	The strong coupling constant	197
7.1.2	Screening, antiscreening and asymptotic freedom	199
★7.1.3	Exotic hadrons	201
★7.1.4	The quark–gluon plasma	208
7.2	Electron–positron annihilation	210
7.2.1	Two-jet events	211
7.2.2	Three-jet events	213
7.2.3	The total cross-section	214
	Problems 7	215
<b>8</b>	<b>Quarks and partons</b>	<b>217</b>
8.1	Elastic electron scattering: the size of the proton	217
8.1.1	Static charge distributions	218

8.1.2	Proton form factors	219
*8.1.3	The basic cross-section formulas	221
8.2	Inelastic electron and muon scattering	222
8.2.1	Bjorken scaling	224
8.2.2	The parton model	226
8.2.3	Parton distributions and scaling violations	228
8.3	Inelastic neutrino scattering	231
8.3.1	Quark identification and quark charges	234
*8.4	Other processes	236
8.4.1	Lepton pair production	239
8.4.2	Jets in pp collisions	242
8.5	Current and constituent quarks	243
Problems 8		246
<b>9</b>	<b>Weak interactions: quarks and leptons</b>	<b>248</b>
9.1	Charged current reactions	250
9.1.1	$W^\pm$ -lepton interactions	250
9.1.2	Lepton–quark symmetry and mixing	254
9.1.3	$W$ boson decays	258
*9.1.4	Selection rules in weak decays	259
9.2	The third generation	262
9.2.1	More quark mixing	263
9.2.2	Properties of the top quark	265
*9.2.3	Discovery of the top quark	267
Problems 9		274
<b>10</b>	<b>Weak interactions: electroweak unification</b>	<b>276</b>
10.1	Neutral currents and the unified theory	277
10.1.1	The basic vertices	277
10.1.2	The unification condition and the $W^\pm$ and $Z^0$ masses	279
10.1.3	Electroweak reactions	281
10.1.4	$Z^0$ formation: how many generations are there?	284
10.2	Gauge invariance and the Higgs boson	287
10.2.1	Unification and the gauge principle	289
10.2.2	Particle masses and the Higgs field	290
10.2.3	Properties of the Higgs boson	294
10.2.4	The discovery of the Higgs boson	297
Problems 10		305
<b>11</b>	<b>Discrete symmetries: <math>C</math>, <math>P</math>, <math>CP</math> and <math>CPT</math></b>	<b>308</b>
11.1	$P$ violation, $C$ violation and $CP$ conservation	308
11.1.1	Muon decay symmetries	310
11.1.2	Left-handed neutrinos and right-handed antineutrinos	312
11.1.3	Pion and muon decays revisited	314
11.2	$CP$ violation and particle–antiparticle mixing	316

---

11.2.1	<i>CP</i> eigenstates of neutral kaons	316
11.2.2	The discovery of <i>CP</i> violation	319
*11.2.3	<i>CP</i> -violating $K_L^0$ decays	321
11.2.4	Flavour oscillations and the <i>CPT</i> theorem	324
11.2.5	Direct <i>CP</i> violation in decay rates	328
11.2.6	$B^0 - \bar{B}^0$ mixing	329
*11.2.7	<i>CP</i> violation in interference	335
*11.2.8	Derivation of the mixing formulas	338
11.3	<i>CP</i> violation in the standard model	340
Problems 11		343
<b>*12</b>	<b>Beyond the standard model</b>	<b>346</b>
12.1	Grand unification	347
12.1.1	Quark and lepton charges	349
12.1.2	The weak mixing angle	349
12.1.3	Proton decay	350
12.2	Supersymmetry	354
12.2.1	The search for supersymmetry	356
12.3	Strings and things	358
12.4	Particle physics and cosmology	360
12.4.1	Dark matter	360
12.4.2	Matter–antimatter asymmetry	367
12.4.3	<i>CP</i> violation and electric dipole moments	369
12.4.4	Axions and the strong <i>CP</i> problem	371
12.5	Dirac or Majorana neutrinos?	373
12.5.1	Double beta decay	375
Problems 12		381
<b>A</b>	<b>Relativistic kinematics</b>	<b>383</b>
A.1	The Lorentz transformation for energy and momentum	383
A.2	The invariant mass	385
A.2.1	Beam energies and thresholds	385
A.2.2	Masses of unstable particles	387
*A.3	Transformation of the scattering angle	388
Problems A		390
<b>B</b>	<b>Amplitudes and cross-sections</b>	<b>392</b>
B.1	Rates and cross-sections	392
B.2	The total cross-section	394
B.3	Differential cross-sections	395
*B.4	The scattering amplitude	397
B.5	The Breit–Wigner formula	400
*B.5.1	Decay distributions	401
*B.5.2	Resonant cross-sections	404
Problems B		406

<b>*C The isospin formalism</b>	<b>408</b>
C.1 Isospin operators	409
C.2 Isospin states	411
C.3 Isospin multiplets	411
C.3.1 Hadron states	412
C.4 Branching ratios	414
C.5 Spin states	416
Problems C	416
<b>*D Gauge theories</b>	<b>418</b>
D.1 Electromagnetic interactions	419
D.2 Gauge transformations	420
D.3 Gauge invariance and the photon mass	421
D.4 The gauge principle	423
D.5 The Higgs mechanism	425
D.5.1 Charge and current densities	425
D.5.2 Spin-0 bosons	427
D.5.3 Spontaneous symmetry breaking	428
D.6 Quantum chromodynamics	429
D.7 Electroweak interactions	434
D.7.1 Weak isospin	434
D.7.2 Gauge invariance and charged currents	436
D.7.3 The unification condition	437
D.7.4 Spin structure and parity violation	440
Problems D	441
<b>E Answers to selected questions</b>	<b>443</b>
<i>References</i>	448
<i>Index</i>	451

---

# Editors' preface to the Manchester Physics Series

---

The Manchester Physics Series is a set of textbooks at first degree level. It grew out of the experience at the University of Manchester, widely shared elsewhere, that many textbooks contain much more material than can be accommodated in a typical undergraduate course; and that this material is only rarely so arranged as to allow the definition of a short self-contained course. The plan for this series was to produce short books so that lecturers would find them attractive for undergraduate courses, and so that students would not be frightened off by their encyclopaedic size or price. To achieve this, we have been very selective in the choice of topics, with the emphasis on the basic physics together with some instructive, stimulating and useful applications.

Although these books were conceived as a series, each of them is self-contained and can be used independently of the others. Several of them are suitable for wider use in other sciences. Each Author's Preface gives details about the level, prerequisites, etc., of that volume.

The Manchester Physics Series has been very successful since its inception over 40 years ago, with total sales of more than a quarter of a million copies. We are extremely grateful to the many students and colleagues, at Manchester and elsewhere, for helpful criticisms and stimulating comments. Our particular thanks go to the authors for all the work they have done, for the many new ideas they have contributed, and for discussing patiently, and often accepting, the suggestions of the editors.

Finally we would like to thank our publishers, John Wiley & Sons, Ltd., for their enthusiastic and continued commitment to the Manchester Physics Series.

J. R. Forshaw  
H. F. Gleeson  
F. K. Loebinger  
*August 2014*



---

## Authors' preface

---

*Particle Physics* is the study of the fundamental constituents of matter and the forces between them. For almost 40 years these have been described by the so-called *standard model of particle physics*, which provides, at least in principle, a basis for understanding most particle interactions, except gravity. The purpose of this book is to provide a short introduction to particle physics, which emphasises the foundations of the standard model in experimental data, rather than its more formal and theoretical aspects. The book is intended as an introduction to the subject for undergraduate students who have previously taken introductory courses in non-relativistic quantum mechanics and special relativity. No prior knowledge of particle physics is assumed.

The introductory nature of the book and the need to keep it reasonably short have influenced both the level of the treatment and the choice of material. We have tried to take a direct approach throughout, focusing on the interpretation of experimental data in terms of the basic properties of quarks and leptons, and extensive use has been made of symmetry principles and Feynman diagrams, which are introduced early in the book. We have not given any detailed calculations using the Feynman rules, since these are beyond the scope of a typical introductory undergraduate course.

Since publication of the Third Edition of this book, there have been many substantial developments in particle physics. These include: the long-awaited discovery of the Higgs boson, the first definitive identification of exotic states not predicted in the simple quark model and substantial progress in neutrino physics, *CP* violation and the exploration of the interface between particle physics and cosmology. However, these are only the highlights and our aim in producing this new edition is again to bring the book up to date throughout, while leaving its basic philosophy unchanged.

Fred Loebinger of Manchester University has read the book in its entirety and we are grateful to him for many helpful suggestions that have improved the presentation.

Brian Martin and Graham Shaw  
April 2016



---

## Suggested Short Course

---

A brief overview is given in Chapters 1, 2 and 3. For a short course, follow these chapters plus Chapters 4 and 5, Sections 6.1, 6.2, 6.4, 7.1, 7.2, 8.1, 8.2, Chapters 9 and 10, plus Sections 11.1 and 11.2.1, omitting all ‘starred’ material throughout.

The starred items contain more specialised or advanced material that is not required elsewhere in the book. They may be omitted at a first reading and used for reference when you want to know about particular topic.



---

# Notes

---

## Starred material

Some sections of this book are marked with a star. These contain more specialised or advanced material that is not required elsewhere in the book and may be omitted at first reading.

## Website

Any misprints or other necessary corrections brought to our notice will be listed on [www.wiley.com/go/martin/particlephysics4](http://www.wiley.com/go/martin/particlephysics4). We would also be grateful for any other comments about the book.

## Data

Except where otherwise stated, all data are from the latest edition of the biannual publication of the Particle Data Group (PDG), which is available on-line at [pdg.lbl.gov](http://pdg.lbl.gov). This also contains useful topical reviews of specific topics in particle physics, to which we often refer.

## Problems

Problems are provided for all chapters and appendices (except Appendix E). They are an integral part of the text. Some of the problems are numerical and require values of physical constants that are given on the inside rear cover. Brief ‘one-line’ answers to selected problems are given in Appendix E. Readers may access the full solutions to the odd-numbered problems at [www.wiley.com/go/martin/particlephysics4](http://www.wiley.com/go/martin/particlephysics4). Full solutions to all problems are available to instructors at the same website, which also contains electronic versions of the figures.

## Illustrations

Some illustrations in the text have been adapted from diagrams that have been published elsewhere. In a few cases they have been reproduced exactly as previously published. We acknowledge, with thanks, permission from the relevant copyright holders to use such illustrations and this is recorded in the captions.



## PHYSICAL CONSTANTS

**Note.** All entries are given to five significant figures where these are known. Otherwise the error is given in brackets.

<i>Quantity</i>	<i>Symbol</i>	<i>Value</i>
Speed of light in vacuum	$c$	$2.9979 \times 10^8$ m/s
Planck's constant	$\hbar$	$4.1356 \times 10^{-24}$ GeV s
Planck's constant reduced	$\hbar \equiv h/2\pi$	$6.5821 \times 10^{-25}$ GeV s
	$\hbar c$	$1.9733 \times 10^{-16}$ GeV m
	$(\hbar c)^2$	$3.8938 \times 10^{-32}$ GeV <sup>2</sup> m <sup>2</sup>
Electron charge (magnitude)	$e$	$1.6022 \times 10^{-19}$ C
Fine structure constant	$\alpha \equiv e^2/4\pi\varepsilon_0 \hbar c$	$7.2974 \times 10^{-3} = 1/137.04$
Electron mass	$m_e$	$0.51100$ MeV/c <sup>2</sup>
Neutron mass	$m_n$	$0.93957$ GeV/c <sup>2</sup>
Proton mass	$m_p$	$0.93827$ GeV/c <sup>2</sup>
$W$ boson mass	$M_W$	$80.385(0.015)$ GeV/c <sup>2</sup>
$Z$ boson mass	$M_Z$	$91.188(0.002)$ GeV/c <sup>2</sup>
Higgs boson mass	$M_H$	$125.09(0.3)$ GeV/c <sup>2</sup>
Fermi coupling constant	$G_F/(\hbar c)^3$	$1.1664 \times 10^{-5}$ GeV <sup>-2</sup>
Strong coupling constant	$\alpha_s(M_Z c^2)$	$0.1185(0.0006)$

## CONVERSION FACTORS

$$\begin{aligned} 1 \text{ eV} &= 1.6022 \times 10^{-19} \text{ J} & 1 \text{ eV/c}^2 &= 1.7827 \times 10^{-36} \text{ kg} \\ 1 \text{ fermi} &= 1 \text{ fm} \equiv 10^{-15} \text{ m} & 1 \text{ barn} &= 1 \text{ b} \equiv 10^{-28} \text{ m}^2 \end{aligned}$$

## NATURAL UNITS

Throughout the book we have used natural units,  $\hbar = c = 1$ . They can be converted back to mks units by the method discussed in Section 1.5. The following conversion factors are often useful.

Time	$1 \text{ GeV}^{-1} = 6.5821 \times 10^{-25} \text{ s}$
Length	$1 \text{ GeV}^{-1} = 0.19733 \text{ fm}$
Cross-section	$1 \text{ GeV}^{-2} = 0.38938 \text{ mb}$



# 1

---

## Some basic concepts

---

### 1.1 Introduction

Particle physics is the study of the fundamental constituents of matter and their interactions. However, which particles are regarded as fundamental has changed with time as physicists' knowledge has improved. Modern theory – called the *standard model* – attempts to explain all the phenomena of particle physics in terms of the properties and interactions of a small number of particles of four distinct types: two spin-1/2 families of fermions called *leptons* and *quarks*; one family of spin-1 bosons – called *gauge bosons* – which act as ‘force carriers’ in the theory; and a spin-0 particle, called the *Higgs boson*, which explains the origin of mass within the theory, since without it, leptons, quarks and gauge bosons would all be massless. All the particles of the standard model are assumed to be *elementary*: that is they are treated as point particles, without internal structure or excited states.

The most familiar example of a lepton is the *electron*  $e^-$  (the superscript denotes the electric charge), which is bound in atoms by the *electromagnetic interaction*, one of the four fundamental forces of nature. A second well-known lepton is the *electron neutrino*  $\nu_e$ , which is a light, neutral particle observed in the decay products of some unstable nuclei (the so-called  $\beta$  decays). The force responsible for the  $\beta$  decay of nuclei is called the *weak interaction*.

Another class of particles called *hadrons* is also observed in nature. Familiar examples are the neutron  $n$  and proton  $p$  (collectively called *nucleons*) and the three *pions* ( $\pi^+$ ,  $\pi^-$ ,  $\pi^0$ ), where the superscripts again denote the electric charges. These are not elementary particles, but are made of quarks bound together by a third force of nature, the *strong interaction*. The theory is unusual in that

the quarks themselves are not directly observable, only their bound states. Nevertheless, we shall see in later chapters that there is overwhelming evidence for the existence of quarks and we shall discuss the reason why they are unobservable as free particles. The strong interaction between quarks gives rise to the observed strong interaction between hadrons, such as the nuclear force that binds nucleons in nuclei. There is an analogy here with the fundamental electromagnetic interaction between electrons and nuclei that also gives rise to the more complicated forces between their bound states, that is between atoms.

In addition to the strong, weak, and electromagnetic interactions, there is a fourth force of nature – gravity. However, the gravitational interaction between elementary particles is so small compared to those from the other three interactions that it can be neglected at presently accessible energies. Because of this, we will often refer in practice to the *three* forces of nature.

The standard model also specifies the origin of these three forces. Consider firstly the electromagnetic interaction. In classical physics this is propagated by electromagnetic waves, which are continuously emitted and absorbed. While this is an adequate description at long distances, at short distances the quantum nature of the interaction must be taken into account. In quantum theory, the interaction is transmitted discontinuously by the exchange of spin-1 photons, which are the ‘force carriers’, or *gauge bosons*, of the electromagnetic interaction and, as we shall see presently, the long-range nature of the force is related to the fact that photons have zero mass. The use of the word ‘gauge’ refers to the fact that the electromagnetic interaction possesses a fundamental symmetry called *gauge invariance*. This property is common to all the three interactions of nature and has profound consequences, as we shall see.

The weak and strong interactions are also associated with the exchange of spin-1 particles. For the weak interaction, they are the charged  $W^\pm$  and the neutral  $Z^0$  *bosons*, with masses about 80–90 times the mass of the proton. The resulting force is very short range, and in many applications may be approximated by an interaction at a point. The equivalent particles for the strong interaction are called *gluons*  $g$ . There are eight gluons, all of which have zero mass and are electrically neutral, like the photon. Thus, by analogy with electromagnetism, the basic strong interaction between quarks is long range. The ‘residual’ strong interaction between the quark bound states (hadrons) is not the same as the fundamental strong interaction between quarks (but is a consequence of it) and is short range, again as we shall see later.

In the standard model, which will play a central role in this book, the main actors are the leptons and quarks, which are the basic constituents of matter; the ‘force carriers’ (the photon, the  $W$  and  $Z$

bosons, and the gluons) that mediate the interactions between them; and the Higgs boson, which gives mass to the elementary particles of the standard model. In addition, because not all these particles are directly observable, quark bound states (i.e. hadrons) also play a very important role.

## 1.2 Antiparticles

In particle physics, high energies are needed both to create new particles and to explore the structure of hadrons. The latter requires projectiles whose wavelengths  $\lambda$  are at least as small as hadron radii, which are of order  $10^{-15}$  m. It follows that their momenta  $p = h/\lambda$  must be several hundred MeV/c (1 MeV =  $10^6$  eV), and hence their energies  $E$  must be several hundred MeV. Because of this, any theory of elementary particles must combine the requirements of both special relativity and quantum theory. This has the startling consequence that for every charged particle of nature, whether it is one of the elementary particles of the standard model, or a hadron, there exists an associated particle of the same mass, but opposite charge, called its *antiparticle*. This important theoretical prediction was first made for spin-1/2 particles by Dirac in 1928, and follows from the solutions of the equation he first wrote down to describe relativistic electrons. We therefore start by considering how to construct a relativistic wave equation.

### 1.2.1 Relativistic wave equations

We start from the assumption that a particle moving with momentum  $\mathbf{p}$  in free space is described by a de Broglie wave function<sup>1</sup>

$$\Psi(\mathbf{r}, t) = N e^{i(\mathbf{p} \cdot \mathbf{r} - Et)/\hbar}, \quad (1.1)$$

with frequency  $\nu = E/h$  and wavelength  $\lambda = h/p$ . Here  $p \equiv |\mathbf{p}|$  and  $N$  is a normalisation constant that is irrelevant in what follows. The corresponding wave equation depends on the assumed relation between the energy  $E$  and momentum  $\mathbf{p}$ . Non-relativistically,

$$E = p^2/2m \quad (1.2)$$

and the wave function (1.1) obeys the non-relativistic Schrödinger equation

$$i\hbar \frac{\partial \Psi(\mathbf{r}, t)}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi(\mathbf{r}, t). \quad (1.3)$$

---

<sup>1</sup> We use the notation  $\mathbf{r} = (x_1, x_2, x_3) = (x, y, z)$ .

Relativistically, however,

$$E^2 = p^2 c^2 + m^2 c^4, \quad (1.4)$$

where  $m$  is the rest mass,<sup>2</sup> and the corresponding wave equation is

$$-\hbar^2 \frac{\partial^2 \Psi(\mathbf{r}, t)}{\partial t^2} = -\hbar^2 c^2 \nabla^2 \Psi(\mathbf{r}, t) + m^2 c^4 \Psi(\mathbf{r}, t), \quad (1.5)$$

as is easily checked by substituting (1.1) into (1.5) and using (1.4). This equation was first proposed by de Broglie in 1924, but is now more usually called the *Klein–Gordon equation*.<sup>3</sup> Its most striking feature is the existence of solutions with negative energy. For every plane wave solution of the form

$$\Psi(\mathbf{r}, t) = N \exp [i(\mathbf{p} \cdot \mathbf{r} - E_p t)/\hbar], \quad (1.6a)$$

with momentum  $\mathbf{p}$  and positive energy

$$E = E_p \equiv +(p^2 c^2 + m^2 c^4)^{1/2} \geq mc^2$$

there is also a solution

$$\tilde{\Psi}(\mathbf{r}, t) \equiv \Psi^*(\mathbf{r}, t) = N^* \exp [i(-\mathbf{p} \cdot \mathbf{r} + E_p t)/\hbar], \quad (1.6b)$$

corresponding to momentum  $-\mathbf{p}$  and negative energy

$$E = -E_p = -(p^2 c^2 + m^2 c^4)^{1/2} \leq -mc^2.$$

Other problems also occur, indicating that the Klein–Gordon equation is not, in itself, a sufficient foundation for relativistic quantum mechanics. In particular, it does not guarantee the existence of a positive-definite probability density for position.<sup>4</sup>

The existence of negative energy solutions is a direct consequence of the quadratic nature of the mass–energy relation (1.4) and cannot be avoided in a relativistic theory. However, for spin-1/2 particles the other problems were resolved by Dirac in 1928, who looked for an equation of the familiar form

$$i\hbar \frac{\partial \Psi(\mathbf{r}, t)}{\partial t} = H(\mathbf{r}, \hat{\mathbf{p}})\Psi(\mathbf{r}, t), \quad (1.7)$$

where  $H$  is the Hamiltonian and  $\hat{\mathbf{p}} = -i\hbar\nabla$  is the momentum operator. Since (1.7) is first order in  $\partial/\partial t$ , Lorentz invariance requires that

---

<sup>2</sup> From now on, the word *mass* will be used to mean the rest mass.

<sup>3</sup> These authors incorporated electromagnetic interactions into the equation, in a form now known to be appropriate for charged spin-0 bosons.

<sup>4</sup> For a discussion of this point, see for example pp. 467–468 of Schiff (1968).

it also be first order in spatial derivatives. Dirac therefore proposed a Hamiltonian of the general form

$$H = -i\hbar c \sum_{i=1}^3 \alpha_i \frac{\partial}{\partial x_i} + \beta mc^2 = c \boldsymbol{\alpha} \cdot \hat{\mathbf{p}} + \beta mc^2, \quad (1.8)$$

in which the coefficients  $\beta$  and  $\alpha_i (i = 1, 2, 3)$  are determined by requiring that solutions of the Dirac equation (1.8) are also solutions of the Klein–Gordon equation (1.5). Acting on (1.7) with  $i\hbar\partial/\partial t$  and comparing with (1.5), leads to the conclusion that this is true if, and only if,

$$\alpha_i^2 = 1, \beta^2 = 1, \quad (1.9a)$$

$$\alpha_i \beta + \beta \alpha_i = 0 \quad (1.9b)$$

and

$$\alpha_i \alpha_j + \alpha_j \alpha_i = 0 \quad (i \neq j). \quad (1.9c)$$

These relations cannot be satisfied by ordinary numbers, and the simplest assumption is that  $\beta$  and  $\alpha_i (i = 1, 2, 3)$  are matrices, which must be hermitian so that the Hamiltonian is hermitian. The smallest matrices satisfying these requirements have dimensions  $4 \times 4$  and are given in many books,<sup>5</sup> but are not required below. We thus arrive at an interpretation of the Dirac equation

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi = -i\hbar c \sum_i \alpha_i \frac{\partial \Psi}{\partial x_i} + \beta mc^2 \Psi \quad (1.10)$$

as a four-dimensional matrix equation in which the  $\Psi$  are four-component wave functions

$$\Psi(\mathbf{r}, t) = \begin{pmatrix} \psi_1(\mathbf{r}, t) \\ \psi_2(\mathbf{r}, t) \\ \psi_3(\mathbf{r}, t) \\ \psi_4(\mathbf{r}, t) \end{pmatrix}, \quad (1.11)$$

called *spinors*. Plane wave solutions take the form

$$\Psi(\mathbf{r}, t) = u(\mathbf{p}) \exp [i(\mathbf{p} \cdot \mathbf{r} - Et)/\hbar], \quad (1.12)$$

where  $u(\mathbf{p})$  is also a four-component spinor satisfying the eigenvalue equation

$$H_{\mathbf{p}} u(\mathbf{p}) \equiv (c \boldsymbol{\alpha} \cdot \mathbf{p} + \beta mc^2) u(\mathbf{p}) = Eu(\mathbf{p}), \quad (1.13)$$

obtained by substituting (1.11) into (1.10). This equation has four solutions:<sup>6</sup> two with positive energy  $E = +E_p$  corresponding to the two possible spin states of a spin-1/2 particle (called ‘spin up’ and

<sup>5</sup> See, for example, pp. 473–475 of Schiff (1968).

<sup>6</sup> A proof of these results is given in, for example, pp. 475–477 of Schiff (1968).

‘spin down’, respectively) and two corresponding negative energy solutions with  $E = -E_p$ .

The problem of the negative-energy solutions will be resolved in the next section. Here we note that the positive-energy solutions of the Dirac equation lead to many predictions that have been verified experimentally to a very high precision. Notable among these are relativistic corrections in atomic spectroscopy, including spin-orbit effects and the prediction that point-like spin-1/2 particles of mass  $m$  and charge  $q$  have a Dirac magnetic moment

$$\mu_D = q \mathbf{S}/m, \quad (1.14)$$

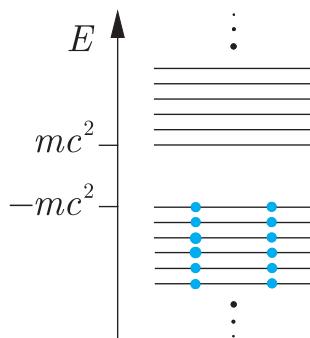
where  $\mathbf{S}$  is the spin vector. This is a key result. It not only yields the correct value for the electron but provides a simple test for the point-like nature of any other spin-1/2 fermion. For the proton and neutron, the experimental values are

$$\mu_p = 2.79 e\mathbf{S}/m_p \quad \text{and} \quad \mu_n = -1.91 e\mathbf{S}/m_n, \quad (1.15)$$

in disagreement with Equation (1.14). Historically, the measurement of the proton magnetic moment by Frisch and Stern in 1933 was the first indication that the proton was not a point-like elementary particle.

### 1.2.2 Hole theory and the positron

The problem of the negative energy states remains. They cannot be ignored, since their existence leads to unacceptable consequences. For example, if such states are unoccupied, then transitions from positive to negative energy states could occur, leading to the prediction that atoms such as hydrogen would be unstable. This problem was resolved by Dirac, who postulated that the negative energy states are almost always filled. For definiteness, consider the case of electrons. Since they are fermions, they obey the Pauli exclusion principle, and the Dirac picture of the vacuum is a so-called ‘sea’ of negative energy states, each with two electrons (one with spin ‘up’, the other with spin ‘down’), while the positive energy states are all unoccupied (see Figure 1.1). This state is indistinguishable from the usual vacuum with  $E_V = 0$ ,  $\mathbf{p}_V = \mathbf{0}$ , etc. This is because for each state of momentum  $\mathbf{p}$  there is a corresponding state with momentum  $-\mathbf{p}$ , so that the momentum of the vacuum  $\mathbf{p}_V = \sum \mathbf{p} = \mathbf{0}$ . The same argument applies to spin, while, since energies are measured relative to the vacuum,  $E_V \equiv 0$  by definition. Similarly, we may define the charge  $Q_V \equiv 0$ , because the constant electrostatic potential produced by the negative energy sea is unobservable. Thus this state has all the *measurable* characteristics of the naive vacuum, and the ‘sea’ is unobservable.



**Figure 1.1** Dirac picture of the vacuum. The sea of negative energy states is totally occupied with two electrons in each level, one with spin ‘up’ and one with spin ‘down’. The positive energy states are all unoccupied.

Dirac's postulate solves the problem of unacceptable transitions from positive energy states, but has other consequences. Consider what happens when an electron is added to, or removed from, the vacuum. In the former case, the electron is confined to the positive energy region since all the negative energy states are occupied. In the latter case, *removing* a negative energy electron with  $E = -E_p < 0$ , momentum  $-\mathbf{p}$ , spin  $-\mathbf{S}$  and charge  $-e$  from the vacuum (which has  $E_V = 0, \mathbf{p}_V = 0, \mathbf{S}_V = 0, Q_V = 0$ ) leaves a state (the sea with a 'hole' in it) with positive energy  $E = E_p > 0$ , momentum  $\mathbf{p}$ , spin  $\mathbf{S}$  and charge  $+e$ . This state cannot be distinguished by any measurement from a state formed by *adding* to the vacuum a particle with momentum  $\mathbf{p}$ , energy  $E = E_p > 0$ , spin  $\mathbf{S}$  and charge  $+e$ . The two cases are equivalent descriptions of the same phenomena. Using the latter, Dirac predicted the existence of a spin-1/2 particle  $e^+$  with the same mass as the electron, but opposite charge. This particle is called the *positron* and is referred to as the *antiparticle* of the electron.<sup>7</sup>

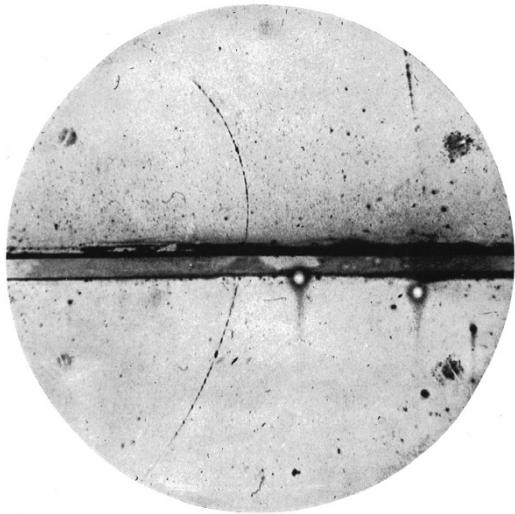
The positron was subsequently discovered by Anderson, and by Blackett and Ochialini, in 1933. The discovery was made using a device of great historical importance, called a *cloud chamber*. When a charged particle passes through matter, it interacts with it, losing energy. This energy can take the form of radiation or of excitation and ionisation of the atoms along the path.<sup>8</sup> It is the aim of *track chambers* – of which the cloud chamber is the earliest example – to produce a visible record of this trail and hence of the particle that produced it.

The cloud chamber was devised by C.T.R. Wilson, who noticed that the condensation of water vapour into droplets goes much faster in the presence of ions. It consisted of a vessel filled with air almost saturated with water vapour and fitted with an expansion piston. When the vessel was suddenly expanded, the air cooled and became supersaturated. Droplets were then formed preferentially along the trails of ions left by charged particles passing through the chamber. The chamber was illuminated by a flash of light immediately after expansion, and the tracks of droplets so revealed were photographed before they had time to disperse.

Figure 1.2 shows one of the first identified positron tracks observed by Anderson in 1933. The band across the centre of the picture is a 6 mm lead plate inserted to slow particles down. The track is curved due to the presence of a 1.5 T applied magnetic field

<sup>7</sup> This prediction, which is now regarded as one of the greatest successes of theoretical physics, was not always so enthusiastically received at the time. For example, in 1933 Pauli wrote: 'Dirac has tried to identify holes with antielectrons. We do not believe that this explanation can be seriously considered' (Pauli, 1933).

<sup>8</sup> This will be discussed in detail in Chapter 4.



**Figure 1.2** One of the first positron tracks observed by Anderson in a Wilson cloud chamber (see text for details). (Anderson 1933. Reproduced with permission from the American Physical Society.)

**B**, and since the curvature of such tracks increases with decreasing momentum, we can conclude that the particle enters at the bottom of the picture and travels upwards. The sign of the particle's charge  $q$  then follows from the direction of the Lorentz force  $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$ , where  $\mathbf{v}$  is the particle's velocity, and hence of the curvature; it is positive.

That the particle is a positron and not a proton follows essentially from the range of the upper track. The rate of energy loss of a charge particle in matter depends on its charge and velocity. (This will be discussed in Chapter 4.) From the curvature of the tracks, one can deduce that the momentum of the upper track is 23 MeV/ $c$ , corresponding to either a slow-moving proton with speed  $v \ll c$  or a relativistic ( $v \approx c$ ) positron. The former would lose energy rapidly, coming to rest in a distance of about 5 mm, comparable with the thickness of the lead plate. The observed track length is more than 5 cm, enabling a limit  $m_+ \leq 20m_e \ll m_p$  to be set on the mass  $m_+$  of the particle, which Anderson suggested was a positron. Many other examples were found, especially by Blackett and Ochialini, and by 1934 Blackett, Ochialini and Chadwick had established that  $m_+ = m_e$  within experimental errors of order 10%. The interpretation of the light positive particles as positrons was thus established beyond all reasonable doubt.

The Dirac equation applies to any spin-1/2 particle, and hole theory predicts that *all* charged spin-1/2 particles, whether they are elementary or hadrons, have distinct antiparticles with opposite charge, but the same mass. The argument does not extend to bosons, because they do not obey the exclusion principle on which hole theory depends, and to show that charged bosons also have antiparticles of opposite charge requires the formal apparatus of

relativistic quantum field theory.<sup>9</sup> We shall not pursue this here, but note that the basic constituents of matter – the leptons and quarks – are not bosons, but are spin-1/2 fermions. The corresponding results on the antiparticles of hadrons, irrespective of their spin, can then be found by considering their quark constituents, as we shall see in Chapter 3. For neutral particles, there is no general rule governing the existence of antiparticles, and while some neutral particles have distinct antiparticles associated with them, others do not. The photon, for example, does not have an antiparticle (or, rather, the photon and its antiparticle are identical) whereas the neutron does. Although the neutron has zero charge, it has a non-zero magnetic moment, and distinct antineutrons exist in which the sign of this magnetic moment is reversed relative to the spin direction. The neutron is also characterised by other quantum numbers (that we will meet later) that change sign between particle and antiparticle.

In what follows, if we denote a particle by  $P$ , then the antiparticle is in general written with a bar over it, that is  $\bar{P}$ . For example, the antiparticle of the proton  $p$  is the antiproton  $\bar{p}$ , with negative electric charge, and associated with every quark,  $q$ , is an antiquark,  $\bar{q}$ . However, for some common particles the bar is usually omitted. Thus, for example, in the case of the positron  $e^+$ , the superscript denoting the charge makes explicit the fact that the antiparticle has the opposite electric charge to that of its associated particle, the electron.

### 1.3 Interactions and Feynman diagrams

By analogy with chemical reactions, interactions involving elementary particles and/or hadrons are conveniently summarised by ‘equations’, in which the different particles are represented by symbols. Thus, in the reaction  $\nu_e + n \rightarrow e^- + p$ , an electron neutrino  $\nu_e$  interacts with a neutron  $n$  to produce an electron  $e^-$  and a proton  $p$ ; while  $e^- + p \rightarrow e^- + p$  represents an electron and proton interacting to give the same particles in the final state, but in general travelling in different directions. The latter, where the initial and final particles are the same, is an example of *elastic scattering*, whereas  $\nu_e + n \rightarrow e^- + p$ , where the initial and final particles differ, is an example of an *inelastic reaction*. The forces producing the above interactions are due to the exchange of particles and a convenient way of illustrating this is to use the pictorial technique of *Feynman diagrams*. These were introduced by Feynman in the 1940s and are now one of the cornerstones of elementary particle physics.

---

<sup>9</sup> For an introduction, see for example Mandl and Shaw (2010).

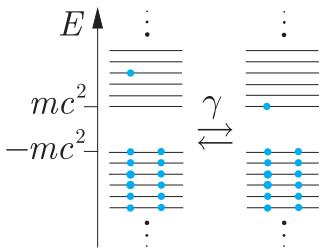
Associated with them are mathematical rules and techniques that enable the calculation of the quantum mechanical probabilities for given reactions to occur. Here we will avoid the mathematical detail but use the diagrams to understand the main features of particle reactions. We will introduce Feynman diagrams by firstly discussing electromagnetic interactions.

### 1.3.1 Basic electromagnetic processes

The electromagnetic interactions of electrons and positrons can all be understood in terms of eight basic processes. In hole theory, they arise from transitions in which an electron jumps from one state to another, with the emission or absorption of a single photon. The interpretation then depends on whether the states are both of positive energy, both of negative energy or one of each.

The basic processes whereby an electron either emits or absorbs a photon are

$$(a) e^- \rightarrow e^- + \gamma \quad \text{and} \quad (b) \gamma + e^- \rightarrow e^-.$$



**Figure 1.3** Hole theory representation of the processes  $e^- \rightleftharpoons e^- + \gamma$ .

They correspond in hole theory to transitions between positive energy states of the electron and are represented pictorially in Figure 1.3. They may also be represented diagrammatically by Figures 1.4(a) and (b), where by convention time runs from left to right. These are examples of Feynman diagrams.

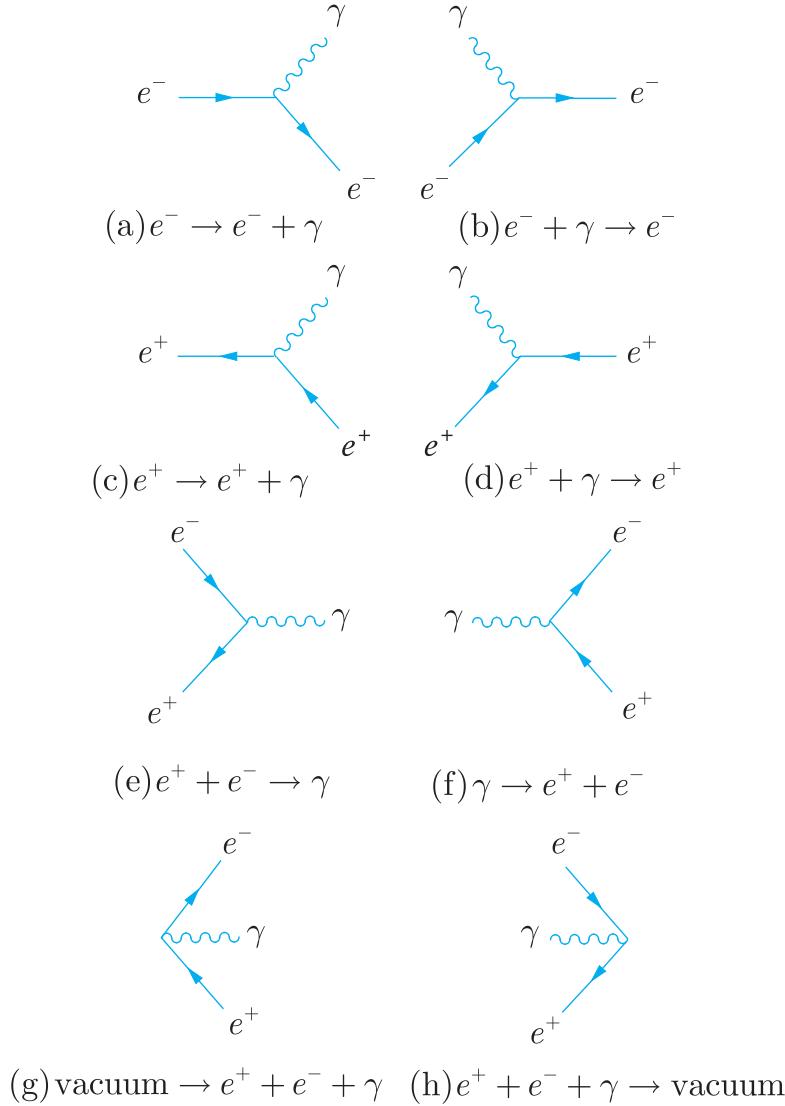
Similar diagrams may be drawn for the corresponding positron processes

$$(c) e^+ \rightarrow e^+ + \gamma \quad \text{and} \quad (d) \gamma + e^+ \rightarrow e^+,$$

and are shown in Figures 1.4(c) and (d). Time again flows to the right, and we have used the convention that an arrow directed towards the right indicates a particle (in this case an electron) while one directed to the left indicates an antiparticle (in this case a positron). The corresponding hole theory diagram, analogous to Figure 1.3, is left as an exercise for the reader. Finally, there are processes in which an electron is excited from a negative energy state to a positive energy state, leaving a ‘hole’ behind, or in which a positive energy electron falls into a vacant level (hole) in the negative energy sea. These are illustrated in Figure 1.5, and correspond to the production or annihilation of  $e^+e^-$  pairs. In both cases, a photon may either be absorbed from the initial state or emitted to the final state, giving the four processes

$$(e) e^+ + e^- \rightarrow \gamma, \quad (f) \gamma \rightarrow e^+ + e^-, \\ (g) \text{vacuum} \rightarrow \gamma + e^+ + e^-, \quad (h) \gamma + e^+ + e^- \rightarrow \text{vacuum},$$

represented by the Feynman diagrams of Figures 1.4(e) to (h).



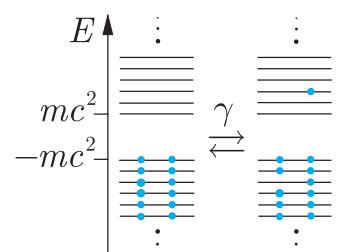
**Figure 1.4** Feynman diagrams for the eight basic processes whereby electrons and positrons interact with photons. In all such diagrams, time runs from left to right, while a solid line with its arrow pointing to the right (left) indicates an electron (positron).

This exhausts the possibilities in hole theory, so that there are just eight basic processes represented by the Feynman diagrams, Figures 1.4(a) to (h). Each of these processes has an associated probability of the order of the electromagnetic fine structure constant

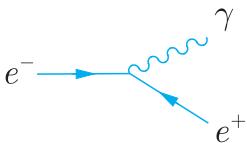
$$\alpha \equiv \frac{1}{4\pi\varepsilon_0}\frac{e^2}{\hbar c} \approx \frac{1}{137}. \quad (1.16)$$

### 1.3.2 Real processes

In each of the diagrams in Figures 1.4(a) to (h), each vertex has a line corresponding to a single photon being emitted or absorbed,



**Figure 1.5** Hole theory representation of the production or annihilation of  $e^+ e^-$  pairs.



**Figure 1.6** The forbidden vertex  $e^- \rightarrow e^+ + \gamma$ .

while one fermion line has the arrow pointing toward the vertex and the other away from the vertex, implying charge conservation at the vertex.<sup>10</sup> For example, a vertex like Figure 1.6 would correspond to a process in which an electron emitted a photon and turned into a positron. This would violate charge conservation and is therefore forbidden.

Momentum and angular momentum are also assumed to be conserved at the vertices. However, in free space, energy conservation is violated. For example, if we use the notation  $(E, \mathbf{k})$  to denote the total energy and three-momentum of a particle, then in the rest frame of the electron, reaction (a), is

$$e^-(E_0, \mathbf{0}) \rightarrow e^-(E_k, -\mathbf{k}) + \gamma(ck, \mathbf{k}), \quad (1.17)$$

where  $k \equiv |\mathbf{k}|$  and momentum conservation has been imposed. In free space,  $E_0 = mc^2$ ,  $E_k = (k^2 c^2 + m^2 c^4)^{1/2}$  and  $\Delta E \equiv E_k + kc - E_0$  satisfies

$$kc < \Delta E < 2kc \quad (1.18)$$

for all finite  $\mathbf{k}$ .

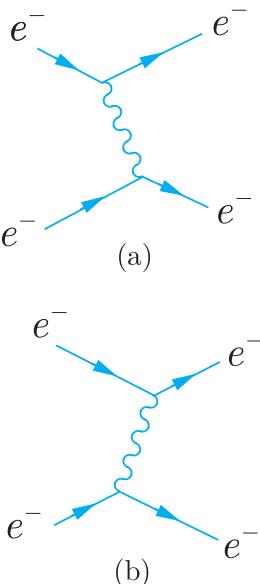
Similar arguments show that energy conservation is violated for all the basic processes. They are called *virtual* processes to emphasise that they cannot occur in isolation in free space. To make a real process, two or more virtual processes must be combined in such a way that energy conservation is only violated for a short period of time  $\tau$  compatible with the energy-time uncertainty principle

$$\tau \Delta E \approx \hbar. \quad (1.19)$$

In particular, the initial and final states – which in principle can be studied in the distant past ( $t \rightarrow -\infty$ ) and future ( $t \rightarrow +\infty$ ), respectively – must have the same energy. This is illustrated by Figure 1.7(a), which represents a process whereby an electron emits a photon that is subsequently absorbed by a second electron. Although energy conservation is violated at the first vertex, this can be compensated by a similar violation at the second vertex to give exact energy conservation overall. Figure 1.7(a) represents a contribution to the physical elastic scattering process

$$e^- + e^- \rightarrow e^- + e^-$$

from a single-photon exchange. There is also a second contribution, represented by Figure 1.7(b) in which the other electron emits the exchanged photon. Both processes contribute to the observed scattering.



**Figure 1.7** Single-photon exchange contributions to electron–electron scattering. Time as usual runs from left to right.

<sup>10</sup> Compare Kirchhoff's laws in electromagnetism.

Scattering can also occur via multi-photon exchange and, for example, one of the diagrams corresponding to two-photon exchange is shown in Figure 1.8. The contributions of such diagrams are, however, far smaller than the one-photon exchange contributions. To see this, we consider the number of vertices in each diagram, called its *order*. Since each vertex represents a basic process whose probability is of order  $\alpha \approx 1/137 \ll 1$ , any diagram of order  $n$  gives a contribution whose probability is of order  $\alpha^n$ . On comparing Figures 1.7 and 1.8, we see that single-photon exchange is of order  $\alpha^2$ , two-photon exchange is of order  $\alpha^4$  and, more generally,  $n$ -photon exchange is of order  $\alpha^{2n}$ . To a good approximation multi-photon exchanges can be neglected, and we would expect the familiar electromagnetic interactions used in atomic spectroscopy, for example, to be accurately reproduced by considering only one-photon exchange. Detailed calculation confirms that this is indeed the case.

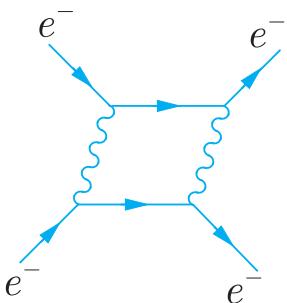
### 1.3.3 Electron–positron pair production and annihilation

Feynman diagrams like Figures 1.7 and 1.8 play a central role in the analysis of elementary particle interactions. The contribution of each diagram to the probability amplitude for a given physical process can be calculated precisely by using the set of mathematical rules mentioned previously (the *Feynman rules*), which are derived from the quantum theory of the corresponding interaction. For electromagnetic interactions, this theory is called *quantum electrodynamics* (or *QED* for short) and the resulting theoretical predictions have been verified experimentally with quite extraordinary precision. The Feynman rules are beyond the scope of this book and our use of Feynman diagrams will be much more qualitative. For this, we need consider only the lowest-order diagrams contributing to a given physical process, neglecting higher-order diagrams like Figure 1.8 for electron–electron scattering.

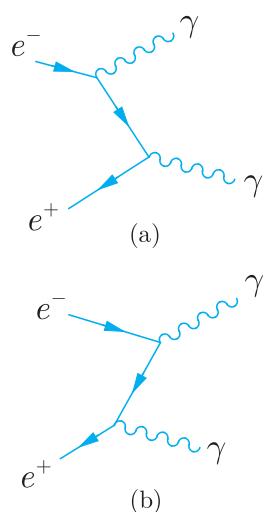
To illustrate this we consider the reactions  $e^+ + e^- \rightarrow p\gamma$ , where to conserve energy at least two photons must be produced,  $p \geq 2$ . In lowest order, to produce  $p$  photons we must combine  $p$  vertices from the left-hand side of Figure 1.4. For  $p = 2$ ,

$$e^+ + e^- \rightarrow \gamma + \gamma$$

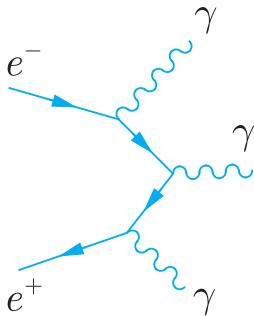
and there are just two such diagrams: Figure 1.9(a), which is obtained by combining Figures 1.4(a) and (e), and Figure 1.9(b), which combines Figures 1.4(c) and (e). These two diagrams are closely related. If the lines of Figure 1.9(a) were made of rubber, we could imagine deforming them so that the top vertex occurred after, instead of before, the bottom vertex, and it became Figure 1.9(b). Figures 1.7(a) and (b) are related in the same way. Diagrams related in this way are called different ‘time orderings’ and, in practice, it is



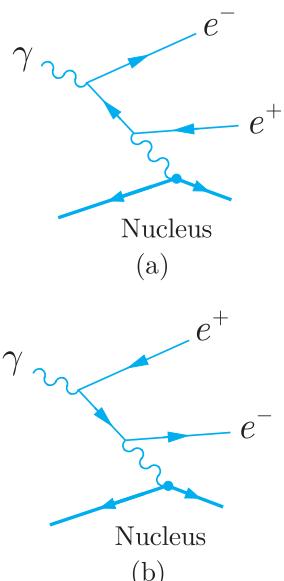
**Figure 1.8** A contribution to electron–electron scattering from two-photon exchange.



**Figure 1.9** Lowest-order contributions to  $e^+ + e^- \rightarrow \gamma + \gamma$ . The two diagrams are related by ‘time ordering’, as explained in the text.



**Figure 1.10** The process  $e^+ + e^- \rightarrow \gamma + \gamma + \gamma$  in lowest order. Only one of the six possible time orderings is shown, leaving the other five implied.



**Figure 1.11** The pair production process  $\gamma + (Z, A) \rightarrow e^+ + e^- + (Z, A)$  in lowest order. The two diagrams represent ‘distinct’ contributions and are not related by time ordering.

usual to draw only one time ordering (e.g. Figure 1.9(a)) leaving the other(s) implied. Thus for  $p = 3$ ,

$$e^+ + e^- \rightarrow \gamma + \gamma + \gamma,$$

and a possible diagram is that of Figure 1.10, obtained by combining Figures 1.4(a), (c) and (e). Since there are three vertices, there are  $3! = 6$  different ways of ordering them in time. We leave it as an exercise for the reader to draw the other five time-ordered diagrams whose existence is implied by Figure 1.10.

In general, the process  $e^+ + e^- \rightarrow p\gamma$  is of order  $p$ , with an associated probability of order  $\alpha^p$ . From just the order of the diagrams (i.e. the number of vertices), we therefore expect that many-photon annihilation is very rare compared to few-photon annihilation, and that<sup>11</sup>

$$R \equiv \frac{\text{Rate}(e^+ e^- \rightarrow 3\gamma)}{\text{Rate}(e^+ e^- \rightarrow 2\gamma)} = \mathcal{O}(\alpha). \quad (1.20)$$

For very low energy  $e^+ e^-$  pairs, this prediction can be tested by measuring the  $2\gamma$  and  $3\gamma$  decay rates of positronium, which is a bound state of  $e^+$  and  $e^-$  analogous to the hydrogen atom.<sup>12</sup> In this case, the experimental value of  $R$  is  $0.9 \times 10^{-3}$ . This is somewhat smaller than  $\alpha = 0.7 \times 10^{-2}$ , and indeed some authors argue that  $\alpha/2\pi = 1.2 \times 10^{-3}$  is a more appropriate measure of the strength of the electromagnetic interaction. We stress that (1.20) is only an order-of-magnitude prediction, and we will in future use  $10^{-2} - 10^{-3}$  as a rough rule-of-thumb estimate of what is meant by a factor of order  $\alpha$ .

As a second example, consider the pair production reaction  $\gamma \rightarrow e^+ + e^-$ . This basic process cannot conserve both energy and momentum simultaneously, but is allowed in the presence of a nucleus, that is

$$\gamma + (Z, A) \rightarrow e^+ + e^- + (Z, A),$$

where  $Z$  and  $A$  are the charge and mass number (the number of nucleons) of the nucleus, respectively. Diagrammatically, this is shown in Figure 1.11, where the two diagrams are not different time orderings. (There is no way that (a) can be continuously deformed into (b).) Since one of the vertices involves a charge  $Ze$ , the corresponding factor  $\alpha \rightarrow Z^2\alpha$ , and the expected rate is of order  $Z^2\alpha^3$ . This  $Z$  dependence is confirmed experimentally, and is exploited in devices

<sup>11</sup>We will often use the symbol  $\mathcal{O}$  to mean ‘order’ in the sense of ‘order-of-magnitude’.

<sup>12</sup>This is discussed in Section 5.5.

for detecting high-energy photons by the pair production that occurs readily in the field of a heavy nucleus.<sup>13</sup>

### 1.3.4 Other processes

Although we have introduced Feynman diagrams in the context of electromagnetic interactions, their use is not restricted to this. They can also be used to describe the fundamental weak and strong interactions. This is illustrated by Figure 1.12(a), which shows a contribution to the elastic weak scattering reaction  $e^- + \nu_e \rightarrow e^- + \nu_e$  due to the exchange of a  $Z^0$ , and by Figure 1.12(b), which shows the exchange of a gluon  $g$  (represented by a coiled line) between two quarks, which is a strong interaction.

Feynman diagrams that involve hadrons can also be drawn. As an illustration, Figure 1.13 shows the decay of a neutron via an intermediate charged  $W$  boson. In later chapters we will make extensive use of Feynman diagrams in discussing particle interactions.

## 1.4 Particle exchange

In Section 1.1 we said that the forces of elementary particle physics were associated with the exchange of particles. In this section we will explore the fundamental relation between the range of a force and the mass of the exchanged particle and show how the weak interaction can frequently be approximated by a force of zero range.

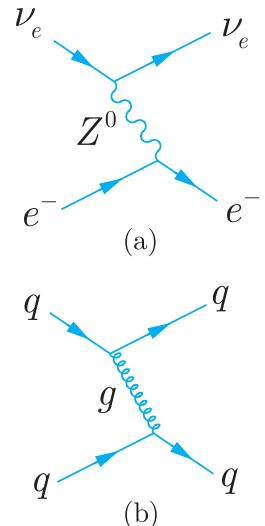
### 1.4.1 Range of forces

The Feynman diagram of Figure 1.14 represents the elastic scattering of two particles  $A$  and  $B$ , of masses  $M_A$  and  $M_B$ , that is  $A + B \rightarrow A + B$ , via the exchange of a third particle  $X$  of mass  $M_X$ , with equal coupling strengths  $g$  to particles  $A$  and  $B$ . In the rest frame of the incident particle  $A$ , the lower vertex represents the virtual process

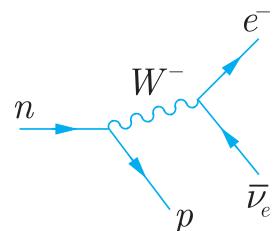
$$A(M_A c^2, \mathbf{0}) \rightarrow A(E_A, \mathbf{p}) + X(E_X, -\mathbf{p}), \quad (1.21)$$

where

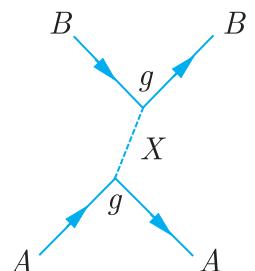
$$E_A = \left( p^2 c^2 + M_A^2 c^4 \right)^{1/2} \quad \text{and} \quad E_X = \left( p^2 c^2 + M_X^2 c^4 \right)^{1/2}, \quad (1.22)$$



**Figure 1.12** Contributions of (a)  $Z^0$  exchange to the elastic weak scattering reaction  $e^- + \nu_e \rightarrow e^- + \nu_e$  and (b) gluon exchange contribution to the strong interaction  $q + q \rightarrow q + q$ .



**Figure 1.13** The decay  $n \rightarrow p + e^- + \bar{\nu}_e$  via an intermediate  $W$  meson.



**Figure 1.14** Contribution to the reaction  $A + B \rightarrow A + B$  from the exchange of a particle  $X$ .

<sup>13</sup>These detectors will be discussed in Chapter 4.

and  $p = |\mathbf{p}|$ . The energy difference between the final and initial states is given by

$$\Delta E = E_X + E_A - M_A c^2 \rightarrow 2pc, \quad p \rightarrow \infty, \\ \rightarrow M_X c^2, \quad p \rightarrow 0, \quad (1.23)$$

and thus  $\Delta E \geq M_X c^2$  for all  $p$ . By the uncertainty principle, such an energy violation is allowed, but only for a time  $\tau \approx \hbar/\Delta E$ , so we immediately obtain

$$r \approx R \equiv \hbar/M_X c \quad (1.24)$$

as the maximum distance over which  $X$  can propagate before being absorbed by particle  $B$ . This constant  $R$  is called the *range* of the interaction and this was the sense of the word used in Section 1.1.

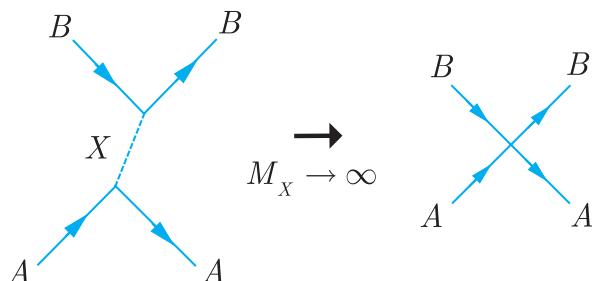
The electromagnetic interaction has an infinite range because the exchanged particle is a massless photon. The strong force between quarks also has an infinite range because the exchanged particles are massless gluons.<sup>14</sup> In contrast, the weak interaction is associated with the exchange of very heavy particles, the  $W$  and  $Z$  bosons, with masses

$$M_W = 80.4 \text{ GeV}/c^2 \text{ and } M_Z = 91.2 \text{ GeV}/c^2 \quad (1 \text{ GeV} = 10^9 \text{ eV}) \quad (1.25)$$

corresponding to ranges that from (1.24) are of order

$$R_{W,Z} \equiv \frac{\hbar}{M_W c} \approx 2 \times 10^{-3} \text{ fm} \quad (1 \text{ fm} = 10^{-15} \text{ m}). \quad (1.26)$$

In many applications, this range is very small compared with the de Broglie wavelengths of all the particles involved. The weak interaction can then be approximated by a zero-range or point interaction, corresponding to the limit  $M_X \rightarrow \infty$ , as shown in Figure 1.15.



**Figure 1.15** The zero-range or point interaction resulting from the exchange of a particle  $X$  in the limit  $M_X \rightarrow \infty$ .

<sup>14</sup>The force between hadrons is much more complicated, because it is not due to single-gluon exchange. It has a range of approximately (1–2) fm, where 1 fm =  $10^{-15}$  m.

### 1.4.2 The Yukawa potential

In the limit that  $M_A$  becomes large, we can regard  $B$  as being scattered by a static potential of which  $A$  is the source. This potential will in general be spin dependent, but its main features can be obtained by neglecting spin and considering  $X$  to be a spin-0 boson, in which case it will obey the Klein–Gordon equation,

$$-\hbar^2 \frac{\partial^2 \phi(\mathbf{r}, t)}{\partial t^2} = -\hbar^2 c^2 \nabla^2 \phi(\mathbf{r}, t) + M_X^2 c^4 \phi(\mathbf{r}, t). \quad (1.27)$$

The static solution of this equation satisfies

$$\nabla^2 \phi(\mathbf{r}) = \frac{M_X^2 c^2}{\hbar^2} \phi(\mathbf{r}), \quad (1.28)$$

where  $\phi(\mathbf{r})$  is interpreted as a static potential. For  $M_X = 0$  this equation is the same as that obeyed by the electrostatic potential, and for a point charge  $-e$  interacting with a point charge  $+e$  at the origin, the appropriate solution is the Coulomb potential

$$V(r) = -e \phi(r) = -\frac{e^2}{4\pi\epsilon_0} \frac{1}{r}, \quad (1.29)$$

where  $r = |\mathbf{r}|$  and  $\epsilon_0$  is the dielectric constant. The corresponding solution in the case where  $M_X^2 \neq 0$  is easily verified by substitution to be

$$V(r) = -\frac{g^2}{4\pi} \frac{e^{-r/R}}{r}, \quad (1.30)$$

where  $R$  is the range defined in (1.24) and we assume equal coupling constants  $g$  for particle  $X$  to the particles  $A$  and  $B$ . It is conventional to introduce a dimensionless strength parameter

$$\alpha_X = \frac{g^2}{4\pi\hbar c} \quad (1.31)$$

that characterises the strength of the interaction at short distances, in analogy to the fine structure constant of QED.

The form of  $V(r)$  in (1.30) is called a *Yukawa potential*, after the physicist who first introduced the idea of forces due to massive particle exchange in 1935. As  $M_X \rightarrow 0$ ,  $R \rightarrow \infty$  and the Coulomb potential is recovered from the Yukawa potential, while for very large masses the interaction is approximately point-like (zero range). The effective coupling strength in this latter approximation and its range of validity are best understood by considering the corresponding scattering amplitude.

### 1.4.3 The zero-range approximation

In lowest-order perturbation theory, the probability amplitude for a particle with initial momentum  $\mathbf{q}_i$  to be scattered to a final state with momentum  $\mathbf{q}_f$  by a potential  $V(\mathbf{r})$  is proportional to the *scattering amplitude*<sup>15</sup>

$$\mathcal{M}(\mathbf{q}) = \int V(\mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r}/\hbar) d^3\mathbf{r}, \quad (1.32)$$

where  $\mathbf{q} \equiv \mathbf{q}_i - \mathbf{q}_f$  is the momentum transfer. The integration may be done using polar coordinates  $(r, \theta, \phi)$ . Taking  $\mathbf{q}$  in the  $z$  direction, gives

$$\mathbf{q} \cdot \mathbf{r} = |\mathbf{q}| r \cos \theta \quad (1.33)$$

and

$$d^3\mathbf{r} = r^2 \sin \theta \, dr \, d\theta \, d\phi, \quad (1.34)$$

where  $r = |\mathbf{r}|$ . For the Yukawa potential, the integral (1.32) gives

$$\mathcal{M}(\mathbf{q}) = \frac{-g^2 \hbar^2}{|\mathbf{q}|^2 + M_X^2 c^2}. \quad (1.35)$$

In using (1.35) for the scattering amplitude we have assumed potential theory, treating the particle  $A$  as a static source. The particle  $B$  then scatters through some angle without loss of energy, so that  $|\mathbf{q}_i| = |\mathbf{q}_f|$  and the initial and final energies of particle  $B$  are equal,  $E_i = E_f$ . While this is a good approximation at low energies, at higher energies the recoil energy of the target particle cannot be neglected, so that the initial and final energies of  $B$  are no longer equal. A fully relativistic calculation taking account of this is more difficult, but the result is surprisingly simple. Specifically, in lowest-order perturbation theory, one obtains

$$\mathcal{M}(q^2) = \frac{g^2 \hbar^2}{q^2 - M_X^2 c^2}, \quad (1.36)$$

where

$$q^2 \equiv (E_f - E_i)^2 - (\mathbf{q}_f - \mathbf{q}_i)^2 c^2 \quad (1.37)$$

is called the squared energy-momentum transfer, or squared four-momentum transfer. In the low-energy limit,  $E_i = E_f$  and (1.36) reduces to (1.35). However, in contrast to (1.35), which was derived in the rest frame of particle  $A$ , the form (1.36) is explicitly Lorentz invariant and holds in all inertial frames of reference.

---

<sup>15</sup>This is called the Born approximation. For a discussion, see for example pp. 397–399 of Gasiorowicz (1974).

The amplitude (1.36) corresponds to the exchange of a single particle, as shown for example in Figure 1.14. (Multiparticle exchange corresponds to higher orders in perturbation theory and higher powers of  $g^2$ .) In the zero-range approximation, (1.36) reduces to a constant. To see this, we note that this approximation is valid when the range  $R = \hbar/M_X c$  is very small compared with the de Broglie wavelengths of all the particles involved. In particular, this implies  $q^2 \ll M_X^2 c^2$ , and neglecting  $q^2$  in (1.36) gives

$$\mathcal{M}(q^2) = -G, \quad (1.38)$$

where the constant  $G$  is given by

$$\frac{G}{(\hbar c)^3} = \frac{1}{\hbar c} \left( \frac{g}{M_X c^2} \right)^2 = \frac{4\pi\alpha_X}{(M_X c^2)^2} \quad (1.39)$$

and has the dimensions of inverse energy squared. Thus we see that in the zero-range approximation, the resulting point interaction between  $A$  and  $B$  (see Figure 1.15) is characterised by a single dimensioned coupling constant  $G$  and not  $g$  and  $M_X$  separately. As we shall see in the next chapter, this approximation is extremely useful in weak interactions, where the corresponding *Fermi coupling constant*  $G_F$ , measured for example in nuclear  $\beta$ -decay, is given by

$$G_F / (\hbar c)^3 = 1.166 \times 10^{-5} \text{ GeV}^{-2}. \quad (1.40)$$

## 1.5 Units and dimensions

In previous sections, we have quoted numerical values for some constants, for example  $G_F$  above, and given formulas for others. Before continuing our discussion, we consider more carefully the question of units.

Units are a perennial problem in physics, since most branches of the subject tend to adopt a system that is convenient for their own purpose. Elementary particle physics is no exception, and adopts so-called *natural units* (*nu*), chosen so that the fundamental constants

$$c = 1 \quad \text{and} \quad \hbar = 1. \quad (1.41)$$

In other words,  $c$  and  $\hbar$  are used as fundamental units of velocity and action (or angular momentum), respectively. To complete the definition, a third unit must be fixed, which is chosen to be the unit of energy. This is taken to be the electron volt (eV), defined as the energy required to raise the electric potential of an electron or proton by one volt. The abbreviations keV ( $10^3$  eV), MeV ( $10^6$  eV), GeV ( $10^9$  eV) and TeV ( $10^{12}$  eV) are also in general use.

Quantum mechanics and special relativity play crucial roles in elementary particle physics, so that  $\hbar$  and  $c$  occur frequently in formulas. By choosing natural units, all factors of  $\hbar$  and  $c$  may be omitted from equations using (1.41), which leads to considerable simplifications. For example, the relativistic energy relation

$$E^2 = p^2 c^2 + m^2 c^4$$

becomes

$$E^2 = p^2 + m^2,$$

while the Fermi coupling constant (1.40) becomes

$$G_F = 1.166 \times 10^{-5} \text{ GeV}^{-2}. \quad (1.42)$$

In natural units (nu) *all* quantities have the dimensions of a power of energy, since they can all be expressed in terms of  $\hbar$ ,  $c$  and an energy. In particular, masses, lengths and times can be expressed in the forms

$$M = E/c^2, \quad L = \hbar c/E, \quad T = \hbar/E,$$

so that a quantity with mks dimensions  $M^p L^q T^r$  has the nu dimensions  $E^{p-q-r}$ . Since  $\hbar$  and  $c$  are suppressed in natural units, this is the only dimension that is relevant, and dimensional checks and estimates are very simple. The nu and mks dimensions are listed for some important quantities in Table 1.1. We note that in natural units many different quantities (e.g. mass, energy and momentum) all have the same dimension.

Natural units are very convenient for theoretical arguments, as we will see. However, we must still know how to convert from natural units to the ‘practical’ units in which experimental results are invariably stated. This is done in two steps. We first restore the  $\hbar$

**Table 1.1** The mks dimension  $M^p L^q T^r$  and the nu dimensions  $E^n = E^{p-q-r}$  of some quantities.

Quantity	<i>p</i>	<i>q</i>	<i>r</i>	<i>n</i>
Action ( $\hbar$ )	1	2	-1	0
Velocity ( $c$ )	0	1	-1	0
Mass	1	0	0	1
Length	0	1	0	-1
Time	0	0	1	-1
Momentum	1	1	-1	1
Energy	1	2	-2	1
Fine structure constant ( $\alpha$ )	0	0	0	0
Fermi coupling constant ( $G_F$ )	1	5	-2	-2

and  $c$  factors by dimensional arguments and then use the conversion factors

$$\hbar = 6.582 \times 10^{-22} \text{ MeV s} \quad (1.43a)$$

and

$$\hbar c = 1.973 \times 10^{-13} \text{ MeV m} \quad (1.43b)$$

to evaluate the result.

We illustrate this by using the nu expression for the cross-section  $\sigma$  for Thomson scattering,<sup>16</sup> that is Compton scattering from free electrons when the photon energy is much less than the electron rest energy. This is given by<sup>17</sup>

$$\sigma = \frac{8\pi\alpha^2}{3m_e^2}.$$

To convert it to practical units, we write

$$\sigma = \frac{8\pi\alpha^2}{3m_e^2} \hbar^a c^b$$

and demand that  $\sigma$  has the dimensions of length squared. This gives  $a = 2, b = -2$ , so that

$$\sigma = \frac{8\pi\alpha^2(\hbar c)^2}{3(m_e c^2)^2} = 6.65 \times 10^{-29} \text{ m}^2, \quad (1.44)$$

using (1.43b) and  $m_e = 0.51 \text{ MeV}/c^2$ .

Cross-sections are usually quoted in *barns* b ( $10^{-28} \text{ m}^2$ ). For example, the Thomson cross-section (1.44) is 0.665b while the total cross-section for  $np$  scattering is typically of order 30–40 mb, depending on the energy, where  $1 \text{ mb} = 10^{-3} \text{ b} = 10^{-31} \text{ m}^2$ . Other abbreviations commonly used are *microbarns* mb ( $10^{-34} \text{ m}^2$ ) and *picobarns* ( $10^{-40} \text{ m}^2$ ). Similarly, lengths are often quoted in *fermis* fm ( $10^{-15} \text{ m}$ ) rather than metres. In these units, the radius of the proton is about 0.8 fm, while the range of the weak force is of order  $10^{-3}$  fm. Energies are measured in MeV, GeV, etc., while momenta are measured in  $\text{MeV}/c$ , etc. and masses in  $\text{MeV}/c^2$ , etc., as in Equation (1.25) for the weak boson mass. This should be compared to natural units, where energy, momentum and mass all have the dimension of energy (see Table 1.1) and are all measured in, for example, MeV.

A list of some useful physical constants and conversion factors is given inside the back cover of this book. *From now on natural units will be used in formulas throughout the book unless explicitly stated otherwise.*

<sup>16</sup>Cross-sections and related quantities are formally defined in Appendix B.

<sup>17</sup>See, for example, p. 20 of Mandl and Shaw (2010).

## Problems 1

- 1.1** Write down equations in symbol form that describe the following interactions:
- the elastic scattering of an electron antineutrino and a positron;
  - the elastic scattering of proton and a neutron;
  - the annihilation of a proton and an antiproton to an electron, a positron and a photon.
- 1.2** The matrices  $\alpha_i$  ( $i = 1, 2, 3$ ) and  $\beta$  in the Dirac equation (1.8) can be chosen to have the explicit forms

$$\alpha_i = \begin{pmatrix} \mathbf{0} & \sigma_i \\ \sigma_i & \mathbf{0} \end{pmatrix}, \quad \beta = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix}, \quad (1.45a)$$

where  $\mathbf{1}$  and  $\mathbf{0}$  are the  $2 \times 2$  unit matrix and null matrix, respectively, and the *Pauli spin matrices*  $\sigma_i$  are

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.45b)$$

Check that these forms satisfy the anticommutation relations

$$[\alpha_i, \beta] \equiv \alpha_i \beta + \beta \alpha_i = 0 \text{ and } [\alpha_i, \alpha_j] = 0, \text{ where } i \neq j = 1, 2, 3.$$

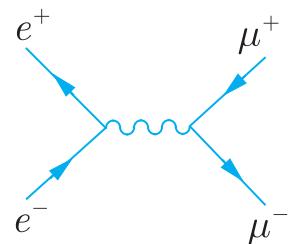
- 1.3** For a free particle of momentum  $\mathbf{p} = (p_x, p_y, p_z)$ , the Dirac equation has four independent solutions (1.12), where the independent spinors can be chosen to be of the forms

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ a_1 \\ b_1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ a_2 \\ b_2 \end{pmatrix}, \quad \begin{pmatrix} a_3 \\ b_3 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} a_4 \\ b_4 \\ 0 \\ 1 \end{pmatrix},$$

where  $a_i, b_i$  are finite coefficients, if we assume the explicit form of the Dirac matrices given in Problem 1.2. Verify that solutions of these forms exist and find the values of the coefficients  $a_i, b_i$  and the corresponding energies.

- 1.4** Draw the topologically distinct Feynman diagrams that contribute to the following process in lowest order:
- $\gamma + e^- \rightarrow \gamma + e^-$ ;
  - $e^+ + e^- \rightarrow e^+ + e^-$ ;
  - $\nu_e \bar{\nu}_e$  elastic scattering.
- (Hint: There are two such diagrams for each reaction.)
- 1.5** Draw one fourth-order diagram for each of the reactions (a)  $\gamma + e^- \rightarrow \gamma + e^-$  and (b)  $e^+ + e^- \rightarrow e^+ + e^-$ .
- 1.6** Show that the Yukawa potential of Equation (1.30) is the only spherically symmetric solution of the static Klein–Gordon equation (1.28) that vanishes as  $r$  goes to infinity.
- 1.7** Verify by explicit integration that substituting the Yukawa potential (1.30) into (1.32) leads to (1.35) for the corresponding scattering amplitude.
- 1.8** In lowest order, the process  $e^+ + e^- \rightarrow \gamma + \gamma$  is given by the Feynman diagrams of Figure 1.9. Show that for electrons and positrons almost at rest, the distance between the two vertices is typically of order  $m^{-1}$

- in natural units, where  $m$  is the electron mass. Convert this result to practical units and evaluate it in fermis.
- 1.9** In lowest order, the process  $e^+ + e^- \rightarrow \mu^+ + \mu^-$  is given by the Feynman diagram of Figure 1.16. Estimate the typical distance between the vertices at energies much larger than the masses of any of the particles in (a) the rest frame of the electron and (b) the centre-of-mass frame. Check the consistency of these estimates by considering the Lorentz contraction in going from the rest frame, in which the initial electron is stationary, to the centre-of-mass frame.
- 1.10** Parapositronium is an unstable bound state of an electron and a positron. Its lifetime is given in natural units by  $\tau = 2/m\alpha^5$ , where  $m$  is the mass of the electron and  $\alpha$  is the fine structure constant. Restore the factors of  $\hbar$  and  $c$  by dimensional arguments and evaluate in seconds.
- 1.11** For total centre-of-mass energies of a few GeV, the cross-section for the reaction  $\nu_\mu + e^- \rightarrow \mu^- + \nu_\mu$  is given by  $\sigma = G_F^2 E^2 / \pi$  in natural units, where  $G_F$  is the Fermi coupling constant (1.42). What is the cross-section in picobarns at an energy  $E = 3$  GeV?
- 1.12** (a) The *classical radius* of an electron is defined as  $r_c = e^2 / 4\pi\epsilon_0 m_e$  in natural units. What is its value in fm?  
 (b) A simple classical model of the electron is to regard it as a sphere of uniform charge density, whose radius  $R$  is such that the electrostatic potential accounts for the whole electron mass. Derive the value of  $R$ , expressing it in terms of the classical radius defined above.



**Figure 1.16** Lowest-order diagram for  $e^+ + e^- \rightarrow \mu^+ + \mu^-$ .

---

## Leptons and the weak interaction

---

In Chapter 1 we introduced some of the ideas underlying the discussion of particle interactions. In doing this, we concentrated mainly on electrons and positrons and their electromagnetic interactions. In this chapter we extend that account to include other leptons, starting with a review of their basic properties, followed by a discussion of the weak interactions between them. The final section gives an account of *neutrino oscillations*, which result in neutrinos changing their properties as they travel over long distances, and the experiments that established this phenomenon. As we shall see, these experiments rule out the possibility that the neutrinos are massless, as was assumed in the original formulation of the standard model.

### 2.1 Lepton multiplets and lepton numbers

Leptons are one of the four classes of fundamental particles in the standard model. They are spin-1/2 fermions without strong interactions. There are six known leptons, and they occur in pairs called *generations*, which we write as doublets:

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}, \quad \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}, \quad \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}. \quad (2.1)$$

The three charged leptons ( $e^-$ ,  $\mu^-$ ,  $\tau^-$ ) are the familiar electron and two new particles, the *mu-lepton* or *muon* and the *tau-lepton* or *tauon*. All have electric charge  $Q = -e$ . Associated with them are three neutral leptons, or *neutrinos*, called the *electron neutrino*, *mu neutrino*, and *tau neutrino*, respectively, all of which have very small

masses. The six distinct types of leptons are also referred to as having different ‘flavours’. In addition to the leptons, there are six corresponding antiparticles (antileptons):

$$\begin{pmatrix} e^+ \\ \bar{\nu}_e \end{pmatrix}, \quad \begin{pmatrix} \mu^+ \\ \bar{\nu}_\mu \end{pmatrix}, \quad \begin{pmatrix} \tau^+ \\ \bar{\nu}_\tau \end{pmatrix}. \quad (2.2)$$

The charged leptons interact via both electromagnetic and weak forces, whereas for neutral leptons only weak interactions have been observed. Because of this, neutrinos can be detected only with considerable difficulty. We will discuss each of these particles in turn presently.

Firstly, however, we note that each generation of leptons has associated with it a quantum number. The first of these *lepton numbers* is the *electron number*, defined for any state by

$$L_e \equiv N(e^-) - N(e^+) + N(\nu_e) - N(\bar{\nu}_e), \quad (2.3a)$$

where  $N(e^-)$  is the number of electrons present, and so on. For single-particle states,  $L_e = 1$  for  $e^-$  and  $\nu_e$ ,  $L_e = -1$  for  $e^+$  and  $\bar{\nu}_e$ , and  $L_e = 0$  for all other particles. In electromagnetic interactions, conservation of  $L_e$  reduces to the conservation of  $N(e^-) - N(e^+)$ , since neutrinos are not involved. This implies that electrons and positrons can only be created or annihilated in pairs. In weak interactions more general possibilities are allowed. For example, an electron could be created together with an antineutrino  $\bar{\nu}_e$ , rather than a positron. Similar remarks apply to *muon number* and *tauon number*, defined by

$$L_\mu \equiv N(\mu^-) - N(\mu^+) + N(\nu_\mu) - N(\bar{\nu}_\mu), \quad (2.3b)$$

and

$$L_\tau \equiv N(\tau^-) - N(\tau^+) + N(\nu_\tau) - N(\bar{\nu}_\tau), \quad (2.3c)$$

respectively. In the standard model, lepton numbers are individually conserved in all known interactions.

### 2.1.1 Electron neutrinos

Neutrinos, they are very small.

They have no charge and have no mass

And do not interact at all.

The earth is just a silly ball

To them, through which they simply pass,  
Like dustmaids down a drafty hall

Or photons through a sheet of glass.

*J. Updike*<sup>1</sup>

---

<sup>1</sup> From Updike (1964).

The existence of the *electron neutrino* was first postulated by Pauli in 1930 in order to understand the observed  $\beta$  decays

$$(Z, A) \rightarrow (Z + 1, A) + e^- + \bar{\nu}_e \quad (2.4a)$$

and

$$(Z', A') \rightarrow (Z' - 1, A') + e^+ + \nu_e, \quad (2.4b)$$

where  $(Z, A)$  denotes a nucleus with atomic number  $Z$  and mass number  $A$ . These reactions are actually decays of bound neutrons and protons via the basic processes

$$n \rightarrow p + e^- + \bar{\nu}_e \quad (2.5a)$$

and

$$p \rightarrow n + e^+ + \nu_e, \quad (2.5b)$$

where only the neutron decay can occur in free space since  $m_n > (m_p + m_e)$ . For bound protons the decay (2.5b) *can* occur because the apparent energy violation can be compensated by changes in the nuclear binding energy. The neutrinos in these decays have been written as  $\nu_e$  or  $\bar{\nu}_e$  by using electron number conservation. They were not initially observed experimentally, but inferred from energy and angular momentum conservation. In the case of energy, if the antineutrino were not present in (2.4a), the reaction would be a two-body decay, and the energy  $E_e$  of the emitted electron would have the unique value

$$E_e = \Delta M = M(Z, A) - M(Z + 1, A),$$

where we have neglected the nuclear recoil energy. However, if the antineutrino is present, the electron energy will not be unique, but lie in the range

$$m_e \leq E_e \leq (\Delta M - m_{\bar{\nu}_e}). \quad (2.6)$$

Experimentally, the observed spectrum spans the whole range (2.6), with  $m_{\bar{\nu}_e} \approx 0$ . Careful study of the spectrum near the end point  $E_e = \Delta M - m_{\bar{\nu}_e}$  allows an upper limit to be set on the neutrino mass. The best results come from the  $\beta$  decay of tritium



which gives

$$m_{\bar{\nu}_e} < 2 \text{ eV}/c^2 \approx 4 \times 10^{-6} m_e. \quad (2.7)$$

Such an extremely small mass can be ignored in most calculations,<sup>2</sup> although we shall see later that there is conclusive evidence that neutrino masses are non-zero.

---

<sup>2</sup> The exceptions occur mostly in astrophysics and cosmology. We will discuss this briefly in Section 2.3.3 and Chapter 12.

Neutrinos and antineutrinos can in principle be detected by observing the *inverse β-decay* processes



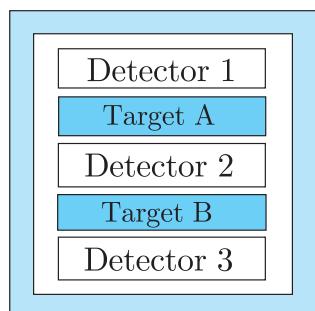
and



However, the cross-sections for these processes are extremely small. For the neutrinos and antineutrinos emitted in β-decays, with energies of order 1 MeV, the cross-sections are of order  $10^{-47} \text{ m}^2$ , corresponding to mean free paths in matter of many millions of kilometres. Nevertheless, if the neutrino flux is intense enough, the reactions can be observed.

This was first done in a classic experiment by Reines and Cowan in 1959 using antineutrinos emitted from a nuclear reactor at Savannah River, North Carolina. Uranium fission fragments are neutron rich and decay by (2.4a) to give an antineutrino flux of order  $10^{17} \text{ m}^{-2} \text{ s}^{-1}$  in the vicinity of the reactor. These occasionally interacted with protons in a large detector to give examples of reaction (2.8b) at rates of approximately two per hour. To distinguish these from ‘background’ (i.e. spurious) events from other sources, a very distinctive signal was required. This was achieved in an ingenious manner. The detector, which is shown schematically in Figure 2.1, comprised two target tanks, each containing an aqueous solution of cadmium chloride, sandwiched between three tanks of liquid scintillator.<sup>3</sup> When the reaction (2.8b) occurred in one of the target tanks, the emitted positron rapidly annihilated with an atomic electron to produce photons, giving a ‘prompt coincidence’ signal in the two adjacent tanks of the scintillator. In contrast, the neutron suffered multiple scattering from protons in the water until, somewhat later, it reached thermal energies. It could then be radiatively captured by a cadmium nucleus, releasing more photons and giving rise to a second ‘delayed coincidence’ in the same two adjacent tanks of the scintillator.

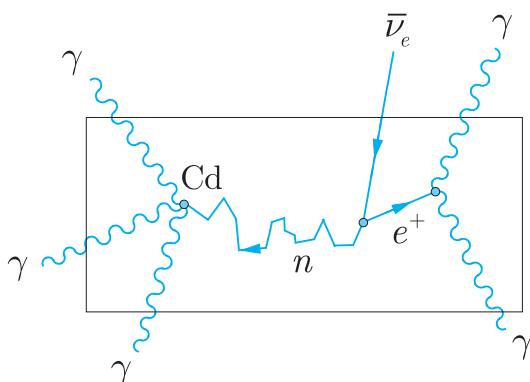
This course of events is illustrated schematically in Figure 2.2. The distinctive signal is thus two coincidences in the same pair of scintillator tanks, separated by a time, which was calculated to be a few microseconds. In their first series of measurements lasting just under 200 h, 567 such events were found, compared to an estimated



**Figure 2.1** Schematic diagram of the apparatus used by Reines and Cowen to detect antineutrinos. The two target tanks, containing an aqueous solution of cadmium chloride, are sandwiched between three detector tanks of the liquid scintillator. The whole apparatus is surrounded by heavy shielding to eliminate all incident particles except neutrinos and antineutrinos.

<sup>3</sup> A charged particle passing through matter will excite atomic electrons and in some substances, called ‘scintillators’, a small fraction of this excitation energy re-emerges as visible light. Scintillators and scintillation counters are discussed in Chapter 4, Section 4.4.4.

**Figure 2.2** Schematic picture of an event corresponding to reaction (2.8b) in the Reines–Cowan experiment. The emitted positron rapidly annihilates with an atomic electron, and the emitted photons are detected in the adjacent tanks of the scintillator as a ‘prompt coincidence’. The neutron is reduced to thermal energies by repeated scattering from protons in the water before being radiatively captured by the cadmium nucleus. The emitted photons are then detected as a ‘delayed coincidence’ in the same pair of scintillator tanks.



209 from background accidentals.<sup>4</sup> Detailed analysis<sup>5</sup> confirmed that the remaining signals were due to inverse  $\beta$ -decay events. Antineutrinos had for the first time been detected.

## 2.1.2 Further generations

In addition to the  $(e^-, \nu_e)$  pair, we have seen that there are two more lepton generations known: the muon  $\mu^-$  and its associated muon neutrino  $\nu_\mu$ ; and the tauon  $\tau^-$  and its associated tau-neutrino  $\nu_\tau$ .

The muon is a very penetrating particle of mass  $105.7 \text{ MeV}/c^2$  that was first identified in cosmic ray experiments by Anderson and Neddermeyer in 1936. Cosmic ray *primaries* are high-energy particles, mostly protons, incident on the Earth’s atmosphere from all directions in space. Other particles, called *secondaries*, are produced when the primaries collide with nuclei in the Earth’s atmosphere, and some penetrate to sea level. It was among these that muons were discovered. Subsequently, they have been copiously produced at accelerator laboratories, enabling their properties to be studied in great detail. Muons have the Dirac magnetic moment

$$\mu = \frac{e}{m_\mu} \mathbf{S} \quad (2.9)$$

appropriate to point-like spin-1/2 particles, and in general their electromagnetic properties are identical with those of electrons, *provided the mass difference is taken into account*. In particular, it is their

<sup>4</sup> There are single coincidences due to unaccompanied positron or neutron counts from other sources. From their measured rates, one can estimate how often two of these will occur separated by a few microseconds, mimicking a genuine neutrino event. These are the so-called ‘accidentals’.

<sup>5</sup> A simple account is given in Trigg (1975).

much greater mass that is the reason for their much greater penetrating power in matter compared to electrons.<sup>6</sup>

The tauon is even heavier ( $m_\tau = 1777 \text{ MeV}/c^2$ ) and was discovered in electron–positron annihilation experiments at high energies in 1975. Its properties have been studied with less precision than those of the muon, but are compatible with a point-like spin-1/2 particle whose electromagnetic interactions are identical with those of the electron and muon.

Because the electron is the lightest charged particle, conservation of electric charge means it is necessarily stable. Both the muon and the tauon, however, are unstable, with lifetimes of  $2.2 \times 10^{-6} \text{ s}$  and  $2.9 \times 10^{-13} \text{ s}$ , respectively. Both decay by weak interactions, and the great difference in their lifetimes is a consequence of the mass difference, as we shall see below. In the case of the muon, the decay is purely leptonic

$$\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu, \quad \mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu \quad (2.10)$$

and conserves both charge and lepton number. For the tauon, many decay modes are observed, most of them involving hadrons in the final state, although the purely leptonic modes

$$\tau^- \rightarrow e^- + \bar{\nu}_e + \nu_\tau \quad (2.11a)$$

and

$$\tau^- \rightarrow \mu^- + \bar{\nu}_\mu + \nu_\tau \quad (2.11b)$$

are also observed. The relative importance of any decay mode is measured by its *branching ratio*  $B$ , defined as the fraction of all decays leading to that particular final state. For the leptonic decays (2.11a) and (2.11b), the measured branching ratios are 0.178 and 0.174, respectively.

The neutrinos emitted in the decays (2.10) and (2.11) were not directly observed, but, like those in  $\beta$  decay, were inferred from energy and angular momentum conservation. Both  $\nu_\mu$  and  $\nu_\tau$  have, however, subsequently been detected in other reactions. Well-defined muon–neutrino beams can be created in the laboratory and used to study reactions like the elastic scattering from electrons

$$\nu_\mu + e^- \rightarrow \nu_\mu + e^-, \quad (2.12)$$

*inverse muon decay*

$$\nu_\mu + e^- \rightarrow \mu^- + \nu_e \quad (2.13)$$

---

<sup>6</sup> High-energy electrons lose energy in matter dominantly by radiative collisions. The rate of energy loss in such collisions is approximately proportional to  $m^{-2}$ . Consequently, it is heavily suppressed for muons, which travel much further. This is discussed in Section 4.3.3.

and reactions on nucleons, such as

$$\nu_\mu + p \rightarrow \nu_\mu + p \quad (2.14a)$$

and

$$\bar{\nu}_\mu + p \rightarrow \mu^+ + n. \quad (2.14b)$$

Well-defined tau-neutrino beams are not available in the laboratory and it was not until 2000 that tau neutrinos were directly detected. Only a few direct interaction events have ever been observed and they supplement evidence based on energy and angular momentum conservation. The masses of both  $\nu_\mu$  and  $\nu_\tau$  can be inferred from the  $e^-$  and  $\mu^-$  energy spectra in the decays (2.10) and (2.11) using energy conservation, although the errors are much larger than those for electron neutrinos. The present limits are

$$m_{\nu_\mu} < 0.19 \text{ MeV}/c^2, \quad m_{\nu_\tau} < 18.2 \text{ MeV}/c^2. \quad (2.15)$$

However, indirect measurements show that the masses of the neutrinos emitted in these decays are very much smaller than these upper limits, as we shall see in Section 2.3.3.

Underlying the whole of this discussion is the assumption that lepton numbers are individually conserved. It is these that fix the identity (as  $\nu_\mu$ ,  $\nu_\tau$ , etc.) of the undetected neutrinos in, for example, muon decay (2.10). One way of testing their validity is to search for reactions that violate lepton number conservation but satisfy all other conservation laws, such as  $\mu^\pm \rightarrow e^\pm + \gamma$  and  $\nu_e + e^- \rightarrow \mu^- + \nu_e$ . Examples of lepton number violating decays of this type, and the experimental upper limits on the branching ratios above which they would have been detected, are shown in Table 2.1. These tests are impressive evidence for lepton number conservation, and no violations have ever been detected in any particle decays.

**Table 2.1** Examples of leptonic decays that would violate conservation of lepton numbers, and the experimental upper limits on their branching ratios  $B$ .

Decay	Violates	B
$\mu^- \rightarrow e^- + e^+ + e^-$	$L_\mu, L_e$	$< 1.0 \times 10^{-12}$
$\mu^- \rightarrow e^- + \gamma$	$L_\mu, L_e$	$< 5.7 \times 10^{-13}$
$\tau^- \rightarrow e^- + \gamma$	$L_\tau, L_e$	$< 3.3 \times 10^{-8}$
$\tau^- \rightarrow \mu^- + \gamma$	$L_\tau, L_\mu$	$< 4.4 \times 10^{-8}$
$\tau^- \rightarrow e^- + \mu^+ + \mu^-$	$L_\tau, L_e$	$< 2.0 \times 10^{-8}$

## 2.2 Leptonic weak interactions

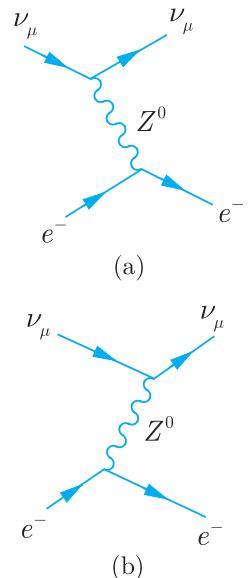
We begin the study of the dynamics of weak interactions by focussing for the moment on simple leptonic reactions like (2.10)–(2.13). Our discussion will be based on the general ideas of particle exchange introduced in Sections 1.3 and 1.4, and in particular on the zero-range approximation of Figure 1.15.

### 2.2.1 $W^\pm$ and $Z^0$ exchange

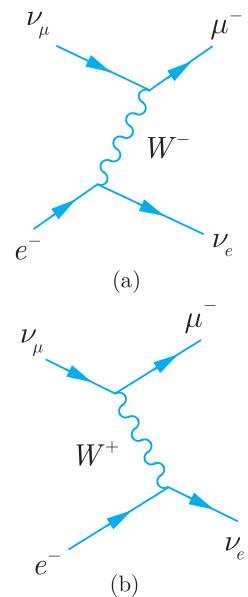
In Section 1.3, we saw that elastic electron–electron scattering is described by the exchange diagrams of Figures 1.7(a) and (b), in which a photon is emitted by one electron and absorbed by another. In the same way, weak reactions involving only leptons are described by exchange processes in which a  $W^\pm$  or  $Z^0$  is emitted by one lepton and absorbed by another. In drawing such diagrams, we must remember that these absorption and emission processes conserve the lepton numbers (2.3), as well as electric charge, where the  $W^\pm$  and  $Z^0$  have zero lepton numbers.

A simple example is elastic  $\nu_\mu e^-$  scattering (2.12), for which the two possible time-ordered Feynman diagrams are shown in Figure 2.3. These are clearly closely analogous to the two time-ordered diagrams in Figures 1.7(a) and (b) for  $e^-e^-$  scattering by photon exchange and as in that case it is conventional to draw one such diagram, leaving the other implied.<sup>7</sup>

The corresponding diagrams for inverse muon decay (2.13) are shown in Figure 2.4. In Figure 2.4(a), the initial process corresponds to the lower vertex, in which an electron converts into an electron neutrino by emitting a  $W^-$  boson, conserving both electric charge and lepton numbers. The  $W^-$  is then absorbed at the upper vertex, again conserving both charge and lepton numbers. Figure 2.4(a) therefore corresponds to  $W^-$  exchange. On the other hand, application of the conservation laws to the second time-ordered diagram in Figure 2.4(b) shows that it corresponds to  $W^+$  exchange. Nevertheless, it is still conventional to draw just one such diagram, leaving the other implied. Thus inverse muon decay is described by just one of the Feynman diagrams of Figure 2.4, while muon decay itself is described by the analogous diagram in Figure 2.5. In the latter, the first vertex corresponds to a muon emitting a  $W^-$  boson, followed by the conversion of the  $W^-$  boson into a lepton pair in the lepton-number conserving process  $W^- \rightarrow e^- \bar{\nu}_e$ , giving the muon decay reaction (2.10) overall.

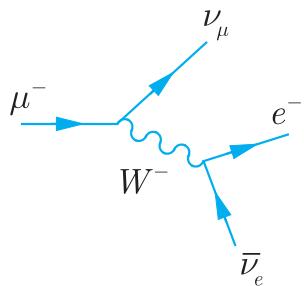


**Figure 2.3** Two time-ordered diagrams for elastic  $\nu_\mu e^-$  scattering (2.12), where as usual time flows from left to right. By convention, only one of these diagrams is usually drawn, leaving the existence of the other implied.

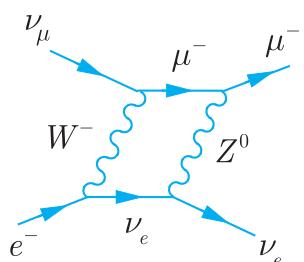


**Figure 2.4** Dominant Feynman diagrams for inverse muon decay.

<sup>7</sup> This convention was introduced for electromagnetic interactions in Section 1.3.3.



**Figure 2.5** Dominant Feynman diagram for muon decay.



**Figure 2.6** Example of a higher-order contribution to inverse muon decay.

These processes can also occur via higher-order diagrams (i.e. ones containing more vertices) like that shown in Figure 2.6 for inverse muon decay, involving the exchange of two vector bosons. In such diagrams, each vertex corresponds to a basic process whose intrinsic probability is much less than unity, as we shall see shortly. Thus the contribution of Figure 2.6, which has four vertices, is expected to be much smaller than that of Figure 2.5, which has only two.<sup>8</sup> Detailed calculations confirm the dominance of Figure 2.5 and from now on we shall only consider the dominant lowest-order diagrams contributing to a given process.

At low energies, the de Broglie wavelengths of all the particles involved are large compared with the range  $R_W \approx 2 \times 10^{-3}$  fm of the  $W$ -exchange interaction. This can then be approximated by a zero-range point interaction whose strength is characterised by the Fermi coupling constant (1.40), that is

$$G_F = 1.166 \times 10^{-5} \text{ GeV}^{-2}. \quad (2.16)$$

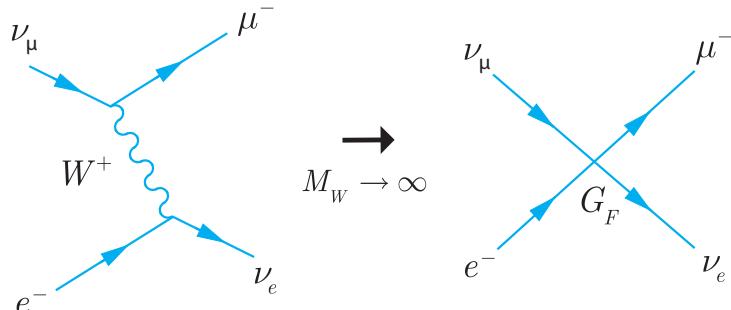
The origin of this approximation is illustrated for inverse muon decay in Figure 2.7, and the corresponding diagram for muon decay itself is shown in Figure 2.8. Figure 2.7 is similar to Figure 1.15 that was discussed for spin-0 particles in Section 1.4, and it can be shown that when spin complications are taken into account, the coupling constant relation (1.39) derived for spinless particles becomes

$$\frac{G_F}{\sqrt{2}} = \frac{g_W^2}{M_W^2} = \frac{4\pi\alpha_W}{M_W^2} \quad (2.17)$$

in natural units. Here,  $g_W$  is the coupling constant associated with the  $W$ -boson-lepton vertices, and

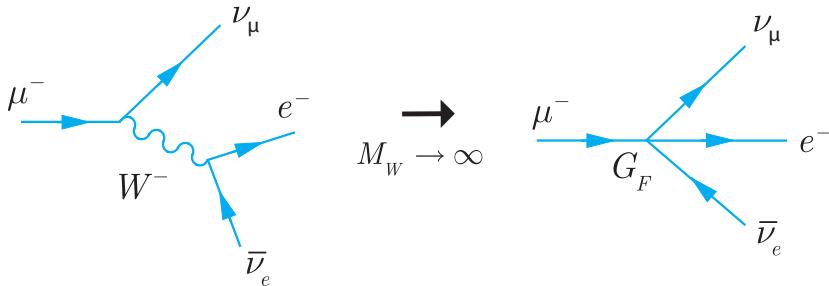
$$\alpha_W \equiv g_W^2 / 4\pi \quad (2.18)$$

is the dimensionless strength parameter introduced by analogy with the fine structure constant  $\alpha = e^2 / 4\pi\epsilon_0$  of electromagnetism. On



**Figure 2.7** Origin of the low-energy zero-range interaction in inverse muon decay (2.13).

<sup>8</sup>This should be compared with the discussion of multi-photon exchange at the end of Section 1.3.3.



**Figure 2.8** Origin of the low-energy zero-range interaction in muon decay (2.10).

substituting the measured values for  $M_W = 80.4 \text{ GeV}/c^2$  (given on the inside rear cover) and  $G_F$  into (2.17), we find that

$$\alpha_W = 4.26 \times 10^{-3} = 0.58\alpha. \quad (2.19)$$

Hence the weak and electromagnetic interactions are of comparable intrinsic strength, and weak interaction rates are only small at low energies because of the large mass of the  $W$ -boson, which enters the low-energy effective coupling constant (2.17) as the inverse square.

### 2.2.2 Lepton decays and universality

All known experimental data are consistent with the assumption that the interactions of the electron and its neutrino are identical with those of the muon and its associated neutrino and the tauon and its neutrino, provided the mass differences are taken into account. This fundamental assumption is called the *universality* of lepton interactions and was discussed briefly, for the case of electromagnetic interactions, in Section 2.1.2. It was also implicit in the previous section, where we assumed that the coupling strengths  $g_W$  and  $G_F$  were independent of the nature of the leptons to which the  $W$  boson coupled.

In the rest of this subsection, we will illustrate universality by considering the leptonic decays (2.10) and (2.11) of the muon and tauon at rest. As usual, we will work to lowest order only and we will use the zero-range approximation, since the masses of the leptons, and hence the energies of all the decay products, are very small compared with the rest energy of the  $W$  bosons. We shall also assume that we can, to a good approximation, neglect the masses of the leptons in the final state relative to the mass of the decaying  $\mu$  or  $\tau$ . Since  $m_e \ll m_\mu \ll m_\tau$ , we would naively expect this approximation to give rise to errors of order  $(m_e/m_\mu) \approx 5 \times 10^{-3}$  in the decay (2.10), of order  $(m_e/m_\tau) \approx 2 \times 10^{-5}$  in the electronic decay of the tauon (2.11a) and of order  $(m_\mu/m_\tau) \approx 5 \times 10^{-2}$  in the muonic decay of the tau (2.11b). In fact the errors are a little smaller than this, but this cannot be seen without a full calculation.

We start by considering muon decay (2.10). In the zero-range approximation, the amplitude is of order  $G_F$ , which has a natural dimension  $[E]^{-2}$ . The only other dimensional constant involved is the muon mass, because we are setting the electron and neutrino masses to zero. It follows that the decay rate must be given by an expression of the form

$$\Gamma(\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu) = K G_F^2 m_\mu^5, \quad (2.20a)$$

because the rate has a natural dimension  $[E]$ . Here,  $K$  is a dimensionless constant whose value will depend on the precise form of the interaction. If we assume this is the same for muons and tauons ( $\mu - \tau$  universality) the same argument gives

$$\Gamma(\tau^- \rightarrow e^- + \bar{\nu}_e + \nu_\tau) = K G_F^2 m_\tau^5, \quad (2.20b)$$

where  $K$  is the same constant, while  $e - \mu$  universality gives

$$\Gamma(\tau^- \rightarrow e^- + \bar{\nu}_e + \nu_\tau) = \Gamma(\tau^- \rightarrow \mu^- + \bar{\nu}_\mu + \nu_\tau). \quad (2.21)$$

This explains why the experimental branching ratios for the two leptonic decay modes of the tauon (2.11a) and (2.11b) are, to a good approximation, equal.<sup>9</sup> It also gives a relation between the  $\mu$  and  $\tau$  lifetimes. The lifetimes  $\tau_\ell$  are related to decay rates by<sup>10</sup>

$$\tau_\ell = \frac{1}{\Gamma_{\text{tot}}} = \frac{B(l^- \rightarrow e^- \bar{\nu}_e \nu_l)}{\Gamma(l^- \rightarrow e^- \bar{\nu}_e \nu_l)},$$

where  $\Gamma_{\text{tot}}$  is the total decay rate and

$$B(l^- \rightarrow e^- \bar{\nu}_e \nu_l) = \frac{\Gamma(l^- \rightarrow e^- \bar{\nu}_e \nu_l)}{\Gamma_{\text{tot}}}$$

is the branching ratio previously defined. Experimentally,  $B = 1$  and  $0.1783 \pm 0.0004$  for  $\ell = \mu$  and  $\tau$ , respectively. Thus from (2.20a) and (2.20b) we have

$$\frac{\tau_\tau}{\tau_\mu} = \frac{B(\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau)}{B(\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu)} \left( \frac{m_\mu}{m_\tau} \right)^5 = (1.326 \pm 0.003) \times 10^{-7},$$

which is consistent with the ratio of the experimental lifetimes  $(1.321 \pm 0.002) \times 10^{-7}$ .

This agreement, involving lifetimes that differ by seven orders of magnitude, is impressive evidence of the universality of lepton interactions. More generally, all known experimental data are consistent with the assumption that the interactions of the electron and its associated neutrino are identical with those of the muon and its

---

<sup>9</sup> A full calculation, taking into account final state masses, gives the ratio  $\Gamma(\tau^- \rightarrow \mu^- + \bar{\nu}_\mu + \nu_\tau)/\Gamma(\tau^- \rightarrow e^- + \bar{\nu}_e + \nu_\tau) = 0.973$ , whereas the experimental value is  $0.976 \pm 0.003$ .

<sup>10</sup> See Equations (B.31a) and (B.31c) of Appendix B.

associated neutrino and of the tauon and its neutrino. The three generations of leptons tell not three stories but, in all essential points, one story three times.

## 2.3 Neutrino masses and neutrino mixing

At present we have no understanding of why the observed leptons have the masses they do. In the standard model it was originally assumed that neutrinos had zero masses, but the results (2.7) and (2.15) do not rule out the possibility that they have very small but finite masses compatible with the experimental bounds given in (2.7) and (2.15). This question is important, because non-zero neutrino masses can lead to non-trivial effects that are not possible for zero masses, and because these effects, and the non-zero masses themselves, can have important consequences in astrophysics and cosmology. This will be discussed briefly in Chapter 12.

Firstly, however, it will be convenient to slightly extend our quantum mechanical notation. So far we have described a single-particle state by naming the particle, for example  $\pi^+$ , and specifying its wavefunction  $\Psi$ , which determines its other properties (spin, momentum, etc.). We now combine these in a single notation, that is  $|\pi^+, \Psi\rangle$ . Alternatively, instead of a wavefunction, we may just designate those observables that are sufficient to uniquely determine the wavefunction. For a spinless particle, momentum is sufficient, so that  $|\pi^+, \mathbf{p}\rangle$  specifies both the particle and its wavefunction up to an arbitrary overall phase.<sup>11</sup> Multi-particle states are specified by, for example,

$$|\pi^+, \Psi_1; \pi^-, \Psi_2\rangle = |\pi^+, \Psi_1\rangle |\pi^-, \Psi_2\rangle ,$$

by analogy with product wavefunctions

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = \Psi_1(\mathbf{r}_1)\Psi_2(\mathbf{r}_2).$$

### 2.3.1 Neutrino mixing

One of the new phenomena that can occur if neutrinos have non-zero masses is *neutrino mixing*. This is the assumption that the neutrino states  $\nu_e, \nu_\mu$  and  $\nu_\tau$  that couple to electrons, muons and tauons, respectively, do not have definite masses; instead they are linear combinations of three other states  $\nu_1, \nu_2$  and  $\nu_3$  that do have definite masses  $m_1, m_2$  and  $m_3$ .

---

<sup>11</sup>This is part of the ‘Dirac notation’ in quantum mechanics. However, no knowledge of this notation is assumed or required beyond that stated above.

Instead of considering mixing between all three ‘flavour’ states  $\nu_e$ ,  $\nu_\mu$  and  $\nu_\tau$ , it is simpler, and often a good approximation, to consider the mixing between just two of them, which we will denote  $\nu_\alpha$  and  $\nu_\beta$ . Then, in order to preserve the orthonormality of the states, we can write

$$\nu_\alpha = \nu_i \cos \theta_{ij} + \nu_j \sin \theta_{ij} \quad (2.22a)$$

and

$$\nu_\beta = -\nu_i \sin \theta_{ij} + \nu_j \cos \theta_{ij}, \quad (2.22b)$$

where  $\nu_i$  and  $\nu_j$  are the two mass eigenstates involved. Here  $\nu_\alpha$  is shorthand for  $|\nu_\alpha, \psi\rangle$ , etc., and  $\theta_{ij}$  is a *mixing angle* that must be determined from experiment.

If  $\theta_{ij} = 0$ , then  $\nu_\alpha = \nu_i$ ,  $\nu_\beta = \nu_j$  and there is no mixing. However, if  $\theta_{ij} \neq 0$ , then some interesting predictions follow. When, for example, a  $\nu_\alpha$  neutrino is produced with momentum  $\mathbf{p}$  at time  $t = 0$ , the  $\nu_i$  and  $\nu_j$  components will have slightly different energies  $E_i$  and  $E_j$  due to their slightly different masses. In quantum mechanics, their associated waves will therefore have slightly different frequencies, giving rise to a phenomenon somewhat akin to the ‘beats’ heard when two sound waves of slightly different frequency are superimposed. As a result of this, one finds that the original beam of  $\nu_\alpha$  particles develops a  $\nu_\beta$  component whose intensity oscillates as it travels through space, while the intensity of the  $\nu_\alpha$  neutrino beam itself is correspondingly reduced. These are called *neutrino oscillations* and their existence follows from simple quantum mechanics.

To illustrate this we will consider a  $\nu_\alpha$  produced with momentum  $\mathbf{p}$  at time  $t = 0$ , so that the initial state (2.22a) is

$$|\nu_\alpha, \mathbf{p}\rangle = |\nu_i, \mathbf{p}\rangle \cos \theta_{ij} + |\nu_j, \mathbf{p}\rangle \sin \theta_{ij}. \quad (2.23a)$$

After time  $t$  this will become

$$a_i(t)|\nu_i, \mathbf{p}\rangle \cos \theta_{ij} + a_j(t)|\nu_j, \mathbf{p}\rangle \sin \theta_{ij} \quad (2.24)$$

where

$$a_i(t) = e^{-iE_i t} \quad \text{and} \quad a_j(t) = e^{-iE_j t} \quad (2.25)$$

are the usual oscillating time factors associated with any quantum mechanical stationary state.<sup>12</sup> For  $t \neq 0$ , the linear combination (2.24) does not correspond to a pure  $\nu_\alpha$  state, but can be written as a linear combination

$$A(t)|\nu_\alpha, \mathbf{p}\rangle + B(t)|\nu_\beta, \mathbf{p}\rangle \quad (2.26)$$

of  $\nu_\alpha$  and  $\nu_\beta$  states, where the latter is given by (2.22b), that is

$$|\nu_\beta, \mathbf{p}\rangle = -|\nu_i, \mathbf{p}\rangle \sin \theta_{ij} + |\nu_j, \mathbf{p}\rangle \cos \theta_{ij}. \quad (2.23b)$$

---

<sup>12</sup>See, for example, Chapter 1 of Mandl (1992).

The functions  $A(t)$  and  $B(t)$  are found by solving (2.23a) and (2.23b) for  $|\nu_i, \mathbf{p}\rangle$  and  $|\nu_j, \mathbf{p}\rangle$ , substituting the results into (2.24) and comparing with (2.26). This gives

$$A(t) = a_i(t) \cos^2 \theta_{ij} + a_j(t) \sin^2 \theta_{ij} \quad (2.27a)$$

and

$$B(t) = \sin \theta_{ij} \cos \theta_{ij} [a_j(t) - a_i(t)]. \quad (2.27b)$$

The probability of finding a  $\nu_\beta$  state is therefore, using (2.25),

$$P(\nu_\alpha \rightarrow \nu_\beta) = |B(t)|^2 = \sin^2(2\theta_{ij}) \sin^2 \left[ \frac{1}{2}(E_j - E_i)t \right] \quad (2.28)$$

and thus oscillates with time, while the probability of finding a  $\nu_\alpha$  particle is reduced by a corresponding oscillating factor. Irrespective of which neutrino states  $\nu_\alpha$  and  $\nu_\beta$  are involved, the oscillations vanish if the mixing angle is zero or if the corresponding mass eigenstates  $\nu_{i,j}$  have equal masses, and hence equal energies, as can be seen explicitly from (2.28). In particular, such oscillations are not possible if  $\nu_i$  and  $\nu_j$  both have zero masses.

In practice, the time  $t$  is determined by measuring the distance  $L$  that the neutrinos have travelled from their source, since their momenta are always very much greater than their possible masses and they travel extremely close to the speed of light. Hence, to a very good approximation,  $t = L$  (recall that  $c = 1$  in natural unit) and

$$E_j - E_i = (m_j^2 + p^2)^{1/2} - (m_i^2 + p^2)^{1/2} = \frac{m_j^2 - m_i^2}{2p}, \quad (2.29)$$

where  $p = |\mathbf{p}|$ , so that (2.28) may be written

$$P(\nu_\alpha \rightarrow \nu_\beta) = \sin^2(2\theta_{ij}) \sin^2 [L/L_0], \quad (2.30a)$$

where  $L_0$  is the ‘oscillation length’, given by

$$L_0 = \frac{4E}{(m_j^2 - m_i^2)}, \quad (2.30b)$$

and  $E = p$ , with

$$P(\nu_\alpha \rightarrow \nu_\alpha) = 1 - P(\nu_\alpha \rightarrow \nu_\beta). \quad (2.30c)$$

As we shall see, typical oscillation lengths range from a few hundred metres to a few hundred kilometres, so that neutrino oscillations can be safely neglected under most laboratory conditions. Nevertheless, neutrino oscillations have now been observed in many experiments.

These formulas assume that the neutrinos are propagating in a vacuum. This is usually a very good approximation, because of the enormous mean free paths for neutrinos to interact with matter. However, if the neutrinos travel very long distances through dense matter, their propagation can be significantly modified by their interactions with the dense material they encounter. This is called the MSW effect, after Mikheyev, Smirnov and Wolfenstein, who first

discussed it. It arises from the fact that electron neutrinos interact with electrons in a different way from muon and tauon neutrinos (see Problem 2.5), and hence have different forward scattering amplitudes. When this is taken into account, then the expressions  $\sin(2\theta_{ij})$  and  $\Delta m_{ij}^2 \equiv m_j^2 - m_i^2$  that occur in the oscillation formulas (2.30) are modified as follows:<sup>13</sup>

$$\Delta m_{ij}^2 \rightarrow C \Delta m_{ij}^2 \quad \text{and} \quad \sin(2\theta_{ij}) \rightarrow \sin(2\theta_{ij})/C, \quad (2.31a)$$

where

$$\text{and} \quad C = \sqrt{[\cos(2\theta_{ij}) - A]^2 + \sin^2(2\theta_{ij})} \quad (2.31b)$$

$$A = \pm \frac{2\sqrt{2}G_F N_e E}{\Delta m_{ij}^2}. \quad (2.31c)$$

In these formulas,  $N_e$  is the electron density in matter,  $G_F$  is the Fermi weak interaction coupling and the positive (negative) sign refers to neutrinos (antineutrinos). From them, we see that the magnitude of the effect depends upon the sign of  $A$ , and hence on whether neutrinos or antineutrinos are observed. The sign of  $A$  also depends on the sign of the squared mass difference. This is important, because on substituting (2.31c) into (2.31b) one finds that oscillations in vacuum depend only on the magnitude of the squared mass difference and not its sign. Hence matter effects, if they can be detected, can be used to determine the sign of the mass difference, which is otherwise undetermined. They can also, in special circumstances, be very large because, if  $A$  is positive and the electron density  $N_e$  and neutrino energy  $E$  are such that  $\cos(2\theta_{ij}) - A = 0$ , a resonant enhancement occurs, giving large effects even for small mixing angles.

The MSW effect is an important factor when solar neutrino data are analysed, as we shall see in the following Section 2.3.2. However, in the other experiments discussed in Section 2.3.2, it is too small to be important.

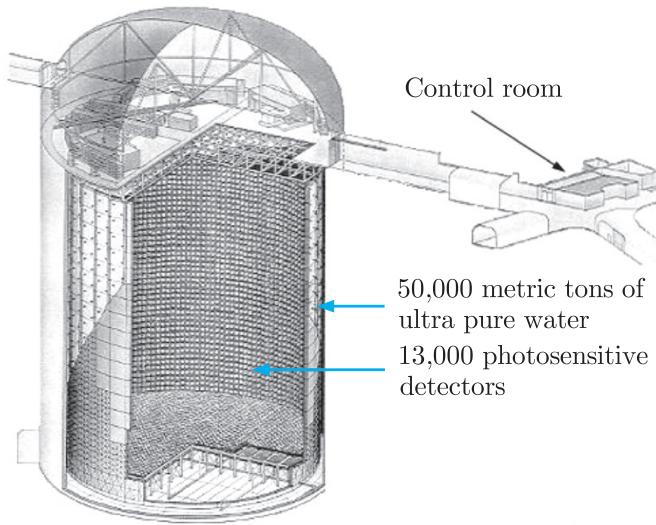
### 2.3.2 Neutrino oscillations

This section is devoted to describing just some of the experiments that established the existence of neutrino oscillations.<sup>14</sup> These experiments fall into four main categories: (a) *atmospheric neutrino experiments*, (b) *neutrino beam experiments*, (c) *solar neutrino experiments* and (d) *reactor experiments*. The first two study ‘muon neutrino oscillations’, in which the initial state is a muon neutrino and the second two are used to study ‘electron neutrino oscillations’.

---

<sup>13</sup>For a derivation, see, for example, Chapter 7 of Giunti and Kim (2007).

<sup>14</sup>For more detailed and complete discussions, see the reviews by Balantekin and Haxton (2013) and by Nakamura and Petcov (2014).



**Figure 2.9** A schematic diagram of the SuperKamiokande detector. (Adapted, with permission, from an original University of Hawaii, Manoa, illustration.)

In all cases, they rely upon the use of lepton number conservation to identify the neutrinos emitted or absorbed in any given reaction as electron, muon or tau-neutrinos.<sup>15</sup> For example, electron neutrinos can produce electrons via reactions like  $\nu_e + n \rightarrow e^- + p$  but cannot produce muons, whereas muon neutrinos can produce muons via reactions like  $\nu_\mu + n \rightarrow \mu^- + p$ , but not electrons.

### 2.3.2(a) Atmospheric neutrino experiments

The first experiment to produce definitive evidence for neutrino oscillations was that of the SuperKamiokande group, who in 1998 used a giant detector to study atmospheric neutrinos. When cosmic ray protons collide with nuclei in the upper atmosphere they create many particles called pions and, as we shall see in Section 3.3, the decay sequence of pions in turn creates neutrinos, with two muon neutrinos being produced for every electron neutrino.

The *SuperKamiokande* detector is shown in Figure 2.9. It consists of a stainless steel cylindrical tank roughly 40 m in diameter and 40 m in height, containing about 50000 metric tons of very pure water. The detector is situated deep underground in the Japanese Alps, at a depth equivalent to 2700 m of water. This is to use the rocks above to shield the detector from cosmic ray muons. The volume is separated into a large inner region, the walls of which are lined with 11200 light-sensitive devices called photomultipliers. When neutrinos with energies above 1 GeV interact with nuclei in the water, the velocities of the electrons and muons produced are greater than the speed of

<sup>15</sup>We emphasise again that it is  $\nu_e$ ,  $\nu_\mu$  and  $\nu_\tau$  that have definite values of the lepton numbers and not the mass eigenstates  $\nu_1$ ,  $\nu_2$  and  $\nu_3$ .

light in water. Because of this, a shock wave of light, called Čerenkov radiation, is emitted. This is analogous to the shock wave emitted when an aircraft exceeds the speed of sound in air. This Čerenkov radiation is detected by the photomultipliers and used to infer properties of the particles that produced it.<sup>16</sup> The outer region of water acts as a shield against low-energy particles entering the detector from outside. An additional 1200 photomultipliers are located there to detect muons that enter or exit the detector.

As stated above, in the absence of neutrino oscillations, one would expect twice as many muon neutrinos as electron neutrinos to be detected. However, the measured ratio was about 1.3 on average, suggesting the presence of oscillations. This was confirmed by exploiting the fact that the detector could measure the direction of the detected neutrinos to study the azimuthal dependence of the effect. Since the flux of cosmic rays that lead to neutrinos with energies above about 1 GeV is isotropic, the production rate for neutrinos should be the same all around the Earth. In particular, one can compare the measured flux from neutrinos produced in the atmosphere directly above the detector, which have a relatively short flight path before detection, with those incident from directly below, which have travelled a long way through the Earth before detection, and so have had more time to oscillate. Experimentally, it was found that the yield of electron neutrinos from above and below were the same within errors and consistent with expectation for no oscillations. However, while the yield of muon neutrinos from above accorded with the expectation for no significant oscillations, the flux of muon neutrinos from below was a factor of about two lower. This is clear evidence for muon neutrino oscillations.

In a later development of the experiment, the flux of muon neutrinos was measured as a function of  $L/E$  by estimating  $L$  from the reconstructed neutrino direction. Values of  $L$  range from 15 km to 13000 km. The results are shown in Figure 2.10 in the form of the ratio of  $P(\nu_\mu \rightarrow \nu_\mu)$  of the observed number of events to the theoretical expectation if there were no oscillations. The data show clear evidence for a deviation of this ratio from unity, particularly at large values of  $L/E$ . Other experiments set limits on  $P(\nu_\mu \rightarrow \nu_e)$ , and taking these into account the most plausible hypothesis is that muon neutrinos are changing into tau neutrinos, which for the neutrino energies concerned could not be detected by SuperKamiokande. As can be seen in Figure 2.10, the data are well described by the two-state mixing formalism of the previous sections and if we define

$$\Delta(m_{ij}^2) \equiv m_i^2 - m_j^2$$

---

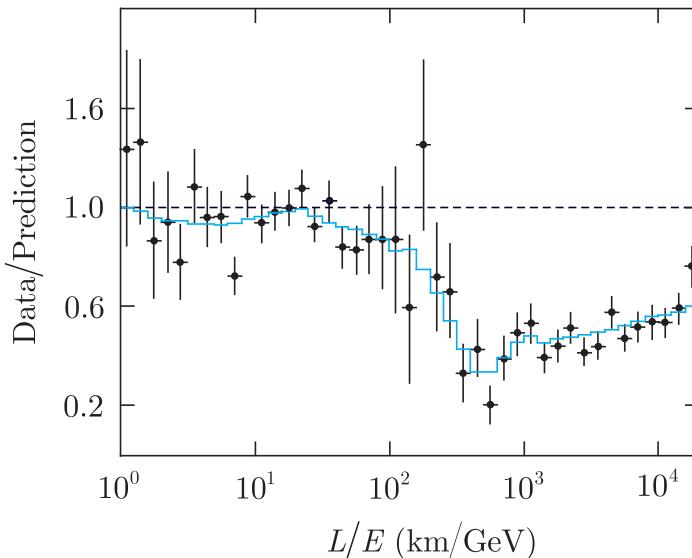
<sup>16</sup>Čerenkov radiation and other aspects of particle detection are discussed in more detail in Chapter 4.

then the experiment yields the values<sup>17</sup>

$$1.9 \times 10^{-3} \lesssim |\Delta(m_{32}^2)| \lesssim 3.0 \times 10^{-3} (\text{eV}/c^2)^2 \quad (2.32\text{a})$$

and

$$\sin^2(2\theta_{23}) \gtrsim 0.9, \quad \text{i.e. } 36^\circ \gtrsim \theta_{23} \gtrsim 54^\circ. \quad (2.32\text{b})$$



**Figure 2.10** Data from the SuperKamiokande detector showing evidence for neutrino oscillations in atmospheric neutrinos. The blue line is the fit from the model of two-neutrino mixing with parameters given by (2.32a) and (2.32b). (Adapted from Ashie *et al.* 2004. Reproduced with permission from the American Physical Society.)

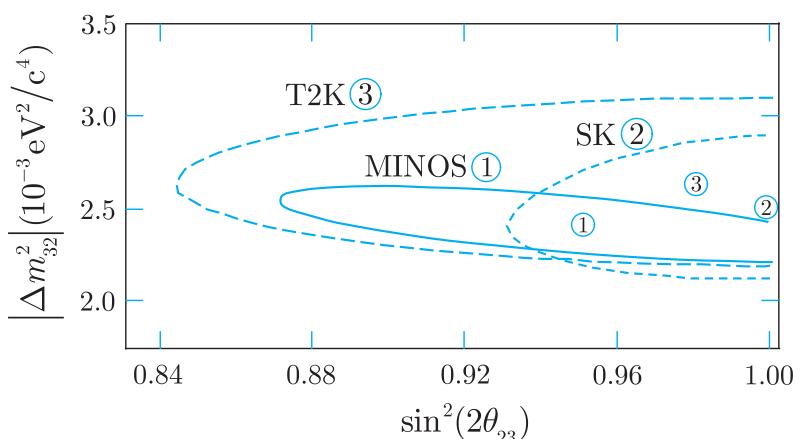
### 2.3.2(b) Neutrino beam experiments

The above results were confirmed by two experiments using muon neutrino beams produced at accelerators.<sup>18</sup> Since the mass difference (2.32a) implies oscillation lengths of order 400 km for neutrinos of energy 1 GeV, this requires the beam intensity to be measured by two detectors a considerable distance apart. In the T2K (Tokyo to Kamioka) experiment, a neutrino beam with a mean energy of 1.3 GeV produced at the KEK accelerator laboratory in Japan was directed to the SuperKamiokande detector 250 km away, with the near detector at 300 m from the source; in the MINOS experiment, a neutrino beam with a peak energy of 3 GeV produced at Fermi National Accelerator Laboratory (FNAL) near Chicago was directed through a near detector at the laboratory to a far detector in the Soudan mine in Minnesota about 735 km away. Both experiments confirmed the existence of neutrino oscillations, with parameters consistent with those obtained in SuperKamiokande, as illustrated in Figure 2.11.

<sup>17</sup>We label the two neutrinos involved in this experiment as  $\nu_2$  and  $\nu_3$ , rather than  $\nu_1$  and  $\nu_2$  to conform with the convention used by Olive *et al.* (Particle Data Group) (2014).

<sup>18</sup>The production of such beams is briefly discussed in Section 4.2.4.

**Figure 2.11** The allowed regions for the parameters  $|\Delta m_{32}^2|$  and  $\sin^2(2\theta_{23})$  at 90% confidence level obtained by two-neutrino mixing fits to later atmospheric neutrino results from SuperKamiokande (SK) and neutrino data from T2K and MINOS. (Adapted from Olive *et al.* (Particle Data Group) (2014) and based on Abe *et al.* (2013)). The smaller circled numbers correspond to the best-fit parameters from the respective experiments, identified by the same numbers.



### 2.3.2(c) Solar neutrinos

Electron neutrino oscillations were first established by a series of observations on neutrinos emitted by the Sun. The energy of the Sun is due to various nuclear reactions and these produce a huge flux of low-energy ( $E \leq 15$  MeV) electron neutrinos, called *solar neutrinos*, that can be detected at the surface of the Earth. This was first done by Davis and co-workers in an experiment that began in 1968 and detected neutrinos via the reaction



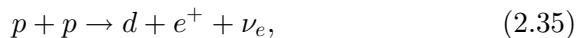
The principal component of this experiment is a huge tank containing tetrachloroethylene ( $\text{C}_2\text{Cl}_4$ ) in which a single argon-37 atom is produced on average every few days by the reaction (2.33). These atoms are unstable, with a half-life of 35 days. They are extracted by flushing the tank with helium gas every few weeks and counted by observing their decays. The experiment was located deep underground, in a gold mine in South Dakota, to reduce the number of background events in which argon-37 atoms are produced by reactions involving cosmic ray muons rather than neutrinos. The inferred neutrino flux is expressed in terms of the ‘solar neutrino unit’ (SNU, pronounced ‘snew’), defined as one capture event per second for every  $10^{36}$  target atoms. After more than twenty years of running, the neutrino flux was measured to be  $2.55 \pm 0.17 \pm 0.18$  SNU, where the first error is statistical and the second systematic. This measured rate is much smaller than the expected rate,  $7.3 \pm 2.3$  SNU, which is predicted by the accepted model of the Sun, called the ‘standard solar model’. The discrepancy between these two values constituted the ‘solar neutrino problem’ and was confirmed by a second experiment, called Kamiokande II, which was a smaller precursor of the SuperKamiokande detector described above. Unlike the

$^{37}\text{Cl}$  experiment, this experiment is sensitive to the direction of the incoming neutrinos, and verifies that they come from the direction of the Sun.

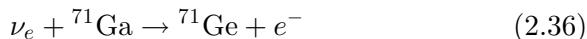
The solution of the solar neutrino problem could be neutrino oscillations or errors in the solar model. In regard to the latter, it is worth noting that both the argon-37 experiment and Kamiokande II could only detect neutrinos if they had an energy in excess of approximately 1 MeV. Such neutrinos come predominantly from the  $\beta$  decay



with an average energy of about 7 MeV. However, this decay contributes only about  $10^{-4}$  to the total solar neutrino flux. It is therefore important to detect neutrinos from other reactions and in particular from the reaction



where  $d$  is the deuteron. This is the primary reaction that produces the energy of the Sun and contributes approximately 90% of the flux of solar neutrinos. The neutrinos in this reaction have average energies of about 0.26 MeV and so cannot be detected in the above experiments. Instead, the reaction



has been used, which has a threshold of 0.23 MeV. Just as for the original experiments of Davis *et al.*, there were formidable problems in identifying the radioactive products from this reaction, which produces only about 1 atom of  ${}^{71}\text{Ge}$  per day in a target of 30 tons of gallium. Nevertheless, results from these experiments, called SAGE, GNO and GALLEX, confirmed the deficit of electron neutrinos and found between 60 and 70% of the flux expected in the absence of oscillations. Since the predicted neutrino flux arising from (2.35) is insensitive to the details of the standard solar model, it is unlikely that shortcomings in the latter could be the source of this discrepancy.<sup>19</sup>

That neutrino oscillations really are the solution to the solar neutrino problem was definitively established by an experiment at the Sudbury Neutrino Observatory (SNO) in Canada in 2002. The experiment consisted of a stainless steel sphere of radius 9 m located 2.1 km underground, and is shown during its construction in coloured Plate 1. It used a water Čerenkov detector, like Kamiokande and

---

<sup>19</sup>For a simple discussion of energy and neutrino generation in stars, see, for example, Section 4.2 of Phillips (1994).

SuperKamiokande, but instead of normal water it used heavy water D<sub>2</sub>O and was therefore able to study the reactions

$$(a) \nu_e + d \rightarrow e^- + p + p, \quad (b) \nu_x + d \rightarrow \nu_x + p + n, \\ (c) \nu_x + e^- \rightarrow \nu_x + e^-, \quad (2.37)$$

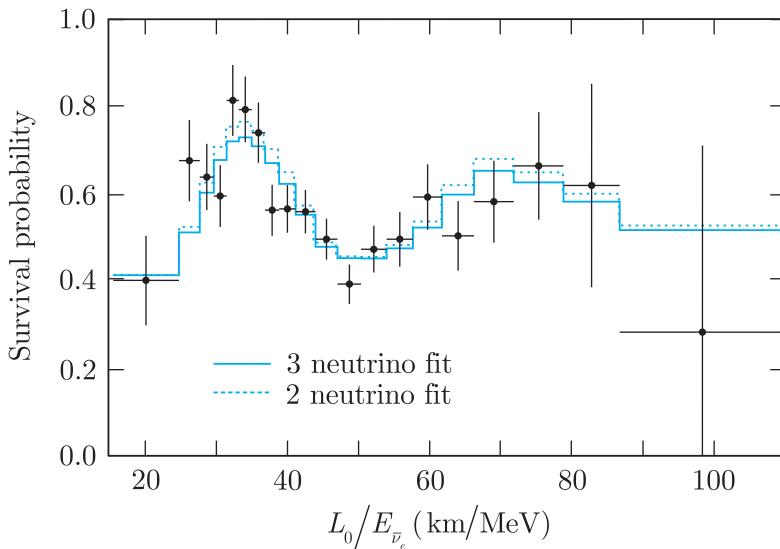
where  $x$  denotes any lepton flavour ( $e, \mu, \tau$ ) and  $d$  is the deuteron. The cross-section for (b) is independent of the lepton flavour (this is a consequence of ‘lepton universality’ discussed in Section 2.2.2) and hence independent of any possible oscillations. Since the observed flux was consistent with expectations, this confirms the correctness of the solar model. On the other hand, the observed flux from (a) was only about 1/3 of expectations, implying that about 2/3 of the electron neutrinos originally produced had transformed to  $\mu$  and/or  $\tau$  neutrinos before being detected at the surface of Earth. The flux for (c) would then be due to a mixture of approximately 1/3 electron neutrinos and 2/3  $\mu/\tau$  neutrinos. Because the cross-section for  $\nu_x e^- \rightarrow \nu_x e^-$  is different for  $\nu_e$  and  $\nu_{\mu,\tau}$  (see Problem 2.5), the rate for (2.37c) should be below what would be expected if there were no oscillations, and the data are consistent with this prediction.

### 2.3.2(d) Reactor experiments

As we have seen in Section 2.1.1, nuclear reactors are a copious source of electron antineutrinos. These have typical energies of a few MeV, and when the solar neutrino results are analysed, taking into account their interactions with matter, they imply that a substantial fraction of the antineutrinos from a nuclear reactor would change to antineutrinos of other flavours after travelling a distance of order 100 km from their source. This prediction has been confirmed by the KamLAND group using a very large detector filled with liquid scintillator and surrounded by phototubes, which detects antineutrinos with energies above 1.8 MeV via the inverse beta decay reaction  $\bar{\nu}_e + p \rightarrow e^+ + n$ .<sup>20</sup> The emitted positron gives rise to scintillation light and, because the kinetic energy of the recoil neutron is typically only a few electron volts, the positron energy can be used to infer the approximate energy of the incident neutrino using energy conservation. The neutron typically loses energy by multiple collisions until it is essentially at rest. It can then undergo the neutron capture reaction  $n + p \rightarrow d + \gamma$  to give rise to a ‘delayed coincidence’ photon of energy 2.2 MeV, which is detected by the photomultipliers about 200 microseconds after the positron is detected. The detector is used to measure the  $\bar{\nu}_e$  flux arriving from more than 60 reactors in Japan and South Korea after the neutrinos have travelled distances

---

<sup>20</sup>This detector is described in more detail in Section 4.6.



**Figure 2.12** The ratio of the observed  $\bar{\nu}_e$  spectrum to that predicted without oscillations (the ‘survival probability’) observed in the KamLAND experiment, showing the best fits obtained using two-neutrino and three-neutrino mixing models. (Gando *et al.* 2011. Reproduced with permission from the American Physical Society.)

of between 150 and 200 km to reach the Kamioka mine, where the detector is located. The first results, published in 2003, showed that the  $\bar{\nu}_e$  flux was only about 60% of that expected from the known characteristics of the reactors, and subsequent data clearly showed the oscillatory data of the effect, as can be seen in Figure 2.12. A simultaneous analysis of the data shown in this figure, together with the solar neutrino data, using two-component mixing yields the results

$$\Delta m_{21}^2 = 7.50^{+0.19}_{-0.20} \times 10^{-5} (\text{eV}/c^2)^2 \quad (2.38a)$$

and

$$\tan^2 \theta_{12} = 0.444^{+0.036}_{-0.030} \quad (2.38b)$$

In this analysis, it is essential to include matter effects (i.e. the MSW effect discussed earlier) in the description of the solar neutrino data, in order to obtain a consistent account, and because these are far from negligible, both the sign and magnitude of the squared mass difference are determined in this case.

From (2.32) and (2.38), we see that the splitting between the squared masses of neutrinos  $\nu_1$  and  $\nu_2$  is very small compared to the corresponding splitting between neutrinos  $\nu_2$  and  $\nu_3$ . Hence

$$|\Delta m_{13}^2| \approx |\Delta m_{23}^2| \approx 2 \times 10^{-3} (\text{eV}/c^2)^2$$

so that, for a neutrino of a given energy, the oscillation length (2.31) between states 1 and 3 is much shorter than the oscillation length between states 1 and 2. For reactor neutrinos, this gives rise to oscillation lengths of order kilometres rather than hundreds of kilometres. To investigate this, the KamLAND and solar neutrino data were re-analysed taking into account the possibility of mixing between all

three neutrinos. The resulting fit is shown in Figure 2.12. It yields parameter values

$$\Delta m_{21}^2 = 7.50_{-0.20}^{+0.19} \times 10^{-5} (\text{eV}/c^2) \quad \text{and} \quad \tan^2 \theta_{12} = 0.452_{-0.033}^{+0.035},$$

which are essentially unchanged from those obtained using two-component mixing, together with a very small value

$$\sin^2 \theta_{13} = 0.020 \pm 0.016$$

for the third mixing angle, which is not convincingly different from zero.

The first clear evidence for non-zero oscillations between  $\nu_1$  and  $\nu_3$  was obtained in 2012 by three reactor experiments with the detectors sufficiently close to the reactors for the contribution from the long baseline oscillations due to  $\nu_1 - \nu_2$  mixing to be very small. These are: the Daya Bay collaboration, which observed neutrinos from six reactors at the Daya Bay nuclear power plant in China using six detectors distributed between three sites at distances of approximately 470, 580 and 1650 m; the RENO collaboration, which measured the neutrino fluxes from six reactors at the Yonggwang nuclear site in Korea using identical detectors at approximately 300 and 1400 m from the reactor centre; and the Double Chooz experiment in France, which measured the flux from two reactors with a detector approximately 1 km from the cores. Recent results for  $\sin^2(2\theta_{13})$  obtained by the three experiments are:

$$\begin{aligned} 0.089 \pm 0.010 \pm 0.005 & \quad (\text{Daya Bay, An } et al., 2013); \\ 0.100 \pm 0.010 \pm 0.012 & \quad (\text{RENO, Seo, 2013}); \\ 0.109 \pm 0.030 \pm 0.025 & \quad (\text{Double Chooz, Abe } et al., 2012), \end{aligned}$$

where in each case the two errors are statistical and systematic, respectively. The results are in excellent agreement and clearly demonstrate the existence of oscillations between neutrino species 1 and 3. The current best value, obtained by combining these and other experiments, is

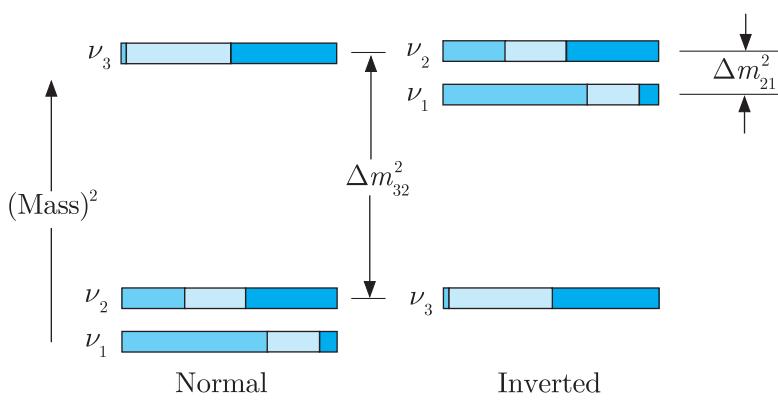
$$\sin^2(2\theta_{13}) = 0.093 \pm 0.008. \quad (2.39)$$

### 2.3.3 Neutrino masses

In order to extract reliable information from all the oscillation data in the most general way, it is necessary to extend the discussion of Section 2.3.1 to include mixing between all three neutrino mass

states  $\nu_1$ ,  $\nu_2$  and  $\nu_3$ , rather than just between two. We will not discuss this in detail,<sup>21</sup> except to say that a global fit to all the data, including the experiments described above, yields values of  $\Delta m_{32}^2$ ,  $\theta_{23}$ ,  $\Delta m_{21}^2$  and  $\theta_{12}$ , which are consistent with those given in equations (2.32), (2.38) and (2.39). We will discuss the implications of these results for neutrino masses in what follows.

We first reiterate that, for the solar neutrino data, the interactions with matter play an important role. As a bonus, this enables the sign of  $\Delta m_{21}^2$  to be measured, whereas oscillations in free space only determine the magnitude of  $\Delta m_{ij}^2$ , as can be seen from (2.30a, b, c). Matter effects can also be used to determine the sign of  $\Delta m_{32}^2$ . In muon neutrino oscillation experiments using neutrino beams generated at accelerators, like the MINOS and T2K experiments discussed in Section 2.3.2(a) above, the neutrinos travel large distances through the Earth's crust before reaching the detector, and the sign of  $\Delta m_{32}^2$  can in principle also be determined when neutrino interactions with matter are taken into account. However, the effect is small and the experiments are not yet precise enough to determine the sign of  $\Delta m_{32}^2$  in this way. Because of this, two solutions for the mass hierarchy are compatible with the current data: the so-called ‘normal’ mass hierarchy,  $m_3 > m_2 > m_1$ , and the ‘inverted’ mass hierarchy,  $m_2 > m_1 > m_3$ . Both cases are shown in Figure 2.13, where we also show the approximate flavour decomposition of the mass eigenstates resulting from three-component mixing with parameters compatible with (2.32), (2.38) and (2.39). As can be seen, the lighter of the two



**Figure 2.13** Three-neutrino squared-mass spectra in both the ‘normal’ and ‘inverted’ mass hierarchies, which are consistent with the values for squared mass differences and mixing angles given in (2.32), (2.38) and (2.39). Also shown are the approximate fractional flavour compositions of  $\nu_1$ ,  $\nu_2$  and  $\nu_3$ :  $\nu_e$  (left shaded area),  $\nu_\mu$  (middle shaded area),  $\nu_\tau$  (right shaded area).

<sup>21</sup>The general form of the mixing will be discussed in Section 11.3. As we shall see there, in addition to the three mixing angles, the mixing matrix contains a phase angle  $\delta$  that can be non-zero if a symmetry called ‘ $CP$  invariance’, which is discussed in Chapter 5, is violated in neutrino oscillations. No such violation has, so far, been detected.

states  $\nu_1$  and  $\nu_2$  that dominate solar neutrino oscillations is predominately an electron neutrino, while  $\nu_3$  has large  $\nu_\mu$  and  $\nu_\tau$  components, but a very small electron neutrino component.

We can now return to the interpretation of the bound (2.7) and consider its consequences for the ‘mass’ of the electron neutrino. The point here is that neutrinos with definite flavours, like the electron neutrino, are superpositions of the mass eigenstates  $\nu_1$ ,  $\nu_2$  and  $\nu_3$  and do not themselves have definite masses. Rather, in accord with the standard theory of measurement in quantum mechanics, a measurement of the mass of the electron neutrino can yield any one of the three values  $m_1$ ,  $m_2$  or  $m_3$ . However, if  $m_i^2 > m_j^2$ , one can easily show that  $(m_i - m_j)^2 < m_i^2 - m_j^2$ . Hence, (2.32) and (2.38) set upper limits on the neutrino mass differences that are approximately

$$m_2 - m_1 \lesssim 10^{-2} \text{ eV}/c^2, \quad |m_3 - m_2| \lesssim 5 \times 10^{-2} \text{ eV}/c^2. \quad (2.40)$$

In other words, the neutrino masses are almost equal compared to a mass scale of  $1 \text{ eV}/c^2$ . Hence it is safe to interpret (2.7) as implying

$$m_i \leq 2 \text{ eV}/c^2 \quad (2.41)$$

for all three neutrinos  $\nu_1$ ,  $\nu_2$  and  $\nu_3$ . This limit is very small compared to the mass of the lightest charged particle, the electron, and is a great improvement on the limits (2.15) obtained from muon and tauon decay experiments.

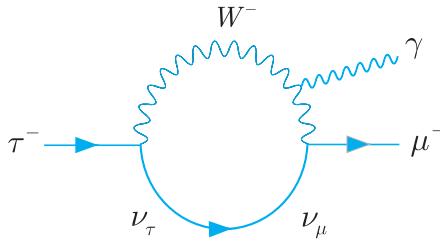
Finally, it is worth remarking that it is possible to obtain bounds on neutrino masses from cosmology. The most exacting of these comes from using the current standard cosmological model to analyse the large-scale structure of the universe and the latest bound, from the Planck collaboration (Ade *et al.*, 2014), is

$$\sum_{i=1,3} m_i < 0.23 \text{ eV}/c^2. \quad (2.42)$$

This is more stringent than (2.41), but unlike the latter is not a direct measurement of mass. We will return to these cosmological arguments briefly in Section 12.4.1.

### 2.3.4 Lepton numbers revisited

In our general discussion of the weak interaction, we have assumed that lepton number conservation holds and can be used to identify the neutrino flavour emitted or absorbed in any weak reaction. However, in principle, lepton number violation can be induced in such reactions by the existence of neutrino oscillations. In practice, such effects are totally negligible in the standard model due to the short range of the weak interaction and can indeed be safely ignored.



**Figure 2.14** A Feynman diagram contributing to the decay  $\tau^- \rightarrow \mu^- + \gamma$ . There are two other diagrams, where the photon is emitted by either the  $\mu^-$  or the  $\tau^-$ .

To illustrate this, we will make a very rough estimate of the branching ratio for the lepton number violating decay<sup>22</sup>

$$\tau^- \rightarrow \mu^- + \gamma, \quad (2.43)$$

which can proceed via the mechanism of Figure 2.14, in which the tau neutrino emitted at the first vertex oscillates into a muon neutrino before being re-absorbed at the second vertex. This diagram involves two weak interaction vertices, the emission of a photon and an oscillation, while the leptonic decay modes (2.11) are purely weak interactions. Therefore, the branching fraction of the decay (2.43) is expected to be of the order of magnitude

$$B(\tau^- \rightarrow \mu^- + \gamma) = O[\alpha P(\nu_\tau \rightarrow \nu_\mu) B(\tau^- \rightarrow e^- \nu_\tau \bar{\nu}_e)], \quad (2.44)$$

where  $P(\nu_\tau \rightarrow \nu_\mu)$  is the probability of oscillation and  $\alpha$  is the fine structure constant. Thus, using  $B(\tau^- \rightarrow e^- \nu_\tau \bar{\nu}_e) \sim O(10^{-1})$ , we have

$$B(\tau^- \rightarrow \mu^- + \gamma) = O[10^{-3} P(\nu_\tau \rightarrow \nu_\mu)]. \quad (2.45)$$

The term  $P(\nu_\tau \rightarrow \nu_\mu)$  can be estimated from the two-component mixing model of  $\nu_\mu - \nu_\tau$  oscillations used to describe the atmospheric neutrino data in Section 2.3.2. From (2.30), we then have

$$P(\nu_\tau \rightarrow \nu_\mu) \approx \sin^2(2\theta_{23}) \sin^2(L/L_0) \approx (L/L_0)^2, \quad (2.46a)$$

where we have used  $\sin^2(2\theta_{23}) \gtrsim 0.9$  and assumed  $L \ll L_0$ . Here  $L$  is the typical distance travelled by the neutrino and is of the same order as the range  $M_W^{-1}$  of the weak interaction. The oscillation length is given by

$$L_0 = 4E/\Delta m_{32}^2, \quad (2.46b)$$

where  $E$  is the typical energy of the neutrinos in Figure 2.14. Substituting (2.46a,b) into (2.45) gives

$$B(\tau^- \rightarrow \mu^- + \gamma) = O\left[\frac{10^{-3}(m_3^2 - m_2^2)^2}{16E^2 M_W^2}\right]. \quad (2.47)$$

<sup>22</sup>The details of the argument that follows is not required for later work.

For  $E$  we make the crude dimensional estimate  $m_\mu \leq E \leq m_\tau$  for a tauon decaying at rest. Therefore, finally, using the experimental value (2.32) for  $\Delta m_{32}^2$ , we have the rough estimate

$$B(\tau \rightarrow \mu + \gamma) \approx 10^{-48} - 10^{-50}. \quad (2.48)$$

This is utterly negligible for all practical purposes and should be compared to the experimental upper limit given in Table 2.1.

The above conclusion is confirmed by a more detailed treatment of both this and other reactions. Lepton number violation in weak interactions is completely negligible within the standard model, but could occur if one abandons the standard model. Some particularly interesting possibilities and the experiments required to investigate them will be briefly discussed in Chapter 12.

## Problems 2

- 2.1** Which of the following reactions are allowed, and which are forbidden, by the conservation laws appropriate to weak interactions?
- (a)  $\nu_\mu + p \rightarrow \mu^+ + n$ ,
  - (b)  $\nu_e + p \rightarrow e^- + \pi^+ + p$ ,
  - (c)  $\nu_\tau + e^- \rightarrow \tau^- + \nu_e$ ,
  - (d)  $\tau^+ \rightarrow \mu^+ + \bar{\nu}_\mu + \nu_\tau$ .
- 2.2** Figure 2.5 shows one of the time-ordered diagrams for muon decay  $\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu$ . What are the basic processes associated with the two vertices? What are the basic processes associated with the other possible time-ordered diagram?
- 2.3** Draw the lowest-order Feynman diagram for  $\nu_e \nu_\mu$  elastic scattering.
- 2.4** Draw the two fourth-order Feynman diagrams for the weak reaction  $e^- + \mu^+ \rightarrow \nu_e + \bar{\nu}_\mu$ .
- 2.5** In Section 2.3.1 it is stated that electron neutrinos interact with electrons in a different way from muon and tauon neutrinos. Justify this remark by considering the lowest-order Feynman diagrams for  $\nu_e + e^- \rightarrow \nu_e + e^-$  and  $\nu_\mu + e^- \rightarrow \nu_\mu + e^-$ .
- 2.6** Show that the oscillation length in (2.30b) may be written  $L_0 = E/(1.27\Delta m_{ij}^2)$ , where  $L_0$  is expressed in km,  $E$  in GeV and  $\Delta m_{ij}^2$  in  $(\text{eV}/c^2)^2$ .
- 2.7** In the KamLAND experiment, antineutrinos are detected via the reaction  $\bar{\nu}_e + p \rightarrow e^+ + n$ , where the target protons are at rest. Show that for incident antineutrinos of energy  $E_\nu = 3 \text{ MeV}$ , the positron produced has an energy

$$E_+ = E_\nu + m_p - m_n = 1.71 \text{ MeV} \quad (2.49)$$

to within an error of 1%, irrespective of the scattering angle.

- 2.8** An electron antineutrino is produced at the centre of the Sun. Somewhere on its journey to the surface it passes through a region where the density is such that matter effects lead to a resonant enhancement of its oscillations. Estimate the minimum energy required for this to occur if electron neutrino oscillations *in vacuo* are described by a two-component mixing model with the parameters

$\Delta m_{12}^2 = -7.5 \times 10^{-5} (\text{eV}/c^2)^2$  and  $\tan^2 \theta_{12} = 0.444$ . The electron density at the centre of the Sun is approximately  $N_0 = 3 \times 10^{31} \text{ m}^{-3}$  and decreases steadily towards the surface.

- 2.9** Calculate the oscillation lengths  $L_{13}$ ,  $L_{12}$  associated with  $\nu_1 \rightarrow \nu_3$  and  $\nu_1 \rightarrow \nu_2$  oscillations for electron antineutrinos of energy 3 MeV. Why is a two-component mixing model a good first approximation for neutrino oscillations at distances of order 1 km and of order 150–200 km? Would it be a reasonable approximation for distances of order 10 km?
- 2.10** A hypothetical KamLAND-type experiment detects  $\bar{\nu}_e$  neutrinos at a distance of 300 m from a nuclear reactor and finds that the flux is  $(90 \pm 10)\%$  of that expected if there were no oscillations. Assuming a two-component model with maximal mixing ( $\theta = 45^\circ$ ) and a mean neutrino energy of 2.5 MeV, estimate the squared mass difference of the electron antineutrino and its oscillating partner.
- 2.11** If the Sun is assumed to be a uniform spherical plasma consisting of nucleons, with radius  $7 \times 10^5 \text{ km}$  and total mass  $2 \times 10^{30} \text{ kg}$ , calculate the mean free path  $\lambda = 1/n\sigma$  of solar neutrinos from the dominant reaction (2.35). Here  $n$  is the number of nucleons per unit volume and  $\sigma$ , the neutrino-nucleon cross-section, may be written  $\sigma = 0.7E_L \times 10^{-42} \text{ m}^2$ , where  $E_L$  is the neutrino laboratory energy in GeV.

# 3

---

## Quarks and hadrons

---

We turn now to the strongly interacting particles – the quarks and their bound states, the hadrons. These also interact by the weak and electromagnetic interactions, although such effects can often be neglected compared to the strong interactions. To this extent we are entering the realm of ‘strong interaction physics’.

Strong interactions are most familiar in nuclear physics, where the interactions of neutrons and protons are studied at relatively low energies of a few tens of MeV. However, in 1947 new types of hadrons, not present in ordinary matter, were discovered in cosmic rays by groups from the universities of Bristol and Manchester. To create these new particles required high energies, in accordance with Einstein’s mass–energy relation  $E = mc^2$ , and as intense beams of particles of increasingly higher energies became available at accelerator laboratories, more and more hadrons were discovered. By the late 1960s several dozen were known, and some unifying theoretical framework was urgently needed to interpret this multitude of states if any progress was to be made. The result was the *quark model*. In 1964, Gell-Mann, and independently Zweig, noted that all the observed hadrons could be simply interpreted as bound states of just three fundamental spin-1/2 particles, together with their antiparticles. These particles were required to have fractional electric charges of 2/3 and  $-1/3$ , in units of e, and were called *quarks* by Gell-Mann, who later cited the now famous quotation ‘Three quarks for Muster Mark’ in James Joyce’s book *Finnegans Wake*.<sup>1</sup>

In the following years, the success of the quark model grew ever more impressive as more and more states were discovered.

---

<sup>1</sup> In fact more than three quarks are known, but this was not discovered until ten years later (see below).

Nonetheless, the existence of quarks as real particles, rather than convenient mathematical entities, was seriously doubted because all attempts to detect free quarks, or any other fractionally charged particles, met with failure. These doubts were subsequently removed in two ways. Firstly, a series of experimental results, starting in 1968 with the scattering of high-energy electrons from protons, showed the dynamical effects of individual quarks within the proton. These experiments will be described in Chapters 7 and 8. Secondly, a detailed theory of strong interactions was constructed, which both successfully described the experimental results and offered an explanation of why isolated free quarks could not be observed. This theory is called *Quantum Chromodynamics (QCD)*, and will also be discussed in Chapter 7. As a result of these developments, the quark hypothesis is now universally accepted, and is central to the interpretation of a wide range of phenomena in particle physics.

In what follows, we take the existence and properties of quarks as our basic assumptions, interpreting hadrons in terms of them. We begin by briefly introducing the quarks, together with a few of their best-known hadronic bound states, leaving more detailed considerations to later chapters.

### 3.1 Quarks

Although only three quarks were initially proposed, six are now known to exist. Like the leptons, these six distinct types, or *flavours*, occur in pairs, or generations, denoted

$$\begin{pmatrix} u \\ d \end{pmatrix}, \quad \begin{pmatrix} c \\ s \end{pmatrix}, \quad \begin{pmatrix} t \\ b \end{pmatrix}. \quad (3.1a)$$

Each generation consists of a quark with charge  $+2/3$  ( $u$ ,  $c$  or  $t$ ) together with a quark of charge  $-1/3$  ( $d$ ,  $s$ ,  $b$ ), in units of  $e$ . They are called the *down* ( $d$ ), *up* ( $u$ ), *strange* ( $s$ ), *charmed* ( $c$ ), *bottom* ( $b$ ) and *top* ( $t$ ) quarks. The corresponding antiquarks are denoted

$$\begin{pmatrix} \bar{d} \\ \bar{u} \end{pmatrix}, \quad \begin{pmatrix} \bar{s} \\ \bar{c} \end{pmatrix}, \quad \begin{pmatrix} \bar{b} \\ \bar{t} \end{pmatrix}, \quad (3.1b)$$

with charges  $+1/3$  ( $\bar{d}, \bar{s}, \bar{b}$ ) and  $-2/3$  ( $\bar{u}, \bar{c}, \bar{t}$ ). There is good experimental evidence for the existence of all six flavours ( $d, u, s, c, b$  and  $t$ ). Approximate quark masses are given in Table 3.1. For the  $d$ ,  $u$ ,  $s$ ,  $c$  and  $b$  quarks, these are the *constituent quark masses* inferred indirectly from the observed masses of their hadron bound states, together with simple models of quark binding. These masses are somewhat analogous to the effective masses used to describe conduction electrons in metals and semiconductors, which are not the same as the masses of free electrons, but are very useful nonetheless.

**Table 3.1** Approximate values of quarks masses in units of  $\text{GeV}/c^2$ , as discussed in the text, and their electric charges  $Q$  in units of  $e$ . Also shown are the values of the quantum numbers: baryon number  $B$ , strangeness  $S$ , charm  $C$ , bottom  $\tilde{B}$  and top  $T$ , which will be defined in Section 3.2. Their values, and those of  $Q$ , for the corresponding antiquarks are equal in magnitude, but opposite in sign.

Name	Symbol	Mass	$Q$	$B$	$S$	$C$	$\tilde{B}$	$T$
Down	$d$	$m_d \approx 0.3$	$-1/3$	$1/3$	0	0	0	0
Up	$u$	$m_u \approx m_d$	$2/3$	$1/3$	0	0	0	0
Strange	$s$	$m_s \approx 0.5$	$-1/3$	$1/3$	$-1$	0	0	0
Charmed	$c$	$m_c \approx 1.5$	$2/3$	$1/3$	0	1	0	0
Bottom	$b$	$m_b \approx 4.5$	$-1/3$	$1/3$	0	0	$-1$	0
Top	$t$	$m_t \approx 173$	$2/3$	$1/3$	0	0	0	1

The top quark is too short-lived to form observable hadron states and its mass is again inferred indirectly, from observations on the decay products of  $t\bar{t}$  pairs, as we shall see.<sup>2</sup> None of the masses can be obtained from measurements on free quarks because free quarks have never been seen, despite many experiments to find them.

Free quarks would be most readily identified via their fractional electric charges. One consequence of this is that the lightest quark would be stable, as electric charge has to be conserved in any decay. In matter, such stable quarks would give rise to ‘exotic atoms’ with fractional charges that could be identified by techniques like mass spectroscopy. Many searches for pre-existing quarks in matter have been made and many strange materials have been investigated (including moon rock, crushed oyster shells and deep sea sludge!), all with null results. The best upper limit on a possible quark density is approximately  $10^{-24}$  quarks/nucleon in sea water.

Another technique for detecting possible pre-existing free quarks is to do experiments similar to the classic Millikan oil drop experiment that first measured precisely the charge on the electron. The first experiment of this type was by Gallinaro and Morpurgo in 1966 and failed to find evidence for fractional charged particles. Later work by Larue, Fairbank and Hubbard in 1977, using niobium balls at low temperatures to replace Millikan’s oil drops, claimed evidence for several examples of fractional charges consistent with  $\pm e/3$ . However, further experiments using a variety of different materials, including niobium balls, failed to confirm the original results. One can conclude that the possible concentration of free quarks in such metals is less than  $10^{-20}$  quarks per nucleon.

There have also been attempts to find quarks that might have been created in high-energy collisions in cosmic rays, or at accelerator

<sup>2</sup> This will be discussed in Section \*9.2.3.

laboratories. The aim here is to use the fact that the ionisation produced by quarks would be either  $1/9$  or  $4/9$  of that produced by a particle of charge  $e$ . In cosmic ray searches it is concluded that the concentration of quarks must be less than one per  $10^{10}$  primary cosmic rays. More than this would have produced a detectable signal in the secondary cosmic rays observed at sea level. At accelerators, a wide range of reactions has been studied, all with null results.

In summary, there is no evidence for the existence of isolated free quarks, or any other fractionally charged particles, despite great efforts to find them.

## 3.2 General properties of hadrons

Although no isolated quarks have been found, well over two hundred of their bound states have been discovered, all with integer electric charges. The reason for the latter is closely associated with a new degree of freedom that exists for quarks, but not for leptons, called *colour*. We shall postpone discussion of this for the moment and follow the *simple quark model* in assuming that only the three simplest types of quark bound states with integer electric charge need be considered. These are the *baryons*, which have half-integer spin and are assumed to be bound states of three quarks ( $qqq$ ); the *antibaryons*, which are their antiparticles and assumed to be bound states of three antiquarks ( $\bar{q}\bar{q}\bar{q}$ ); and the *mesons*, which have integer spin and are assumed to be bound states of a quark and an antiquark ( $q\bar{q}$ ). Some examples of baryons and mesons that we shall meet later in this chapter, together with their quark compositions, are shown in Tables 3.2 and 3.3.

Tables 3.2 and 3.3 also show the values of several quantum numbers that are associated with any state and that refer to its quark content. For example, the *strangeness*  $S$  is defined by

$$S \equiv -N_s \equiv -[N(s) - N(\bar{s})], \quad (3.2a)$$

**Table 3.2** Some examples of baryons, with their quark compositions and the corresponding values of their electric charge  $Q$ , strangeness  $S$ , charm  $C$  and bottom  $\tilde{B}$ .

Particle		Mass (MeV/ $c^2$ )	$Q$	$S$	$C$	$\tilde{B}$
$p$	$uud$	938	1	0	0	0
$n$	$udd$	940	0	0	0	0
$\Lambda$	$uds$	1116	0	-1	0	0
$\Lambda_c$	$udc$	2286	1	0	1	0
$\Lambda_b$	$udb$	5620	0	0	0	-1

**Table 3.3** Some examples of mesons, with their quark compositions and the corresponding values of their electric charge  $Q$ , strangeness  $S$ , charm  $C$  and bottom  $\tilde{B}$ .

Particle		Mass (MeV/ $c^2$ )	$Q$	$S$	$C$	$\tilde{B}$
$\pi^+$	$u\bar{d}$	140	1	0	0	0
$K^-$	$s\bar{u}$	494	-1	-1	0	0
$D^-$	$d\bar{c}$	1870	-1	0	-1	0
$D_s^+$	$c\bar{s}$	1968	1	1	1	0
$B^-$	$b\bar{u}$	5279	-1	0	0	-1
$\Upsilon$	$b\bar{b}$	9460	0	0	0	0

where  $N(s)$  and  $N(\bar{s})$  are the number of  $s$  quarks and  $\bar{s}$  antiquarks present in the state. Clearly  $S = -1$  for an  $s$  quark,  $S = 1$  for an  $\bar{s}$  antiquark and  $S = 0$  for all other quarks and antiquarks. The *charm* ( $C$ ), *bottom* ( $\tilde{B}$ ), and *top* ( $T$ ) quantum numbers are similarly defined by<sup>3</sup>

$$C \equiv N_c \equiv N(c) - N(\bar{c}), \quad (3.2b)$$

$$\tilde{B} \equiv -N_b \equiv -[N(b) - N(\bar{b})] \quad (3.2c)$$

and

$$T \equiv N_t \equiv N(t) - N(\bar{t}), \quad (3.2d)$$

respectively, giving values for the individual quarks, which are summarised for convenience in Table 3.1.

The resulting values of  $S$ ,  $C$  and  $\tilde{B}$  for our specimen hadrons are given in Tables 3.2 and 3.3 together with the hadron's electric charge  $Q$ , which is just the sum of the appropriate quark charges. The top quantum number  $T = 0$  for all known hadrons. The remaining quark numbers,

$$N_u = N(u) - N(\bar{u}) \quad (3.3a)$$

and

$$N_d = N(d) - N(\bar{d}), \quad (3.3b)$$

do not have special names since, given  $S, C, \tilde{B}, T$  and  $Q$ , their values can be inferred from the *baryon number*  $B$  of the state, defined by

$$B \equiv \frac{1}{3}[N(q) - N(\bar{q})], \quad (3.4a)$$

where  $N(q)$  and  $N(\bar{q})$  are the total numbers of quarks and antiquarks present, irrespective of their flavours. Baryon number  $B = 1$

<sup>3</sup> *Bottom* was originally called *beauty* and *top* was called *truth*, but these names are seldom used now.

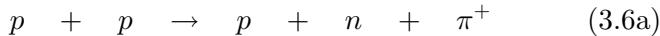
for baryons,  $B = -1$  for antibaryons and  $B = 0$  for mesons. In terms of the quark numbers (3.2) and (3.3)

$$\begin{aligned} B &= \frac{1}{3}(N_u + N_d + N_s + N_c + N_b + N_t) \\ &= \frac{1}{3}(N_u + N_d - S + C - \tilde{B} + T), \end{aligned} \quad (3.4b)$$

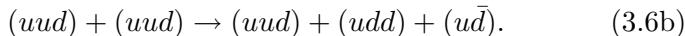
while the electric charge is given by

$$\begin{aligned} Q &= \frac{2}{3}(N_u + N_c + N_t) - \frac{1}{3}(N_d + N_s + N_b) \\ &= \frac{2}{3}(N_u + C + T) - \frac{1}{3}(N_d - S - \tilde{B}). \end{aligned} \quad (3.5)$$

All these quantum numbers are called *internal quantum numbers* because they are not associated with motion or the spatial properties of wave functions. They are important because in strong and electromagnetic interactions quarks and antiquarks are only created or destroyed in particle–antiparticle pairs. For example, the quark description of the strong interaction process



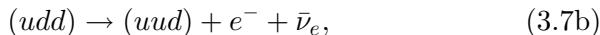
is



On counting the quarks of each flavour, we see that the final state contains the same number of quarks of each flavour as the initial state, plus an additional  $(d\bar{d})$ , so that the quark numbers  $N_u$  and  $N_d$  are separately conserved. This is characteristic of strong and electromagnetic processes, in which all the quark numbers (3.2)–(3.5) are separately conserved. However, in neutron  $\beta$  decay,

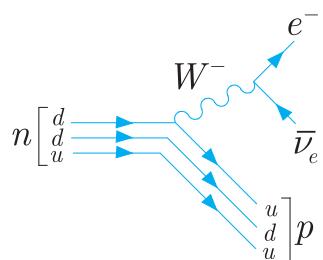


whose quark interpretation is



where a  $d$ -quark is replaced by a  $u$ -quark. This decay is conveniently represented by the *quark diagram* of Figure 3.1. This corresponds to the Feynman diagram of Figure 1.13, except that each hadron is represented by lines corresponding to its constituent quarks and antiquarks, with quarks and antiquarks labelled by arrows pointing to the right and left, respectively. The fact that  $N_u$  and  $N_d$  are not conserved is characteristic of the weak interactions, in which the quark flavours can change and only baryon number (3.4a) and the total electric charge are in general conserved.

The quark numbers (3.2)–(3.5) play an important role in understanding the long lifetimes of some hadrons. Hadrons have typical effective radii  $r$  of order 1 fm, with an associated time scale  $r/c$  of order  $10^{-23}$  s. The vast majority are highly unstable, and decay to lighter hadrons by the strong interaction with lifetimes of this order. However, each hadron is characterised by a set of values for



**Figure 3.1** Quark diagram for the decay  $n \rightarrow p + e^- + \bar{\nu}_e$ .

**Table 3.4** Typical lifetimes of hadrons decaying by the three interactions.

Interaction	Lifetime (s)
Strong	$10^{-22} - 10^{-24}$
Electromagnetic	$10^{-16} - 10^{-21}$
Weak*	$10^{-7} - 10^{-13}$

\*The neutron lifetime is an exception, for reasons explained in Section 3.2.

$B$ ,  $Q$ ,  $S$ ,  $C$ ,  $\tilde{B}$  and  $T$ , and in some cases there are no lighter hadron states with the same values of these quantum numbers to which they can decay. These hadrons, which cannot decay by strong interactions, are long-lived on a timescale of order  $10^{-23}$  s and are often called *stable particles*. Here we shall call them *long-lived particles*, because except for the proton they are not absolutely stable, but decay by either the electromagnetic or weak interaction.<sup>4</sup> Electromagnetic decay rates are suppressed by powers of the fine structure constant  $\alpha$  relative to strong decays, leading to observed lifetimes in the range  $10^{-16} - 10^{-21}$  s. Weak decays give longer lifetimes that depend sensitively on the characteristic energy of the decay, as we have seen in the case of the  $\mu$  and  $\tau$  leptons (cf. Section 2.2.2). A useful measure of this characteristic energy is the *Q value*, which is the kinetic energy released in the decay of the particle at rest. For neutron decay (3.7a),

$$Q = m_n - m_p - m_e - m_{\bar{\nu}_e} = 0.79 \text{ MeV} \quad (3.8)$$

and the lifetime is about  $10^3$  s. However,  $Q$  values of order  $10^2 - 10^3$  MeV are more typical, and the observed lifetimes for other weak hadron decays lie in the range  $10^{-7} - 10^{-13}$  s. Thus hadron lifetimes span some 27 orders of magnitude, from about  $10^{-24}$  s to about  $10^3$  s. The typical ranges corresponding to each interaction are summarised in Table 3.4.

### 3.3 Pions and nucleons

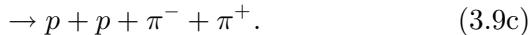
The lightest known mesons are the pions or pi-mesons  $\pi^\pm(140)$ ,  $\pi^0(135)$ , where we indicate the masses in  $\text{MeV}/c^2$  in brackets. These particles are produced copiously in many hadronic reactions that

<sup>4</sup> Some modern theories predict that even the proton is unstable, albeit with a lifetime of order  $10^{29} - 10^{33}$  y. This is discussed in Section 12.1.3.

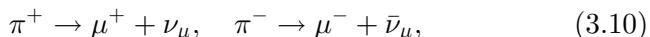
conserve both charge and baryon number, for example in the proton–proton collisions



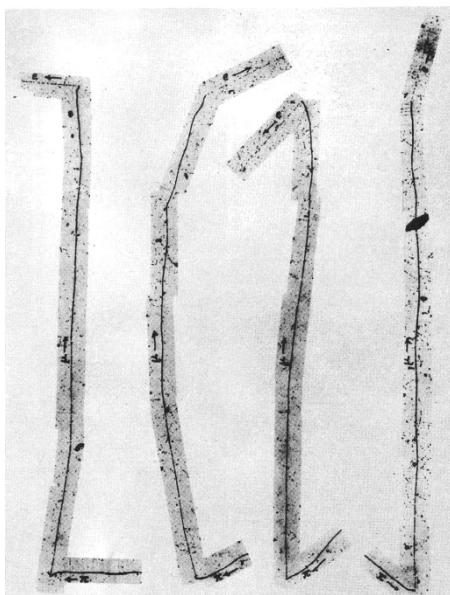
and



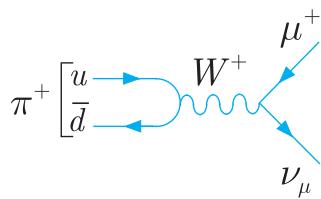
The charged pions decay predominantly by the reactions



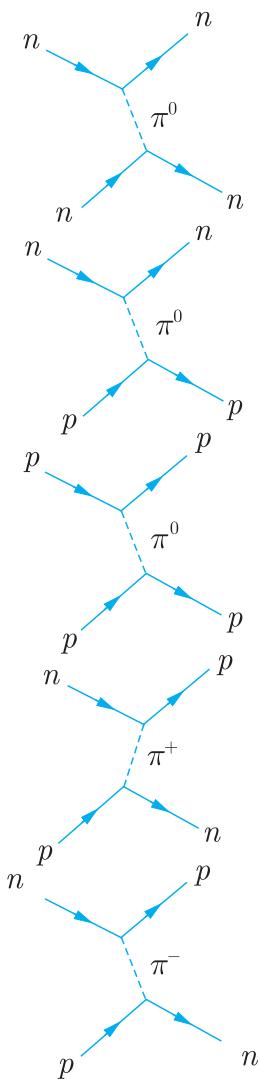
with lifetimes of  $2.6 \times 10^{-8}$  s, typical of weak interactions (cf. Table 3.4). They were first discovered in cosmic rays by a Bristol group in 1947 using photographic emulsions containing a silver halide. The ionisation energy deposited by a particle passing through the emulsion causes the formation of a latent image, and the silver grains resulting from subsequent development form a visual record of the path of the particle. As in a cloud chamber, the range of a stopping particle can be used to estimate its energy. Some historic examples of stopping  $\pi^+$  decays observed in this way are shown in Figure 3.2. Because these are two-body decays, the muons in the final state all have the same energy. Consequently, they all travel approximately the same distance before coming to rest and decaying to give a positron by reaction (3.10). Neutral pions were discovered somewhat later and decay by the electromagnetic interaction



**Figure 3.2** Some early examples of pion decays observed in a photographic emulsion in 1947 by the Bristol group. The positive pions come to rest before decaying by the reaction  $\pi^+ \rightarrow \mu^+ + \nu_\mu$  to give monoenergetic muons. These travel about 600 mm in the emulsion before coming to rest and decaying by the reaction  $\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu$  to give a positron.



**Figure 3.3** Quark diagram for the decay  $\pi^+ \rightarrow \mu^+ + \nu_\mu$ .



**Figure 3.4** The Yukawa model for nuclear forces: direct forces (upper three diagrams); exchange forces (lower two diagrams).

with a lifetime of  $0.8 \times 10^{-16}$  s. Because they are neutral they do not leave tracks and must be detected via their decay photons.

Pions and nucleons are the lightest known mesons and baryons, respectively, and are bound states of the lightest quarks ( $u, d$ ) and their antiquarks ( $\bar{u}, \bar{d}$ ). The only  $3q$  and  $q\bar{q}$  combinations that give the observed electric charges are then

$$p = uud, \quad n = udd \quad (3.12)$$

for the nucleons and

$$\pi^+ = u\bar{d}, \quad \pi^0 = u\bar{u}, d\bar{d}, \quad \pi^- = d\bar{u} \quad (3.13)$$

for the pions, where the ambiguity in the case of the  $\pi^0$  decays will be resolved in Chapter 6. The detailed evidence that this quark picture is correct will also be described in later chapters. For the moment, we shall assume it is, and interpret various reactions in terms of it. The quark interpretation of (3.9a) has already been discussed, following (3.6), and we leave it to the reader to show that the quark numbers  $N_u$  and  $N_d$  are also conserved in (3.9b) and (3.9c). However, in  $\pi^+$  decay (3.10), whose quark interpretation is

$$(u\bar{d}) \rightarrow \mu^+ + \nu_\mu,$$

and is shown in Figure 3.3, a  $u$  quark annihilates with a  $\bar{d}$  antiquark, demonstrating again that individual quark numbers like  $N_u$  and  $N_d$  are not conserved in weak interactions.

Pions play an important role in nuclear forces. In 1935 Yukawa proposed that these were due to the exchange of spin-0 mesons and from the range of the forces (which was not precisely known at that time) predicted that these mesons should have a mass of approximately  $200 \text{ MeV}/c^2$  (cf. the discussion of Section 1.4.2). The discovery of pions was a great triumph for the Yukawa theory. In it, the nuclear forces are given by Figure 3.4, where the nucleons and pions are treated as point particles. Neutral pion exchange gives rise to normal direct forces, while  $\pi^\pm$  exchange gives rise to ‘exchange forces’ whereby the neutron and proton are interchanged. The resulting potential is of the general Yukawa type (1.30), and reproduces the longest range part of the nuclear force  $r \geq 2 \text{ fm}$  very accurately, including spin effects when these are taken into account. At short ranges  $r \leq 1 \text{ fm}$ , comparable with the sizes of the nucleons and pions themselves, this agreement breaks down completely. At these distances, the internal quark structure of the hadrons can no longer be neglected and the situation is much more complicated.

Finally, we note that the atmospheric neutrinos discussed in Section 2.3.2 are produced almost entirely from pion decays. When cosmic ray protons collide with atoms in the upper atmosphere, they create many charged pions in reactions like (3.9). These decay almost entirely by (3.10), followed by muon decay (2.10) to give two muon

neutrinos or antineutrinos for each electron neutrino or antineutrino, as noted in Section 2.3.2.

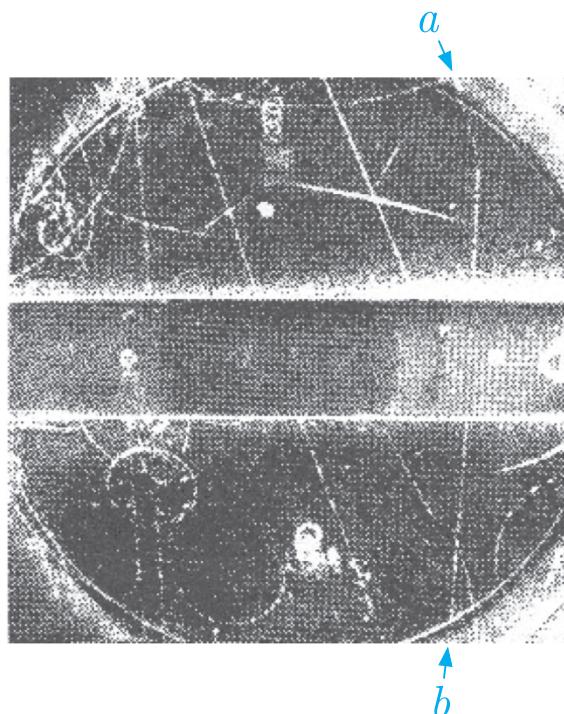
### 3.4 Strange particles, charm and bottom

Soon after the discovery of the pion, members of the Manchester cosmic ray group discovered other hadrons that were produced in strong interactions, but decayed by weak interactions. This was unexpected, as there was apparently no reason why they should not decay by the strong interaction with lifetimes of order  $10^{-23}$  s. For this reason they were christened ‘strange particles’. Figure 3.5 shows one of the first observed events, in which a positively charged particle of mass around  $500 \text{ MeV}/c^2$  decays to a muon and an unseen neutrino. The decaying particle is actually a  $K^+$  meson (or kaon), with mass  $494 \text{ MeV}/c^2$  and a lifetime of  $1.2 \times 10^{-8} \text{ s}$ . Charged kaons have many decay modes, but the principle ones and their branching ratios,  $B$ , are

$$K^+ \rightarrow \mu^+ + \nu_\mu, \quad B = 0.63, \quad (3.14a)$$

as observed in Figure 3.5, and

$$K^+ \rightarrow \pi^+ + \pi^0, \quad B = 0.21. \quad (3.14b)$$



**Figure 3.5** One of the first strange particle events observed in a cloud chamber. A  $K^+$  meson (track *a*) enters the top of the chamber before decaying by the reaction  $K^+ \rightarrow \mu^+ + \nu_\mu$  to give a muon (track *b*) and an unseen neutrino. The dark band across the chamber is a 3 cm thick lead plate, and track *b* is identified by its ability to penetrate this plate with little loss of energy. (Rochester & Butler 1947. Reproduced with permission from Macmillan Publishers Ltd.)

Another example of a strange particle is the  $\Lambda$  (lambda) baryon, which has a mass of  $1116 \text{ MeV}/c^2$  and decays mainly into pions and nucleons

$$\Lambda \rightarrow \pi^- + p, \quad B = 0.64 \quad (3.15a)$$

and

$$\rightarrow \pi^0 + n, \quad B = 0.36, \quad (3.15b)$$

with a lifetime of  $2.6 \times 10^{-10} \text{ s}$ .

It is clear from the long lifetimes of the  $K^+$  and  $\Lambda$  that they both decay via the weak interaction. This strongly suggests that these particles are not made of  $u$  and  $d$  quarks alone, because if this were the case then, for example, the neutral  $\Lambda$  would be a  $(udd)$  state just like the neutron. At the quark level, the decay (3.15a) would then be

$$(udd) \rightarrow (d\bar{u}) + (uud),$$

which conserves the quark numbers  $N_u$  and  $N_d$ . We would therefore expect (3.15a) to be a strong decay, with a lifetime of order  $10^{-23} \text{ s}$ , in contradiction to experiment. The solution is to assign the quark structure  $uds$  to the  $\Lambda$ , so that the decay is

$$\begin{array}{ccc} (uds) & \rightarrow & (d\bar{u}) + (uud) \\ S = -1 & & 0 & 0 \end{array}$$

and neither the quark number  $N_d$  nor the strangeness  $S$  are conserved. As both the strong and electromagnetic interactions conserve all quark numbers, the decay can only proceed by the weak interaction, in which such quark numbers are not conserved.

Strange particles are now defined as any particle with a non-zero value of the strangeness quantum number. Most of them, like most hadrons with  $S = 0$ , decay by the strong interactions. However, conservation of quark numbers in strong and electromagnetic interactions means that if a particle is the lightest state with a given non-zero set of  $B$ ,  $Q$  and  $S$  values,<sup>5</sup> it can only decay by weak interactions, and so will be relatively long-lived. From its quark structure,  $\Lambda$  has  $B = 1$ ,  $Q = 0$  and  $S = -1$ . It is the lightest strange baryon. The lightest strange mesons are the kaons. In addition to the  $K^+$ , which we have already met, there are negative and neutral kaons, which also decay via the weak interaction. Their quark structures are

$$K^+(494) = u\bar{s}, \quad K^0(498) = d\bar{s} \quad (3.16a)$$

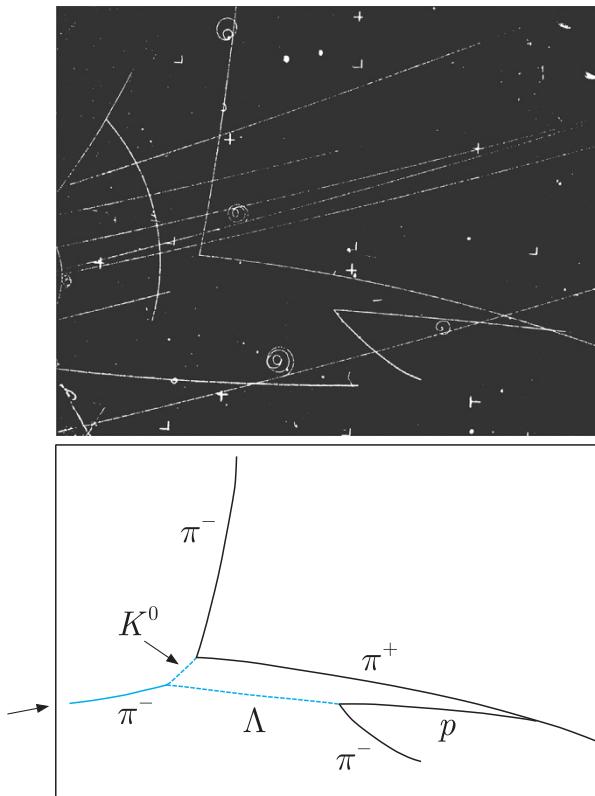
and

$$K^-(494) = s\bar{u}, \quad \bar{K}^0(498) = s\bar{d}, \quad (3.16b)$$

where  $K^+$  and  $K^0$  have  $S = +1$  and  $K^-$  and  $\bar{K}^0$  have  $S = -1$ .

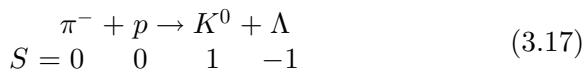
---

<sup>5</sup>The lightest hadron with  $B = Q = S = 0$  is the  $\pi^0$  meson, which decays to photons by the electromagnetic interaction, as we have seen.



**Figure 3.6** A bubble chamber picture of the associated production reaction  $\pi^- + p \rightarrow K^0 + \Lambda$ . The incoming pion is indicated by the arrow, and the unseen neutrals (shown as dotted lines) are detected by their decays  $K^0 \rightarrow \pi^+ + \pi^-$  and  $\Lambda \rightarrow p + \pi^-$ . This picture was taken in the 10 in (25 cm) bubble chamber at the Lawrence Berkeley National Laboratory. (Reproduced with permission from Lawrence Berkeley National Laboratory.)

The production of strange particles in strong interactions is an example of *associated production*. In such processes, more than one strange particle is produced, giving strangeness conservation overall. A beautiful example of such an event is given in Figure 3.6, which shows the reaction



together with the subsequent decays of the kaon and lambda. This picture was taken in a bubble chamber, which is a detection device similar in principle to a cloud chamber, but where the working medium is a liquid rather than a gas. Bubble chambers were invented by Glaser in 1952 and largely superceded cloud chambers in the late 1950s. The liquid is initially held in a vessel at a pressure above that of the equilibrium vapour pressure at that temperature. When the pressure is suddenly reduced by an expansion, the liquid is left in a ‘superheated’ state, where its temperature is above the boiling point at the reduced pressure. The passage of a charged particle through the liquid produces ion pairs and bubbles form preferentially along the ionisation trails. The liquid in the chamber (hydrogen in the case of Figure 3.6) is both target and detector, and by placing the whole

chamber in a magnetic field the particles' momenta can be obtained. The curvatures of the tracks are clearly visible in Figure 3.6.

In the thirty years following the discovery of the pions and kaons, a great many hadrons were discovered. Until 1974 all could be accounted for as bound states of just the three quarks,  $u$ ,  $d$  and  $s$ , originally proposed by Gell-Mann and Zweig. However, in that year a relatively heavy particle was discovered in two independent experiments, one at the Brookhaven National Laboratory (BNL) and the other at the Stanford Linear Accelerator Laboratory (SLAC). The BNL group, led by Ting, named this new particle  $J$ , while the SLAC group, led by Richter, chose  $\psi$ . It is now known as  $J/\psi$ , and its properties show that it is one of the lightest of a family of particles that are bound states of a charmed quark and its antiparticle, that is

$$J/\psi(3097) = c\bar{c} \quad (C = 0). \quad (3.18)$$

Such  $c\bar{c}$  states are collectively known as *charmonium* by analogy with positronium, the bound state of an electron and positron that we will discuss in Section 5.5. Since  $C = 0$ , these states are often said to contain ‘hidden charm’.

Particles with ‘naked charm’  $C \neq 0$  were also discovered at SLAC shortly after the discovery of the  $J/\psi$ . Because charm is a quark number, like strangeness, it should be conserved in strong and electromagnetic interactions, and the lightest charmed particles should decay by weak interactions. This is indeed the case. For example, the lightest charmed mesons are the  $D$  mesons with quark structures

$$D^+(1870) = c\bar{d}, \quad D^0(1865) = c\bar{u} \quad (C = +1) \quad (3.19a)$$

and

$$D^-(1870) = d\bar{c}, \quad \bar{D}^0(1865) = u\bar{c} \quad (C = -1) \quad (3.19b)$$

and  $D_s$  mesons with quark structures

$$D_s^+(1968) = c\bar{s} \quad (C = +1, S = +1) \quad (3.20a)$$

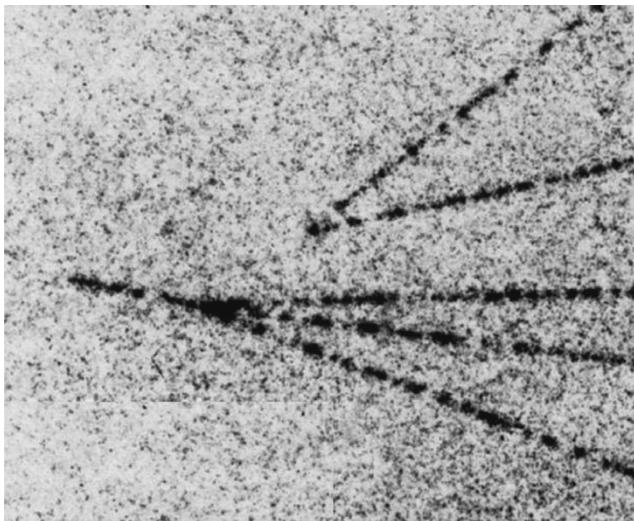
and

$$D_s^-(1968) = s\bar{c} \quad (C = -1, S = -1), \quad (3.20b)$$

while the lightest charmed baryon is

$$\Lambda_c^+(2286) = udc \quad (C = +1). \quad (3.21)$$

These particles all have lifetimes of order  $10^{-13}$  s, which is in the expected range for weak decays. Charmed particles can be produced in strong and electromagnetic interactions by associated production reactions, just like strange particles, and a bubble chamber picture of such an event is shown in Figure 3.7. However, because the charmed particles have much shorter lifetimes than the strange particles (3.15) and (3.16), they travel much shorter distances before decaying, and very good spatial resolution is needed to observe their tracks. This was achieved in the present case by only allowing the bubbles to



**Figure 3.7** An example of the associated production of two charmed particles observed in a high-resolution bubble chamber at SLAC (Abe *et al.*, 1982). An incident unseen photon enters from the left and interacts with a proton in the liquid hydrogen of the chamber to produce a charged, charmed particle, which travels about 2 mm before decaying to give three charged particles, and an unseen, neutral charmed particle, which travels about 3 mm before decaying to give two charged particles. (Abe *et al.* 1982. Reproduced with permission from the American Physical Society.)

develop to a size of about 0.05 mm before photographing them, compared to sizes of about 0.5 mm used in a typical bubble chamber.

Historically, the discovery of strange particles caused great excitement because they clearly represented a new form of matter that was completely unexpected at the time. The discovery of charmed particles caused equally great excitement because their existence was expected, having been predicted from the then newly formulated theory of electroweak interactions. Their discovery was a decisive event in confirming the essential correctness of this theory, which is a unified theory of both weak and electromagnetic interactions. In its present form it requires that the number of leptons and quarks should be the same, implying that there should be six quarks to match the six known leptons. Evidence for the fifth quark – the bottom quark  $b$  with its associated quantum number  $\tilde{B}$  – came from the discovery in 1977 of one of the lightest *bottomonium* states

$$\Upsilon(9460) = b\bar{b} \quad (\tilde{B} = 0), \quad (3.22)$$

which is a hidden bottom state called the upsilon. Subsequently mesons with ‘naked bottom’  $\tilde{B} \neq 0$  were discovered, the lightest examples of which are the  $B$ -mesons

$$B^+(5279) = u\bar{b}, \quad B^0(5280) = d\bar{b} \quad (\tilde{B} = +1) \quad (3.23a)$$

and

$$B^-(5279) = b\bar{u}, \quad \bar{B}^0(5280) = b\bar{d} \quad (\tilde{B} = -1) \quad (3.23b)$$

with lifetimes of order  $10^{-12}$  s, consistent with their decay via the weak interaction. Examples of baryons with  $\tilde{B} \neq 0$ , including

$$\Lambda_b^0(5620) = udb \quad (3.24)$$

and the doublet

$$\Xi_b^0(5793) = usb, \quad \Xi_b^-(5795) = dsb, \quad (3.25)$$

which decay via the weak interaction, with lifetimes of order  $10^{-12}$  s, have also been discovered. Research in the field of particle spectroscopy is ongoing. For example, recent experiments have found evidence for the first examples of a ‘charmed-strange’ and a ‘bottom-strange’ baryon, that is ones containing a strange quark in combination with a charmed or bottom quark, respectively.

### 3.5 Short-lived hadrons

The early history of hadron physics was dominated by cosmic ray studies, in which new particles, such as pions and kaons, were produced in collisions between cosmic ray protons and atoms in the upper atmosphere. Such processes are random and hard to observe, making detailed studies of hadronic interactions extremely difficult. However, by the 1960s the situation had been transformed by the creation of high-energy beams of pions and kaons in the laboratory, leading to the discovery of large numbers of hadrons that decayed by the strong interaction, with lifetimes of order  $10^{-23}$  s. This profusion of states was the ‘last nail in the coffin’ for the idea that hadrons were in any sense fundamental and led rapidly to the development of the quark model to explain their properties. In this section we will firstly explain how such states are detected and then discuss their interpretation within the quark model.

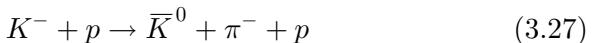
Hadrons that decay by strong interactions are called *resonances*. They are far too short-lived to be directly observed, even by electronic methods, and their existence must be inferred from observations on the more stable hadrons to which they decay. A typical example is the production of a boson resonance  $X^-$  in the reaction



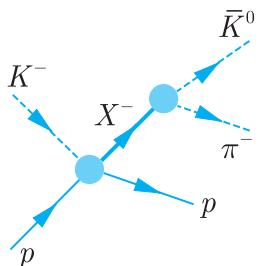
and its subsequent decay by the reaction



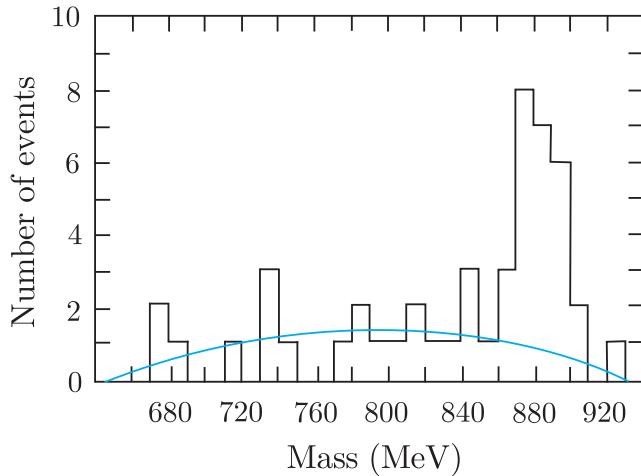
where  $\bar{K}^0(498)$  is the neutral boson with  $S = -1$  given in (3.16b). This is illustrated in Figure 3.8, where the distance between the points at which the resonance is produced and decays is too small to be measured (see Problem 3.10). The observed reaction is therefore



and the fleeting existence of the  $X^-$  must be inferred from measurements of the energies  $E_K$  and  $E_\pi$  and the momenta  $\mathbf{p}_K$  and  $\mathbf{p}_\pi$  of the  $\bar{K}^0$  and  $\pi^-$  mesons to which it decays. If the decaying particle



**Figure 3.8** Production of a negatively charged resonance  $X^-$  in the reaction  $K^- + p \rightarrow X^- + p$  and its subsequent decay,  $X^- \rightarrow \bar{K}^0 + \pi^-$ .



**Figure 3.9** Invariant mass distribution of the  $\bar{K}^0\pi^-$  pairs in the reaction  $K^- + p \rightarrow \bar{K}^0 + \pi^- + p$  for an initial beam momentum of 1.15 GeV/c. (Alston *et al.* 1961. Reproduced with permission from the American Physical Society.)

has mass  $M$ , energy  $E$  and momentum  $\mathbf{p}$ , then by energy–momentum conservation the invariant mass<sup>6</sup>  $W$  of the  $\bar{K}^0\pi^-$  pair is given by

$$W^2 = (E_K + E_\pi)^2 - (\mathbf{p}_K + \mathbf{p}_\pi)^2 = E^2 - \mathbf{p}^2 = M^2. \quad (3.28)$$

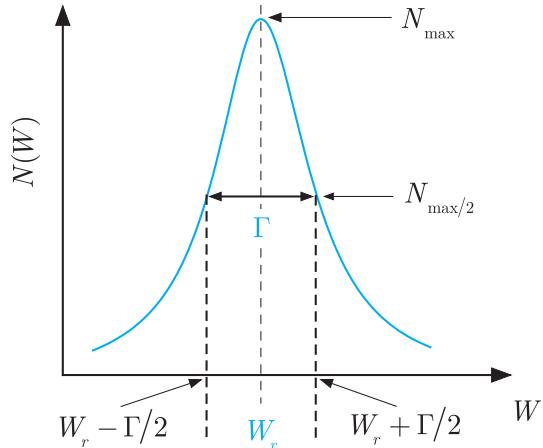
Thus if (3.27) proceeds by the mechanism of Figure 3.8, a plot of the invariant mass distribution will show a sharp peak at the resonance mass  $M$ . On the other hand, if uncorrelated  $\bar{K}^0$  and  $\pi^-$  particles are produced by some other mechanism, a smooth distribution would be expected.

Figure 3.9 shows a histogram of the invariant mass distribution obtained in an early experiment to detect a meson resonance. In this experiment, a 1.15 GeV/c  $K^-$  beam was directed into a hydrogen bubble chamber and 47 events corresponding to the reaction (3.27) were identified. Despite the limited statistics, a clear peak is seen with a mass corresponding to a resonance called the  $K^{*-}(890)$ . The peak sits on a background (shown by the solid line) arising from uncorrelated pairs produced by some other, non-resonant, mechanism, and has a width of about 50 MeV. Widths of this order are characteristic of particles that decay by the strong interaction. For a particle at rest,  $W = E$  by (3.28) and the energy–time uncertainty principle leads to a natural line width

$$\Delta W = \Delta E \approx \Gamma \equiv 1/\tau, \quad (3.29)$$

where  $\tau$  is the mean lifetime at rest. The constant  $\Gamma$  is called the *decay width* (or just the *width*) of the state. If  $\tau$  is of order  $10^{-23}$  s, as expected for a strong decay, then  $\Gamma = O(100 \text{ MeV})$ , which is similar to the width of the resonance peak in Figure 3.9.

<sup>6</sup> See Appendix A, Section A.2.



**Figure 3.10** Plot of the Breit–Wigner formula (3.30a).

The shape of an isolated resonance peak is conveniently approximated by the *Breit–Wigner formula*

$$N(W) = \frac{K}{(W - W_r)^2 + \Gamma^2/4}, \quad (3.30a)$$

derived in Appendix B, Section B.5.<sup>7</sup> Here  $K$  is a constant that depends on the total number of decays observed and  $W_r$  is the position of the maximum, as shown in Figure 3.10.

Since the invariant mass  $W$  is identical to the rest mass for a single particle,<sup>8</sup> this can be regarded as a mass distribution corresponding to a mean value  $W_r$  and an uncertainty  $\Gamma/2$ . In practice, it is more usual to redefine the mass  $M$  as being equal to  $W_r$ , so that the Breit–Wigner formula becomes (cf. Equation (B.41))

$$N(W) = \frac{K}{(W - M)^2 + \Gamma^2/4}. \quad (3.30b)$$

<sup>7</sup> This formula is closely analogous to that used to describe the natural line width of an excited state of an atom, which is an unstable particle made of a nucleus and electrons, rather than quarks and antiquarks. In this context, we note that (3.30a) becomes

$$N(E) = \frac{K}{(E - E_r)^2 + \Gamma^2/4}$$

in the rest frame of the decaying particle, and putting  $E = \omega$  (remember  $\hbar = 1$  here) and  $\Gamma = 1/\tau$  we obtain

$$N(\omega) = \frac{K}{(\omega - \omega_r)^2 + 1/4\tau^2}.$$

This is just the Lorentzian line shape used in atomic spectroscopy. (A simple semiclassical discussion of its origin in this context is given in Sections 3.14–3.26 of Wichmann (1967)).

<sup>8</sup> See Appendix A, Equation (A.10b).

The masses and widths of resonances are determined by fitting this formula to the experimentally observed peaks. In doing this, careful account must be taken of non-resonant backgrounds. When this is done for the  $K^{*-}$ , using later data with much higher statistics than those shown in Figure 3.10, a mass and width of  $892 \text{ MeV}/c^2$  and  $51 \text{ MeV}$ , respectively, are obtained. A similar resonance is also observed in the reaction sequence

$$K^- + p \rightarrow \bar{K}^{*0} + n \quad (3.31\text{a})$$

and

$$\bar{K}^{*0} \rightarrow K^- + \pi^+ \quad (3.31\text{b})$$

and found to have mass and width values of  $896 \text{ MeV}/c^2$  and  $51 \text{ MeV}$ , respectively, which are almost identical to those of the  $K^{*-}$ .

Resonances are characterised not only by their masses and widths, but also by their various quantum numbers. These are determined from observations on their decay products by exploiting the appropriate conservation laws. In particular, their spins can be determined from observations on the angular distributions of their decay products, using angular momentum conservation. We shall not discuss this here<sup>9</sup> but merely quote the result that both the  $K^{*-}(896)$  and the  $\bar{K}^{*0}(892)$  have spin 1, like the photon. Their internal quantum numbers are obtained much more easily from (3.26) and (3.31), using the known values for the other particles given in Tables 3.2 and 3.3. In this way, one obtains

$$B = 0, \quad S = -1, \quad C = \tilde{B} = T = 0 \quad (3.32)$$

for both resonances. Within the quark model, the only allowed combinations compatible with (3.32) are  $s\bar{u}$  and  $s\bar{d}$ , with charges  $Q = -1$  and  $0$ , respectively. We thus arrive at the unique quark assignments

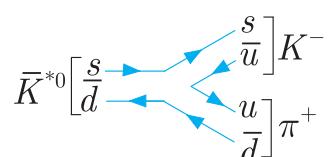
$$K^{*-}(896) = s\bar{u} \quad \text{and} \quad \bar{K}^{*0}(892) = s\bar{d}, \quad S = -1, \quad (3.33)$$

so that the quark description of the strong decay (3.31b) is

$$(s\bar{d}) \rightarrow (s\bar{u}) + (u\bar{d}),$$

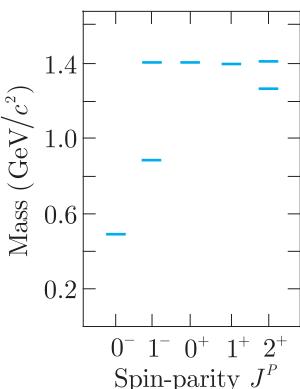
involving the creation of a  $u\bar{u}$  pair. This is summarised by the quark diagram of Figure 3.11. The continuity of quark lines is a general feature of quark diagrams for strong decays and reflects the fact that quarks and antiquarks can only be created or destroyed as particle-antiparticle pairs in strong interactions. A similar quark diagram, involving the creation of a  $d\bar{d}$  pair, describes the decay (3.26b) (where  $X^- \equiv K^{*-}$ ) and is left as an exercise for the reader.

At this point we note that the quark assignments (3.33) are identical to those of the kaons given in (3.16). Subsequently, many

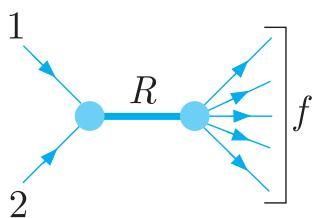


**Figure 3.11** Quark diagram for the decay  $\bar{K}^{*0} \rightarrow K^- + \pi^+$ .

<sup>9</sup> This topic is discussed, for example, in Chapter 9 of Hughes (1985).



**Figure 3.12** Observed bound states of the  $s\bar{u}$  system with masses below  $1.5 \text{ GeV}/c^2$ , together with values of their spin-parities<sup>10</sup>  $J^P$ . The ground state is the  $K^-$ (494) and the others can be interpreted as its excited states.



**Figure 3.13** Formation and decay of a resonance  $R$  in the reaction  $1 + 2 \rightarrow f$ .

other states with these quantum numbers have been discovered and Figure 3.12 shows the spectrum of states corresponding to the  $s\bar{u}$  system observed to date. The lightest state (the ground state) is the long-lived  $K^-$  meson discussed in Section 3.4, which decays by the weak interaction processes  $K^- \rightarrow \mu^- + \bar{\nu}_\mu, K^- \rightarrow \pi^- + \pi^0$ . The heavier states (excited states) are resonances that decay by the strong interaction with widths typically of order 50–250 MeV.

This picture is not restricted to strange mesons, but applies qualitatively to all quark systems,  $u\bar{d}, u\bar{u}d, u\bar{d}s$ , etc. Each system has a ground state, which is usually a long-lived particle decaying by weak or electromagnetic interactions, and a number of excited (resonance) states. The resulting spectra are qualitatively similar to those shown in Figure 3.12 and the analogy with energy-level diagrams of other composite systems, like atoms and nuclei, is obvious.

Finally, we note that the factor  $K$  in the Breit–Wigner formula (3.30b) depends on the mechanism by which the resonance is produced and is often very difficult to calculate. However, an exception is the class of *resonance formation* reactions

$$1 + 2 \rightarrow R \rightarrow f \quad (3.34a)$$

corresponding to the mechanism of Figure 3.13, in which two particles 1, 2 collide to form a resonance  $R$ , which subsequently decays to a final state  $f$ . In this case, if the initial particles are unpolarised, the cross-section for reaction (3.34a) is given by<sup>11</sup> (cf. Section B.5.2 and Equation (B.44b))

$$\sigma_{if} = \frac{\pi}{q_i^2} \frac{(2S_R + 1)}{(2S_1 + 1)(2S_2 + 1)} \frac{\Gamma_i \Gamma_f}{(W - M)^2 + \Gamma^2/4}, \quad (3.34b)$$

where  $S_1, S_2$  and  $S_R$  are the spins of particles 1, 2 and  $R$ , respectively;  $\Gamma$  is the total decay width;  $\Gamma_i$  and  $\Gamma_f$  are the partial decay widths to the initial state  $i = (1 + 2)$  and the final state  $f$ , respectively; and  $q_i = |\mathbf{q}_i|$  is the magnitude of the centre-of-mass momentum of the particles in the initial state.<sup>12</sup> This formula is often written in the alternative form

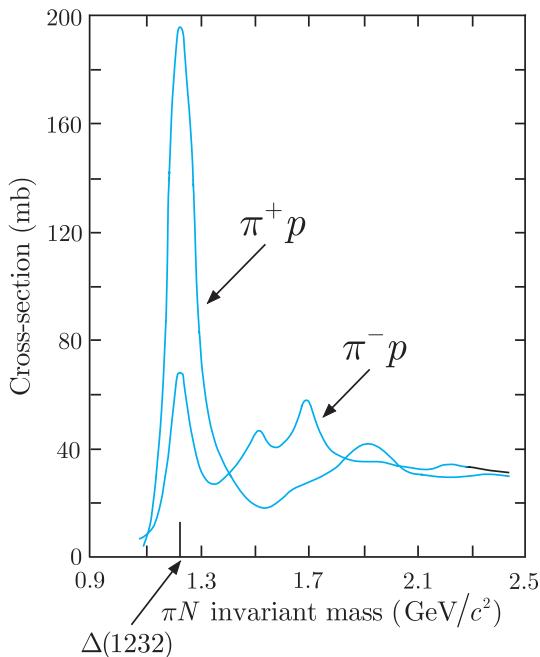
$$\sigma_{if} = \frac{\pi}{q_i^2} \frac{(2S_R + 1)}{(2S_1 + 1)(2S_2 + 1)} \frac{4M^2 \Gamma_i \Gamma_f}{(W^2 - M^2)^2 + M^2 \Gamma^2}, \quad (3.34c)$$

which is equivalent to (3.34b) in the vicinity of the resonance provided  $\Gamma \ll M$ .

<sup>10</sup>Parity is discussed in Chapter 5, but is included here for later convenience.

<sup>11</sup>Reaction cross-sections are formally defined in Sections B.1 and B.2, and partial widths in B.5.

<sup>12</sup>Although we have expressed the cross-section in terms of the centre-of-mass momentum and the invariant mass  $W$ , the resulting values also hold in the laboratory frame, since reaction cross-sections are invariant under a Lorentz transformation along the beam direction.



**Figure 3.14** Total cross-sections for  $\pi^\pm p$  scattering.

These formulas have many applications, as we shall see in later chapters, but they are only a good approximation when a single isolated resonance dominates the cross-section. In many cases, non-resonant scattering also occurs and other terms have to be added to the amplitude to extract reliable resonance parameters. These terms may be calculable from known non-resonant contributions to the reaction, or included in an *ad hoc* way. In either case they will vary with energy less rapidly than (3.34b) throughout the resonance region and are referred to as ‘background’. These points are illustrated by the fits to data on the total cross-sections for  $\pi^\pm p$  scattering shown in Figure 3.14, obtained by summing over all final states  $f$ . At invariant masses of about 1230  $\text{MeV}/c^2$ , the cross-sections are both dominated by a single Breit–Wigner resonance corresponding to a mass of 1232  $\text{MeV}/c^2$  and a width of approximately 120 MeV. These resonances are denoted  $\Delta^{++}(1232)$  and  $\Delta^0(1232)$  in the  $\pi^+ p$  and  $\pi^- p$  cases, respectively, and one easily deduces that they have the quark compositions

$$\Delta^{++} = uuu \quad \text{and} \quad \Delta^0 = udd.$$

However, at higher invariant masses, several other broad peaks are observed, on a smooth background that is slowly varying with energy. These peaks usually contain more than a single resonance and a more detailed analysis, using angular momentum distributions and polarised scattering data as well as total cross-section data, is required to obtain reliable resonance parameters. Here we merely

note that when this is done, a rich spectrum of states is revealed, corresponding to the excited states of the  $uuu$  and  $udd$  systems, in analogy to those shown in Figure 3.12 for the  $s\bar{u}$  system.

### 3.6 Allowed and exotic quantum numbers

So far we have emphasised the quark interpretation of the internal quantum numbers

$$Q, B, C, S \quad \text{and} \quad \tilde{B}, \quad (3.35)$$

but when a hadron is discovered the identities of its quark constituents are not known. Here we will show how the quantum numbers (3.35) can be determined for hadrons without assuming anything about their quark contents and how the resulting values can be used to test the simple quark model assumption that mesons, baryons and antibaryons have the quark compositions

$$q\bar{q}, \quad qqq \quad \text{and} \quad \bar{q}\bar{q}\bar{q} \quad (3.36)$$

respectively, rather than more complicated possibilities like

$$qq\bar{q}\bar{q} \quad \text{and} \quad qqqq\bar{q}. \quad (3.37)$$

We first consider the allowed values of charge and strangeness corresponding to the usual quark compositions (3.36), starting with hadrons for which<sup>13</sup>

$$C = \tilde{B} = 0. \quad (3.38)$$

For baryons, the possible  $qqq$  combinations are then

$$sss, \quad ssq_i, \quad sq_iq_j, \quad q_iq_jq_k, \quad (3.39)$$

where  $s$  is the strange quark and  $q_i = u, d$ , because of (3.38). The allowed strangeness values are therefore  $S = -3, -2, -1$  and  $0$ , and the allowed charges are obtained by adding the quark charges. Thus a baryon  $sss$  with  $S = -3$  must have charge  $Q = -1$ , while a baryon

$$ssq_j \equiv ssu, ssd$$

with  $S = -2$  can have  $Q = 0$  or  $Q = -1$ . The remaining possibilities are

$$sq_iq_j \equiv suu, sud, sdd,$$

with  $S = -1$  and  $Q = 1, 0$  or  $-1$ , and

$$q_iq_jq_k \equiv uuu, uud, udd, ddd,$$

---

<sup>13</sup>The known particles all have top  $T = 0$ , as noted in Section 3.1, and we ignore top in what follows.

**Table 3.5** The combinations of charge and strangeness allowed for hadrons in the simple quark model for states with  $C = \tilde{B} = 0$ .

(a) Baryons		(b) Mesons	
$S$	$Q$	$S$	$Q$
0	2, 1, 0, -1	1	1, 0
-1	1, 0, -1	0	1, 0, -1
-2	0, -1	-1	0, -1
-3	-1		

with  $S = 0$  and  $Q = 2, 1, 0$  or  $-1$ . These values are collected together in Table 3.5(a), and the corresponding values for meson combinations  $q\bar{q}$ , left as an exercise for the reader, are given in Table 3.5(b). For hadrons with  $C = \tilde{B} = 0$ , combinations of  $S$  and  $Q$  that do not occur in these tables are called *exotic*, since they can arise only from quark combinations like (3.37) that are not included in the simple quark model. Examples of hadrons with exotic quantum numbers<sup>14</sup> would be a meson  $d\bar{d}\bar{s}\bar{s}$  with  $S = 1$  and  $Q = -1$  or a baryon  $u\bar{u}u\bar{d}\bar{s}$  with  $S = 1$  and  $Q = 2$ . Furthermore, the discussion is easily extended to include non-zero charm and bottom. For example, if we consider hadrons in which both  $C$  and  $\tilde{B}$  are non-zero, the only possible meson configurations are  $c\bar{b}$  with  $Q = C = \tilde{B} = 1$  and  $S = 0$ , and their antiparticles  $b\bar{c}$ , with  $Q = C = \tilde{B} = -1$  and  $S = 0$ . All other combinations are exotic.

From this we see that, in the simple quark model, the internal quantum numbers of hadrons are restricted to relatively few allowed combinations, while all other combinations – the vast majority – are exotic and should not be observed. This is a highly non-trivial prediction, since, when a new hadron is identified, its quark content is unknown, but despite this, its internal quantum numbers can easily be measured. This is done by starting from the values of the quantum numbers  $B, S, C$  and  $\tilde{B}$  for the particles<sup>15</sup>

$$p(938), n(940), K^-(494), D^-(1870), B^-(5279) \quad (3.40)$$

shown in Tables 3.2 and 3.3. Given these values, those for all other hadrons can be determined by exploiting the associated conservation laws. For example, by applying strangeness conservation to the

<sup>14</sup>It is important to distinguish between the terms ‘hadrons with exotic quantum numbers’ and the term ‘exotic hadrons’, which may already be familiar to some readers. The latter term is not introduced in this book until Section 7.1.3 and in this chapter the word ‘exotic’ is only used with reference to quantum numbers.

<sup>15</sup>This choice of reference hadrons is not unique, but is convenient.

observed strong interaction processes

$$p + p \rightarrow p + n + \pi^+, \quad (3.41a)$$

$$p + p \rightarrow p + p + \pi^0, \quad (3.41b)$$

$$\pi^- + p \rightarrow \pi^0 + n, \quad (3.41c)$$

$$K^- + p \rightarrow \pi^0 + \Lambda \quad (3.41d)$$

and

$$\pi^- + p \rightarrow K^+ + \pi^- + \Lambda \quad (3.41e)$$

and assuming the strangeness values of Table 3.2, we find that the pions have strangeness  $S = 0$ , the  $K^+$  has  $S = 1$  and the  $\Lambda$  has  $S = -1$  without using their quark composition. Similar arguments apply to the other quantum numbers and enable us to construct a catalogue of their values for all known hadrons by considering a sufficient number of observed reactions like (3.41).

At present, well over two hundred hadrons with internal quantum numbers compatible with the expectations of the simple quark model have been clearly established. In addition, from the 1960s onwards there have been numerous searches for hadrons with exotic internal quantum numbers, that is internal quantum numbers that are incompatible with the simple quark model and could only arise from more complicated quark structures, like the four and five quark states given in (3.37) for mesons and baryons. From time to time, experiments have claimed evidence for states with exotic internal quantum numbers, but as more experiments were performed and far greater quantities of data produced, the evidence for their existence receded, and by 2015 the consensus was that there was no convincing evidence for the existence of states with exotic internal quantum numbers. However, in 2016 a new candidate emerged. Specifically, the D0 collaboration measured the invariant mass distribution of  $B_s^0\pi^\pm$  pairs produced in proton–antiproton collisions at the Tevatron, where  $B_s^0(5367) = sb$  is a long-lived pseudoscalar meson with a lifetime of about  $1.5 \times 10^{-12}$  s. The results (Abasov *et al.*, 2016) show a clear peak at an invariant mass of  $5568 \text{ MeV}/c^2$ , with a width of about 20 MeV, which they interpreted as resulting from the strong interaction decay

$$X^\pm \rightarrow B_s^0 + \pi^\pm \quad (3.42a)$$

of an unknown meson  $X^\pm(5586)$ , with the quantum numbers

$$Q = \pm 1, B = 0, S = -1, \tilde{B} = +1. \quad (3.42b)$$

These quantum numbers are exotic, with the simplest possible tetraquark assignments  $s\bar{b}u\bar{d}$  and  $s\bar{b}d\bar{u}$  for the positively and negatively charged states, respectively. On the other hand, the LHCb Collaboration (2016) has subsequently studied the invariant mass distribution of  $B_s^0\pi^\pm$  pairs produced in proton–proton collisions at the

LHC, with a view to confirming the above result, but failed to see any  $X^\pm(5568)$  signal, despite very high statistics. While these results do not directly contradict each other, since they are obtained using different accelerator beams at different energies, the contrast is surprising, and it is clear that further experimental work is required before the  $X^\pm(5568)$  can be regarded as a well-established state.

To summarise, it is clear that the quantum numbers of the overwhelming majority of known hadrons conform to the values required by the simple quark model, and while the proposed  $X^\pm(5568)$  state discussed above remains a possibility, not a single example of a hadron with exotic internal quantum numbers can yet be regarded as firmly established. Nonetheless, multi-quark states like (3.37) do not necessarily give rise to exotic quantum numbers, and  $q\bar{q}$  and  $3q$  states are not the whole story, as we shall see in Sections 6.4.1 and 7.1.3.

## Problems 3

**3.1** Which of the following reactions are allowed and which are forbidden by the conservation laws appropriate to weak interactions?

- (a)  $\Lambda \rightarrow \pi^+ + e^- + \bar{\nu}_e$ ;      (b)  $K^+ \rightarrow \pi^+ + e^- + \bar{\nu}_e$ ;
- (c)  $K^+ \rightarrow \pi^0 + \mu^+ + \bar{\nu}_\mu$ ;    (d)  $K^- \rightarrow \pi^- + e^+ + e^-$ .

**3.2** Classify the following experimentally observed processes into strong, electromagnetic and weak reactions by considering the particles involved and the appropriate selection rules.

- (a)  $\pi^- \rightarrow \pi^0 + e^- + \bar{\nu}_e$       (b)  $\gamma + p \rightarrow \pi^+ + n$ ;
- (c)  $p + \bar{p} \rightarrow \pi^+ + \pi^- + \pi^0$ ;    (d)  $D^- \rightarrow K^+ + \pi^- + \pi^-$ ;
- (e)  $\Lambda + p \rightarrow K^- + p + p$ ;      (f)  $\pi^- + p \rightarrow n + e^+ + e^-$ .

**3.3** A large number of stopped  $K^+(494)$  mesons are observed to decay to the  $\pi^+\pi^0$  final state. Calculate the mean distance travelled by the  $\pi^+$  mesons before they decay, in the approximations that both pions have a mass of  $140 \text{ MeV}/c^2$ . The mean lifetime at rest of the  $\pi^+$  is  $\tau_0 = 2.6 \times 10^{-8} \text{ s}$ .

**3.4** The particles  $X^0(1193)$  and  $Y^-(1321)$  can be produced in the strong interaction processes  $K^- + p \rightarrow \pi^0 + X^0$  and  $K^- + p \rightarrow K^+ + Y^-$  respectively. Deduce the baryon number, strangeness, charm and bottom quantum numbers in each case and using these give their quark content. The  $Y^-(1321)$  decays by the reaction  $Y^- \rightarrow \Lambda + \pi^-$ . Use the ‘typical values’ of Table 3.5 to give a rough estimate of its lifetime.

**3.5** Six observed hadrons have the quantum numbers  $(Q, B, S, C, \tilde{B}) = (2, 1, 0, 1, 0), (0, 1, -2, 1, 0), (0, 0, 1, 0, -1), (0, -1, 1, 0, 0), (0, 1, -1, 1, 0)$  and  $(-1, 1, -3, 0, 0)$ , respectively. Identify their quark constituents.

- 3.6** The  $\Sigma^+(1189)$ ,  $\Sigma^0(1193)$  and  $\Sigma^-(1197)$  baryons have the quark compositions  $uus$ ,  $uds$  and  $dds$ , respectively. Which of the following reactions are allowed and which are forbidden in strong interactions?
- (a)  $\pi^- + p \rightarrow K^- + \Sigma^+$ ; (b)  $\pi^- + p \rightarrow K^0 + \Sigma^0$ ;  
(c)  $\pi^+ + p \rightarrow K^0 + \Sigma^+$ ; (d)  $\pi^+ + p \rightarrow K^+ + \pi^+ + \Sigma^0$ .
- 3.7** Apart from the  $p$ ,  $n$  and  $\Lambda$  described in the text, the only baryon lighter than the  $\Sigma^0(1193) = uds$  is its charged companion  $\Sigma^+(1189) = uus$ . Use this fact to identify the dominant decay mode of the  $\Sigma^0(1193)$  and make a rough estimate of its lifetime.
- 3.8** The  $\Sigma^+(1189) = uus$  decays almost entirely to states composed solely of hadrons. Identify these states and make a rough estimate of the particle lifetime.
- 3.9** Draw a lowest-order Feynman diagram at the quark level for each of the following decays:
- (a)  $D^- \rightarrow K^0 + \pi^-$ ; (b)  $\Lambda \rightarrow p + e^- + \bar{\nu}_e$ .
- 3.10** Estimate the mean distance travelled by a resonance produced in a reaction like (3.26a) with  $\gamma = E/m \approx 10$ , before it decays by a process like (3.26b). Compare this with the shortest distance that can be resolved experimentally (of order 1 mm) and with the range of the strong interactions. Why is this latter comparison significant?
- 3.11** Draw the quark diagram for: (a) the decay  $\Delta^{++} \rightarrow \pi^+ + p$  of the  $\Delta(1232)$  resonance discussed in Section 3.5 and (b) the two corresponding diagrams for the process  $\pi^- p \rightarrow \Delta^0 \rightarrow \pi^0 n$ .
- 3.12** Derive the allowed combinations of charm  $C$  and electric charge  $Q$  for mesons and baryons in the simple quark model.
- 3.13** Which of the following possible hadronic states with quantum numbers  $(Q, B, S, C, \tilde{B}) = (1, 0, 0, 1, 1)$ ,  $(-1, 1, -2, 0, -1)$ ,  $(0, 0, 1, 0, 1)$  and  $(-1, 1, 0, 1, -1)$  are compatible with the simple quark model and which are not?

# 4

---

## Experimental methods

---

It may be so, there is no arguing against facts and experiment.

*Isaac Newton<sup>1</sup>*

In this chapter we will take a brief look at experimental methods. This is a very extensive subject and the aim will not be to give a comprehensive review, but rather to emphasise the physical principles behind the methods. More details may be found in specialised texts.<sup>2</sup>

### 4.1 Overview

In particle physics, high energies are needed both to create new and unstable particles and to explore the structure of hadrons. Until the early 1950s the only source of such high-energy particles was cosmic rays, and studies of them led to many notable discoveries, including positrons and pions. However, most subsequent experiments have been conducted using beams of particles produced by machines called *accelerators*. This has the advantage that interactions take place at high frequency at a known position where the detector can be located. Another big advantage is that the projectiles are of a single type and have energies that may be controlled by the experimenter. For example, beams that are essentially monoenergetic

---

<sup>1</sup> In reply to a friend on being told of observations that were thought – wrongly – to contradict the Newtonian system. See p. 59 of Andrade (1961).

<sup>2</sup> See, for example, Kleinknecht (1998), Knoll (1999), Green (2000) and Grupen and Schwartz (2011). There are also reviews in Sections 33 and 34 of Olive *et al.* (Particle Data Group) (2014).

may be prepared, and can be used to study the energy dependence of interactions.

The beam in an accelerator, once established, is directed on to a target so that interactions may be produced. In a *fixed-target* experiment the target is stationary in the laboratory. Although historically of great importance, the main disadvantage of this type of experiment is apparent with the need to work at higher centre-of-mass energies. The centre-of-mass energy is important because it is a measure of the energy available to create new particles. In the laboratory frame at least some of the final-state particles must be in motion to conserve momentum. Consequently, at least some of the initial beam energy must reappear as kinetic energy of the final-state particles, and is therefore unavailable for particle production. The centre-of-mass energy  $E_{CM}$  is given in terms of the projectile's laboratory energy  $E_L$  by<sup>3</sup>

$$E_{CM} = \left( m_b^2 c^4 + m_t^2 c^4 + 2m_t c^2 E_L \right)^{1/2}, \quad (4.1a)$$

where  $m_b$  and  $m_t$  are the masses of the beam and target particles, respectively. At high energies this increases only like  $E_L^{1/2}$  and so a decreasing fraction of the beam energy is available for particle production, the rest going to impart kinetic energy to the target.

In contrast, if we could work in the centre-of-mass system, all the available energy would in principle be available for particle production. This is achieved in a *colliding-beam* experiment, where two beams of particles travelling in almost opposite directions are made to collide at a small, or zero, crossing angle. If for simplicity we assume the particles in the two beams have the same mass and laboratory energy and collide at a zero crossing angle, then the total centre-of-mass energy is

$$E_{CM} = 2E_L. \quad (4.1b)$$

This increases linearly with the energy of the accelerated particles, and is a significant improvement on the fixed-target result. Colliding beam experiments are not, however, without their own disadvantages: the colliding particles have to be stable, which limits the interactions that can be studied, and the reaction rates in such collisions are far smaller than those achieved in fixed-target experiments, because the particle densities in the beams are low compared to a solid or liquid target.

The rate  $W_r$  at which any particular reaction  $r$  occurs is proportional to the *partial cross-section*  $\sigma_r$ , which depends on the collision

---

<sup>3</sup> This formula is derived as Equation (A.12) in Appendix A, where a brief summary of relativistic kinematics may be found. In this chapter, since we are concerned with practical matters, we use practical units, retaining factors of  $\hbar$  and  $c$ .

energy and the identities of the interacting particles.<sup>4</sup> The constant of proportionality is called the *luminosity*  $L$ , that is

$$W_r = L \sigma_r \quad (4.2a)$$

where the luminosity contains all the dependencies on the densities and geometries of the beam and target. In a collider, where the target is one of the accelerated beams, the luminosity therefore characterises the accelerator. (The explicit dependence of the luminosity on the characteristics of the accelerator is discussed for both the fixed-target and collider case in Appendix B, Section B.1.) A related quantity is the *integrated luminosity*  $L_I$ , which is the integral of  $L$  over time, that is

$$L_I = \int L dt. \quad (4.2b)$$

Finally, details of the particles produced in the collision (e.g. their energies and angular distributions) are deduced by observing their interactions with the material of *detectors*, which are placed in the vicinity of the interaction region. A wide range of detectors is available. Some have a very specific characteristic; others serve more than one purpose. Modern experiments in particle physics typically use several types in a single experiment.

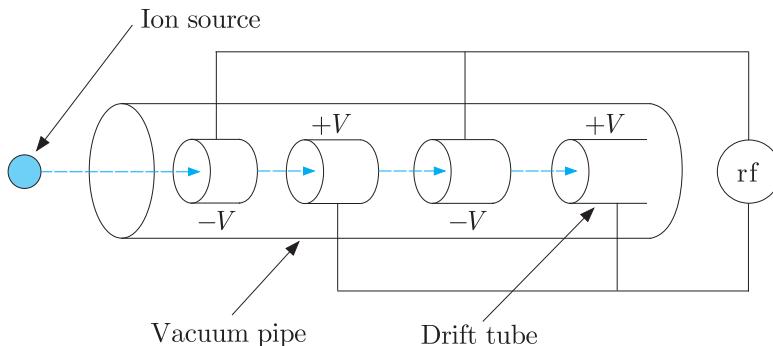
In this chapter we start by describing some of the different types of accelerator that have been built, the beams that they can produce and also how beams of neutral and unstable particles can be prepared. Then we discuss the ways in which particles interact with matter and review how these mechanisms are exploited in the construction of a range of particle detectors. In Section 4.5 we will illustrate how individual detectors are combined into modern complex detector systems and illustrate this with examples from accelerator experiments. We then conclude with a brief discussion of detector systems in non-accelerator experiments. The latter have been important historically and are becoming increasingly so again, particularly in the field of neutrino physics, as was illustrated in Chapter 2.

## 4.2 Accelerators and beams

All accelerators in particle physics use radio frequency (rf) electric fields to boost the energy of stable charged particles. The latter are injected into the machine from a device that provides a high-intensity source of low-energy particles, for example an electron gun (a hot filament) or a proton ion source. Accelerators are conveniently divided into *linear* and *cyclic* varieties and we will discuss each in turn.

---

<sup>4</sup> Cross-sections are formally defined and discussed in Appendix B.



**Figure 4.1** Acceleration in a linear proton accelerator.

#### 4.2.1 Linear accelerators

In a linear accelerator (or *linac*) for accelerating protons, the particles pass through a series of metal pipes called *drift tubes* that are located in a vacuum vessel and connected successively to alternate terminals of an rf oscillator, as shown in Figure 4.1. The positively charged particles accelerated by the field move towards the first drift tube. If the alternator can change its direction as the particles pass through that tube, then they will be accelerated again on their way between the exit of the first and entry to the second tube, and so on. Thus the particles will form bunches and their final energy will be the sum of the energies received at the gaps between the drift tubes. Because the particles are accelerating their speed is increasing, and hence the lengths of the drift tubes have to increase to ensure continuous acceleration. To produce a useful beam, the particles must keep in phase with the rf field and remain focused. Proton linacs of this type are often used as *injectors*; that is, they produce protons of moderate energy that are then injected into a more powerful machine, often a synchrotron (discussed in the next section), where they are accelerated to much higher energies. An example is the linac at J-PARC (Japan Proton Accelerator Research Complex), which accelerates protons to an energy of 3 GeV before they are injected into the main J-PARC ring, where they are further accelerated to 50 GeV.

For electrons, a variation of this method is used. In this case the accelerator consists of a straight tube in the form of a series of cylindrical metal cavities. Power is fed to the accelerator from a series of devices called klystrons, which produce electromagnetic radiation in the form of microwave pulses that are transported via waveguides to the accelerator. There they generate an oscillating electric field pointing along the direction of the metal tube and a magnetic field in a circle around the interior of the accelerating tube. The magnetic field helps to keep the beam focused, and the frequency of the microwaves is adjusted so that the electrons arrive at each cavity of the accelerator at the optimal time to receive the maximum energy

boost from the electric field. As long as this phase relationship can be maintained, the particles will be continuously accelerated. The largest electron linac built was at the SLAC laboratory in Stanford, USA. This accelerator consisted of 80 000 copper cavities separated by copper discs with a small hole at the centre to transmit the beam and was over 3 km long. It could accelerate electrons and positrons to a maximum energy of 50 GeV, but was closed after a series of historically important experiments in 2008.

An ingenious way of reducing the enormous lengths of high-energy linacs has been developed at the Continuous Electron Beam Accelerator Facility (CEBAF) at the Jefferson Laboratory in the USA. This utilises the fact that, above about 50 MeV, electron velocities are very close to the speed of light and thus electrons of very different energies can be accelerated in the same device. Instead of a single long linac, the CEBAF machine consists of two much shorter linacs and the beam from one is bent and passed through the other. This can be repeated for up to four cycles. Even with the radiation losses inherent in bending the beams, very intense beams can be produced with energies between 0.5 and 6.0 GeV. CEBAF has proved to be an important machine for studying the details of nucleon structure.

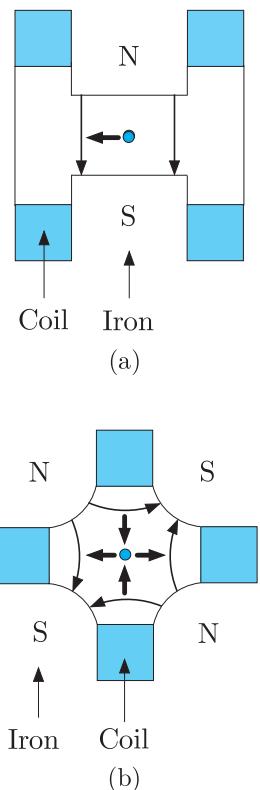
#### 4.2.2 Cyclic accelerators

Cyclic accelerators used in particle physics are of a type called *synchrotrons*. The principle of a synchrotron is analogous to that of a linear accelerator, but the acceleration takes place in a near circular orbit rather than in a straight line. The beam of particles travels in an evacuated tube called the *beam pipe* and is constrained in a circular or near circular path by an array of dipole magnets called bending magnets (see Figure 4.2(a)).

Acceleration is achieved as the beam repeatedly traverses a number of cavities placed in the ring where energy is given to the particles. Since the particles travel in a circular orbit they continuously emit radiation, called in this context *synchrotron radiation*. The amount of energy radiated per turn by a relativistic particle of mass  $m$  and charge  $q$  is given by<sup>5</sup>

$$\Delta E = \frac{4\pi q^2 \beta^3 \gamma^4}{3\epsilon_0 \rho}, \quad (4.3)$$

where  $\beta \equiv v/c$ ,  $v$  is the particle's velocity,  $\gamma \equiv (1 - \beta^2)^{-1/2}$  and  $\rho$  is the radius of curvature of the orbit. For relativistic particles ( $\beta \sim 1$ ) of a given energy  $\gamma mc^2$ , the energy loss is inversely proportional



**Figure 4.2** Cross-section of: (a) a typical bending (dipole) magnet; (b) a focusing (quadrupole) magnet. The thin arrows indicate field directions; the thick arrows indicate the force on a positive particle travelling into the paper.

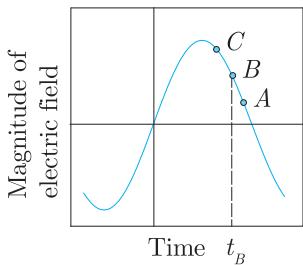
<sup>5</sup> See, for example, p. 661 of Jackson (1975). His results must be divided by  $4\pi\epsilon_0$  to convert from Gaussian to SI units.

to the fourth power of the particle's mass. The losses for electrons are thus very severe, and the need to compensate for these by the input of large amounts of rf power limits the energies of electron synchrotrons. This problem can be somewhat reduced by using a very large ring; the highest energy achieved so far was at the LEP facility, located at the CERN laboratory, Geneva. This had a radius of just over 4 km and accelerated electrons and positrons to an energy of approximately 100 GeV. It operated from 1989 to 2000, when it was decommissioned to make way for the Large Hadron Collider (LHC) discussed below.

The momentum in  $\text{GeV}/c$  of an orbiting particle assumed to have unit charge is given by  $p = 0.3B\rho$ , where  $B$  is the magnetic field in tesla and  $\rho$ , the radius of curvature, is measured in metres. Because  $p$  is increased during acceleration,  $B$  must also be steadily increased if  $\rho$  is to remain constant, and the final momentum is limited both by the maximum field available and by the size of the ring. With conventional electromagnets, the largest field attainable over an adequate region is about 1.5 tesla, and even with superconducting coils it is only of order 10 tesla. Hence, even for protons, when radiation losses are very small, the radius of the ring must be very large to achieve very high energies. For example, the Tevatron, which operated from 1983 to 2011 at the Fermi National Accelerator Laboratory (FNAL) near Chicago, accelerated protons to an energy of  $1 \text{ TeV} = 10^3 \text{ GeV}$  and had a radius of 1 km.

In the course of its acceleration, a beam may make typically  $10^5$  traversals of its orbit before reaching its maximum energy. Consequently stability of the orbit is vital, both to ensure that the particles continue to be accelerated and that they do not strike the sides of the vacuum tube. In practice, the particles are accelerated in bunches, each being synchronised with the rf field. In equilibrium, a particle increases its momentum just enough to keep the radius of curvature constant as the field  $\mathbf{B}$  is increased during one rotation, and the circulation frequency of the particle is in step with the rf of the field. This is illustrated in Figure 4.3. Particle B is assumed to be in equilibrium orbit, synchronous with the rf field. Particle A, behind the rf phase, receives a lower momentum increase from the field than particle B. This will reduce the radius of its orbit and (since  $v \approx c$ ) increase its rotational frequency relative to particle B. Conversely, a particle C, ahead of the rf phase, receives a greater momentum increase and a decrease in its rotational frequency. With obvious changes, a similar principle is used in linear accelerators.

In practice, the particles remain in the bunch, but their trajectories oscillate about the stable orbits. These oscillations are controlled by a series of focusing magnets, usually of the quadrupole type, which are placed at intervals around the beam and act like optical lenses. A schematic diagram of one of these is shown in Figure 4.2(b). Each



**Figure 4.3** Magnitude of the electric field as a function of time at a fixed point in the rf cavity. Particle B is synchronous with the field and arrives at time  $t_B$ . Particle A (C) is behind (ahead of) B and receives an increase (decrease) in its rotational frequency. Thus particles oscillate about the equilibrium orbit.

focuses the beam in one direction and de-focuses it in the orthogonal direction, so alternate magnets have their field directions reversed to keep the particles in a stable orbit.

In addition to the energy of the beam, one is also concerned to produce a beam of high intensity, so that interactions will be plentiful. The intensity is ultimately limited by de-focusing effects, for example the mutual repulsion of the particles in the beam, and a number of technical problems have to be overcome, which are outside the scope of this brief account.

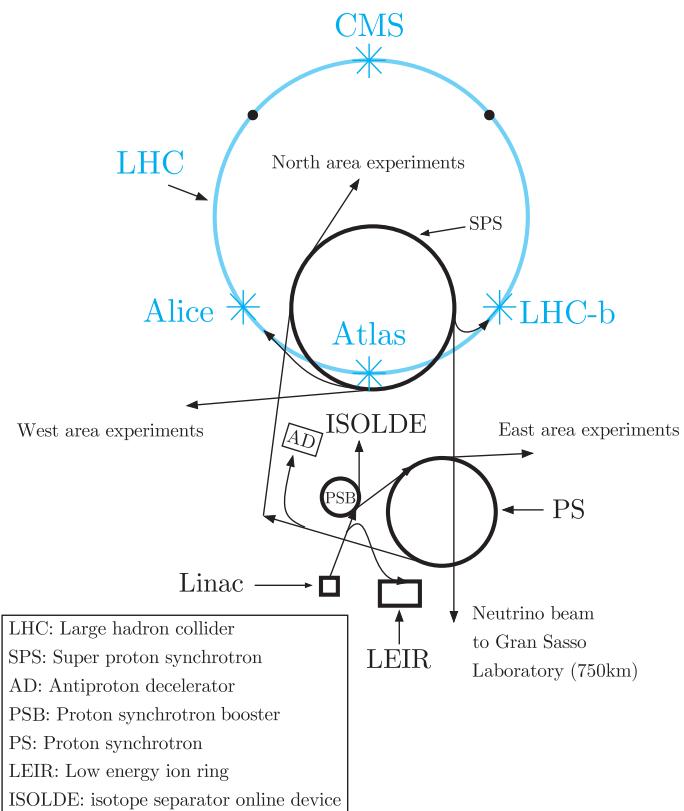
#### 4.2.3 Fixed-target machines and colliders

Both linear and cyclic accelerators can be divided into *fixed-target* and *colliding-beam* machines. The latter are also known as *colliders*, or sometimes, in the case of cyclic machines, *storage rings*.<sup>6</sup> In fixed-target machines, particles are accelerated to the highest operating energy and then the beam is extracted from the machine and directed on to a stationary target, which is usually a solid or liquid. Much higher energies have been achieved for protons than electrons, because of the large radiation losses inherent in electron machines mentioned earlier. The intensity of the beam is such that large numbers of interactions can be produced, which can either be studied in their own right or used to produce secondary beams, as discussed in Section 4.2.4.

The main disadvantage of fixed-target machines for particle physics has been mentioned earlier: the need to achieve large centre-of-mass energies to produce new particles. For the last four decades, almost all new machines for particle physics have therefore been colliders, although some fixed-target machines for specialised purposes are still constructed. The largest collider currently operational is the LHC at the CERN laboratory near Geneva, mentioned above, which commenced data taking in 2008.<sup>7</sup> This is a massive proton–proton accelerator of circumference 27 km, with a luminosity of  $(1 - 2) \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  and with each beam having an energy of 6.5–7 TeV. A schematic diagram of the CERN site showing the LHC and some of its other accelerators is shown in Figure 4.4. The acceleration process starts with a linac, whose beam is boosted in energy in the PSB (Proton Synchrotron Booster) and passed to the PS (Proton Synchrotron), a machine that is still the source of beams for lower-energy experiments. The beam energy is increased still further in the SPS (Super Proton Synchrotron), which also provides

<sup>6</sup> The use of the terms *storage rings* and *colliders* as synonymous is not strictly correct, because the former can also describe a machine that stores a single beam for use on both internal and external fixed targets.

<sup>7</sup> The beam line in the LHC tunnel is shown in coloured Plate 2.



**Figure 4.4** A schematic diagram of the CERN site showing the LHC and some of its other accelerators.

beams for a range of experiments, as well as the injection beams for the LHC itself. The two beams circulate in opposite directions and are made to intersect at the four points shown in the LHC where experiments (ALICE, CMS, LHC-b and ATLAS) are located.<sup>8</sup> The extracted neutrino beam<sup>9</sup> shown at the bottom of the diagram is sent to the Gran Sasso laboratory, which is a large underground laboratory 750 km away in Italy, where neutrino and other experiments are performed.

Other large colliders currently operational include the Relativistic Heavy Ion Collider (RHIC), located at Brookhaven National Laboratory, USA, which can accelerate and collide polarised proton beams. This machine, which began operation in 2000, was also the first collider capable of accelerating heavy ions to high energies. Like the LHC, there are several stages before the ions are injected into the main machine. There they form two counter-circulating beams controlled by two 4-km rings of superconducting magnets and are accelerated to an energy of 100 GeV/nucleon. Thus the total

<sup>8</sup> The CMS and ATLAS detectors during their construction phases are shown in coloured Plates 3 and 4.

<sup>9</sup> Production of neutral particle beams is discussed in the next section.

centre-of-mass energy is 200 GeV/nucleon. Collisions occur at six intersection points, where major experiments are sited. RHIC primarily accelerates ions of gold and is used to study matter at extreme energy-densities, where a new state of matter called a ‘quark–gluon plasma’ is predicted to occur. We will return to this briefly in Section \*7.1.4.

Although these colliders all have beams of equal energies, there are circumstances where having unequal energies is important. Thus,  $e^+e^-$  colliders using beams of unequal energies have made important contributions to the physics of particles containing a quark with the bottom quantum number. This is discussed further in Section 11.2.6.

All the above colliders are cyclic machines. Linear colliders have also been built, although fewer in number. For example, the SLC machine at the SLAC laboratory was an  $e^+e^-$  linear collider with a centre-of-mass energy of 90 GeV and a luminosity of  $2.5 \times 10^{30} / \text{cm}^{-2} \text{ s}^{-1}$  that did important early work on the  $Z^0$  boson. There is also a large international project underway to design and build a high-luminosity  $e^+e^-$  accelerator (the International Linear Collider – ILC) capable of producing centre-of-mass energies of 250–1000 GeV.

#### 4.2.4 Neutral and unstable particle beams

The particles used in accelerators must be stable and charged, but one is also interested in the interaction of neutral particles, for example photons and neutrinos, as well as those of unstable particles (such as charged pions). Beams appropriate for performing such experiments are produced in a number of ways.

Beams of unstable particles can be formed, provided the particles live long enough to travel appreciable distances in the laboratory. One way of doing this is to direct an extracted primary beam on to a heavy target. In the resulting interactions with the target nuclei, many new particles are produced, which may then be analysed into secondary beams of well-defined momentum. Such beams will ideally consist predominantly of particles of one type, but if this cannot be achieved, then the wanted species may have to be identified by other means. In addition, if these secondary beams are composed of unstable particles, they can themselves be used to produce further beams formed from their decay products. Two examples will illustrate how in principle such secondary particle beams can be formed.

Consider firstly the construction of a  $\pi^+$  beam from a primary beam of protons. By allowing the protons to interact with a heavy target, secondary particles, most of which will be pions, will be produced. A collimator can be used to select particles in a particular direction, and the  $\pi^+$  component can subsequently be removed and focused into a monoenergetic beam by selective use of electrostatic fields and bending and focusing magnets.

The pion beam may also be used to produce a beam of neutrinos. For example, the  $\pi^-$  is unstable and, as we have seen, its dominant decay mode is  $\pi^- \rightarrow \mu^- + \bar{\nu}_\mu$ . So if pions are passed down a long vacuum pipe, many will decay in flight to give muons and antineutrinos, which will mostly travel in essentially the same direction as the initial beam. The muons and any remaining pions can then be removed by passing the beam through a very long absorber, leaving the neutrinos. In this case, the final neutrino beam will not be monoenergetic, but instead will have a momentum spectrum reflecting the initial momentum spectrum of the pions, and since neutrinos are neutral, no further momentum selection using magnets is possible. Experiments using neutrino beams are discussed in Sections 2.3.2 and 8.3.

### 4.3 Particle interactions with matter

In order to be detected, a particle must undergo an interaction with the material of a detector. In this section we discuss these interactions, but only in sufficient detail to be able to understand the detectors themselves. For more details, we again refer the reader to the references in Footnote 2.

The first possibility is that the particle interacts with an atomic nucleus. For example, this could be via the strong nuclear interaction if it is a hadron or by the weak interaction if it is a neutrino. We know from the work of Chapter 1 that both are *short-range interactions*. If the energy is sufficiently high, new particles may be produced, and such reactions are often the first step in the detection process. In addition to these short-range interactions, a charged particle will also excite and ionise atoms along its path, giving rise to *ionisation energy losses*, and emit radiation, leading to *radiation energy losses*. Both of these processes are due to the long-range electromagnetic interaction. They are important because they form the basis of most detectors for charged particles. Photons are also directly detected by electromagnetic interactions, and at high energies their interactions with matter lead predominantly to the production of  $e^+e^-$  pairs via the *pair production* process  $\gamma \rightarrow e^+ + e^-$ , which has to occur in the vicinity of a nucleus to conserve energy and momentum. (Recall the discussion of pair production in the presence of a nucleus in Section 1.3.3.) All these types of interactions are described in the following sections.

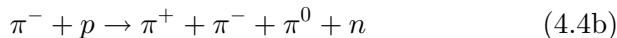
#### 4.3.1 Short-range interactions with nuclei

For hadrons, the most important short-range interactions with nuclei are due to the strong nuclear force, which unlike the electromagnetic interaction is as important for neutral particles as for charged ones.

For the simplest nucleus, the proton, the resulting reactions are of two types: *elastic scattering* such as

$$\pi^- + p \rightarrow \pi^- + p, \quad (4.4a)$$

where the initial and final particles are the same and for illustration we have taken the incident particle to be a  $\pi^-$ , and *inelastic reactions*, such as



and

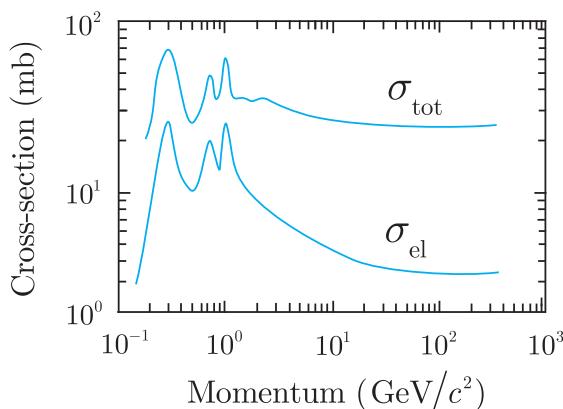
$$\pi^- + p \rightarrow K^0 + \Lambda, \quad (4.4c)$$

in which the final-state particles differ from those in the initial state. At high energies, many inelastic reactions are possible, most of them involving the production of several particles in the final state. The total cross-section

$$\sigma_{\text{tot}} = \sigma_{\text{el}} + \sigma_{\text{inel}} \quad (4.5)$$

is the sum of the cross-section for elastic scattering  $\sigma_{\text{el}}$  and that for inelastic reactions  $\sigma_{\text{inel}}$ , where the latter is itself a sum over all possible inelastic processes allowed by conservation laws.<sup>10</sup> The behaviours of the total and elastic cross-sections for  $\pi^-p$  scattering are shown in Figure 4.5.

There is considerable structure at low energies due to the production of resonances, but for momenta above about  $3 \text{ GeV}/c$  the total cross-section is slowly varying and much larger than the elastic cross-section. The same general features are found for other incident hadrons at high energies, with total cross-sections typically in the range 10–50 mb. This is of the same order-of-magnitude as the ‘geometrical’ cross-section  $\pi r^2 \approx 30 \text{ mb}$ , where  $r \approx 1 \text{ fm}$  is the approximate range of the strong interaction between hadrons. Total



**Figure 4.5** Total and elastic cross-sections for  $\pi^-p$  scattering as functions of the pion laboratory momentum.

<sup>10</sup> The various types of cross-section and related quantities are formally defined in Appendix B.

**Table 4.1** Nuclear cross-sections and the associated collision lengths  $l_c$  and absorption lengths  $l_a$  for incident neutrons with energies in the range 80–300 GeV. The values for protons are approximately the same.

Element	$Z$	Nuclear cross-section (b)		Interaction length (cm)	
		$\sigma_{\text{tot}}$	$\sigma_{\text{inel}}$	$l_c$	$l_a$
H*	1	0.039	0.033	687	806
C	6	0.331	0.231	26.6	38.1
Al	13	0.634	0.421	26.1	39.4
Fe	26	1.120	0.703	10.5	16.8
Pb	82	2.960	1.770	10.2	17.1

\*Liquid hydrogen at 26 K. The other materials are solids.

cross-sections on neutrons are similar to those on protons, but total cross-sections on nuclei are much larger, increasing roughly like the square of the nuclear radius.

The probability of a hadron–nucleus interaction occurring as the hadron traverses a small thickness  $dx$  of material is given by  $n\sigma_{\text{tot}} dx$ , where  $n$  is the number of nuclei per unit volume in the material. Consequently, the mean distance travelled before an interaction occurs is given by

$$l_c = 1/n\sigma_{\text{tot}}. \quad (4.6)$$

This is called the *collision length*.<sup>11</sup> An analogous quantity is the *absorption length*, defined by

$$l_a = 1/n\sigma_{\text{inel}}, \quad (4.7)$$

which governs the probability of an inelastic collision. This is illustrated in Table 4.1, where  $l_a$  and  $l_c$  are given for incident neutrons on several common materials.

Neutrinos and antineutrinos can also be absorbed by nuclei, leading to reactions of the type

$$\bar{\nu}_\ell + p \rightarrow \ell^+ + X, \quad (4.8)$$

where  $\ell$  is a lepton and  $X$  denotes any hadron or set of hadrons allowed by the conservation laws. Such processes are weak interactions (because they involve neutrinos) and the associated cross-sections are extremely small compared to the cross-sections for strong interaction processes. The corresponding interaction lengths are therefore enormous. Nonetheless, in the absence of other possibilities such reactions are the basis for detecting neutrinos. Finally, photons can be absorbed by nuclei, giving *photoproduction* reactions

<sup>11</sup>There is a further discussion of this in Section B.2, Appendix B.

such as  $\gamma + p \rightarrow X$ . However, these electromagnetic interactions are only used to detect photons at low energies, because at higher energies there is a far larger probability for  $e^+e^-$  pair production in the Coulomb field of the nucleus. We will return to this in Section 4.3.4 below.

### 4.3.2 Ionisation energy losses

Ionisation energy losses are important for all charged particles, and for particles other than electrons and positrons they dominate over radiation energy losses at all but the highest attainable energies. The theory of such losses, which are due dominantly to Coulomb scattering from the atomic electrons,<sup>12</sup> was worked out by Bethe, Bloch and others in the 1930s.<sup>13</sup> The result is called the Bethe–Bloch formula, and for spin-0 bosons with charge  $\pm q$  (in units of  $e$ ), mass  $M$  and velocity  $v$ , it takes the approximate form

$$-\frac{dE}{dx} = \frac{D q^2 n_e}{\beta^2} \left[ \ln \left( \frac{2m_e c^2 \beta^2 \gamma^2}{I} \right) - \beta^2 - \frac{\delta(\gamma)}{2} \right], \quad (4.9a)$$

where  $x$  is the distance travelled through the medium. In this expression,

$$D = \frac{4\pi \alpha^2 \hbar^2}{m_e} = 5.1 \times 10^{-25} \text{ MeV cm}^2, \quad (4.9b)$$

$m_e$  is the electron mass,  $\beta = v/c$  and  $\gamma = (1 - \beta^2)^{-1/2}$ . The other constants refer to the properties of the medium:  $n_e$  is the electron density,  $I$  is the mean ionisation potential of the atoms averaged over all electrons, which is given approximately by  $I = 10Z$  eV for  $Z$  greater than 20, and  $\delta$  is a dielectric screening correction that is important only for highly relativistic particles. The corresponding formula for spin-1/2 particles differs from this, but in practice the differences are small and may be neglected in discussing the main features of ionisation energy losses.

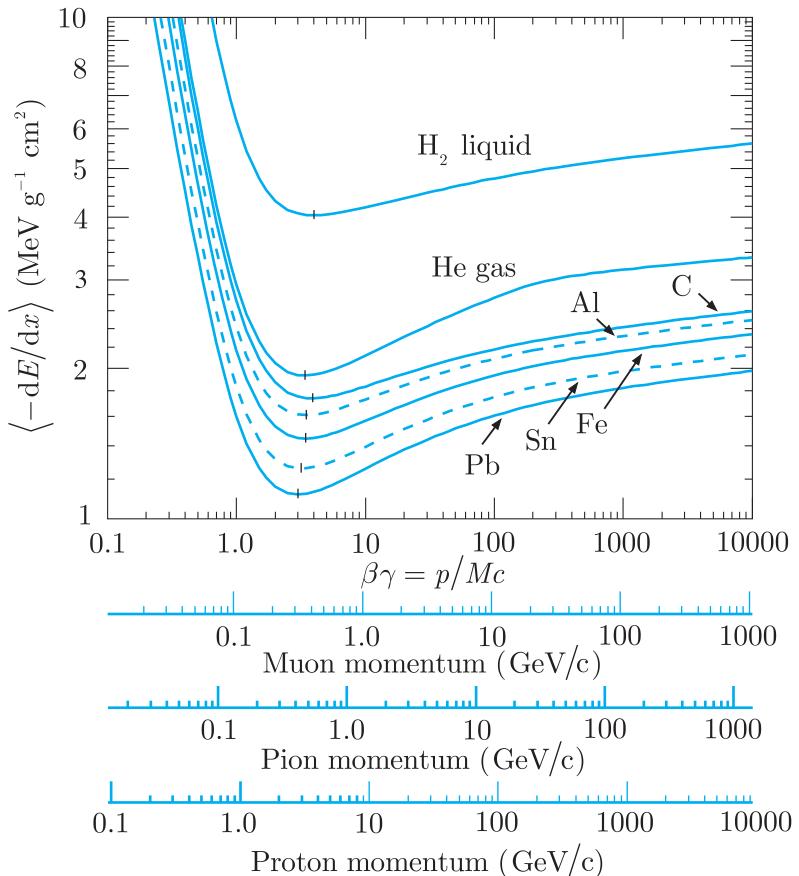
It is common practice to divide (4.9a) by the mass density  $\rho$  and to redefine  $x$  as the density multiplied by the distance travelled, so that

$$-\frac{1}{\rho} \frac{dE}{dx} \rightarrow -\frac{dE}{dx}, \quad (4.10)$$

where  $x$  is now measured in g/cm<sup>2</sup>. Examples of the behaviour of (4.10) for muons, pions and protons traversing a range of materials are shown in Figure 4.6, using this convention. As can be seen,

<sup>12</sup>The analogous scattering from the nucleus gives a negligible contribution to the energy loss since the recoil energy of the nucleus is much smaller because of its much greater mass.

<sup>13</sup>For a semi-classical discussion, see, for example, Chapter 13 of Jackson (1975).



**Figure 4.6** Ionisation energy loss for muons, pions and protons in a variety of materials. The units of  $dE/dx$  are explained in the text. (From Olive *et al.* (Particle Data Group), 2014.)

$-dE/dx$  falls rapidly as the velocity increases from zero because of the  $1/\beta^2$  factor in the Bethe–Bloch equation. All particles have a region of ‘minimum ionisation’ for  $\beta\gamma$  in the range 3 to 4. Beyond this,  $\beta$  tends to unity, and the logarithmic factor in the Bethe–Bloch formula gives a ‘relativistic rise’ in  $-dE/dx$ .

The magnitude of the energy loss depends on the medium. The electron density is given by  $n_e = \rho N_A Z/A$ , where  $N_A$  is Avogadro’s number,  $\rho$  is the density in grams per unit volume and  $A$  is the atomic weight of the medium, so the mean energy loss is proportional to the density of the medium. The remaining dependence on the medium is relatively weak because  $Z/A \approx 0.5$  for all atoms except the very light and the very heavy elements, and because the ionisation energy  $I$  only enters the Bethe–Bloch formula logarithmically. In the ‘minimum ionisation’ region, where  $\beta\gamma \approx 3 – 4$ , the minimum value of  $-dE/dx$  can be calculated from (4.9), and for a particle with unit charge is given approximately (in the units of Figure 4.6) by

$$\frac{1}{\rho} \left( -\frac{dE}{dx} \right)_{\min} \approx 3.5 \frac{Z}{A} \text{ MeV g}^{-1} \text{ cm}^2, \quad (4.11)$$

**Table 4.2** The minimum ionisation energy losses  $(-\mathrm{d}E/\mathrm{dx})_{\min}$  for various materials and their dependence on the density  $\rho$  in  $\mathrm{g}/\mathrm{cm}^3$ .

Element	$Z$	$\rho$	$(-\mathrm{d}E/\mathrm{dx})_{\min}$ (MeV/cm)	$-\frac{1}{\rho}(\mathrm{d}E/\mathrm{dx})_{\min}$ (MeV $\mathrm{g}^{-1} \mathrm{cm}^2$ )
H*	1	0.063	0.26	4.12
C	6	2.26	4.02	1.78
Al	13	2.70	4.37	1.62
Fe	26	7.87	11.6	1.48
Pb	82	11.35	12.8	1.13

\*Liquid hydrogen at 26 K. The other materials are solids.

where we revert to using  $x$  to denote the distance travelled, as usual.<sup>14</sup> Typical values for a minimum ionising particle with unit charge are given in Table 4.2 and are in good agreement with (4.11) for a wide range of elements except the very lightest. Also shown are values of  $(-\mathrm{d}E/\mathrm{dx})_{\min}$  in units of MeV/cm and the corresponding value of the density.

Ionisation losses are proportional to the squared charge of the particle, so that a fractionally charged particle with  $\beta\gamma \geq 3$  would have a much lower rate of energy loss than the minimum energy loss of any integrally charged particle. This has been used as a means of searching for free quarks, but without success.

From the rate of energy loss as a function of  $\beta$ , we can calculate the rate of energy loss as a function of the distance  $x$  travelled in the medium. This is called the *Bragg curve*. Because of the factor  $\beta^{-2}$  in (4.9), most of the ionisation loss occurs near the end of the path where the speed is smallest and the curve has a pronounced peak close to the end point before falling rapidly to zero at the end of the particle's path length. For particles whose energy loss is dominated by ionisation, the *range R*, that is the mean distance a particle travels before it comes to rest, is given by

$$R = \int_0^R \mathrm{d}x = \int_0^{\beta_i} \left( -\frac{\mathrm{d}E}{\mathrm{d}x} \right)^{-1} \left[ \frac{\mathrm{d}E}{\mathrm{d}\beta} \mathrm{d}\beta \right] = \frac{M}{q^2 n_e} F(\beta_i), \quad (4.12)$$

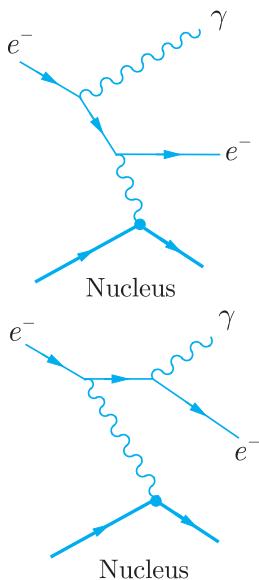
where  $F$  is a function of the initial velocity  $\beta_i$  and we have used the relation  $E = \gamma M c^2$  to show the dependence on the projectile mass  $M$ . For non-relativistic particles  $\beta_i \ll 1$ , the rate of energy loss is very high and the particle comes to rest very quickly. This was the basis for excluding the possibility that the upper track in Figure 1.2

<sup>14</sup>In this book we only use  $x$  in  $\mathrm{g}/\mathrm{cm}^2$  in the right-hand side of (4.10) and in Figure 4.6.

could be due to a proton of momentum  $23 \text{ MeV}/c$ . Such a particle would be highly non-relativistic ( $\beta \approx 0.02$ ) and would come to rest in a short distance, which detailed calculation using (4.9) shows is of order 5 mm, much less than the length of the observed track.

Finally, we note that the range as given by (4.12) is actually an average value because scattering is a statistical process and there will therefore be a spread of values for individual particles. This is illustrated by the muon tracks in Figure 3.2. Here the four muons all have the same energy, since they all arise in the decay of pions at rest, and all travel approximately the same distances in the emulsion before themselves coming to rest and subsequently decaying.

### 4.3.3 Radiation energy losses



**Figure 4.7** Dominant Feynman diagrams for the bremsstrahlung process  $e^- + (Z, A) \rightarrow e^- + \gamma + (Z, A)$ .

When a charged particle traverses matter it can also lose energy by radiative collisions, especially with nuclei. The electric field of a nucleus will accelerate and decelerate the particles as they pass, causing them to radiate photons, and hence lose energy. This process is called *bremssstrahlung* (literally ‘braking radiation’ in German) and is a particularly important contribution to the energy loss for electrons and positrons.

The dominant Feynman diagrams for electron bremsstrahlung in the field of a nucleus, that is

$$e^- + (Z, A) \rightarrow e^- + \gamma + (Z, A), \quad (4.13)$$

are shown in Figure 4.7 and are of order  $Z^2\alpha^3$ , like the pair creation diagrams of Figure 1.11.

There are also contributions from bremsstrahlung in the fields of the atomic electrons, each of order  $\alpha^3$ . Since there are  $Z$  atomic electrons for each nucleus, these give a total contribution of order  $Z\alpha^3$ , which is small compared to the contribution from the nucleus for all but the lightest elements. A detailed calculation shows that for relativistic electrons with  $E \gg mc^2/\alpha Z^{1/3}$ , the average rate of energy loss is given by

$$-dE/dx = E/L_R. \quad (4.14)$$

The constant  $L_R$  is called the *radiation length* and is given by

$$\frac{1}{L_R} = 4 \left( \frac{\hbar}{mc} \right)^2 Z(Z+1)\alpha^3 n_a \ln \left( \frac{183}{Z^{1/3}} \right), \quad (4.15)$$

where  $n_a$  is the number density of atoms/cm<sup>3</sup> in the medium. Integrating (4.14) gives

$$E = E_0 \exp(-x/L_R), \quad (4.16)$$

where  $E_0$  is the initial energy. It follows that the radiation length is the average thickness of material that reduces the mean energy by a factor  $e$ . Values of  $L_R$  for various materials are shown in Table 4.3.

**Table 4.3** The radiation length  $L_R$  and the energy  $E_C$ , at which electron radiation energy losses equal their ionisation energy losses, for various elements. Except for the very lightest elements,  $E_C \approx 600/Z$  MeV.

Element	$Z$	$L_R$ (cm)	$E_C$ (MeV)
H*	1	≈1000	340
C	6	18.8	103
Al	13	8.9	47
Fe	26	1.8	24
Pb	82	0.56	7

\*Liquid hydrogen at 26 K. The other materials are solids.

From these results, we see that at high energies the radiation losses are proportional to  $E/m_P^2$  for an arbitrary charged particle  $P$  of mass  $m_P$ . On the other hand, the ionisation energy losses are only weakly dependent on the projectile mass and energy at very high energies. Consequently, radiation losses completely dominate the energy losses for electrons and positrons at high enough energies, but are much smaller than ionisation losses for all particles other than electrons and positrons at all but the highest energies. In Table 4.3 we show the energy  $E_C$  at which radiation losses become as large as ionisation losses for electrons in various materials. For the next lightest particles, muons,  $E_C$  will be increased by a factor of approximately  $(m_\mu/m)^2 \approx 2 \times 10^4$ , which justifies our earlier statement that radiation losses are much smaller than ionisation losses for all particles other than electrons and positrons at all but the highest energies.

#### 4.3.4 Interactions of photons in matter

In contrast to heavy charged particles, photons have a high probability of being absorbed or scattered through large angles by the atoms in matter. Consequently, a collimated monoenergetic beam of  $I$  photons per second traversing a thickness  $dx$  of matter will lose

$$dI = -I \frac{dx}{\lambda} \quad (4.17)$$

photons per second, where

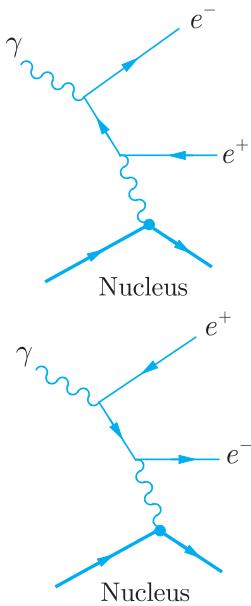
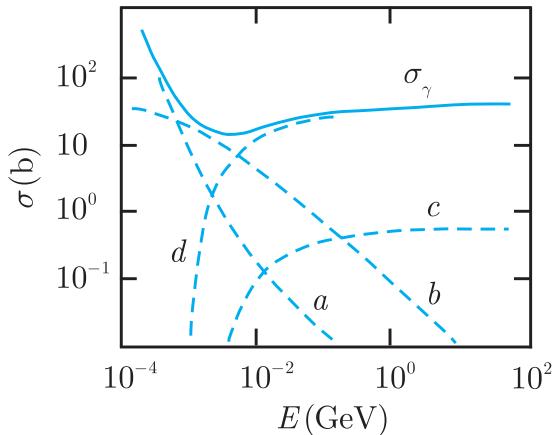
$$\lambda = (n_a \sigma_\gamma)^{-1} \quad (4.18)$$

is the mean free path before absorption or scattering out of the beam and  $\sigma_\gamma$  is the total photon interaction cross-section with an atom. The mean free path  $\lambda$  is analogous to the collision length for hadronic reactions. Integrating (4.17) gives

$$I(x) = I_0 e^{-x/\lambda} \quad (4.19)$$

for the intensity of the beam as a function of distance, where  $I_0$  is the initial intensity.

**Figure 4.8** Total experimental photon cross-section  $\sigma_\gamma$  on a lead atom, together with the contributions from: (a) the photoelectric effect, (b) Compton scattering, (c) pair production in the field of the atomic electrons, (d) pair production in the field of the nucleus. (Adapted from Olive *et al.* (Particle Data Group), 2014.)



**Figure 4.9** The pair production process  $\gamma + (Z, A) \rightarrow e^- + e^+ + (Z, A)$ .

The main processes contributing to  $\sigma_\gamma$  at the energies relevant for particle physics are: the *photoelectric effect*, in which the photon is absorbed by the atom as a whole with the emission of an electron; *Compton scattering*, where the photon scatters from an atomic electron; and *electron–positron pair production* in the field of a nucleus or of an atomic electron. The corresponding cross-sections on lead are shown in Figure 4.8, where it can be seen that above a few MeV the cross-section is dominated by pair production from the nucleus.<sup>15</sup>

The pair production process is closely related to electron bremsstrahlung, as can be seen by comparing the Feynman diagrams shown in Figures 4.7 and 4.9. The cross-section for pair production rises rapidly from threshold, and is given to a good approximation by

$$\sigma_{pair} = \frac{7}{9} \frac{1}{n_a L_R}, \quad (4.20)$$

for  $E_\gamma \gg mc^2/\alpha Z^{1/3}$ , where  $L_R$  is the radiation length. Substituting these results into (4.19) gives

$$I(x) = I_0 \exp(-7x/9L_R), \quad (4.21)$$

so that, at high energies, photon absorption, like electron radiation loss, is characterised by the radiation length  $L_R$ .

### 4.3.5 Ranges and interaction lengths

For non-relativistic charged particles ( $\beta \ll 1$ ), ionisation energy losses are very high and the particles come to rest relatively quickly,

<sup>15</sup>The photoelectric, Compton and pair production cross-sections are roughly proportional to  $Z^5/E_\gamma$ ,  $Z/E_\gamma$  and  $Z^2$ , respectively, at high energies. Hence, for smaller atomic numbers, the cross-sections are smaller and the region of pair production dominance is shifted to somewhat higher energies.

as we saw in Section 4.3.2. However, as  $\beta$  becomes larger, the range (4.12) becomes comparable with the interaction lengths (4.6) and (4.7) and the radiation length (4.15). The situation is complicated, but for highly relativistic particles, a fairly simple picture emerges for any given material. We shall briefly summarise the situation here by discussing classes of particles in order of their increasing penetration power in condensed matter, since this is an important factor in the design of experiments at high energies, as we shall see in Section 4.5.

#### 4.3.5(a) Electrons and photons

For energies above 1 GeV, these are the least penetrating particles, whose interactions are controlled by the radiation lengths listed in Table 4.3.

#### 4.3.5(b) Hadrons

Radiation energy losses are unimportant for charged hadrons, because they are suppressed, relative to those of electrons, by the square of their mass. Their strong interactions with atomic nuclei are controlled by the interaction lengths of Table 4.1, which are typically much larger than the radiation lengths for electrons in the same material.

#### 4.3.5(c) Muons

Again, radiation energy losses are suppressed by the square of their mass and muons have no strong interactions with nuclei. They are easily the most penetrating of all charged particles.

#### 4.3.5(d) Neutrinos

As these only interact by the weak interaction, they can pass through enormous distances of matter without interacting, as discussed earlier.

### 4.4 Particle detectors

The detection of a particle means more than simply its localisation. To be useful this must be done with a resolution sufficient to enable particles to be separated in both space and time in order to determine which are associated with a particular event. We also need to be able to identify each particle and measure its energy and momentum. No single detector is optimal with respect to all these requirements and modern particle physics experiments commonly use very large multi-component detector systems that integrate many different

sub-detectors in a single device. Such systems rely heavily on fast electronics and computers to monitor and control the sub-detectors, and to co-ordinate, classify and record the vast amount of information flowing in from different parts of the apparatus. Such multi-component detectors will be considered in Section 4.5. Here we will briefly introduce some of the most important individual detectors of which they are composed. Detector development is a rapidly moving major area of research and new devices are frequently developed, so the list below is by no means exhaustive.<sup>16</sup>

#### 4.4.1 Introduction

We start by discussing the important family of *gaseous detectors*. These convert the ionisation produced by the passage of a charged particle through a gas into an electronic signal. Such detectors are primarily used to provide accurate measurements of a particle's position or, by a sequence of such measurements, a record of the particle's trajectory. In this context they are also called *tracking detectors*.

Tracking detectors are very often placed in a magnetic field, in which case they can provide a measurement of the particle's momentum from the resulting curvature of its track. An apparatus that is dedicated to measuring momentum is called a *spectrometer*. It consists of a magnet and a series of detectors to track the passage of the particles. The precise design depends on the nature of the experiment being undertaken. For example, in a fixed-target experiment at high energies, the reaction products are usually concentrated in a narrow cone about the initial beam direction, whereas in colliding beam experiments spectrometers must completely surround the interaction region to obtain full angular coverage. However, in this case the beam will also be deflected, and so at colliders so-called 'compensating magnets' are added to correct for this.

Next we discuss four more types of charged particle detectors: *solid-state detectors*, *scintillation counters*, *Čerenkov counters* and *transition radiation detectors*. Solid-state detectors exploit the properties of semiconductors. They are in some respects the solid-state analogue of gas ionisation detectors and have, to some extent, replaced the latter in current experiments. Scintillation counters have excellent time resolution and are sometimes used for 'triggering' other devices in multi-component detector systems, that is to decide whether or not to activate other detectors, or whether to record the information from a particular event. Čerenkov counters measure the velocity of a charged particle and can be used to distinguish

---

<sup>16</sup>More detailed discussions of particle detectors are given in the references in Footnote 2.

between different particles having a given very high momentum by using their velocities to determine their masses. Transition radiation detectors utilise the fact that electromagnetic radiation is emitted when a charged particle crosses the boundary between two materials, and the probability of emission depends on the relativistic  $\gamma$ -factor. For a given energy, this allows lighter particles to be distinguished from heavier ones of the same energy.

All the above detectors can only detect charged particles, and all leave the nature of the particle unchanged as it passes through the detector. In contrast, the final detectors we discuss, *calorimeters*, can detect both neutral and charged particles. They totally absorb the detected particle to yield a measurement of its energy. Calorimeters can also have good spatial and time resolutions, which is particularly important for neutral particles that often cannot be detected in any other way.

#### 4.4.2 Gaseous ionisation detectors

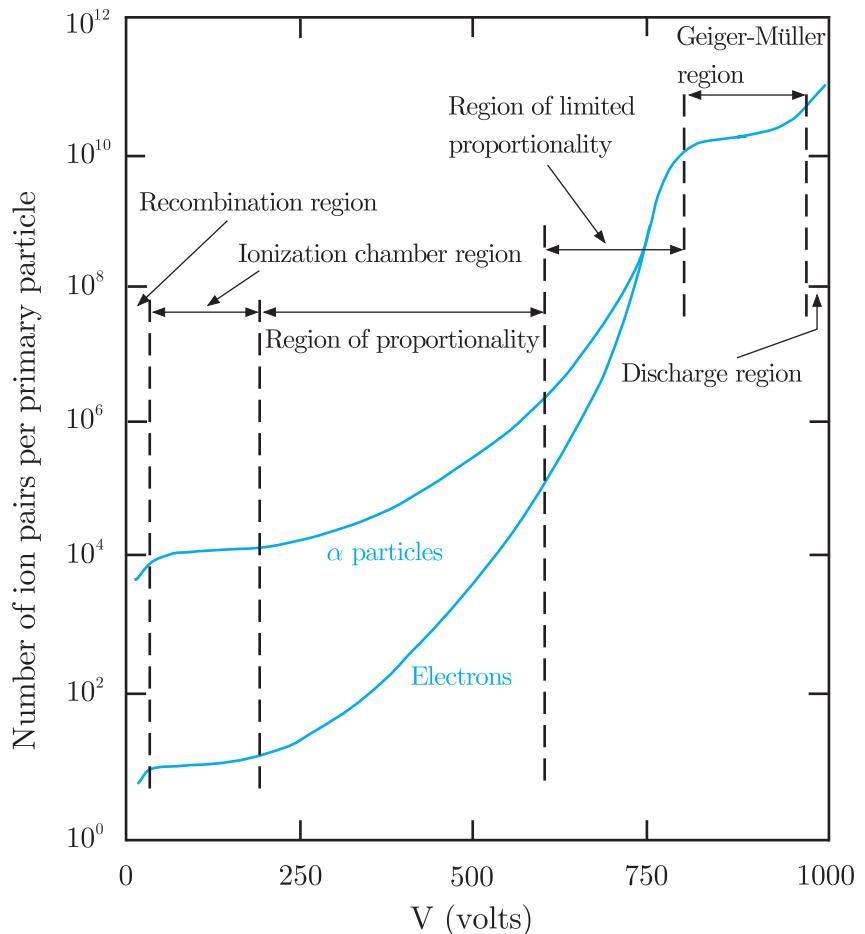
Most gaseous detectors detect the ionisation produced by the passage of a charged particle through a gas, typically an inert one such as argon, either by collecting the ionisation products or induced charges on to electrodes, or (historically) by making the ionisation track visible in some form. The average energy needed to produce an electron–ion pair is  $30 \pm 10$  eV, with a weak dependence on the gas used and the energy of the incident particle. In practice, the output is a pulse at the anode, which is amplified by electronic means.

For a certain range of applied voltages – the so-called ‘proportional region’ (see below) – these devices are primarily used to provide accurate measurements of a particle’s position. As position detectors, gaseous detectors largely replaced earlier detectors that used visual techniques, such as cloud chambers, bubble chambers and stacks of photographic emulsions (these detectors were mentioned in Chapter 2). Although historically important, they have been superceded by electronic detectors and none of these visual devices are now in general use.

To understand the principles of gaseous detectors we refer to Figure 4.10, which shows the number of ion pairs produced per incident charged particle (the *gas amplification factor*) as a function of the applied voltage  $V$  for two cases: a heavily ionising particle (e.g. an alpha particle) and a lightly ionising particle (e.g. an electron).

##### 4.4.2(a) Ionisation chamber

At low applied voltages, the output signal is very small because electron–ion pairs recombine before reaching the electrodes, but as the voltage increases the number of pairs increases to a saturation



**Figure 4.10** Gas amplification factor as a function of voltage  $V$  applied in a typical single-wire gas detector, with a wire radius typically  $20\text{ }\mu\text{m}$ , for a strongly ionising particle (alpha particle) and a weakly ionising particle (electron).

level representing complete collection. This is the region of the *ionisation chamber*. The simplest type of chamber is a parallel plate condenser filled with an inert gas and having an electric field  $E = V/d$ , where  $d$  is the distance between the plates. In practice the gas mixture must contain at least one ‘quenching’ component that absorbs ultraviolet light and stops a plasma forming and spreading throughout the gas.

Another arrangement is cylindrical with an inner anode of radius  $r_a$  and an outer cathode of radius  $r_c$ , giving an electric field

$$E(r) = \frac{V}{r \ln(r_c/r_a)} \quad (4.22)$$

at a radial distance  $r$  from the centre of the anode wire, where  $V$  is the potential difference between the anode and the cathode. The output signal is proportional to the number of ions formed and hence the energy deposited by the radiation, but is independent of the applied voltage. However, the signal is very small compared to the noise

of all but the slowest electronic circuits and requires considerable amplification to be useful. Overall, the energy resolution and the time resolution of the chamber are relatively poor and ionisation chambers are of very limited use in recording individual pulses. They are used, for example, as beam monitors, where the particle flux is very large.

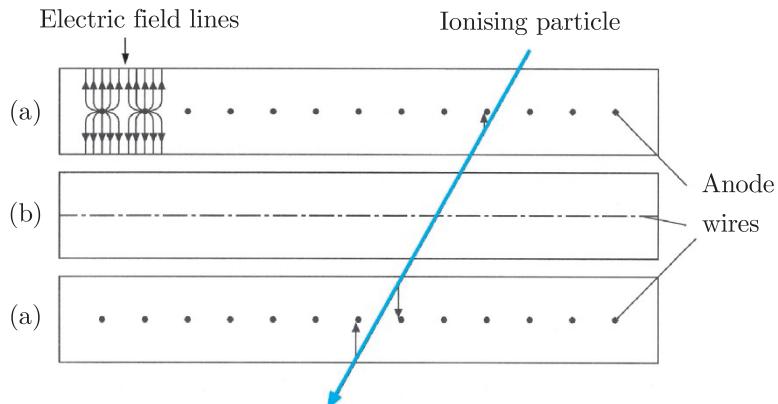
#### 4.4.2(b) Wire chambers

If the voltage is increased beyond the region of operation of the ionisation chamber, we move into the *proportional region*. In this region, a cylindrical arrangement as used in the ionisation chamber will produce electric field strengths or order ( $10^4$ – $10^5$ ) V/cm near the wire and this is strong enough for electron–ion pairs released in the primary ionisation to gain sufficient energy to cause secondary ionisation. The rapid increase in amplification due to secondary ionisation is called a *Townsend avalanche*. The output signal at the anode is still proportional to the energy lost by the original particle. There are a number of different types of device working in the proportional region and they are sometimes generically referred to as *track chambers* or simply *wire chambers*.

The earliest detector using this idea was the *proportional counter*, which consists of a cylindrical tube filled with gas (again a quenching component in the gas is required) and maintained at a negative potential, and a fine central anode wire at a positive potential. Subsequently, the resolution of proportional counters was greatly improved as a result of the discovery that if many anode wires were arranged in a plane between a common pair of cathode plates, each wire acts as an independent detector. This device is called a *multi-wire proportional chamber* (MWPC) and was introduced in 1968. An MWPC can achieve spatial resolutions 50–200 µm and has a typical time resolution of about 2 ns.

A schematic diagram of an MWPC is shown in Figure 4.11. The planes (a) have anode wires into the page and those in plane (b) are at right angles. The wire spacing is typically 2 mm. The cathodes are the faces of the chambers. A positive voltage applied to the anode wires generates a field as shown in the upper corner. A particle crossing the chamber ionises the gas and the electrons drift along the field lines to the anode wires. In this particular example, there would be signals from one wire in the upper (a) chamber and two in the lower (a) chamber.

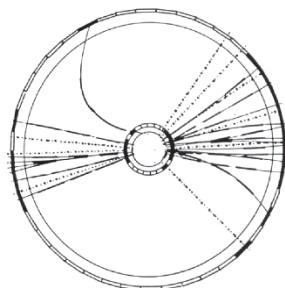
A related device is the *drift chamber*, which has largely replaced the MWPC as a general detector. This uses the fact that the liberated electrons take time to drift from their point of production to the anode. Thus the time delay between the passage of a charged particle through the chamber and the creation of a pulse at the anode



**Figure 4.11** A group of three planes of an MWPC (see text for details). (Povh *et al.* 1999. Reproduced with permission from Springer Science and Business Media.)

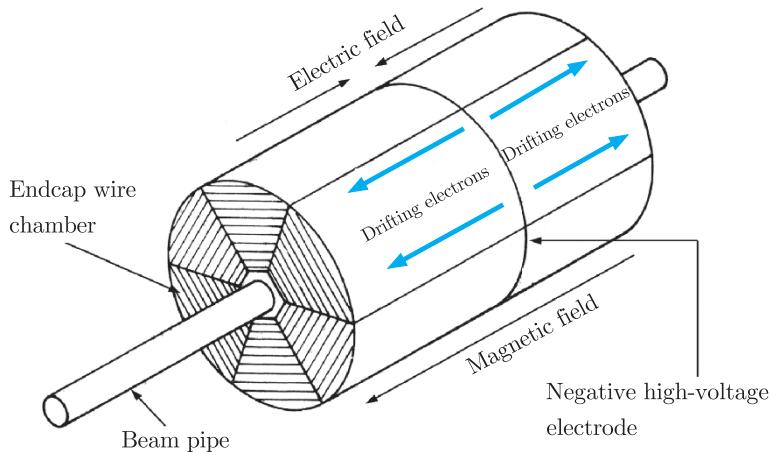
is related to the distance between the particle trajectory and the anode wire. In practice, additional wires are incorporated to provide a relatively constant electric field in each cell in a direction transverse to normal incidence. A reference time has to be defined, which, for example, could be done by allowing the particle to pass through a scintillator positioned elsewhere in the experiment. (Scintillation counters are discussed in Section 4.4.4 below.) The electrons drift for a time and are then collected at the anode, thus providing a signal that the particle has passed. If the drift time can be measured accurately (to within a few ns) and if the drift velocity is known, then spatial and temporal resolutions similar to those of the MWPC can easily be achieved.

Drift chambers are constructed in a variety of geometries to suit the nature of the experiment, and arrangements where the wires are in planar, radial or cylindrical configurations have all been used. The latter type is also called a ‘jet chamber’. Such a chamber was used by the JADE collaboration at an  $e^+e^-$  collider at DESY, Hamburg. It was a cylindrical array of drift chambers with the beam direction as the axis, with the collision region, at which the  $e^+e^-$  interactions occurred, at the centre. The anode wires ran parallel to the axis and the whole detector was divided into 24 segments, with 64 anode wires in each. Figure 4.12 shows an end view along the beam direction of a computer reconstruction of an event observed in the chamber in which the  $e^-$  and  $e^+$  annihilate to produce two ‘jets’ of hadrons. The solid lines indicate the reconstructed charged particle trajectories taking into account the known magnetic field, which is also parallel to the beam direction; the dotted lines indicate the reconstructed trajectories of neutral particles, which were detected outside the chamber by other means.



**Figure 4.12** Computer reconstruction of a typical ‘two-jet’ event observed in the JADE chamber. The solid lines indicate the trajectories of charged particles and the dotted lines those of neutral particles, which were detected outside the chamber by other means. (Adapted from Orito 1979.)

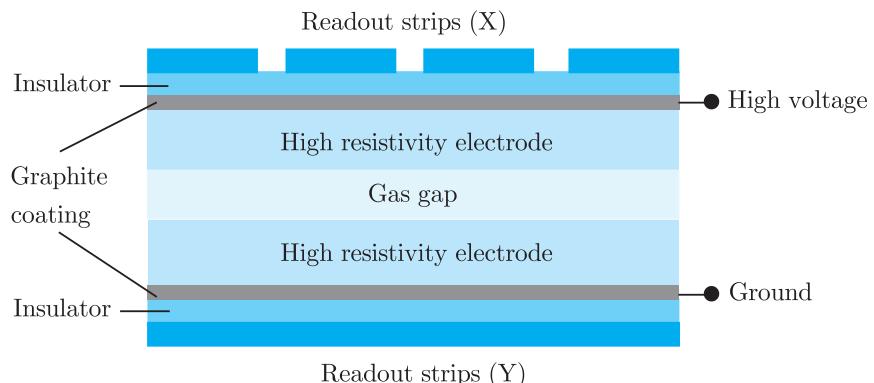
One of the most important applications of proportional and drift chamber principles is embodied in the *time projection chamber* (TPC), which is a device consisting of a chamber, typically a few



**Figure 4.13** Schematic diagram of a time projection chamber. (Kleinknecht 1986. Reproduced with permission from Cambridge University Press.)

metres long with a similar cross-sectional area, containing either a liquid or a gas. A schematic diagram of such a chamber is shown in Figure 4.13. In this example, there is a uniform electric field **E** within the chamber and parallel to its axis, due to a negative high-voltage electrode plane at the centre of the chamber, and a strong magnetic field **B** that is aligned parallel and antiparallel to **E** in the two sections of the chamber with reference to the axis of the cylinder. Electrons formed along the track of an ionising particle traversing the medium drift along helical trajectories whose direction is parallel to the axis of the chamber towards an endcap, which is a segmented layer of proportional counters. The signal amplitudes and arrival times are recorded to provide full three-dimensional measurements of the particle trajectories. Gaseous TPCs are often designed to work in a strong magnetic field so that it operates as a spectrometer, and particle momenta can be estimated from the track curvature. Because of the long drift distances involved, a very precise knowledge of the parameters of TPCs, such as the values of the fields and the purity of the active medium, is essential. TPCs have excellent spatial resolution and have been successfully used in many experiments, such as those involving  $e^+e^-$  annihilation, the STAR detector at the RHIC heavy-ion collider and the ALICE experiment at the LHC.

All wire chambers are limited in their particle localisation performance by basic diffusion processes and space charge effects. In many applications this can be greatly improved by modern advances in photolithographic techniques in which the wires are replaced by conductive metal strips on a printed circuit board. These *microstrip gas chambers* are robust, have good resistance to radiation damage and have enabled spatial resolutions of  $\sim 30\ \mu\text{m}$  to be achieved. In addition, the absence of space charge effects mean that, unlike conventional MWPCs, the gains obtained do not fall off at high particle rates. They have been incorporated in many modern particle physics



**Figure 4.14** Schematic diagram of a resistive-plate chamber.

experiments and have replaced wires in some modern time projection chambers.

#### 4.4.2(c) Resistive-plate chambers

In a single-gap resistive-plate chamber, gas at atmospheric pressure is contained in the narrow space between two parallel plates of high resistivity, typically glass or bakelite, and a high potential difference is maintained between the two plates. As shown in Figure 4.14, the backs of the two electrodes are coated with a lower-resistivity material beyond which are readout strips. When a charged particle passes through the detector it produces ionisation that is drifted towards the electrodes, producing an avalanche. The high resistivity of the electrodes ensures that the discharge does not spread out too much and the readout strips on the two electrodes are orientated at right angles to each other, enabling the particle to be located in the plane of the detector. Resistive-plate chambers have good time resolution. With gaps of order 0.2 mm, resolutions of order 100 ps can be achieved in multi-gap detectors, while detectors with larger gaps, of order 2 mm, are used to trigger other detectors; that is its signal is used to decide whether or not to record information from them. In this respect it is a cheaper alternative to the scintillation counters discussed below when large areas need to be covered.

#### 4.4.2(d) Beyond the region of proportionality

Referring again to Figure 4.10, by increasing the external voltage still further one moves into a region where the output signal ceases to be proportional to the number of ion pairs produced and hence the incident energy. This is the region of *limited proportionality*. Eventually the process runs out of control and we enter the *Geiger–Müller region* where the output signal is independent of the energy lost by the incident particle. In this region a quenching agent is not used.

Detectors working in this region are called Geiger–Müller counters. Physically they are similar to the simple cylindrical proportional counter and are widely used as portable radiation monitors in the context of health physics.

For completeness, we mention that if the gas amplification factor is taken beyond the Geiger–Müller region, the avalanche develops moving plasmas or streamers. Eventually complete breakdown occurs and a spark is emitted as the incident particle traverses the gas. Detectors in this region, called streamer and spark chambers (these were usually of parallel plate construction, rather than cylindrical), were widely used in the 1960s and 1970s and played an important role in hadron physics, but are no longer in general use.

#### 4.4.3 Semiconductor detectors

Solid-state detectors operate through the promotion of electrons from the valence band of a solid to the conduction band as a result of the entry of the incident particle into the solid. The resulting absence of an electron in the valence band (a ‘hole’) behaves like a positron. Semiconductor detectors are essentially solid-state ionisation chambers with the electron–hole pairs playing the role of electron–ion pairs in gas detectors, with the number of pairs being proportional to the energy loss. In the presence of an externally applied electric field, the electrons and holes separate and collect at the electrodes, giving a signal proportional to the energy loss of the incident charged particle.

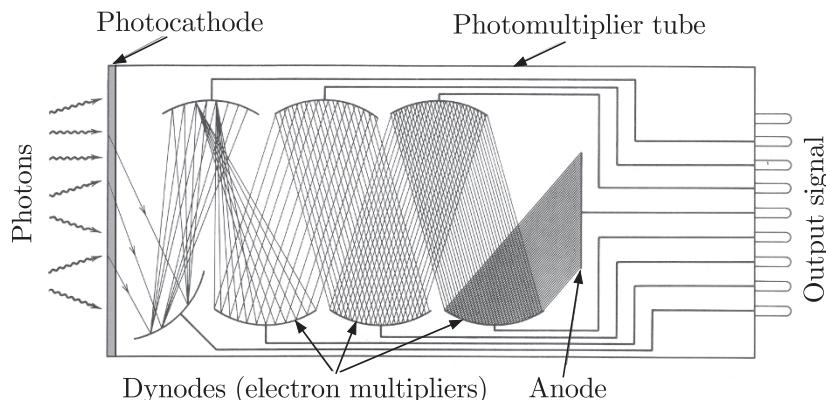
Most semiconductors are manufactured from either silicon or germanium and use the principle of the junction diode. The band gap in some solids is as small as 1 eV and the energy loss required to produce a pair is only 3–4 eV on average, compared to the 30 eV ionisation energy required in a gas detector, or the approximately 300 eV to extract an electron from a photocathode coupled to a plastic scintillator (to be discussed in Section 4.4.4). Thus a very large number of electron–hole pairs with only a small statistical fluctuation will be produced by a low-energy particle. Solid-state detectors are therefore very useful in detecting low-energy particles. Semiconductors are used as a compromise between materials that have residual conductivity sufficient to enable conduction pulses due to single particles to be distinguished above background and those in which the charge carriers are not rapidly trapped in impurities in the material. Such detectors have excellent energy resolution and linearity coupled with a fast response time and have long been used in nuclear physics, but thin planar detectors only became important in particle physics at a later date because of the expense of covering large areas. Nevertheless, more than 200 m<sup>2</sup> of semiconductor detector have been used in experiments at the LHC.

One example of a solid-state detector is a *silicon microstrip detector*, where narrow strips of active detector are etched on to a thin slice of silicon, with gaps of order 10 m, to give a small analogue of an MWPC. Arrays of such strips can then be used to form detectors with resolutions of order 10 mm. These are often placed close to the interaction vertex in a colliding beam experiment, with a view to studying events involving the decay of very short-lived particles. Another example is the *pixel detector*. A single-plane strip detector only gives position information in one dimension (orthogonal to the strip). A pixel detector improves on this by giving information in two dimensions from a single plane. Solid-state ‘vertex detectors’ have become increasingly important in particle physics and have been incorporated in most of the multi-component detectors used at the largest colliders. Their main advantage is their superb spatial resolution; a disadvantage is their limited ability to withstand radiation damage, although newer designs are improving this aspect.

#### 4.4.4 Scintillation counters

For charged particles we have seen that energy losses occur due to excitation and ionisation of atomic electrons in the medium of the detector. In suitable materials, called *scintillators*, a small fraction of the excitation energy re-emerges as visible light (or sometimes in the UV region) during de-excitation. In a scintillation counter this light passes down the scintillator and on to the face of a *photodetector* – a device that converts a weak photon signal to a detectable electric impulse. Short-wavelength light may initially be collected on the material of a *wavelength shifter*, which is a device that shifts the wavelength to higher values that are better matched to the frequency sensitivity of the photodetector.

An important example of a photodetector is the *photomultiplier tube*, a schematic diagram of which is shown in Figure 4.15. Electrons



**Figure 4.15** Schematic diagram of the main elements of a photomultiplier tube. (Adapted from Krane 1988. Reproduced with permission from John Wiley and Sons.)

are emitted from the cathode of the photomultiplier by the photoelectric effect and strike a series of focusing dynodes. These amplify the electrons by secondary emission at each dynode and accelerate the particles to the next stage. The final signal is extracted from the anode at the end of the tube. The electronic pulse can be as short as 100 ps if the scintillator has a short decay time. The scintillation counter is thus an ideal timing device and it is widely used for ‘triggering’ other detectors. They can also be used to distinguish between different types of particle with a common momentum by measuring the ‘time-of-flight’, and hence the velocity, between two scintillation counters a known distance apart. However, the method is limited to relatively low-energy particles because all high-energy particles have velocities close to the speed of light.

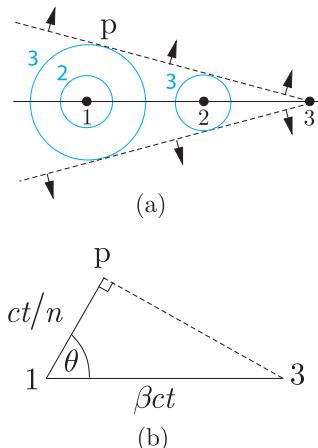
Commonly used scintillators are inorganic single crystals (e.g. cesium iodide) or organic liquids and plastics, with the latter two generally used in particle physics, one reason being that crystal scintillators suffer from radiation damage. Some complex detector systems have used several tons of detector in combination with thousands of photomultiplier tubes. The robust and simple nature of the scintillation counter, together with its low cost and the ease with which a plastic scintillator can be fabricated, has made it a mainstay of experimental particle physics since the earliest days of the subject.

#### 4.4.5 Čerenkov counters and transition radiation

When a charged particle traverses a dispersive medium with a velocity greater than the speed of light in the medium, it produces radiation, called Čerenkov radiation, which is somewhat analogous to the shock wave produced by an aircraft travelling faster than the speed of sound in the atmosphere. Also, when a highly relativistic particle crosses from one medium to another with different optical properties *transition radiation*, usually in the X-ray region, is produced. Although neither of these processes is important in the context of energy loss, both are used in high-energy particle detectors to detect and, particularly, to identify charged particles.

##### 4.4.5(a) Čerenkov counters

An important identification method for high-energy particles is based on the Čerenkov effect. If a charged particle with velocity  $v$  passes through a dispersive medium of refractive index  $n$ , then if  $v$  is greater than the speed of light in the medium  $c/n$ , excited atoms in the vicinity of the particle become polarised and a part of the excitation energy reappears as coherent radiation emitted at a characteristic angle  $\theta$  to the direction of motion. The necessary condition  $v > c/n$



**Figure 4.16** Huygen's construction for the Čerenkov radiation emitted by a particle travelling with speed  $c/n$ , greater than the speed of light in the medium.

implies  $\beta n > 1$ . The angle is found by using Huygen's principle, as used in optics, applied to the construction of Figure 4.16. In Figure 4.16(a), the points 1, 2 and 3 on the axis represent three successive positions of the particle at equally spaced times  $t_1$ ,  $t_2$  and  $t_3$ ; the blue curves labelled 2 and 3 represent the positions of the wavelets emitted from those points at times  $t_2$  and  $t_3$ . The resulting wave-front is indicated by the dashed line and moves in the direction of the arrows. Figure 4.16(b) shows the triangle 13p, where  $t = t_3 - t_1$  and the Čerenkov radiation is emitted at an angle  $\theta$ . It follows that  $\cos \theta = 1/\beta n$  for the angle  $\theta$ , where  $\beta = v/c$  as usual. A determination of  $\theta$  is thus a direct measurement of the velocity.

Čerenkov radiation appears as a continuous spectrum and may be collected on to a photosensitive detector. (A wavelength shifter may also be used for the same reason as in scintillation counters.) Its main limitation from the point of view of particle detection is that very few photons are produced. The number of photons  $N(\lambda) d\lambda$  radiated per unit path length for a particle with charge  $ze$  in a wavelength interval  $d\lambda$  is

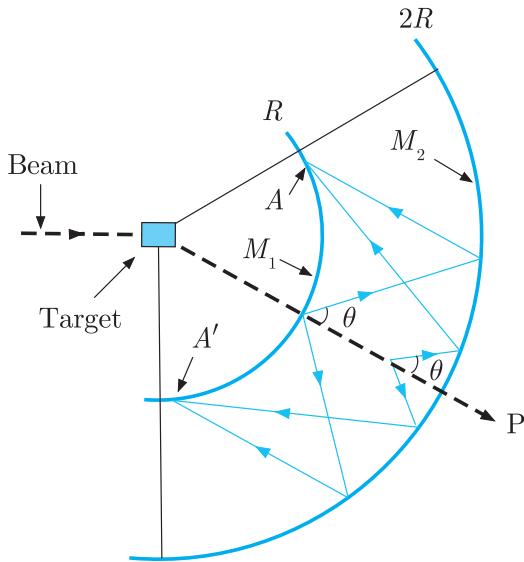
$$N(\lambda) d\lambda = 2\pi\alpha z^2 \left(1 - \frac{1}{\beta^2 n^2}\right) \frac{d\lambda}{\lambda^2} < 2\pi\alpha z^2 \left(1 - \frac{1}{n^2}\right) \frac{d\lambda}{\lambda^2} \quad (4.23)$$

and so vanishes rapidly as the refractive index approaches unity. The maximum value occurs for  $\beta = 1$  and is given for visible Čerenkov radiation in various media in Table 4.4. These numbers should be compared to the  $10^4$  photons/cm emitted by a typical scintillator. Because the yield is so small, appreciable lengths are needed to give enough photons, and gas Čerenkov counters in fixed-target experiments can be several metres long.

Čerenkov counters are used in two different modes. The first is as a *threshold counter* to detect the presence of particles whose velocities exceed some minimum value. Suppose that two particles with  $\beta$  values  $\beta_1$  and  $\beta_2$  at some given momentum  $p$  are to be distinguished. If a medium can be found such that  $\beta_1 n > 1 \geq \beta_2 n$ , then particle 1 will

**Table 4.4** The refractive index  $n$  and the threshold value of  $\gamma_{th}$  for some commonly used Čerenkov radiators, together with the number of photons/cm emitted in the visible region 300–700 nm by a particle with unit charge and  $\beta \approx 1$ .

Medium	$n - 1$	$\gamma_{th}$	photon/cm
Helium (STP)	$3.5 \times 10^{-5}$	120	0.03
$\text{CO}_2$ (STP)	$4.1 \times 10^{-4}$	35	0.40
Silica aerogel	0.025–0.075	4.6 – 2.7	24–66
Water	0.33	1.52	213
Glass	0.46–0.75	1.37 – 1.22	261–331



**Figure 4.17** A particle  $P$ , produced from the target, emits Čerenkov radiation at an angle  $\theta$  on traversing a medium contained between two spheres of radius  $R$  and  $2R$ . The mirror  $M_2$  on the outer sphere focuses the radiation into a ring image at  $A$  and  $A'$  on the inner detector sphere  $M_1$ . The radius of the ring image depends on the angle of emission of the Čerenkov radiation and hence on the velocities of the particles.

produce Čerenkov radiation but particle 2 will not. Clearly, to distinguish between highly relativistic particles with  $\gamma \gg 1$  also requires  $n \approx 1$ , so that from (4.23) very few photons are produced. Nevertheless, common charged particles can be distinguished in this way up to at least  $30 \text{ GeV}/c$ .

Another device is the so-called *ring-image* Čerenkov detector, which is a very important device at both fixed-target machines and colliders. If we assume that the particles are not all travelling parallel to a fixed axis, then the radiating medium can be contained within two concentric spherical surfaces of radii  $R$  and  $2R$  centred on the target or interaction region where the particles are produced, as illustrated in Figure 4.17. The outer surface is lined with a mirror, which focuses the Čerenkov radiation into a ring at the inner detector surface. The radius of this ring depends on the angle  $\theta$  at which the Čerenkov radiation is emitted, and hence on the particle velocity  $v$ . It is determined by constructing an image of the ring electronically. This was the technique used in the SuperKamiokande detector discussed in Chapter 2 to detect relativistic electrons and muons produced by neutrino interactions. In that experiment the radiating medium was pure water.

#### 4.4.5(b) Coherent Čerenkov radiation

When two charged particles traverse a dispersive medium, their radiations will normally add incoherently. However, if the distance between them is smaller than the wavelengths emitted, the Čerenkov radiation emitted by the two particles will be coherent so that interference can occur. Hence if the electron and positron of a pair

produced by a high-energy photon in the field of a nucleus are sufficiently close together, the Čerenkov signal will be suppressed relative to the sum of the two separate signals that would be seen if they were incoherent. However, in some circumstances a shower of electrons and positrons can be produced with a significant net charge. Suppose, for example, that an extremely high-energy photon is incident on a dense material that is transparent to microwaves or radiowaves and produces an initial electron–positron pair. Each of these particles may produce a high-energy photon by bremsstrahlung, which can in turn produce another pair, or alternatively convert an atomic electron to a high-energy electron by Compton scattering, creating a charge asymmetry. These processes can be repeated, generating a shower of many charged particles with an excess of electrons typically of order 20%. Of course, as the shower develops it will spread laterally, but if the Čerenkov radiation is observed in the longer wavelength microwave or radio wave region, an enhanced coherent signal will be produced. This effect was first predicted by Askaryan in 1965 and is called the *Askaryan effect*.

The Askaryan effect was first observed in 2000 and has been used, for example, in an experiment called ANITA based in the Antarctic. The experiment, which has been in operation in various forms since 2006, is designed to shed light on the origin of ultra-high-energy cosmic rays that have been observed with energies up to  $10^{20}$  eV. It does this by using the Askaryan effect to detect the ultra-high-energy neutrinos, up to energies of  $10^{18}$  eV, that result from collisions of the cosmic rays with photons in the cosmic microwave background radiation that pervades the universe. The resulting neutrinos pass through the Antarctic ice sheet and produce charged particles, leading to a strong electromagnetic pulse that propagates through the ice, because the latter is transparent to radio waves up to a frequency of 1.5 GHz. The ice sheet thus converts neutrino energy to radio waves. The ANITA detector system is mounted on a platform suspended from a helium balloon at a height of about 35 km above the Antarctic ice shelf. It traverses a circumpolar flight path due to the continuous wind circulation around the South Pole and ‘sees’ the ice below out to the horizon at about 700 km, resulting in an effective telescope ‘lens’ with an area of approximately  $1.5 \times 10^6$  m<sup>2</sup>.

#### 4.4.5(c) Transition radiation

When a highly relativistic charged particle of mass  $m$  and energy  $\gamma mc^2$  crosses the boundary between a vacuum and a dispersive medium, electromagnetic radiation called *transition radiation* is emitted at a characteristic angle of  $1/\gamma$  to the particle direction. If the particle has charge  $ze$ , the total energy emitted is given by

$\alpha z^2 \gamma \hbar \omega_p / 3$ , where the plasma frequency<sup>17</sup>  $\omega_p$  is given by

$$\hbar \omega_p = \frac{1}{\alpha} (4\pi N_e r_e^3)^{1/2} m_e c^2. \quad (4.24a)$$

Here  $N_e$  is the electron number density and  $r_e = e^2 / (4\pi \epsilon_0 m_e c^2)$  is the classical radius of the electron, so that numerically

$$\hbar \omega_p = [\rho (\text{g/cm}^3) \langle Z/A \rangle]^{1/2} \times 28.81 \text{ eV}, \quad (4.24b)$$

where  $\rho$  is the density of the medium having atomic number  $Z$  and mass number  $A$ .

For typical low- $Z$  materials like styrene,  $\hbar \omega_p \approx 20 \text{ eV}$ , and for  $\gamma \approx 1000$ , the radiation takes the form of soft X-rays with energies between 2 and 40 keV. As can be seen from Figure 4.8, photons of this energy can be detected in high- $Z$  materials by the photoelectric effect. However, the radiation yield from a single interface is small so that simple transition radiation detectors (TRDs) typically consist of a series of low- $Z$  foil radiators separated by high- $Z$  detector layers made of proportional counters filled with a gas mixture containing xenon atoms. The radiator materials, for example mylar or polypropylene, are chosen to minimise absorption. In such multiple-layer detectors interference effects cause the amount of energy emitted, which is proportional to the Lorentz factor  $\gamma$  for a single boundary crossing, to saturate eventually for  $\gamma$  values of order 2000 in a typical detector. Below this limit, the radiation from electrons will be more than two orders of magnitude greater than that from heavier particles like muons and pions, because their much smaller masses lead to much larger values of  $\gamma = E/mc^2$ . Hence such detectors can be operated in ‘threshold mode’ to distinguish electrons with momenta between about 1 and 150 GeV/ $c$  from heavier particles with the same momenta. Such devices have been used in many detectors, such as the ALICE detector at the LHC. Finally, we note that transition radiation can also be used as part of a tracking detector by, for example, distributing a large number of small-diameter ‘straw tube’ detectors through the radiating material, as is done in the ATLAS detector at LHC.

#### 4.4.6 Calorimeters

Calorimeters are an important class of detector used for measuring the energy and direction of a particle (or a collection of particles) by its total absorption. They differ from most other detectors in that the nature of the particle is changed by the detector and by

---

<sup>17</sup>This is actually the plasma frequency that would pertain if all the electrons could move freely through the material.

the fact that they can detect neutral as well as charged particles. A calorimeter may be a homogeneous absorber/detector, for example a scintillator such as CsI. Alternatively, it can be a sandwich construction with separate layers of absorber (e.g. a metal such as lead) and detector (scintillator, MWPC, etc.). The latter are also known as ‘sampling calorimeters’. During the absorption process, the particle will interact with the material of the absorber, generating secondary particles, which will themselves generate further particles and so on, so that a cascade or shower develops. For this reason calorimeters are also called ‘shower counters’. The shower is predominantly in the longitudinal direction due to momentum conservation, but will be subject to some transverse spreading due both to multiple Coulomb scattering and the transverse momentum of the produced particles. Eventually all, or almost all, of the primary energy is deposited in the calorimeter, and gives a signal in the detector part of the device.

There are several reasons why calorimeters are important, especially at high energies:

1. They can detect neutral particles, by detecting the charged secondaries.
2. The absorption process is statistical (and governed by the Poisson distribution), so that the relative precision of energy measurements  $\Delta E/E$  varies like  $E^{-1/2}$  for large  $E$ , which is a great improvement on high-energy spectrometers where  $\Delta E/E$  varies like  $E^2$ .
3. The signal produced can be very fast, of order 10–100 ns, and is ideal for making triggering decisions.

Since the characteristics of electromagnetic and hadronic showers are somewhat different it is convenient to describe each separately. In practice, it is common to have both types in one experiment with the hadron calorimeter stacked behind the electromagnetic one, although it is possible to design calorimeters that detect both classes of particle.

#### 4.4.6(a) *Electromagnetic showers*

When a high-energy electron or positron interacts with matter we have seen that the dominant energy loss is due to bremsstrahlung, and for the photons produced the dominant absorption process is pair production. Thus the initial electron will, via these two processes, lead to a cascade of  $e^\pm$  pairs and photons, and this will continue until the energies of the secondary electrons fall below the critical energy  $E_C$  where ionisation losses equal those from bremsstrahlung. This energy is roughly given by  $E_C \approx 600 \text{ MeV}/Z$ .

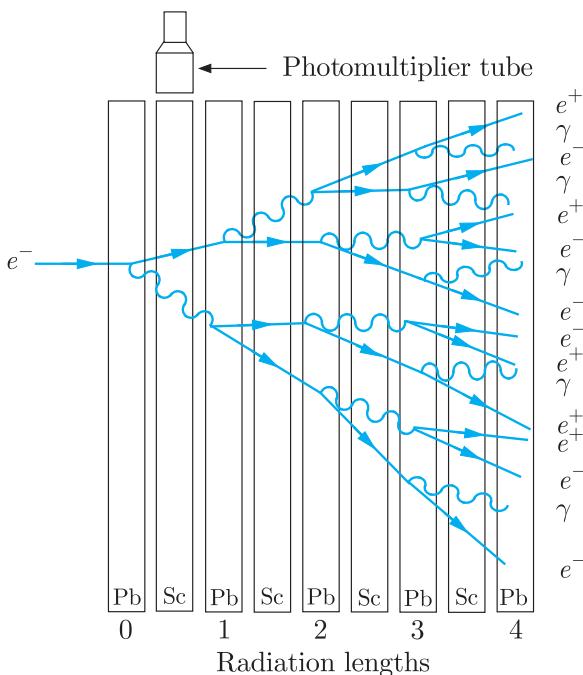
Most of the correct qualitative features of shower development may be obtained from the following very simple model. We assume:

1. Each electron with  $E > E_C$  travels one radiation length and then gives up half of its energy to a bremsstrahlung photon.
2. Each photon with  $E > E_C$  travels one radiation length and then creates an electron–positron pair with each particle having half the energy of the photon.
3. Electrons with  $E < E_C$  cease to radiate and lose the rest of their energy by collisions.
4. Ionisation losses are negligible for  $E > E_C$ .

A schematic diagram of the approximate development of a shower in an electromagnetic calorimeter assuming this simple model is shown in Figure 4.18.

If the initial electron has energy  $E_0 \gg E_C$ , then after  $t$  radiation lengths the shower will contain  $2^t$  particles, which consist of approximately equal numbers of electrons, positrons and photons each with an average energy

$$E(t) = E_0/2^t.$$



**Figure 4.18** Approximate development of an electromagnetic shower in a sampling calorimeter assuming the simple model of the text. The calorimeter consists of alternate layers of lead (Pb) and a scintillator (Sc), the latter attached to photomultipliers (one only shown). In this figure the transverse spread of secondary particles has been exaggerated for reasons of clarity.

The multiplication process will cease abruptly when  $E(t) = E_C$ , that is at  $t = t_{\max}$ , where

$$t_{\max} = t(E_C) \equiv \frac{\ln(E_0/E_C)}{\ln 2} \quad (4.25)$$

and the number of particles at this point will be

$$N_{\max} = \exp(t_{\max} \ln 2) = E_0/E_C. \quad (4.26)$$

The main features of this simple model are observed experimentally, and in particular the maximum shower depth increases only logarithmically with primary energy. Because of this, the physical sizes of calorimeters need increase only slowly with the maximum energies of the particles to be detected. The energy resolution of a calorimeter, however, depends on statistical fluctuations, which are neglected in this simple model, but for an electromagnetic calorimeter it is typically  $\Delta E/E \approx 0.05/E^{1/2}$ , where  $E$  is measured in GeV.

#### 4.4.6(b) Hadronic showers

Although hadronic showers are qualitatively similar to electromagnetic ones, shower development is far more complex because many different processes contribute to the inelastic production of secondary hadrons. The scale of the shower is determined by the nuclear absorption length defined earlier. Since this absorption length is larger than the radiation length, which controls the scale of electromagnetic showers, hadron calorimeters are thicker devices than electromagnetic ones. Another difference is that some of the contributions to the total absorption may not give rise to an observable signal in the detector. Examples are nuclear excitation and leakage of secondary muons and neutrinos from the calorimeter. The loss of ‘visible’ or measured energy for hadrons is typically 20–30% greater than for electrons.

The energy resolution of calorimeters is in general much worse for hadrons than for electrons and photons because of the greater fluctuations in the development of the hadron shower. Depending on the proportion of  $\pi^0$  produced in the early stages of the cascade, the shower may develop predominantly as an electromagnetic one because of the decay  $\pi^0 \rightarrow \gamma\gamma$ . These various features lead to an energy resolution typically a factor of 5–10 poorer than in an electromagnetic calorimeter.

## 4.5 Detector systems and accelerator experiments

In this section, we illustrate how the basic detectors we have discussed are combined into large multi-component systems and used in colliding beam experiments. Our first example is the discovery of the  $W^\pm$  and  $Z^0$  bosons at CERN. This is discussed in detail to show

how the resulting data are analysed. Then we look at three more recent detector systems, leaving the physics results to later chapters. These are: the ATLAS detector at CERN, which is the largest experiment at the LHC; the STAR detector at Brookhaven National Laboratory; and the LHCb detector, also at the LHC. Further historic collider experiments, including the discoveries of the top quark and the Higgs boson, are discussed in some detail in later chapters.

Before proceeding, however, we should emphasise that fixed-target experiments played a dominant role in the early history of accelerator physics and they continue to be important for specialised purposes, as illustrated in Chapters 2 and 3. Several other examples will be discussed in later chapters, including the inelastic scattering of electrons and neutrinos in Chapter 8 and the detection of  $CP$  violation in the decays of neutral  $K$ -mesons in Chapter 11.

#### 4.5.1 Discovery of the $W^\pm$ and $Z^0$ bosons

The  $W^\pm$  and  $Z^0$  bosons are highly unstable particles which were first produced in the reactions

$$p + \bar{p} \rightarrow W^+ + X^-, \quad (4.27a)$$

$$p + \bar{p} \rightarrow W^- + X^+ \quad (4.27b)$$

and

$$p + \bar{p} \rightarrow Z^0 + X^0, \quad (4.27c)$$

where  $X^\pm$  and  $X^0$  are arbitrary hadronic states allowed by the conservation laws. The heavy bosons were detected via their subsequent decays

$$W^+ \rightarrow \ell^+ + \nu_\ell, \quad (4.28a)$$

$$W^- \rightarrow \ell^- + \bar{\nu}_\ell \quad (4.28b)$$

and

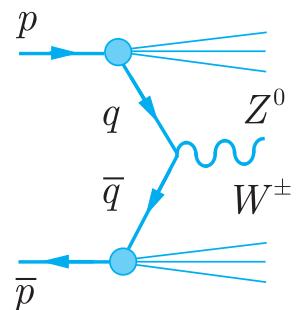
$$Z^0 \rightarrow \ell^+ + \ell^-, \quad (4.28c)$$

where the charged leptons  $\ell^\pm$  were either muons or electrons. (These are not the only possible decay modes, as we shall see later.) The production mechanism that gives rise to the reactions (4.27) is shown in Figure 4.19 and involves quark–antiquark annihilation processes such as<sup>18</sup>

$$u + \bar{d} \rightarrow W^+, \quad d + \bar{u} \rightarrow W^- \quad (4.29a)$$

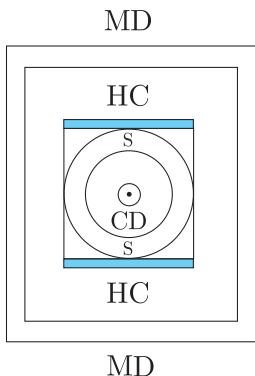
and

$$u + \bar{u} \rightarrow Z^0, \quad d + \bar{d} \rightarrow Z^0 \quad (4.29b)$$



**Figure 4.19** Mechanism for producing  $W^\pm$  and  $Z^0$  bosons in proton–antiproton annihilations.

<sup>18</sup>The interactions of the  $W^\pm$  and  $Z^0$  bosons with leptons and quarks will be discussed more systematically in Chapters 9 and 10.



**Figure 4.20** Schematic diagram showing a cross-section of the UA1 detector seen along the beam direction. The beams collide at the central point, which is surrounded by a central tracking detector CD, shower counters S, hadron calorimeters HC and muon detectors MD. The shaded areas indicate the coils used to produce the magnetic field in the central detector.

for charged and neutral vector bosons, respectively. In the centre-of-mass of the  $q\bar{q}$  system, the total energy of the  $q\bar{q}$  pair must be at least 80 GeV or 91 GeV, corresponding to the production of either a  $W^\pm$  or a  $Z^0$  boson at rest. However, the energy of the  $p\bar{p}$  system has to be considerably higher for a reasonable reaction rate to occur, because each quark (or antiquark) has only a fraction of the parent proton (or antiproton) energy.

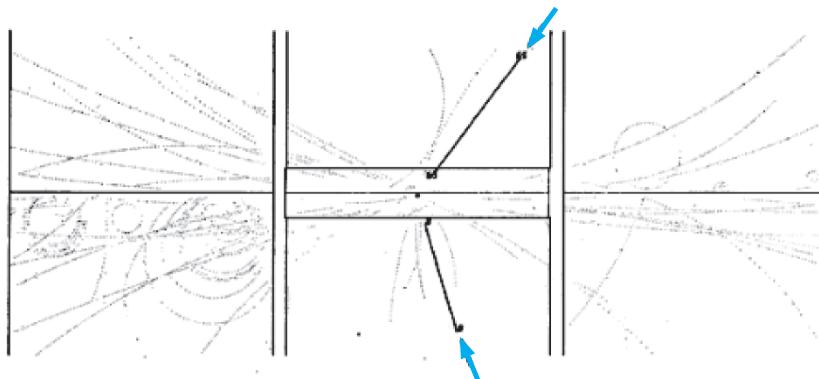
The  $W^\pm$  and  $Z^0$  bosons were first discovered using a proton–antiproton collider at CERN, which was specifically built for this purpose. At the time it had proton and antiproton beams with maximum energies of 270 GeV each, giving a total centre-of-mass energy of 540 GeV. Two independent experiments were mounted, called UA1 and UA2.<sup>19</sup> In each case the proton and antiproton beams were brought together in an intersection region that lay at the centre of a very large and elaborate detector system. We will not describe this in detail, but simply list the main components, which are shown schematically in Figure 4.20 for the case of the UA1 experiment. The order of the detectors reflects the penetrating power of the various types of particles discussed in Section 4.3.5. As one moves out from the centre there is:

1. A central tracking detector used to observe charged particles and to measure their momentum from the curvature of the tracks in an applied magnetic field.
2. A set of electromagnetic shower counters that absorb and detect both electrons, which are also observed in the central detector, and photons, which are not.
3. A set of hadron calorimeters, which absorb and detect both neutral and charged hadrons.
4. A set of counters to identify muons, which are the only charged particles to penetrate the hadron calorimeters.

Thus only neutrinos escape detection. An example of a reconstructed event from the UA1 detector is shown in Figure 4.21. We shall see below that it is in fact a  $Z^0$  production event.

One of the main problems facing the experimenters was that for each event in which a  $W^\pm$  or  $Z^0$  is produced and decays to leptons, there are more than  $10^7$  events in which hadrons alone are produced. The extraction of a ‘signal’ in the presence of such a large ‘background’ is only possible because the leptons arising from the decays (4.28) have large momenta (because the  $W^\pm$  or  $Z^0$  bosons are very heavy) and are often emitted at wide angles to the initial beam

<sup>19</sup>The accelerator lies below ground level, and UA1 and UA2 stand for Underground Areas 1 and 2, where the experiments were performed.



**Figure 4.21** Computer reconstruction of a  $Z^0$  production event observed in the UA1 detector. The dotted lines correspond to hadrons produced together with the  $Z^0$  and the solid lines to the electron and positron produced in its decay. (Adapted from Sadoulet 1983.)

directions. In other words, the leptons often have large ‘transverse momenta’. In contrast, the hadrons produced in proton–antiproton collisions, and the leptons that arise from their decays, very rarely have such large transverse momenta. Hence the background of purely hadronic reactions could almost be entirely eliminated by retaining only those events containing at least one electron or muon with a large transverse momentum.<sup>20</sup> This subset of events was then examined to see if any corresponded to  $W^\pm$  or  $Z^0$  production followed by the decays (4.28).

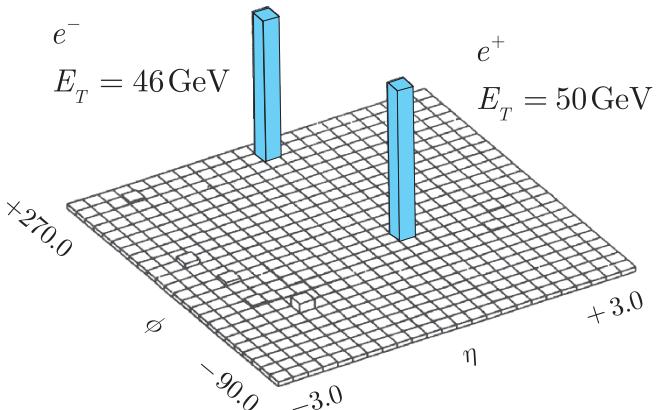
#### 4.5.1(a) $Z^0$ events

A typical event in which a  $Z^0$  was produced and decayed to give an  $e^+e^-$  pair is shown in Figure 4.21. The electron and positron were both detected in the central detector, where they gave rise to tracks that were almost straight, despite the applied magnetic field, because they had very large momenta. They were identified as electrons, and not muons or hadrons, from their signal in the electromagnetic shower counters that surrounded the central detector. The  $e^+e^-$  pair stands out dramatically in the output from them, as shown in Figure 4.22.

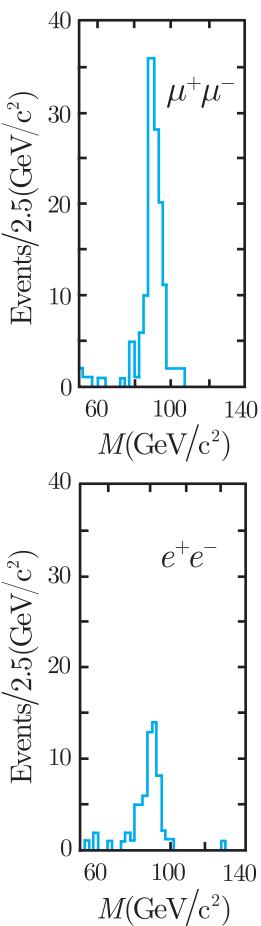
Finally, if the observed  $e^+e^-$  pair indeed arose from the decay of a  $Z^0$ , then its invariant mass should correspond to the mass of the decaying particle, which was predicted in the electroweak theory<sup>21</sup> to be about  $90\text{ GeV}/c^2$ . For the event shown, this is indeed the case within experimental errors. It was also confirmed for the other twelve events observed by the UA1 and UA2 collaborations in their first experiments. Subsequently, the experiment has been repeated with

<sup>20</sup>For example, in the UA1 experiment this lepton was required to be emitted at more than  $5^\circ$  to the beam direction and to have a transverse momentum in excess of  $10\text{ GeV}/c$ .

<sup>21</sup>This prediction is discussed in Section 10.1.2.



**Figure 4.22** Output from the electromagnetic shower counters corresponding to the  $Z^0$  event shown in Figure 4.21. Each square element corresponds to a particular direction in space and the height of the ‘tower’ to the energy of the particle detected. (Adapted from Sadoulet 1983.)



**Figure 4.23**  $Z^0$  peaks observed in the  $e^+e^-$  and  $\mu^+\mu^-$  mass distributions from a Fermilab experiment by the CDF collaboration. (Adapted from Campbell 1989.)

much higher statistics, and in Figure 4.23 we show the  $e^+e^-$  and  $\mu^+\mu^-$  mass distributions obtained in an experiment at Fermilab. A clear peak is seen in both cases, corresponding to a mass of  $90.9 \pm 0.4 \text{ GeV}/c^2$  and a width of order 3 GeV. These parameters have since been measured much more accurately by other methods,<sup>22</sup> leading to a mass

$$M_Z = 91.1876 \pm 0.0021 \text{ GeV}/c^2 \quad (4.30\text{a})$$

and a width

$$\Gamma_Z = 2.4952 \pm 0.0023 \text{ GeV}, \quad (4.30\text{b})$$

corresponding to a lifetime of  $2.6 \times 10^{-25}$  s. The branching ratios to  $e^+e^-$  and  $\mu^+\mu^-$  are both approximately 3%.

#### 4.5.1(b) $W^\pm$ events

In this case, the  $W^\pm$  bosons decay by (4.28a) and (4.28b) to give a charged lepton plus an unseen neutrino. Thus a typical  $W^\pm$  event looks just like a typical  $Z^0$  event (for example that shown in Figure 4.21) except that one of the electron or muon tracks is replaced by an invisible neutrino. Because the neutrino is expected to have a large transverse momentum, its presence is revealed by summing the transverse momenta of all the observed particles. If the sum is not zero, as required by momentum conservation, the ‘missing transverse momentum’ can only be due to a neutrino, as this is the only particle that is not detected. In summary, an event in which a  $W^\pm$  is created and then decays via (4.28a) and (4.28b) is characterised by a charged lepton with a large transverse momentum,<sup>23</sup> and overall there is a large missing transverse momentum.

<sup>22</sup>See Section 10.1.3.

<sup>23</sup>Occasionally a high-momentum  $e^\pm$  or  $\nu^\pm$  from the decay of a  $W^\pm$  will be emitted close to the beam direction. Such events are lost in the background and escape detection.

In all, a total of 43 events were observed in the original UA1 experiment, where the charged lepton was either an electron or a positron. If the  $W^\pm$  were produced with zero transverse momentum, then momentum conservation would imply that the transverse momentum of the  $e^\pm$  produced in the decay would be precisely equal and opposite to the missing transverse momentum ascribed to the neutrino. Since the lepton masses are negligible, this in turn implies that the corresponding transverse energies  $E_T = (p_T^2 c^2 + m^2 c^4)^{1/2} \approx p_T c$  are approximately equal, and this is approximately true for the observed events as shown in Figure 4.24. This result had been predicted theoretically, and the observed events conformed in every way to what was expected for the production and decay of a  $W^\pm$  boson with a mass of approximately  $80 \text{ GeV}/c$ . We shall not discuss this further because the analysis is rather complicated, but merely note that the results have been confirmed by later experiments with higher statistics. These yield a mass value of

$$M_W = 80.403 \pm 0.029 \text{ GeV}/c^2 \quad (4.31\text{a})$$

and a decay width of

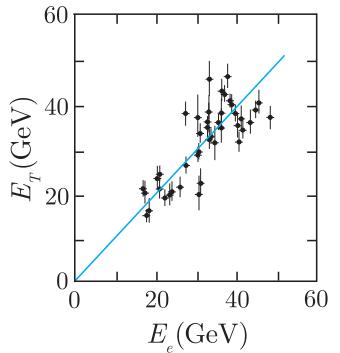
$$\Gamma_W = 2.141 \pm 0.041 \text{ GeV}, \quad (4.31\text{b})$$

which corresponds to a lifetime of  $3.2 \times 10^{-25}$  s. The branching ratio for each of the decay modes  $e\nu_e$  and  $\mu\nu_\mu$  is approximately 10%.

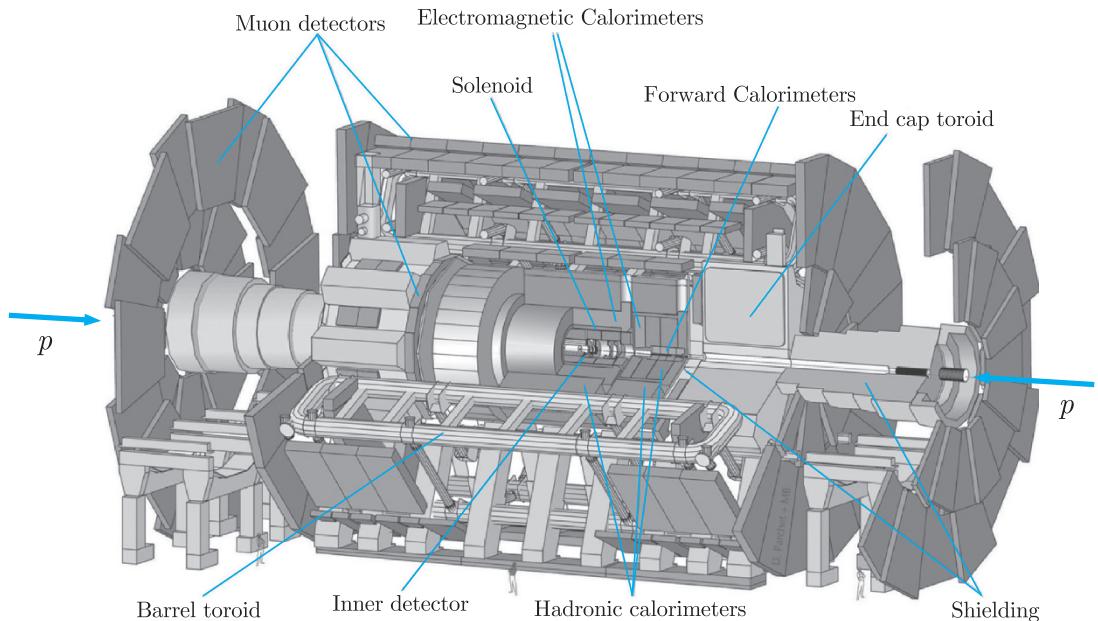
### 4.5.2 Some modern detector systems

The order of the detectors shown schematically in Figure 4.19 is dictated by the penetrating power of the various particles discussed in Section 4.3.5, and is typical of multi-purpose detector systems at high-energy colliders. However, actual detectors are much more complicated when considered in detail, and often include other features, like scintillation counters used for triggering, as discussed in Section 4.4.4. We will illustrate this by some examples.

The first of these is ATLAS, which together with CMS, is one of two broadly similar general-purpose detectors that were built at the LHC and have been operational since 2008. (They are shown during their construction in Plates 3 and 4.) ATLAS is shown in a cutaway view in Figure 4.25. It measures about 25 m in diameter and 46 m in length, with an overall weight of approximately 7000 tonnes. It has the usual basic arrangement of an inner tracking detector, calorimeters and muon counters. The tracking detector has a radius of 1.15 m, a length of 7 m and is composed of three components. In order of increasing radius, the first two are composed of semiconductor pixel detectors and silicon strip detectors, which provide precise measurements close to the vertex region. Beyond this, the third



**Figure 4.24** The ‘missing transverse energy’  $E_T$  plotted as a function of the measured transverse energy  $E_e$  of the electron for the first  $W^\pm$  production events observed in the UA1 experiment. (Adapted from Spiro 1983.)

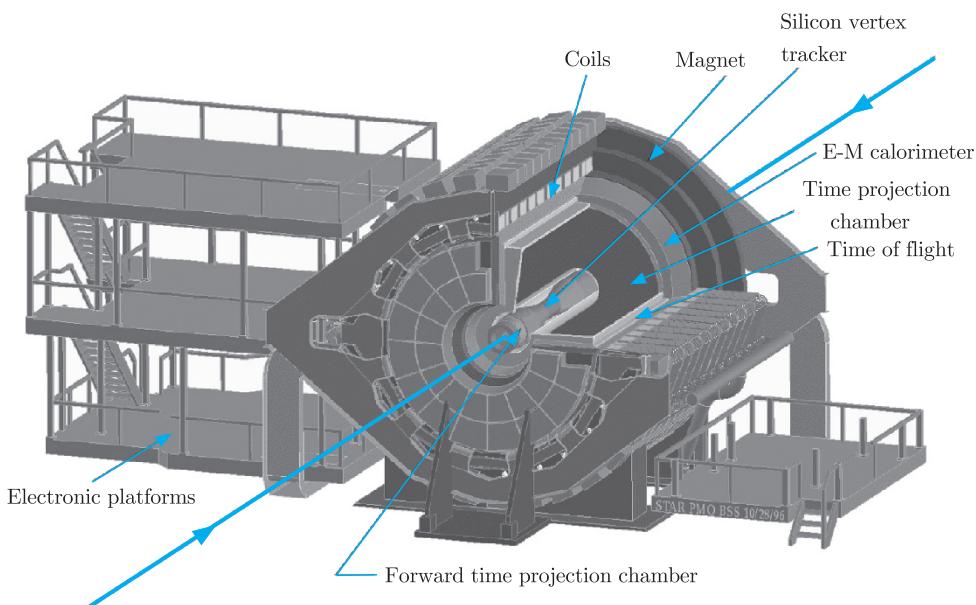


**Figure 4.25** The ATLAS detector used at the  $pp$  collider LHC at CERN, Geneva.  
(Reproduced by permission of CERN.)

component contains a large number of transition radiation ‘straw tube’ detectors, each of diameter 4 mm, which, in addition to tracking the particles, also aid in particle identification. Both the electromagnetic and hadron calorimeters are of the sandwich form, in which the active elements are liquid argon and plastic scintillator, respectively. The muon counters are composed of straw tube transition radiation detectors, but of larger radius than those used in the inner tracking detector. In addition, there is a complex magnetic system, involving both solenoidal and toroidal components, to enable the momenta of the charged particles to be determined from the curvature of their paths.

One of the major aims of both the ATLAS and CMS detectors was to search for the Higgs boson, the discovery of which was finally announced in 2012. (This is discussed in detail in Section 10.2.4). Now that the Higgs boson is established, both ATLAS and CMS will play a major role in measuring its properties, as well as searching for possible new phenomena, including supersymmetry (cf. Chapter 12).

The second example is the STAR detector, shown in Figure 4.26 and colour Plate 5, weighs 1200 tons and is about the size of a two-storey house. It has been in use at the RHIC accelerator at Brookhaven National Laboratory since 2000. Like the ALICE detector at CERN, its main purpose is to detect signals for a new form of matter that might be formed at very high densities and

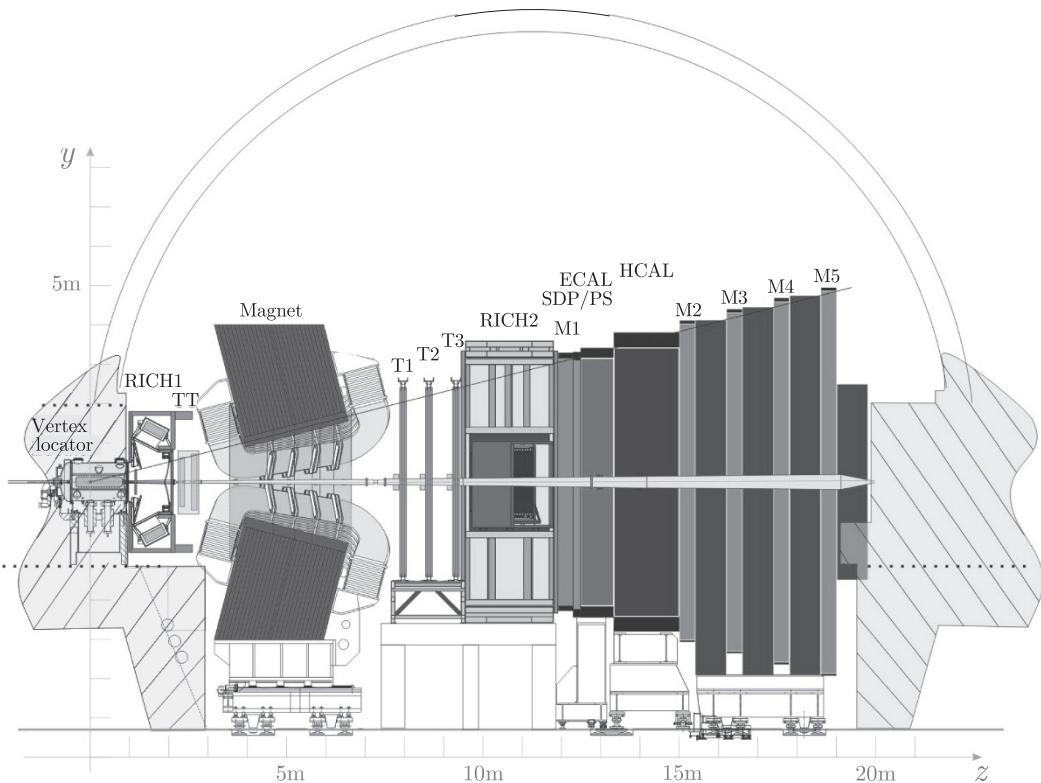


**Figure 4.26** The STAR detector at the RHIC accelerator at Brookhaven National Laboratory, USA. (Courtesy of Brookhaven National Laboratory.)

temperatures.<sup>24</sup> This is done by studying events resulting from the high-energy collisions of heavy ions, typically those of fully stripped gold nuclei, and the detector is designed to handle the many thousands of particles typically produced in such collisions. At its heart is a large time projection chamber (cf. Plate 5), which can not only track many particles simultaneously but also provides partial particle identification using ionisation energy losses. In addition, as can be seen from Figure 4.26, other components include a solid-state vertex detector, an electromagnetic calorimeter and a ‘time-of-flight’ system, composed of several thousand multi-gap resistive plate chambers, which is also used for particle identification.

The third system is the LHCb detector, shown in Figure 4.27. This detector, as its name suggests, is also in the LHC ring, but was designed to study the properties of hadrons containing  $b$  and  $c$  quarks and their decay products. Because these particles are produced almost entirely at small angles to the beam, it has a quite different geometry to a general-purpose detector like ATLAS, which completely surrounds the collision point in order to detect particles produced and emitted in all directions. In contrast, the 5600-tonne LHCb detector is made up mainly of a forward spectrometer of conical shape with maximum dimensions 20 m long, 10 m high and 13 m wide, placed to one side of the collision point and surrounding

<sup>24</sup>This is called a ‘quark–gluon plasma’ and is discussed in Section 7.1.4.



**Figure 4.27** The LHCb detector at the LHC in CERN. (Reproduced by permission of CERN.)

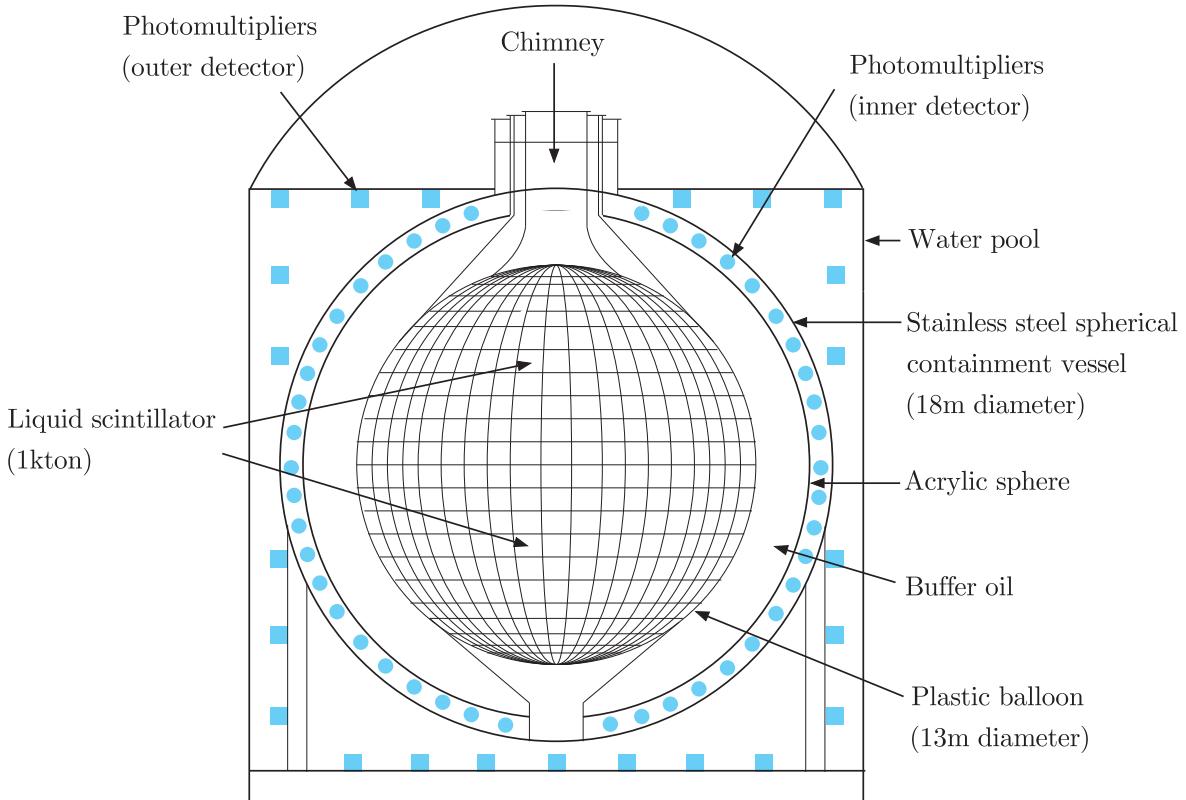
the beamline, as shown in Figure 4.27. In addition, the ‘long-lived’ particles of most interest ( $B$ ,  $B_s$  and  $D$  mesons, for example) have lifetimes of order  $10^{-12}$  s and typically travel only a few millimetres before decaying, so that very good resolution is required close to the interaction point. This is done using a solid-state detector called the VELO (vertex locator).  $B$  mesons produced in the collision, for example, will decay before they reach the detector, but the VELO can locate the primary vertex and  $B$ -decay vertices from very precise measurements of their decay products. The elements of the VELO approach to within 5 mm of the collision point and to protect them from damage they are moved further away while the accelerator beams are injected and stabilised. They are only restored to their operating positions when it is safe to do so. The rest of the detector includes two ring image Čerenkov counters (RICH1-2) for particle identification, required to identify the final-state charged pions, kaons and protons from  $B$  and other decays, together with the usual sequence of tracking detectors (T1-3), electromagnetic and hadron calorimeters (ECAL, HCAL) and muon chambers (M1-5), as shown on the diagram. Close to the beam pipe, the trackers are

composed of silicon microstrip detectors to give high resolution. Farther out they contain thousands of gas-filled straw tubes, and the time taken for the ionisation produced to drift to the central anode wire is used to precisely determine the position. The muon detectors are multi-wire proportional chamber systems containing about 1400 chambers and about 2.5 million wires. The detector has been operating since 2008 and its main aim, among others, is to investigate the small differences between matter and antimatter by comparing the properties of  $B$  and  $\bar{B}$  mesons, as discussed in Chapter 11.

## 4.6 Non-accelerator experiments

Although in previous sections we have concentrated on accelerator experiments, non-accelerator experiments have been important historically and have become increasingly important again, as illustrated in Chapters 2 and 3. In contrast to collider experiments, which usually involve high event rates and where very large number of particles are produced in each event, non-accelerator experiments are frequently designed to detect rare events in which small numbers of particles are produced. This results in marked differences in the nature of the detectors required, which are less elaborate than the multi-component, multi-purpose detectors used at colliders, since relatively few particles are produced. In addition, their design is very dependent on the particular reaction, or reactions, studied. Nonetheless, they often have certain characteristic features. Firstly, in order to increase the event rate, there is a very large ‘active volume’ in which events can occur. Secondly, because event rates are intrinsically small, it is necessary to be especially careful to minimise ‘accidental’ events that might mimic real events. The detectors are therefore often situated deep underground to shield them from cosmic rays. In addition, very pure materials are required in the construction of the detectors themselves, to eliminate traces of radioactive impurities that might give rise to spurious events, and shielding or other methods are required to deal with ambient radioactivity from the surroundings of the detector.

All of these features are realised in the SuperKamiokande detector shown in Figure 2.9, which was used to detect neutrinos with energies above 1 GeV produced by the secondary interactions of cosmic rays in the Earth’s atmosphere, as discussed in Section 2.3.2(a). They are also exhibited by the KamLand detector, which was used to detect the much lower energy antineutrinos from nuclear reactors, yielding the results discussed in Section 2.3.2(c). A schematic diagram of the latter detector, which is located in the Kamioka mine in Japan, is shown in Figure 4.28. At the centre is a plastic balloon filled with a thousand tons of liquid scintillator. This is surrounded by a layer of



**Figure 4.28** Schematic diagram of the KamLAND detector. (Suzuki 2005. Reproduced with permission from the Institute of Physics.)

buffer oil contained between an acrylic sphere and a stainless steel sphere together with about two thousand photomultipliers, as shown. The buffer oil shields the scintillator from radioactivity emanating from the photomultipliers. Antineutrinos are detected by observing the reaction  $\bar{\nu}_e + p \rightarrow e^+ + n$  in the scintillator, which acts as both target and detector. When this occurs, the scintillation light generated by the positron, and by the higher-energy photon produced when the neutron is subsequently captured in the reaction  $n + p \rightarrow d + \gamma$ , passes through the buffer oil and is detected by the photomultipliers. Finally, the steel sphere is itself surrounded by a water Čerenkov counter, which is used to detect any cosmic ray muons that reach the detector, despite its location deep underground.

SuperKamiokande and KamLand are examples of neutrino detectors. In addition, various important theoretical hypotheses predict the existence of other rare processes and their detection, or alternatively their exclusion, at the predicted rate, is needed to either confirm or rule out the associated theories. Several important experiments of this type are non-accelerator experiments, including the

searches for proton decays, neutrinoless double  $\beta$ -decay and dark matter, which are all discussed in Chapter 12. As we shall see, the detectors used, although varied in many ways, exhibit some of the same characteristic features – large active regions, underground locations and very pure materials – as the neutrino detectors described above.

## Problems 4

- 4.1** What is the length  $L$  of the longest drift tube in a linac, which, operating at a frequency  $f = 50$  MHz, is capable of accelerating protons to a maximum energy of  $E = 300$  MeV?
- 4.2** Show that the momentum  $p$  in  $\text{GeV}/c$  for a relativistic particle of charge  $e$  moving in a circular orbit of radius  $\rho$  metres in a magnetic field of magnitude  $B = |\mathbf{B}|$  tesla is given by  $p = 0.3B\rho$ .
- 4.3** At a collider, two high-energy particles  $A$  and  $B$  with energies  $E_A$  and  $E_B$ , which are much greater than either of the masses, collide at a crossing angle  $\theta$ . Show that the total centre-of-mass energy  $E_{\text{CM}}$  is given by

$$E_{\text{CM}}^2 = 2E_A E_B (1 + \cos \theta)$$

if the rest masses are neglected. At the HERA collider (which was operational until 2007) in the DESY laboratory in Hamburg, a 30 GeV electron beam collided with an 820 GeV proton beam at zero crossing angle. Evaluate the total centre-of-mass energy and show that a fixed-target electron accelerator would require a beam energy of approximately  $5 \times 10^4$  GeV to achieve the same total centre-of-mass energy.

- 4.4** An  $e^+e^-$  collider has a diameter of 8 km and produces beams of energy 45 GeV. Each beam consists of 12 bunches each containing  $3 \times 10^{11}$  particles. The bunches have a cross-sectional area of  $0.02 \text{ mm}^2$ . What is the luminosity of the machine in units of  $\text{cm}^{-2} \text{ s}^{-1}$ ?
- 4.5** A liquid hydrogen target of volume  $125 \text{ cm}^3$  and density  $0.071 \text{ g/cm}^3$  is bombarded with a monoenergetic beam of negative pions with a flux  $2 \times 10^7/\text{m}^2 \text{ s}$  and the reaction  $\pi^- + p \rightarrow \pi^0 + n$  observed by detecting the photons from the decay of the  $\pi^0$ . Calculate the number of photons emitted from the target per second if the cross-section is 40 mb.
- 4.6** Charged pions with momentum  $50 \text{ GeV}/c$  are deflected through a collimator slit 3 mm wide by a bending magnet 2 m long that produces a uniform field of 1.5 T. Calculate how far from the magnet the slit should be placed so that it accepts particles with momentum within about 1% of the central value.
- 4.7** Estimate the thickness of iron through which a beam of neutrinos with energy 200 GeV must travel if 1 in  $10^9$  of them is to interact. Assume that at high energies the neutrino–nucleon total cross-section is given approximately by  $\sigma_\nu \approx 10^{-38} E_\nu \text{ cm}^2$ , where  $E_\nu$  is given in GeV. The density of iron is  $\rho = 7.9 \text{ g/cm}^3$ .

- 4.8** A positron with laboratory energy 50 GeV interacts with the atomic electrons in a lead target to produce  $\mu^+ \mu^-$  pairs. If the cross-section for this process is given by

$$\sigma = 4\pi\alpha^2 \hbar^2 c^2 / 3E_{\text{CM}}^2,$$

calculate the positron's effective interaction length. The density of lead is  $1.14 \times 10^7 \text{ kg/m}^3$ .

- 4.9** An electron with an initial energy of 2 GeV traverses 10 cm of water with a radiation length of 36.1 cm. Calculate its final energy. How would the energy loss change if the particle were a muon rather than an electron?
- 4.10** Derive the result of Equation (4.12) for the range  $R$ .
- 4.11** Assuming the Bethe–Bloch formula is valid for low energies, show that the rate of ionisation has a maximum (the Bragg peak) and find the kinetic energy of protons in iron for which this maximum would occur.
- 4.12** Two particles of masses  $m_1$  and  $m_2$ , and common momentum  $p$ , travel between two scintillation counters that are a distance  $L$  apart. Show that the difference in their flight times decreases like  $p^{-2}$  for large momenta. Calculate the minimum flight path necessary to distinguish pions from kaons if they have momentum  $3 \text{ GeV}/c$  and the time-of-flight can be measured with an accuracy of 200 ps.
- 4.13** Estimate the minimum length of a gas Čerenkov counter that could be used in the threshold mode to distinguish between pions and kaons with momentum  $20 \text{ GeV}/c$ . Assume that 200 photons need to be radiated to ensure a high probability of detection. Assume also that the radiation covers the whole visible spectrum and neglect the variation with wavelength of the refractive index of the gas.
- 4.14** A cylindrical proportional chamber has a central anode wire of radius 0.02 mm and an outer cathode of radius 10 mm with a voltage of 500 volts applied between them. What is the electric field at the surface of the anode? If the threshold for ionisation by collision is  $750 \text{ kV/m}$  and the mean free path of the particles being detected is  $4 \times 10^{-6} \text{ m}$ , estimate the number of ion pairs produced per primary particle.
- 4.15** Derive (4.24b) from (4.24a).
- 4.16** A charged particle moving with speed  $v$  through a medium of refractive index  $n$  emits Čerenkov radiation at an angle  $\theta$  to the direction of the particles.
- What is the maximum angle of emission and to what limit does it correspond?
  - If the momentum  $p$  of the particle is known, show that the mass squared  $x$  of the particle is given by

$$x = (mc^2)^2 = p^2 c^2 (n^2 \cos^2 \theta - 1).$$

- (c) If the error on the momentum is negligible, show, by taking derivatives of this expression, that for highly relativistic particles, the standard error  $\sigma_x$  on  $x$  is approximately

$$\sigma_x \approx 2p^2 c^2 \sqrt{(n^2 - 1)} \sigma_\theta,$$

where  $\sigma_\theta$  is the standard error on  $\theta$ .

- 4.17** Use the model for electromagnetic showers of Section 4.4.6 to show that (a) the energy spectrum of all the secondaries contained in the shower falls approximately like  $E^{-2}$  for  $E_0 \gg E > E_C$  and (b) the sum of the track lengths of all charged secondaries is approximately proportional to the primary energy. (*Hint:* Start by writing an expression for the total number of secondaries whose energy is greater than  $E$  and approximate the sum involved by an integral.)
- 4.18** The reaction  $e^+ + e^- \rightarrow \mu^+ + \mu^-$  is studied using a collider with equal beam energies of 3 GeV. The differential cross-section *in natural units* is given by

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E_{CM}^2} (1 + \cos^2 \theta),$$

where  $E_{CM}$  is the total centre-of-mass energy and  $\theta$  is the angle between the initial  $e^-$  and the final  $\mu^-$ . If the detector can only record an event if the  $\mu^+\mu^-$  pair makes an angle of at least  $25^\circ$  relative to the beam line, what fraction of events will be recorded? What is the total cross-section for this reaction in nanobarns? If the reaction is recorded for  $10^7$  s at a luminosity of  $L = 10^{31} / \text{cm}^2 \text{ s}$ , how many events are expected?

- 4.19** The reaction  $e^+e^- \rightarrow \tau^+\tau^-$  is studied using a collider with equal beam energies of 5 GeV and with the collision point at the centre of the detector. The latter is of cylindrical construction and at increasing radii from the beam line there is a drift chamber, an electromagnetic calorimeter, a hadronic calorimeter and finally muon chambers. If in a particular event the tau decays are

$$\tau^- \rightarrow \mu^- + \bar{\nu}_\mu + \nu_\tau \quad \text{and} \quad \tau^+ \rightarrow e^+ + \bar{\nu}_\tau + \nu_e,$$

what signals would be observed in the various parts of the detector?

---

## Space–time symmetries

---

Symmetries and conservation laws are a constant theme in physics, and nowhere more so than in particle physics. In this chapter we concentrate on those laws that are associated with space–time symmetries and their applications in strong and electromagnetic interactions. Such conservation laws are particularly important in spectroscopy. In atomic physics, spectroscopy is a vital ingredient in understanding the structure of atoms in terms of nuclei and electrons, and it plays an analogous role in elucidating the internal structure of hadrons in terms of their constituent quarks. In both cases, each state in the spectrum has not only a specific energy but also well-defined values of ‘good’ quantum numbers associated with conserved observables, like angular momentum, whose quantum mechanical operators commute with the Hamiltonian of the system. In atomic physics these quantum numbers are crucial to understanding the degeneracies of the energy levels, their behaviours in the presence of electric and magnetic fields, and the selection rules that govern transitions between them. Conservation laws have an equally important place in hadron spectroscopy, and the first task in studying hadrons is to determine which are the appropriate conserved quantities and to measure their values for observed states.

Some of the conservation laws discussed in this chapter – those for linear and angular momentum – are universal laws of nature, valid for all interactions. Others we shall meet, like parity, are only conserved in the approximation that weak interactions are neglected. Their violation will be discussed in Chapter 11. Here we shall neglect weak interactions and concentrate on the strong and electromagnetic interactions with which we will be primarily concerned in the next two chapters.

The conservation laws we discuss have their origin in the symmetries and invariance properties of the underlying interactions. To understand the connection between these we consider firstly a simple example, that of translational invariance, and show that it leads directly to the conservation of linear momentum.

## 5.1 Translational invariance

Translational invariance expresses the fact that all positions in space are physically indistinguishable. By this we mean that when a closed system of particles (i.e. one on which no external forces are acting) is moved as a whole from one position in space to another, its physical properties are unaltered. In quantum mechanics this is expressed as an invariance of the Hamiltonian. If the system is displaced a distance  $\mathbf{a}$ , then the position vector  $\mathbf{r}_i$  of a particle  $i$  becomes  $\mathbf{r}'_i = \mathbf{r}_i + \mathbf{a}$ , that is

$$\mathbf{r}_i \rightarrow \mathbf{r}'_i = \mathbf{r}_i + \mathbf{a}, \quad (5.1)$$

where it is sufficient for what follows to consider only infinitesimal displacements,  $\mathbf{a} = \delta\mathbf{r}$ . The Hamiltonian will similarly be replaced by a new Hamiltonian

$$H(\mathbf{r}'_1, \mathbf{r}'_2, \dots) = H(\mathbf{r}_1 + \delta\mathbf{r}, \mathbf{r}_2 + \delta\mathbf{r}, \dots).$$

For the case of a single free particle of mass  $m$ ,

$$H = -\frac{1}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right)$$

and one easily sees that the two Hamiltonians are equal, that is

$$H(\mathbf{r}') = H(\mathbf{r} + \delta\mathbf{r}) = H(\mathbf{r}). \quad (5.2)$$

In general,

$$H(\mathbf{r}'_1, \mathbf{r}'_2, \dots) = H(\mathbf{r}_1, \mathbf{r}_2, \dots) \quad (5.3)$$

for a closed system, and the Hamiltonian is said to be invariant under the translation (5.1).

We now specialise to a system containing just a single particle and introduce a translation operator  $\hat{D}$  defined by

$$\hat{D} \Psi(\mathbf{r}) \equiv \Psi(\mathbf{r} + \delta\mathbf{r}), \quad (5.4)$$

where  $\Psi(\mathbf{r})$  is an arbitrary wavefunction. Expanding the right-hand side of (5.4) to first order in  $\delta\mathbf{r}$  gives

$$\Psi(\mathbf{r} + \delta\mathbf{r}) = \Psi(\mathbf{r}) + \delta\mathbf{r} \cdot \nabla \Psi(\mathbf{r}).$$

Thus

$$\hat{D} = 1 + i\delta\mathbf{r} \cdot \hat{\mathbf{p}}, \quad (5.5)$$

where  $\hat{\mathbf{p}} = -i\boldsymbol{\nabla}$  is the operator for linear momentum. The conservation law for linear momentum follows by first applying  $\hat{D}$  to the wavefunction

$$\Psi'(\mathbf{r}) = H(\mathbf{r}) \Psi(\mathbf{r}),$$

giving

$$\hat{D} \Psi'(\mathbf{r}) = \hat{D} H(\mathbf{r}) \Psi(\mathbf{r}), \quad (5.6)$$

and then comparing with the results obtained from the definition (5.4), that is

$$\begin{aligned} \hat{D} \Psi'(\mathbf{r}) &= \Psi'(\mathbf{r} + \delta\mathbf{r}) = H(\mathbf{r} + \delta\mathbf{r}) \Psi(\mathbf{r} + \delta\mathbf{r}) \\ &= H(\mathbf{r}) \Psi(\mathbf{r} + \delta\mathbf{r}) = H(\mathbf{r}) \hat{D} \Psi(\mathbf{r}), \end{aligned} \quad (5.7)$$

where we have used the invariance property of the Hamiltonian (5.3). Comparing (5.6) and (5.7), we see that

$$[\hat{D} H(\mathbf{r}) - H(\mathbf{r}) \hat{D}] \Psi(\mathbf{r}) = 0,$$

and since  $\Psi(\mathbf{r})$  is an arbitrary wavefunction it follows that  $\hat{D}$  commutes with the Hamiltonian, that is

$$[\hat{D}, H] = 0.$$

Substituting (5.5) into this commutation relation leads to the conservation law for linear momentum

$$[\hat{\mathbf{p}}, H] = 0, \quad (5.8)$$

for a single-particle state whose Hamiltonian is invariant under the translation (5.1). The generalisation to an  $N$ -particle state, described by the wavefunction  $\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ , is straightforward (see Problem 5.1) and leads to conservation of the total linear momentum

$$\mathbf{p} = \sum_{i=1}^N \mathbf{p}_i,$$

where  $\mathbf{p}_i$  is the momentum of the  $i$ th particle.

The chain of arguments leading to the conservation of linear momentum is characteristic of the way conservation laws in general arise in quantum mechanics. Associated with a symmetry of the system is a transformation of the particle coordinates,

$$\mathbf{r}_i \rightarrow \mathbf{r}'_i, \quad (5.9)$$

which leaves the Hamiltonian invariant:

$$H(\mathbf{r}_1, \mathbf{r}_2, \dots) = H(\mathbf{r}'_1, \mathbf{r}'_2, \dots). \quad (5.10)$$

The associated observable is then identified by examining the corresponding transformation of an arbitrary wavefunction  $\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots)$  and the conservation law established by applying the

same transformation to the wavefunction  $\Psi' = H\Psi$ , as illustrated in (5.6) to (5.8).

## 5.2 Rotational invariance

Angular momentum conservation is a consequence of rotational invariance in the same way that linear momentum conservation follows from translational invariance. Rotational invariance expresses the fact that all directions in space are physically indistinguishable. In particular, when a closed system of particles is rotated as a whole about its centre-of-mass to a new orientation in space, its physical properties are unchanged. Under such a rotation, the position vector  $\mathbf{r}_i$  of any particle  $i$  will change to a new value  $\mathbf{r}'_i$ , where, for example, if the rotation is through an angle  $\theta$  about the  $z$ -axis,

$$\begin{aligned} x_i &\rightarrow x'_i = x_i \cos \theta - y_i \sin \theta, \\ y_i &\rightarrow y'_i = x_i \sin \theta + y_i \cos \theta, \\ z_i &\rightarrow z'_i = z_i. \end{aligned} \quad (5.11)$$

Correspondingly, the Hamiltonian  $H(\mathbf{r}_1, \mathbf{r}_2, \dots)$  will be replaced by a new Hamiltonian  $H(\mathbf{r}'_1, \mathbf{r}'_2, \dots)$ , and if the system is rotationally invariant, these two Hamiltonians are the same, that is

$$H(\mathbf{r}_1, \mathbf{r}_2, \dots) = H(\mathbf{r}'_1, \mathbf{r}'_2, \dots). \quad (5.12)$$

This invariance property holds for any closed system and also for the case of a particle moving in a central potential  $V(r)$  with the Hamiltonian

$$H = -\frac{1}{2m} \nabla^2 + V(r),$$

where

$$r = (x^2 + y^2 + z^2)^{1/2}$$

as usual. The derivation of angular momentum conservation from this invariance is similar to the derivation of linear momentum conservation above, but more complicated. We will therefore give the derivation only for the case of spinless particles, and state without proof the corresponding result for particles with spin.

### 5.2.1 Angular momentum conservation

We start by deriving the law of angular momentum conservation for a single spin-0 particle moving in a central potential. To do this, we need only consider infinitesimal rotations, and for an infinitesimal rotation through an angle  $\delta\theta$  about the  $z$ -axis, (5.11) reduces to

$$x' = x - y \delta\theta, \quad y' = x \delta\theta + y, \quad z' = z.$$

We next introduce a rotation operator  $\hat{R}_z$  defined by analogy with the displacement operator (5.4) by

$$\hat{R}_z(\delta\theta)\Psi(\mathbf{r}) = \Psi(\mathbf{r}') = \Psi(x - y\delta\theta, x\delta\theta + y, z). \quad (5.13)$$

Expanding the right-hand side to first order in  $\delta\theta$  gives

$$\Psi(\mathbf{r}') = \Psi(\mathbf{r}) - \delta\theta \left( y \frac{\partial\psi}{\partial x} - x \frac{\partial\psi}{\partial y} \right) = (1 + i\delta\theta \hat{L}_z)\Psi(\mathbf{r}), \quad (5.14)$$

where

$$\hat{L}_z = -i \left( x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right)$$

is the  $z$ -component of the orbital angular momentum operator  $\hat{\mathbf{L}}$ . Comparing (5.13) and (5.14) then gives the explicit form

$$\hat{R}_z(\delta\theta) = (1 + i\delta\theta \hat{L}_z)$$

for the rotation operator to first order in  $\delta\theta$ . There is of course nothing special about the  $z$ -direction, and for a rotation about an arbitrary axis specified by the unit vector  $\mathbf{n}$ , we have similarly

$$\hat{R}_{\mathbf{n}}(\delta\theta)\Psi(\mathbf{r}) \equiv \Psi(\mathbf{r}'), \quad (5.15)$$

where  $\hat{L}_z$  in (5.14) is replaced by the component of  $\hat{\mathbf{L}}$  in the new direction  $\mathbf{n}$ , that is

$$\hat{R}_{\mathbf{n}}(\delta\theta) = (1 + i\delta\theta \hat{\mathbf{L}} \cdot \mathbf{n}) \quad (5.16)$$

to first order in the rotation angle  $\delta\theta$ .

The conservation law for orbital angular momentum can now be obtained in the same way that linear momentum conservation was derived in Equations (5.6) to (5.8). Applying  $\hat{R}_n(\delta\theta)$  to the wavefunction

$$\Psi(\mathbf{r}) = H(\mathbf{r})\Psi(\mathbf{r})$$

gives

$$\hat{R}_{\mathbf{n}}\Psi'(\mathbf{r}) = \hat{R}_{\mathbf{n}}H(\mathbf{r})\Psi(\mathbf{r}), \quad (5.17)$$

whereas the definition (5.15) gives

$$\begin{aligned} \hat{R}_{\mathbf{n}}\Psi'(\mathbf{r}) &= \Psi'(\mathbf{r}') = H(\mathbf{r}')\Psi(\mathbf{r}') \\ &= H(\mathbf{r})\Psi(\mathbf{r}') = H(\mathbf{r})\hat{R}_{\mathbf{n}}\Psi(\mathbf{r}), \end{aligned} \quad (5.18)$$

where we have used the invariance of the Hamiltonian

$$H(\mathbf{r}') = H(\mathbf{r}).$$

Comparing (5.17) and (5.18) gives, for an arbitrary wavefunction  $\Psi(\mathbf{r})$ ,

$$[\hat{R}_{\mathbf{n}}H(\mathbf{r}) - H(\mathbf{r})\hat{R}_{\mathbf{n}}]\Psi(\mathbf{r}) = 0,$$

so that  $\hat{R}_{\mathbf{n}}$  commutes with the Hamiltonian,

$$[\hat{R}_{\mathbf{n}}, H] = 0. \quad (5.19)$$

Finally, substituting (5.16) into this commutation relation leads to the conservation law for orbital angular momentum,

$$[\hat{\mathbf{L}}, H] = 0 \quad (5.20)$$

for a spinless particle whose Hamiltonian is invariant under rotation.

The above argument applies to a spin-0 particle in a central potential. If the particle has spin, then the total angular momentum  $\mathbf{J}$  is the sum of the orbital and spin angular momenta,

$$\mathbf{J} = \mathbf{L} + \mathbf{S},$$

and the wave function  $\Psi$  may be written as the product of a space wavefunction  $\psi(\mathbf{r})$  and a spin wavefunction  $\chi$ , that is

$$\Psi = \psi(\mathbf{r})\chi.$$

For the case of spin-1/2 particles we shall use the notation  $\chi = \alpha$  to denote a spin ‘up’ state ( $S_z = 1/2$ ) and  $\chi = \beta$  to denote a spin ‘down’ state ( $S_z = -1/2$ ). It is convenient to write the spin-1/2 operator  $\hat{\mathbf{S}}$  in terms of the *Pauli spin matrices*  $\boldsymbol{\sigma}$  by

$$\hat{\mathbf{S}} = \frac{1}{2}\boldsymbol{\sigma}, \quad (5.21a)$$

where

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (5.21b)$$

Then  $\alpha$  and  $\beta$  may be represented by

$$\alpha = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad (5.22)$$

which satisfy the eigenvalue equations

$$\hat{S}_z \alpha = \frac{1}{2}\alpha, \quad \hat{S}_z \beta = -\frac{1}{2}\beta,$$

as required. Under a rotation, both spin and space wavefunctions will alter. We will not discuss the rotation properties in detail, but simply state without proof the result that under a rotation through a small angle  $\delta\theta$  about a direction  $\mathbf{n}$ , the rotation operator of (5.16) generalises to

$$\hat{R}_{\mathbf{n}}(\delta\theta) = (1 + i\delta\theta \hat{\mathbf{J}} \cdot \mathbf{n}) \quad (5.23)$$

so that

$$\Psi \rightarrow \Psi' = (1 + i\delta\theta \hat{\mathbf{J}} \cdot \mathbf{n})\Psi. \quad (5.24)$$

On the right-hand side,  $\hat{\mathbf{L}}$  and  $\hat{\mathbf{S}}$  act only on the space and spin wavefunctions respectively, that is

$$\hat{\mathbf{J}}\Psi = (\hat{\mathbf{L}} + \hat{\mathbf{S}})\Psi = [\hat{\mathbf{L}}\psi(\mathbf{r})]\chi + \psi(\mathbf{r})[\hat{\mathbf{S}}\chi].$$

The arguments that led to the conservation of orbital angular momentum in the absence of spin may be applied in the present

case using (5.23) for the rotation operator, and lead directly to conservation of the total angular momentum,

$$[\hat{\mathbf{J}}, H] = 0. \quad (5.25)$$

Rotational invariance does not lead to the conservation of  $\mathbf{L}$  and  $\mathbf{S}$  separately, and in general the orbital and spin angular momenta are not conserved, that is

$$[\hat{\mathbf{L}}, H] = -[\hat{\mathbf{S}}, H] \neq 0,$$

because of the existence of spin-dependent forces.

### 5.2.2 Classification of particles

In this book, we use the word *particle* to describe both elementary particles, which have no internal structure, and composite particles, which are bound states of smaller constituents. In either case, one of the quantum numbers that characterise the particle is its spin, which is defined as the particle's angular momentum in its own rest frame. For a composite particle, this frame corresponds to the centre-of-mass frame of its constituents. For example, mesons are  $q\bar{q}$  bound states, with the rest frame of the meson corresponding to the centre-of-mass frame of the  $q\bar{q}$  pair. Consequently, the spin  $\mathbf{S}_P$  of a particle  $P$  is the total angular momentum  $\mathbf{J}$  of its constituents in this frame. The spin quantum number  $\mathbf{S}_P = \mathbf{J}$  can be either integer or half-integer, while its  $z$ -component can take on any of the  $2J + 1 = 2S_P + 1$  possible values

$$(S_P)_z = J_z = J, J - 1, \dots, -J,$$

depending on the orientation of the spin relative to the  $z$ -axis.

In addition to its spin, a composite particle may have other ‘good’ quantum numbers, related to its internal structure, which take on definite values and therefore characterise the particle. ‘Good’ quantum numbers are defined as those that are constants of the motion, so that the associated operators commute with the Hamiltonian. In other words, they are associated with conserved observables. In the rest frame of the particle, the total angular momentum  $\mathbf{J}$  of the constituents that make up the particle is conserved, but the total orbital angular momentum  $\mathbf{L}$  and the total spin angular momentum  $\mathbf{S}$  are not separately conserved. However, despite this, it is often a very good approximation to assume that  $\mathbf{L}^2$  and  $\mathbf{S}^2$ , defined by

$$\mathbf{L}^2 \equiv L_x^2 + L_y^2 + L_z^2$$

and

$$\mathbf{S}^2 \equiv S_x^2 + S_y^2 + S_z^2,$$

are conserved, that is to assume that

$$[H, \hat{\mathbf{L}}^2] = [H, \hat{\mathbf{S}}^2] = 0. \quad (5.26)$$

This corresponds to assuming that the forces can, for example, flip the direction of the spin, thereby changing its components, but not its magnitude. In this approximation, the orbital angular momentum and spin quantum numbers  $L$  and  $S$  are also good quantum numbers, so that the particle is characterised by  $S_P = J, L$  and  $S$ , while a fourth  $J_z$  depends on the orientation of its spin. This is the basis of the so-called ‘spectroscopic notation’, in which states are denoted by

$$^{2S+1}L_J \quad (5.27)$$

and where, instead of the numerical value of  $L$ , it is conventional to use the historical notation  $S, P, D, F, \dots$  for  $L = 0, 1, 2, 3, \dots$ . Thus a composite particle denoted by  $^3S_1$  has  $L = 0, S = 1$  and  $S_P = J = 1$ , while one denoted  $^2D_{3/2}$  has  $L = 2, S = 1/2$  and  $S_P = J = 3/2$ . In all of this it is of course crucial to distinguish between the spin  $S_P = J$  of the particle and the spin  $S$  of its constituents.

Before applying these ideas to hadrons, it is instructive to consider a simple example from nuclear physics. This is the deuteron, which is a spin-1 bound state of a neutron and a proton. As usual in nuclear physics, we shall neglect the quark substructure of the nucleons and treat them as elementary spin-1/2 particles. Since the deuteron has spin 1, the total angular momentum of the  $np$  system in its centre-of-mass frame must be  $J = 1$ . It is also reasonable to assume that  $L = 0$ , as the lowest lying bound states of any system are usually  $S$ -waves. Thus we deduce that  $S = 1$  for the sum of the proton and neutron spins, so that in the spectroscopic notation of (5.27), the deuteron is an  $^3S_1$  bound state of the  $np$  system.

This assignment may be tested by considering the magnetic moment of the deuteron. If  $L = 0$ , as we have assumed, there will be no contribution from the orbital motion of the proton’s charge. Hence, the deuteron’s magnetic moment must arise solely from the spin magnetic moments of the nucleons, and must be given by

$$\mu_d = \mu_p + \mu_n = 2.793 - 1.913 = 0.880$$

nuclear magnetons, since the proton and neutron spins are parallel ( $S = 1$ ). The experimental value is  $\mu_d = 0.857$  nuclear magnetons, which is close to, but not exactly equal to, the above values. The difference is due to the fact that the deuteron is not a pure  $^3S_1$  state, but has a few percent admixture of the  $^3D_1$  state with  $L = 2$ . Thus, while the assumption that  $L$  is a good quantum number in the presence of nuclear or strong forces is a good approximation and very useful, it is not exact. With this caveat, we now use the spectroscopic notation to classify hadrons in the framework of the quark model.

### 5.2.3 Angular momentum in the quark model

In the simple quark model we assume that it is a good approximation to treat  $L$  and  $S$  as good quantum numbers and that the lightest states for any combinations of  $q\bar{q}$  and  $qqq$  have zero orbital angular momenta. The spectral classification of mesons and baryons is then as follows.

Mesons are  $q\bar{q}$  bound states, with the rest frame of the meson corresponding to the centre-of-mass frame of the  $q\bar{q}$  system. In this frame there is a single orbital angular momentum  $\mathbf{L}$ , but two constituent spins, so that in an obvious notation

$$\mathbf{S} = \mathbf{S}_q + \mathbf{S}_{\bar{q}}.$$

Since both  $q$  and  $\bar{q}$  have spin 1/2, the only possibilities are  $S = 0$  and  $S = 1$ . For  $L = 0$ ,  $J = S$ , and in the spectroscopic notation (5.27) the possible states are

$$^{2S+1}L_J = {}^1S_0, {}^3S_1 \quad (L = 0). \quad (5.28a)$$

For  $L \geq 1$ , the possible  $J$  values are  $J = L$  for singlet states ( $S = 0$ ) and  $J = L \pm 1$ ,  $L$  for triplet states ( $S = 1$ ), giving the four possibilities

$$^{2S+1}L_J = {}^1L_L, {}^3L_{L+1}, {}^3L_L, {}^3L_{L-1} \quad (L \geq 1). \quad (5.28b)$$

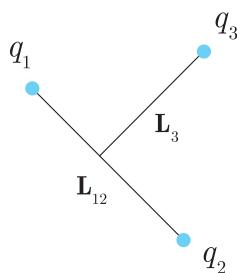
Here  $J$  refers to the total angular momentum in the  $q\bar{q}$  centre-of-mass frame, which from the discussion in Section 5.2.2 we identify with the meson spin

$$\mathbf{S}_M = \mathbf{J}.$$

If the lightest states have  $L = 0$ , then by (5.28a) they can have spin 0 or 1, so that for any flavour combination  $u\bar{d}$ ,  $u\bar{s}$ ,  $c\bar{c}$ , etc., we would expect two low-mass states with spin 0 and spin 1. This is exactly what is found experimentally, as we shall discuss in Chapter 6, and the lower lying of the two states always has spin-0. Examples are the  $\pi$ ,  $K$  and  $D$  mesons discussed in Chapter 3, which are all spin-0 mesons corresponding to  ${}^1S_0$  states of the appropriate  $q\bar{q}$  systems given in (3.13), (3.16) and (3.19).

Baryons are  $qqq$  bound states, so in the rest frame of the baryon, corresponding to the centre-of-mass frame of the  $qqq$  system, there are two orbital angular momenta associated with the relative motion of the three quarks. The first is conveniently taken to be the orbital angular momentum  $\mathbf{L}_{12}$  of a chosen pair of quarks in their mutual centre-of-mass frame. The second is then the orbital angular momentum  $\mathbf{L}_3$  of the third quark about the centre-of-mass of the pair in the overall centre-of-mass frame. This is illustrated in Figure 5.1. The total orbital angular momentum is given by

$$\mathbf{L} = \mathbf{L}_{12} + \mathbf{L}_3,$$



**Figure 5.1** Internal orbital angular momenta of a three-quark state.

while the spin is the sum of the spins of the three quarks

$$\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_3,$$

so that  $S = 1/2$  or  $S = 3/2$ . It is now straightforward to show that the possible baryon states in the spectroscopic notation are

$$^2S_{1/2}, {}^4S_{3/2} \quad (L = 0), \quad (5.29a)$$

$${}^2P_{1/2}, {}^2P_{3/2}, {}^4P_{1/2}, {}^4P_{3/2}, {}^4P_{5/2} \quad (L = 1) \quad (5.29b)$$

$${}^2L_{L+1/2}, {}^2L_{L-1/2}, {}^4L_{L-3/2}, {}^4L_{L-1/2}, {}^4L_{L+1/2}, {}^4L_{L+3/2} \quad (L = 2). \quad (5.29c)$$

The baryon spectrum is clearly very complicated! However, if we restrict ourselves to the lowest-lying states with zero orbital angular momentum  $\mathbf{L}_{12} = \mathbf{L}_3 = 0$ , then only those of (5.29a) can occur, and the baryon spin  $S_B = J$  can be  $1/2$  or  $3/2$ . We therefore expect the lightest baryons to have spin- $1/2$  or spin- $3/2$ . Again this is in accord with experiment, as we shall see in Chapter 6. For the moment, we merely note that we have already met some of the low-lying spin- $1/2$  baryons in Chapter 3, namely the  $p$ ,  $n$ ,  $\Lambda$ ,  $\Lambda_c$  and  $\Lambda_b$ , which we now identify with the  ${}^2S_{1/2}$  states of the appropriate quark constituents given in Table 3.2.

### 5.3 Parity

The parity transformation is

$$\mathbf{r}_i \rightarrow \mathbf{r}'_i = -\mathbf{r}_i, \quad (5.30)$$

so the position vector  $\mathbf{r}_i$  of every particle  $i$  is reflected in the origin. A system is said to be invariant under parity if the Hamiltonian remains unchanged by this transformation, that is if

$$H(\mathbf{r}'_1, \mathbf{r}'_2, \dots) = H(-\mathbf{r}_1, -\mathbf{r}_2, \dots) = H(\mathbf{r}_1, \mathbf{r}_2, \dots). \quad (5.31)$$

Unlike translational and rotational invariance, parity is not an exact symmetry of nature that holds for all closed systems, but is violated by the weak interaction. This will be discussed in Chapter 11. For the present, we will ignore the weak interaction and only consider systems that interact by the strong and electromagnetic interactions. We will assume that these interactions are invariant under the parity transformation and explore its consequences.

We start by specialising to a single particle, and introduce a parity operator  $\hat{P}$  defined by

$$\hat{P} \Psi(\mathbf{r}, t) \equiv P_a \Psi(-\mathbf{r}, t), \quad (5.32a)$$

where the label  $a$  identifies the particle type (e.g.  $a = e^-$  for an electron,  $a = \mu^-$  for a muon, etc.) and  $P_a$  is a constant phase factor. Since two successive parity transformations leave the system unchanged, we require that

$$\hat{P}^2 \Psi(\mathbf{r}, t) = \Psi(\mathbf{r}, t),$$

implying

$$P_a = \pm 1 \quad (5.32b)$$

for the possible values of  $P_a$ . If we consider an eigenfunction of momentum

$$\Psi_p(\mathbf{r}, t) = \exp[i(\mathbf{p} \cdot \mathbf{r} - Et)],$$

then

$$\hat{P} \Psi_p(\mathbf{r}, t) = P_a \Psi_p(-\mathbf{r}, t) = P_a \Psi_{-\mathbf{p}}(\mathbf{r}, t),$$

so that a particle at rest, with  $\mathbf{p} = \mathbf{0}$ , is an eigenstate of the parity operator with eigenvalue  $P_a$ . For this reason  $P_a$  is called the *intrinsic parity* of particle  $a$  or, more usually, just the *parity* of particle  $a$ , with the words ‘at rest’ left implicit. The values of  $P_a$  for various particle types will be determined below; here we merely note that had we tried to set  $P_a = 1$  for all particles, we would be led to contradictions with experiments in which particles are either created or destroyed. For many-particle systems the appropriate generalisation of (5.32a) is

$$\hat{P} \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, t) \equiv P_1 P_2 \dots \Psi(-\mathbf{r}_1, -\mathbf{r}_2, \dots, t),$$

with one intrinsic parity factor occurring for each particle present.

In addition to a particle at rest, a particle with a definite orbital angular momentum is also an eigenstate of parity. The wavefunction for such a particle has the form

$$\Psi_{nlm}(\mathbf{r}) = R_{nl}(r) Y_l^m(\theta, \phi), \quad (5.33a)$$

where  $(r, \theta, \phi)$  are spherical polar coordinates,  $R_{nl}(r)$  is a function of the radial variable  $r$  only and the  $Y_l^m(\theta, \phi)$  are spherical harmonics. For example, the first few spherical harmonics are

$$Y_0^0 = \frac{1}{\sqrt{4\pi}}, \quad Y_1^0 = \sqrt{\frac{3}{4\pi}} \cos \theta, \quad Y_1^{\pm 1} = \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\phi}. \quad (5.33b)$$

Using the relations

$$x = r \sin \theta \cos \phi, \quad y = r \sin \theta \sin \phi, \quad z = r \cos \theta$$

between Cartesian and polar coordinates, the parity transformation becomes

$$r \rightarrow r' = r, \quad \theta \rightarrow \theta' = \pi - \theta, \quad \phi \rightarrow \phi' = \pi + \phi,$$

from which it can be shown that

$$Y_l^m(\theta, \phi) \rightarrow Y_l^m(\pi - \theta, \pi + \phi) = (-1)^l Y_l^m(\theta, \phi).$$

Hence

$$\hat{P} \Psi_{nlm}(\mathbf{r}) = P_a \Psi_{nlm}(-\mathbf{r}) = P_a (-1)^l \Psi_{nlm}(\mathbf{r}), \quad (5.34)$$

and  $\Psi_{nlm}(\mathbf{r})$  is an eigenstate of parity with an eigenvalue  $P_a(-1)^l$ .

If the Hamiltonian is invariant under the parity transformation, then an argument analogous to those used to demonstrate linear and orbital angular momentum conservation in Sections 5.1 and 5.2 leads to the conclusion that parity is conserved, that is

$$[\hat{P}, H] = 0.$$

This has two obvious consequences: in a reaction, the initial and final parities are equal  $P_i = P_f$ , and for bound states parity is a good quantum number. In particular, atomic and nuclear states have been shown to be eigenstates of parity to very high accuracy by searching for parity-violating transitions, verifying that parity invariance indeed holds for strong (nuclear) and electromagnetic interactions.

### 5.3.1 Leptons and antileptons

Our discussion of parity invariance has been based so far on non-relativistic quantum mechanics. In this approximation, electrons and positrons are described by distinct wavefunctions that satisfy the Schrödinger equation and there is no obvious connection between their intrinsic parities defined by

$$\hat{P} \Psi(\mathbf{r}, t) = P_{e^\pm} \Psi(-\mathbf{r}, t)$$

for electron and positron wavefunctions, respectively. However, relativistically a single four-component wavefunction (1.11) describes both electrons and positrons. One would therefore expect their intrinsic parities to be related, and indeed an analysis<sup>1</sup> shows that the Dirac equation is only compatible with parity invariance if

$$P_{e^+} P_{e^-} = -1, \quad (5.35)$$

that is, if electrons and positrons have opposite parities. The same argument applies to other spin-1/2 particles, giving

$$P_f P_{\bar{f}} = -1, \quad (5.36)$$

where  $f$  labels a particular fermion and  $\bar{f}$  the corresponding antifermion.

---

<sup>1</sup> See, for example, p. 372 of Perkins (1987).

The prediction (5.35) can be confirmed experimentally by studying the reaction

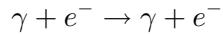
$$e^+ + e^- \rightarrow \gamma + \gamma,$$

where the  $e^+e^-$  are bound together in an  $S$ -wave state called para-positronium.<sup>2</sup> The initial state has zero orbital momentum, so that its parity is just the product of the intrinsic parities  $P_i = P_{e^+}P_{e^-}$ . If the final state has relative orbital angular momentum  $l_\gamma$ , then its parity is  $P_f = P_\gamma^2(-1)^{l_\gamma} = (-1)^{l_\gamma}$ , where  $P_\gamma$  is the intrinsic parity of the photon and  $P_\gamma^2 = 1$  by (5.32b). Parity conservation  $P_i = P_f$  then gives

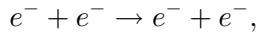
$$P_{e^+}P_{e^-} = (-1)^{l_\gamma},$$

so that a measurement of the orbital angular momentum in the final state  $l_\gamma$  can determine the product  $P_{e^+}P_{e^-}$ . Such a measurement can be made indirectly by studying the polarisations of the emitted photons,<sup>3</sup> and unambiguously confirms the prediction (5.35).

The above argument determines the product  $P_{e^+}P_{e^-}$ , but says nothing about the parity of a single electron or positron. Other reactions, such as



and



contain the same factors of  $P_{e^-}$  in both the initial and final states, so that they cancel on equating initial and final state parities. In general, since single electrons can be neither created nor destroyed in electromagnetic interactions, similar cancellations occur in all electromagnetic processes and therefore  $P_{e^-}$  cannot be determined. Similarly, the intrinsic parities of muons and tauons cannot be measured, since, like electrons, they can only be created or destroyed in electromagnetic interactions as particle–antiparticle pairs. Because of this it is conventional to define

$$P_{e^-} \equiv P_{\mu^-} \equiv P_{\tau^-} \equiv 1. \quad (5.37a)$$

These choices are arbitrary, but having made them it follows from (5.36) that

$$P_{e^+} = P_{\mu^+} = P_{\tau^+} = -1. \quad (5.37b)$$

---

<sup>2</sup> This is the ground state of the  $e^+e^-$  system. Both it and the excited states are discussed further in Section 5.5.

<sup>3</sup> See, for example, pp. 89–90 of Perkins (1987).

### 5.3.2 Quarks and hadrons

In Section 3.2 we saw that quarks could only be created or destroyed in the strong and electromagnetic interactions as particle–antiparticle pairs. Consequently, the quark intrinsic parities, like those of electrons and positrons, are indeterminate and must be fixed by convention. The usual convention is

$$P_u \equiv P_d \equiv P_s \equiv P_c \equiv P_b \equiv P_t \equiv 1, \quad (5.38a)$$

implying

$$P_{\bar{u}} = P_{\bar{d}} = P_{\bar{s}} = P_{\bar{c}} = P_{\bar{b}} = P_{\bar{t}} = -1 \quad (5.38b)$$

for the antiquarks. With these conventions, we can predict the parities of hadrons from their assumed quark structures, remembering that the hadron rest frame is the same as the centre-of-mass frame of the bound-state quark system. Hence the (intrinsic) parity  $P_M$  of a meson  $M = a\bar{b}$  is

$$P_M = P_a P_{\bar{b}} (-1)^L = (-1)^{L+1}, \quad (5.39a)$$

where  $L$  is the orbital angular momentum of the  $q\bar{q}$  pair and the quark labels  $a, b$  can each be  $u, d, s, c$  or  $b$ . Low-lying mesons, corresponding to  $L = 0$ , are thus predicted to have negative parity, and this is consistent with the observed parities of the spin-0 pions (3.13), kaons (3.16) and  $D$  mesons (3.19). The parity of a baryon  $B = abc$  is similarly given by

$$P_B = P_a P_b P_c (-1)^{L_{12}} (-1)^{L_3} = (-1)^{L_{12} + L_3}, \quad (5.39b)$$

where  $L_{12}$  and  $L_3$  are the internal angular momenta illustrated in Figure 5.1. For the corresponding antibaryon  $\bar{B} = \bar{a}\bar{b}\bar{c}$  we have

$$P_{\bar{B}} = P_{\bar{a}} P_{\bar{b}} P_{\bar{c}} (-1)^{L_{12}} (-1)^{L_3} = -(-1)^{L_{12} + L_3} = -P_B, \quad (5.39c)$$

so that baryons and antibaryons have opposite parities, consistent with (5.36). For the low-lying baryons with  $L_{12} = L_3 = 0$ , we predict positive parity, and this is consistent with the observed parities of the spin-1/2 baryons  $p, n, \Lambda$ , and  $\Lambda_c$  of Table 3.2.

In the above discussion, we have fixed the parities of the known hadrons by making the choice (5.38a) for the first five quarks  $u, d, s, c, b$ , since hadrons containing the  $t$  quark have not yet been discovered. We could equally well have defined our conventions by fixing the parities of five suitably chosen hadrons. A convenient choice is

$$P_p \equiv P_n \equiv 1 \quad (5.40a)$$

for the proton and neutron, which are  $uud$  and  $udd$  states with  $L_{12} = L_3 = 0$ , and

$$P_{K^-} \equiv P_{D^-} \equiv P_{B^-} \equiv -1 \quad (5.40b)$$

for the  $K^-$ (494),  $D^-$ (1870) and  $B^-$ (5280), which are  $L = 0$  mesons with the quark compositions  $s\bar{u}$ ,  $d\bar{c}$  and  $b\bar{u}$  given in Table 3.3. The conventions (5.40) are equivalent to the conventions (5.38a) within the context of the simple quark model. However, they have the considerable advantage that they refer to directly observable particles, and can be used to determine experimentally the parity of any hadron without reference to its assumed quark substructure. This is done by using suitably chosen strong and electromagnetic reactions to relate the unknown parity to the values (5.40) by parity conservation. That the resulting values agree with the simple quark model predictions (5.39) is a significant test of the model and of the assignment of the state within it. This will be illustrated immediately for the  $\pi^-$  meson, which should have negative parity if it is an  $L = 0$   $u\bar{d}$  state as we have assumed.

### 5.3.3 Parity of the charged pion

The parity of the  $\pi^-$  meson was first demonstrated to be negative by the observation of the absorption reaction

$$\pi^- + d \rightarrow n + n \quad (5.41)$$

for stopped pions in deuterium. We have already met the deuteron in Section 5.2.2 as an  $S$ -wave bound state of a proton and a neutron with spin 1 resulting from the parallel spins of the two nucleons. The deuteron parity is therefore  $P_d = P_p P_n = 1$ . The pions are captured into atomic orbitals to form ‘pionic deuterium’, and then cascade down through the levels with emission of ‘pionic X-rays’. The absorption takes place from  $S$ -states, because only they have a non-vanishing wavefunction at the origin, and so a reasonable chance of finding the  $\pi^-$  and  $d$  close together. For  $S$ -wave capture the total angular momentum is  $J = 1$ , since the pions have spin 0 and the neutrons have spin 1, and the parity of the initial state is  $P_i = P_{\pi^-} P_d = P_{\pi^-}$ . By parity conservation, we then have

$$P_{\pi^-} = P_f = P_n P_n (-1)^L = (-1)^L, \quad (5.42)$$

where  $L$  is the orbital angular momentum of the final two-neutron state, which must have  $J = 1$  by angular momentum conservation.

In most reactions, the  $L$  values, and hence parities, are found by measuring the angular distributions of the particles involved. In the present case, however, the result can be established by exploiting the Pauli principle, which requires that the final-state wave function (i.e. the product of space and spin parts) must be antisymmetric under the interchange of the two neutrons. The possibilities for the spin state are  $S = 0$  and  $S = 1$ . We examine each in turn.

The spin triplet state  $S = 1$  has the symmetric spin wavefunctions<sup>4</sup>

$$\begin{aligned} \alpha_1\alpha_2 & (S = 1, S_z = 1), \\ \frac{1}{\sqrt{2}}(\alpha_1\beta_2 + \beta_1\alpha_2) & (S = 1, S_z = 0), \\ \beta_1\beta_2 & (S = 1, S_z = -1), \end{aligned} \quad (5.43)$$

where we have used our previous notation, that is  $\alpha_i(\beta_i)$  corresponds to a particle  $i$  having spin up (down) in the  $z$ -direction. By the Pauli principle, the space wavefunction must in this case be antisymmetric, which requires  $L$  to be odd. Since both  $J = 1$  and  $S = 1$ , the only possibility is  $L = 1$ , which would then imply  $P_{\pi^-} = -1$  by (5.42). The other possibility is the spin singlet state

$$\frac{1}{\sqrt{2}}(\alpha_1\beta_2 - \beta_1\alpha_2) \quad (S = S_z = 0), \quad (5.44)$$

which is antisymmetric. Consequently, the space wavefunction must be symmetric with  $L$  even, implying  $J = 0, 2, \dots$  since  $S = 0$ . Thus it is not possible to satisfy the Pauli principle and angular momentum conservation with  $S = 0$ , since the initial state has  $J = 1$ . The only possibility is  $S = 1$ , and consequently  $P_{\pi^-} = -1$ , in agreement with the assignment of the pion to an  $L = 0$  state in the simple quark model.

### 5.3.4 Parity of the photon

In contrast to the intrinsic parity of an electron or a quark, the intrinsic parity of the photon can be deduced theoretically by invoking the correspondence principle. Classically, the electric field obeys Poisson's equation

$$\nabla \cdot \mathbf{E}(\mathbf{r}, t) = \frac{1}{\epsilon_0} \rho(\mathbf{r}, t),$$

where  $\rho(\mathbf{r}, t)$  is the charge density. Under a parity transformation, the value of the charge density at  $\mathbf{r}$  is replaced by its value at  $-\mathbf{r}$ , that is

$$\rho(\mathbf{r}, t) \rightarrow \rho(-\mathbf{r}, t),$$

and since  $\nabla$  changes sign under (5.30), the electric field itself must transform according to

$$\mathbf{E}(\mathbf{r}, t) \rightarrow -\mathbf{E}(-\mathbf{r}, t) \quad (5.45)$$

if Poisson's equation is to remain invariant. Alternatively, the electromagnetic field can be described by the vector and scalar potentials

---

<sup>4</sup> See, for example, Equation (5.80) in Section 5.5.1 of Mandl (1992).

$\mathbf{A}(\mathbf{r}, t)$  and  $\phi(\mathbf{r}, t)$  respectively. These are related to the electric field by<sup>5</sup>

$$\mathbf{E} = -\nabla\phi - \frac{\partial\mathbf{A}}{\partial t} \quad (5.46)$$

and in the absence of charges the scalar potential may be set to zero. In quantum mechanics, it is the vector potential that corresponds to the photon wave function,<sup>6</sup> and a photon propagating freely through space with momentum  $\mathbf{k}$  is described by a wavefunction

$$\mathbf{A}(\mathbf{r}, t) = N \boldsymbol{\varepsilon}(\mathbf{k}) \exp[i(\mathbf{k} \cdot \mathbf{r} - Et)], \quad (5.47)$$

where  $\boldsymbol{\varepsilon}(\mathbf{k})$  is a unit polarisation vector and  $N$  is a normalisation constant, which is irrelevant for what follows. Under the parity transformation

$$\mathbf{A}(\mathbf{r}, t) \rightarrow P_\gamma \mathbf{A}(-\mathbf{r}, t), \quad (5.48a)$$

where the intrinsic parity of the photon  $P_\gamma$  is defined by analogy with (5.32), implying

$$\mathbf{E}(\mathbf{r}, t) \rightarrow P_\gamma \mathbf{E}(-\mathbf{r}, t) \quad (5.48b)$$

on substituting into (5.46). Consistency with (5.45) in the classical limit then requires that the intrinsic parity of the photon  $P_\gamma = -1$ .

## 5.4 Charge conjugation

Charge conjugation ( $\hat{C}$ ) is the operation that replaces all particles by their antiparticles in the same state, so that momenta, positions, etc., are unchanged. Since the charge and magnetic moment of every particle are reversed, the electromagnetic interaction remains invariant under this operation. Charge conjugation is also a symmetry of the strong interaction, but is violated by the weak interaction, as we shall see in Chapter 11. Again, we shall neglect weak interactions for the present and assume invariance of the Hamiltonian and the associated conservation law

$$[\hat{C}, H] = 0 \quad (5.49)$$

for strong and electromagnetic interactions only.

We now distinguish between those particles  $a = \pi^+, K^+, p, n, \dots$  that have distinct antiparticles and those particles  $\alpha = \pi^0, \gamma, \dots$  that do not. Charged particles of all types are obvious examples of the former, but some neutral particles, like the neutron

---

<sup>5</sup> See for example, p. 229 of Grant and Phillips (1990).

<sup>6</sup> A full discussion of the quantisation of electromagnetic radiation may be found in, for example, Chapter 1 of Mandl and Shaw (2010).

(which has a non-zero magnetic moment) also belong to this class. Under  $\hat{C}$  we then have

$$\hat{C} |\alpha \Psi\rangle = C_\alpha |\alpha \Psi\rangle \quad (5.50a)$$

and

$$\hat{C} |a \Psi\rangle = |\bar{a} \Psi\rangle, \quad (5.50b)$$

where  $C_\alpha$  is a phase factor analogous to that introduced in the discussion of parity (cf. (5.32)) and we use the Dirac notation summarised at the beginning of Section 2.3. Since a second transformation turns antiparticles back into particles, we require  $\hat{C}^2 = 1$  and hence

$$C_\alpha = \pm 1. \quad (5.51)$$

A similar phase factor  $C_a$  could have been introduced into (5.50b), but it would have no physical consequences as the relative phase of the two states  $|a\Psi\rangle$  and  $|\bar{a}\Psi\rangle$  cannot be measured (see Problem 5.11). We therefore omit such a factor. For multi-particle states we then have

$$\begin{aligned} \hat{C} &|\alpha_1, \alpha_2, \dots; a_1, a_2, \dots; \Psi\rangle \\ &= C_{\alpha_1} C_{\alpha_2} \cdots |\alpha_1, \alpha_2, \dots; \bar{a}_1, \bar{a}_2, \dots; \Psi\rangle, \end{aligned} \quad (5.50c)$$

where  $\Psi$  denotes the same multi-particle wavefunction on both sides. In future, when the wavefunction is not our primary concern, we shall omit it.

From (5.50a) we see that the particles  $\alpha = \pi^0, \gamma, \dots$  are eigenstates of  $\hat{C}$  with eigenvalues  $C_\alpha = \pm 1$ , called their *C parities*. These values are readily measured using *C* conservation, as we shall see below. Otherwise, eigenstates can only be constructed from particle–antiparticle pairs, when  $\hat{C}$  in effect interchanges particle with antiparticle. If the state is symmetric or antisymmetric under  $a \leftrightarrow \bar{a}$ , then

$$\hat{C} |a\Psi_1, \bar{a}\Psi_2\rangle = |\bar{a}\Psi_1, a\Psi_2\rangle = \pm |a\Psi_1, \bar{a}\Psi_2\rangle$$

and hence  $|a\Psi_1, \bar{a}\Psi_2\rangle$  is an eigenstate of  $\hat{C}$ . A simple example is given by a  $\pi^+ \pi^-$  pair in a state of definite orbital angular momentum  $L$ , when

$$\hat{C} |\pi^+ \pi^-; L\rangle = (-1)^L |\pi^+ \pi^-; L\rangle, \quad (5.52)$$

since interchanging the  $\pi^+$  and  $\pi^-$  reverses their relative position vector in the spatial wavefunction. The same factor occurs for spin-1/2 fermions  $f$ , like quarks and leptons, together with two further factors: a factor  $(-1)^{S+1}$  that arises from interchanging the particles in the spin wavefunctions, which may be verified from the explicit wavefunctions (5.43) and (5.44), and a factor  $(-1)$  that arises whenever

fermions and antifermions are interchanged.<sup>7</sup> Overall, we therefore obtain

$$\hat{C} |f \bar{f}; J, L, S\rangle = (-1)^{L+S} |f \bar{f}; J, L, S\rangle, \quad (5.53)$$

for a fermion–antifermion pair with total, orbital and spin angular momentum quantum numbers  $J$ ,  $L$  and  $S$ , respectively. Thus if the  $\pi^0$ , for example, really is a  ${}^1S_0$  state of  $u\bar{u}$  or  $d\bar{d}$ , then it must have  $C$  parity  $C_{\pi^0} = 1$ .

### 5.4.1 $\pi^0$ and $\eta$ decays

The prediction  $C_{\pi^0} = 1$  of the simple quark model is confirmed experimentally by the observation of the dominant decay mode  $\pi^0 \rightarrow \gamma + \gamma$ . If we represent the states in the Dirac notation by just the particle symbol, we have

$$\hat{C} |\pi^0\rangle = C_{\pi^0} |\pi^0\rangle$$

and

$$\hat{C} |\gamma\gamma\rangle = C_\gamma C_\gamma |\gamma\gamma\rangle = |\gamma\gamma\rangle$$

since  $C_\gamma^2 = 1$  by (5.51). Since the final state has  $C = 1$ , so does the initial state, implying  $C_{\pi^0} = 1$ , in agreement with the predicted value.

The  $C$  parity of the photon can, like its parity, be inferred from the behaviour of the classical electromagnetic field. Under charge conjugation

$$\mathbf{A}(\mathbf{r}, t) \rightarrow C_\gamma \mathbf{A}(\mathbf{r}, t) \quad (5.54a)$$

by analogy with the corresponding result for parity (5.48). However, since all electric charges change sign under charge conjugation, it is clear that the electric field  $\mathbf{E}$  and the scalar potential  $\phi$  must also change sign, that is

$$\mathbf{E}(\mathbf{r}, t) \rightarrow -\mathbf{E}(\mathbf{r}, t), \quad \phi(\mathbf{r}, t) \rightarrow -\phi(\mathbf{r}, t). \quad (5.54b)$$

Substituting (5.54a) and (5.54b) into (5.46) gives the desired result

$$C_\gamma = -1 \quad (5.55)$$

for the  $C$ -parity of the photon. This value is confirmed, and  $C$  conservation tested, by considering the possible decay mode

$$\pi^0 \rightarrow \gamma + \gamma + \gamma. \quad (5.56)$$

Since

$$\hat{C} |\gamma\gamma\gamma\rangle = (C_\gamma)^3 |\gamma\gamma\gamma\rangle = C_\gamma |\gamma\gamma\gamma\rangle,$$

---

<sup>7</sup>This factor can only be properly understood in the context of quantum field theory. See for example, pp. 259–260 of Gottfried and Weisskopf (1986).

the reaction (5.56) is forbidden if  $C_\gamma = -1$  and  $C$  parity is conserved. If, on the contrary,  $C_\gamma = +1$  or if  $C$  parity were not conserved in electromagnetic interactions, we would expect

$$R \equiv \frac{\Gamma(\pi^0 \rightarrow 3\gamma)}{\Gamma(\pi^0 \rightarrow 2\gamma)} = O(\alpha),$$

since no other selection rule forbids this decay. Experimentally, the  $3\gamma$  mode has never been seen and the measured upper limit on this ratio is  $R < 3 \times 10^{-8} \ll \alpha$ . We conclude therefore that (5.56) is indeed forbidden as an electromagnetic interaction, thus confirming both the assignment  $C_\gamma = -1$  and  $C$  conservation.

Further tests of  $C$  invariance can be made in the decays of the eta ( $\eta$ ) particle. This is a neutral spin-0 meson of mass  $558 \text{ MeV}/c^2$  that, like the neutral pion, is detected via its decay products. The most important decay modes, with their branching ratios  $B$ , are

$$\eta \rightarrow \gamma + \gamma, \quad B = 0.39, \quad (5.57a)$$

$$\eta \rightarrow \pi^0 + \pi^0 + \pi^0, \quad B = 0.33, \quad (5.57b)$$

and

$$\eta \rightarrow \pi^+ + \pi^- + \pi^0, \quad B = 0.23. \quad (5.57c)$$

Reaction (5.57a) is obviously electromagnetic, and since the other two reactions have comparable rates, it is reasonable to assume that they too are electromagnetic. The measured lifetime of approximately  $6 \times 10^{-19} \text{ s}$  is consistent with this assumption. The  $C$  parity of the  $\eta$  follows directly from the existence of the  $2\gamma$  mode, and is confirmed by the existence of the  $3\pi^0$  mode; it is  $C_\eta = 1$ . In addition, tests of  $C$  invariance may be made by considering, for example, the reaction

$$\eta \rightarrow \pi^+(\mathbf{p}_1) + \pi^-(\mathbf{p}_2) + \pi^0(\mathbf{p}_3), \quad (5.58)$$

where we have labelled the final-state particles by their momenta. If the interaction is  $C$  invariant, then the charge-conjugate reaction

$$\eta \rightarrow \pi^-(\mathbf{p}_1) + \pi^+(\mathbf{p}_2) + \pi^0(\mathbf{p}_3)$$

should be indistinguishable from (5.58), implying identical momentum distributions for the  $\pi^+$  and  $\pi^-$  mesons. Experimentally, the  $\pi^\pm$  spectra are indeed the same within experimental errors, and  $C$ -violating asymmetries are ruled out to better than 1%.

## 5.5 Positronium

Positronium is a hydrogen-like bound state of an electron and positron. It has much in common with meson systems, which are

bound states of quarks and antiquarks, and provides striking examples of the application of conservation laws. In particular, observations on positronium can be used to verify experimentally the fundamental result (5.53) on the  $C$  parity of fermion–antifermion systems, which has been assumed so far to exist on the basis of theory alone.

A good approximation to the spectrum of positronium can be obtained by firstly considering only the Coulomb interaction between the electron and positron

$$V(r) = -\frac{\alpha}{r}.$$

The energy levels then have the same form as those for the hydrogen atom

$$E_n = -\frac{m_r \alpha^2}{2n^2} = -\frac{Ry}{2n^2}, \quad (5.59)$$

where  $Ry = 13.6$  eV is the Rydberg energy,  $n$  is the principal quantum number and  $m_r = m_e/2$  is the reduced mass, where  $m_e$  is the mass of the electron. The level spacings are thus half those in hydrogen, which is confirmed by detecting photons from the  $2p \rightarrow 1s$  transition with the predicted energy  $3Ry/8 = 5.1$  eV. In each level, the orbital angular momentum  $L$  is restricted to  $L \leq n - 1$  as in hydrogen, while the total spin is the sum of the electron and positron spins

$$\mathbf{S} = \mathbf{S}_+ + \mathbf{S}_-, \quad (5.60)$$

so that  $S = 0$  or  $S = 1$ . The resulting states corresponding to the  $n = 1$  and  $n = 2$  bands are easily deduced, and are summarised in Table 5.1 together with the corresponding values of the parity

$$P = P_{e^+} P_{e^-} (-1)^L = -(-1)^L$$

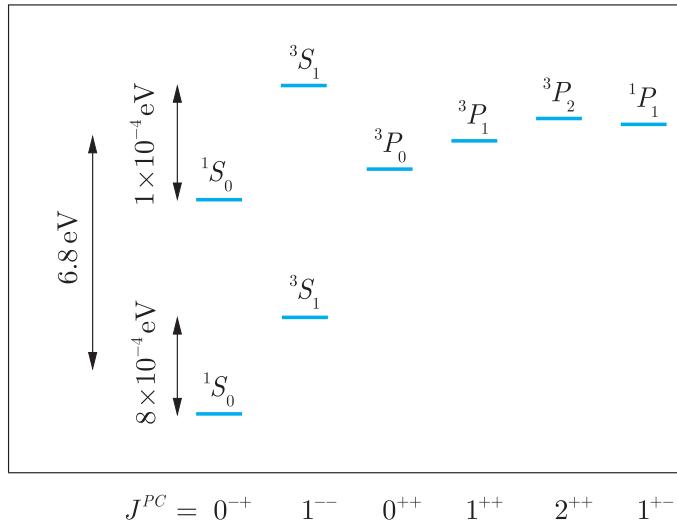
and  $C$  parity

$$C = (-1)^{L+S}$$

which follow from (5.35) and (5.53), respectively. The  ${}^3S_1$  state with  $n = 1$  is called *orthopositronium*, and the  ${}^1S_0$  state with  $n = 1$  is called *parapositronium*.

**Table 5.1** The states of positronium with the principal quantum numbers  $n = 1$  and  $n = 2$ .

		$J$	$P$	$C$
$n = 1$	${}^1S_0$	0	-1	1
	${}^3S_1$	1	-1	-1
$n = 2$	${}^1S_0$	0	-1	1
	${}^3S_1$	1	-1	-1
	${}^1P_1$	1	1	-1
	${}^3P_2$	2	1	1
	${}^3P_1$	1	1	1
	${}^3P_0$	0	1	1



**Figure 5.2** The energy levels of positronium corresponding to principal quantum numbers  $n = 1$  and  $n = 2$ .

### \*5.5.1 Fine structure

In the positronium spectrum, different states corresponding to the same values of  $n$  are not exactly degenerate, but are split due to small spin-dependent interactions as shown in Figure 5.2. In particular, orthopositronium and parapositronium are split by an amount that is measured to be

$$\Delta E = E(n = 1, {}^3S_1) - E(n = 1, {}^1S_0) = 8.45 \times 10^{-4} \text{ eV}, \quad (5.61)$$

so that parapositronium is the ground state. This splitting is a sum of two contributions. The first of these arises because the magnetic moment of the positron  $\mu_+ = e \mathbf{S}_+ / m$  gives rise to a magnetic field that interacts with the magnetic moment of the electron  $\mu_- = -e \mathbf{S}_- / m$ . A detailed calculation shows that the resulting shifts in the  $n = 1$  energy levels are given by

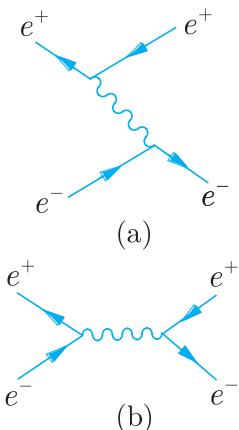
$$\delta E = -\frac{8}{3} \left( \frac{\mu_0}{4\pi} \right) \frac{\langle \boldsymbol{\mu}_+ \cdot \boldsymbol{\mu}_- \rangle}{a_0^3}, \quad (5.62)$$

where  $\mu_0$  is the magnetic permeability of free space ( $\mu_0 = 1/\epsilon_0$  in natural units),  $a_0 = 2/\alpha m$  is the analogue of the Bohr radius and the expectation value is given by

$$\begin{aligned} \langle \boldsymbol{\mu}_+ \cdot \boldsymbol{\mu}_- \rangle &= -\frac{e^2}{m^2} \langle \mathbf{S}_+ \cdot \mathbf{S}_- \rangle = -\frac{e^2}{2m^2} \langle \mathbf{S}^2 - \mathbf{S}_+^2 - \mathbf{S}_-^2 \rangle \\ &= -\frac{e^2}{2m^2} \left[ S(S+1) - \frac{3}{2} \right]. \end{aligned}$$

Straightforward substitution in (5.62) then gives

$$\Delta E_1 = 2\alpha^2 Ry / 3 \quad (5.63)$$



**Figure 5.3** Contributions to the  $e^+e^-$  interaction from (a) one-photon exchange and (b) one-photon annihilation.

as the contribution to the splitting (5.61), which is just over one-half of the measured value. The corresponding effect also exists in hydrogen, but is very much less important because the magnetic moment of the proton is very much smaller than that of the positron.

The second contribution to the energy shift has no analogue in hydrogen. In the latter, the  $e^-p$  interaction is due to one-photon exchange, which gives rise to the spin–spin and spin–orbit interactions as well as the Coulomb force. While the same diagram occurs for  $e^+e^-$  (Figure 5.3a), there is an additional ‘annihilation diagram’ (Figure 5.3b). Because annihilation occurs at a point, the contribution from the latter will be proportional to  $|\psi_{nlm}(0)|^2$  and will vanish except for  $S$ -waves, and since the photon has spin 1 and angular momentum is conserved, only  $e^+e^-$  states with  $J = 1$  will be affected. It thus shifts the energies of the  ${}^3S_1$  states only. A perturbation theory calculation gives a contribution of  $\Delta E_2 = \alpha^2 Ry/2$  for orthopositronium, and adding this to (5.63) gives

$$\Delta E = \Delta E_1 + \Delta E_2 = \frac{2}{3}\alpha^2 Ry + \frac{1}{2}\alpha^2 Ry = \frac{7}{6}\alpha^2 Ry$$

for the splitting (5.61). This prediction is in excellent agreement with the experimental value, thus confirming both the existence of the annihilation term and the Dirac values of the magnetic moments used. The latter are also confirmed by studying  $P$ -wave splittings, which depend on both spin–spin and spin–orbit interactions, and which have been measured for the three  ${}^3P$  states with  $n = 2$ .

### \*5.5.2 C parity and annihilations

The excited states of positronium can decay to lower-lying states by photon emission, as in hydrogen. In contrast to hydrogen, however, the ground state is itself unstable, and can decay via electron–positron annihilation. Because these annihilations only occur when the electron and positron are very close together (see Problem 1.8), the rate is to a good approximation proportional to the squared wavefunction at the origin  $|\Psi_{nlm}(0)|^2$ , which vanishes for all except  $S$ -waves. The main features of the decay can be understood by considering  $C$  conservation. The  $C$  parities of orthopositronium ( ${}^3S_1$ ) and parapositronium ( ${}^3S_0$ ) are  $-1$  and  $+1$ , respectively, while  $C = (-1)^N$  for a final state of  $N$  photons. Thus, while the two-photon decay

$${}^1S_0 \rightarrow \gamma + \gamma \quad (5.64)$$

of parapositronium is allowed, the decay  ${}^1S_0 \rightarrow 3\gamma$  is forbidden. Searches for the latter have been unsuccessful, giving an experimental upper limit on the ratio of decay rates

$$\frac{\Gamma({}^1S_0 \rightarrow 3\gamma)}{\Gamma({}^1S_0 \rightarrow 2\gamma)} < 1 \times 10^{-5}.$$

This should be compared with a value of order  $\alpha$ , which would be expected if both decays were allowed, since two-photon annihilations could then take place via the second-order diagrams of Figure 1.9, while three-photon annihilations can only proceed via third-order diagrams like Figure 1.10. Correspondingly, for orthopositronium the decay  ${}^3S_1 \rightarrow \gamma\gamma$  is forbidden by  $C$  conservation, while the third-order process



is allowed. Thus we would expect the ratio of the lifetimes of para-positronium and orthopositronium to be

$$\frac{\tau({}^1S_0; n=1)}{\tau({}^3S_1; n=1)} = O(\alpha), \quad (5.66)$$

where we allow the right-hand side to be interpreted also as  $O(\alpha/2\pi)$ , as discussed in Section 1.3.3.<sup>8</sup> Experimentally,

$$\frac{\tau({}^1S_0; n=1)}{\tau({}^3S_1; n=1)} = \frac{1.25 \times 10^{-10} \text{ s}}{1.42 \times 10^{-7} \text{ s}} = 0.88 \times 10^{-3},$$

in reasonable agreement with (5.66) and in precise agreement with more detailed theoretical calculations.

## \*5.6 Time reversal

Time-reversal invariance, or  $T$ -invariance, is defined as invariance under the transformation

$$t \rightarrow t' = -t, \quad (5.67)$$

leaving all position vectors unchanged. Like parity and charge conjugation invariance, it is a symmetry of the strong and electromagnetic interactions, but is violated by the weak interactions. However, unlike parity and charge conjugation, there is no associated quantum number that is conserved when weak interactions are neglected. To understand this we consider the transformation of a single-particle wavefunction, which must satisfy

$$|\Psi(\mathbf{r}, t)|^2 \xrightarrow{T} |\Psi'(\mathbf{r}, t)|^2 = |\Psi(\mathbf{r}, -t)|^2 \quad (5.68)$$

if the system is  $T$ -invariant, so that the probability of finding the particle at position  $\mathbf{r}$  at time  $-t$  becomes the probability of finding it at position  $\mathbf{r}$  at time  $t$  in the transformed system. In addition, since linear and angular momentum change sign under (5.67) in classical mechanics, we would expect the same result

$$\mathbf{p} \xrightarrow{T} \mathbf{p}' = -\mathbf{p}; \quad \mathbf{J} \xrightarrow{T} \mathbf{J}' = -\mathbf{J} \quad (5.69)$$

---

<sup>8</sup> See the discussion following Equation (1.20).

to hold in quantum mechanics by the correspondence principle. Hence a free-particle wavefunction

$$\Psi_{\mathbf{p}}(\mathbf{r}, t) = \exp[i(\mathbf{p} \cdot \mathbf{r} - Et)],$$

corresponding to momentum  $\mathbf{p}$  and energy  $E = p^2/2m$ , must transform into a wavefunction corresponding to momentum  $-\mathbf{p}$  and energy  $E$ , that is

$$\Psi_{\mathbf{p}}(\mathbf{r}, t) \xrightarrow{T} \Psi'_{\mathbf{p}}(\mathbf{r}, t) = \Psi_{-\mathbf{p}}(\mathbf{r}, t) = \exp[-i(\mathbf{p} \cdot \mathbf{r} + Et)]. \quad (5.70)$$

A suitable transformation that satisfies both (5.68) and (5.70) is

$$\Psi(\mathbf{r}, t) \xrightarrow{T} \Psi'(\mathbf{r}, t) = \Psi^*(\mathbf{r}, -t) \equiv \hat{T} \Psi(\mathbf{r}, t), \quad (5.71)$$

where we have introduced the time reversal operator  $\hat{T}$  by analogy with the parity operator  $\hat{P}$  introduced in Equation (5.32). However, quantum mechanical operators  $\hat{O}$  that correspond to physical observables must be both linear

$$\hat{O}(\alpha_1 \Psi_1 + \alpha_2 \Psi_2) = \alpha_1 (\hat{O} \Psi_1) + \alpha_2 (\hat{O} \Psi_2) \quad (5.72a)$$

and hermitian

$$\int (\hat{O} \Psi_1)^* \Psi_2 d\mathbf{x} = \int \Psi_1^* (\hat{O} \Psi_2) d\mathbf{x}, \quad (5.72b)$$

where  $\Psi_{1,2}$  are arbitrary wavefunctions and  $\alpha_{1,2}$  are arbitrary complex numbers.<sup>9</sup> In contrast, definition (5.71) implies

$$\hat{T}(\alpha_1 \Psi_1 + \alpha_2 \Psi_2) = \alpha_1^* (\hat{T} \Psi_1) + \alpha_2^* (\hat{T} \Psi_2) \neq \alpha_1 (\hat{T} \Psi_1) + \alpha_2 (\hat{T} \Psi_2)$$

for complex  $\alpha_1$  and  $\alpha_2$ , and one easily verifies that (5.72b) is also not satisfied by  $\hat{T}$ . Thus the time reversal operator does not correspond to a physical observable and there is no observable analogous to parity that is conserved as a consequence of  $T$  invariance.

Although  $T$  invariance does not give rise to a conservation law, it does lead to a relation between any reaction and the ‘time-reversed’ process related to it by (5.67). This is familiar in classical mechanics where, because Newton’s equations are second order in time, they are invariant under time reversal  $t \rightarrow -t$ . Consequently, for every process corresponding to a given set of particle trajectories, there is a corresponding process in which all the trajectories are reversed. A similar result holds in quantum mechanics and relates reactions like

$$a(\mathbf{p}_a, m_a) + b(\mathbf{p}_b, m_b) \rightarrow c(\mathbf{p}_c, m_c) + d(\mathbf{p}_d, m_d) \quad (5.73a)$$

---

<sup>9</sup> Linearity ensures that the superposition principle holds, while hermiticity ensures that the eigenvalues of  $\hat{O}$ , that is the observed values, are real.

to their time-reversed counterparts

$$c(-\mathbf{p}_c, -m_c) + d(-\mathbf{p}_d, -m_d) \rightarrow a(-\mathbf{p}_a, -m_a) + b(-\mathbf{p}_b, -m_b) \quad (5.73b)$$

in which the initial and final states are interchanged and the particle momenta ( $\mathbf{p}_a$ , etc.) and  $z$ -components of their spins ( $m_a$ , etc.) are reversed in accordance with (5.69). In particular, if weak interactions are neglected, the rates for reactions (5.73a) and (5.73b) must be equal.

### 5.6.1 Principle of detailed balance

A more useful relation between reaction rates can be obtained if we combine time reversal with parity invariance. Under the parity transformation (5.30), momenta  $\mathbf{p}$  change sign while orbital angular momenta  $\mathbf{L} = \mathbf{r} \times \mathbf{p}$  do not. If we assume the same behaviour holds for spin angular momenta, then

$$\mathbf{p} \xrightarrow{T} \mathbf{p}' = -\mathbf{p}; \quad \mathbf{J} \xrightarrow{T} \mathbf{J}' = \mathbf{J} \quad (5.74)$$

under parity. The parity-transformed reaction corresponding to (5.73b) is

$$c(\mathbf{p}_c, -m_c) + d(\mathbf{p}_d, -m_d) \rightarrow a(\mathbf{p}_a, -m_a) + b(\mathbf{p}_b, -m_b), \quad (5.73c)$$

so that if both  $P$  and  $T$  invariance holds, all three reactions (5.73a), (5.73b) and (5.73c) must have the same rate. If we average over all spin projections

$$m_i = -s_i, -s_i + 1, \dots, s_i \quad (i = a, b, c, d),$$

where  $s_i$  is the spin of particle  $i$ , then reactions (5.73a) and (5.73c) differ only by the interchange of initial and final states. Consequently, the rates for the reactions

$$i = a(\mathbf{p}_a) + b(\mathbf{p}_b) \leftrightarrow c(\mathbf{p}_c) + d(\mathbf{p}_d) = f \quad (5.75)$$

should be equal, provided that we average over all possible spin states. This relation is called the *principle of detailed balance*, and has been confirmed experimentally in a variety of strong and electromagnetic reactions. Here we will assume its validity and illustrate its use in a classic application.

### 5.6.2 Spin of the charged pion

Although we have assumed throughout that pions have spin-zero, we have not so far discussed any experimental evidence for this assignment. The first evidence was provided by using detailed balance to relate the cross-section for the reaction

$$p + p \rightarrow \pi^+ + d \quad (5.76a)$$

to the time-reversed process

$$\pi^+ + d \rightarrow p + p \quad (5.76b)$$

at the same centre-of-mass energy. The experiments were carried out using unpolarised beams and targets, so that an *average* over the possible initial spin states is required. However, all final-state particles were detected, irrespective of their spin projections, so the final spin states are *summed*. Because of this, the spin-averaged reaction rates must be multiplied by a factor  $(2S_f + 1)$  for each final-state particle of spin  $S_f$ . In addition, flux and phase space factors must be taken into account. When this is done, the differential cross-sections are given by<sup>10</sup>

$$\frac{d\sigma(p + p \rightarrow \pi^+ + d)}{d \cos \theta} = \frac{(2S_\pi + 1)(2S_d + 1)}{2\pi} \frac{p_\pi^2}{v_i v_f} \overline{|\mathcal{M}_{if}|^2} \quad (5.77a)$$

and

$$\frac{d\sigma(\pi^+ + d \rightarrow p + p)}{d \cos \theta} = \frac{(2S_p + 1)^2}{2\pi} \frac{p_p^2}{v_i v_f} \overline{|\mathcal{M}_{fi}|^2}, \quad (5.77b)$$

where  $\theta$  is the scattering angle,  $p_\pi$  and  $p_p$  are the magnitudes of the pion and proton momenta, and  $v_i$ ,  $v_f$  are the relative velocities of the  $pp$  and  $\pi^+d$  pairs, respectively, in the centre-of-mass frame. Here  $\overline{|\mathcal{M}_{fi}|^2}$  is the spin-average of the squared scattering amplitude  $|\mathcal{M}_{fi}|^2$ , which is in turn proportional to the transition rate between specific quantum states  $i$  and  $f$ . In terms of it, the principle of detailed balance takes the form

$$\overline{|\mathcal{M}_{if}|^2} = \overline{|\mathcal{M}_{fi}|^2}, \quad (5.78)$$

and taking the ratio of cross-sections (5.77b) gives

$$\frac{d\sigma(p + p \rightarrow \pi^+ + d)/d \cos \theta}{d\sigma(\pi^+ + d \rightarrow p + p)/d \cos \theta} = 3 \left( \frac{p_\pi}{p_p} \right)^2 (2S_\pi + 1),$$

where we have set  $S_p$  and  $S_d$  equal to their known values of  $1/2$  and  $1$ , respectively. This ratio is sensitive to the pion spin, and measurements of both cross-sections in the early 1950s clearly established the familiar result  $S_\pi = 0$ .

---

<sup>10</sup> These formulas are derived in Appendix B.

## Problems 5

- 5.1** Generalise the discussion leading to the conservation of linear momentum (5.8) to the case of an  $N$ -particle system described by the wavefunction  $\Psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N)$ .
- 5.2** Use the standard commutation relations for angular momentum operators to show that  $L$  and  $S$  remain good quantum numbers if the spin-dependent forces arise from a simple spin-orbit interaction, that is if

$$H = H_0 + \alpha \hat{\mathbf{L}} \cdot \hat{\mathbf{S}},$$

where

$$[H_0, \hat{\mathbf{L}}] = [H_0, \hat{\mathbf{S}}] = 0$$

and  $\alpha$  is a constant.

- 5.3** Show that parity invariance (5.31) leads to parity conservation, that is  $[H, \hat{P}] = 0$ , where

$$\hat{P} \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, t) = P_1 P_2 \dots \Psi(-\mathbf{r}_1, -\mathbf{r}_2, \dots, t).$$

- 5.4** The deuteron is a bound state of a proton and a neutron with spin 1 and positive parity. Show that it may only exist in the  ${}^3S_1$  and  ${}^3D_1$  states of the  $np$  system.
- 5.5** If  $p\bar{p}$  annihilation at rest proceeds via  $S$ -states, explain why the reaction  $p\bar{p} \rightarrow \pi^0\pi^0$  is forbidden as a strong or electromagnetic interaction.
- 5.6** Assuming that exotic states exist in the  $s\bar{s}d\bar{d}$  system, express in spectroscopic notation the lowest states that would be expected and give their corresponding  $J^P$  values.
- 5.7** Why is the reaction  $\pi^- + d \rightarrow n + n + \pi^0$  effectively forbidden for a  $\pi^-$  at rest, but proceeds at a normal rate for a strong reaction at higher energies?
- 5.8** The  $\eta(549)$  meson has spin-0 and is observed to decay to three-pion final states by the electromagnetic processes (5.57b) and (5.57c). Use this information to deduce the parity of the  $\eta$ , and hence explain why the decays  $\eta \rightarrow \pi^+ \pi^-$  and  $\eta \rightarrow \pi^0 \pi^0$  have never been observed.
- 5.9** List the  $J^{PC}$  values of all possible  $q\bar{q}$  states with (a)  $L = 0$  and (b)  $L = 1$ .
- 5.10** A neutral spin-2 meson  $M^0$  can decay via the strong interaction to a  $\pi^+ \pi^-$  final state. Use this to deduce its parity and charge conjugation quantum numbers.
- 5.11** Suppose that an intrinsic  $C$ -parity factor  $C_a$  is introduced into (5.50b), which then becomes  $\hat{C} |a, \Psi\rangle = C_a |\bar{a}, \Psi\rangle$  and analogously  $\hat{C} |\bar{a}, \Psi\rangle = C_{\bar{a}} |a, \Psi\rangle$ . Show that the eigenvalue corresponding to any eigenstate of  $C$  is independent of  $C_a$ , so that  $C_a$  cannot be measured.
- 5.12** Use the conservation of angular momentum and parity to determine possible angular momentum states in the reaction  $p + \bar{p} \rightarrow \pi^+ + \pi^-$  and express the allowed transitions in spectroscopic notation for initial states in which the orbital angular momentum is  $L = 0, 1$  or  $2$ . Does taking account of  $C$  invariance impose any further restrictions on this reaction?

- 5.13** List the  $n = 3$  states of positronium, together with their  $J^{PC}$  values. What are their binding energies, neglecting fine structure?
- 5.14** Consider a hydrogen-like atom composed of a negative muon and a positron. What would be its ground-state energy, ignoring fine structure? What would be the splitting between the corresponding  ${}^3S_1$  and  ${}^3S_0$  states when fine structure is taken into account?
- 5.15** What is the value of the ratio

$$R = \frac{d\sigma(\gamma + n \rightarrow \pi^- + p)/d\Omega}{d\sigma(\pi^- + p \rightarrow \gamma + n)/d\Omega}$$

if the centre-of-mass momentum of the pion is  $q = 200 \text{ MeV}/c$ ?

(*Note:* Despite having spin 1, the photon has only two distinct spin states, corresponding to the two polarisation states of electromagnetic waves.)

- 5.16** What is the behaviour of the electric and magnetic fields  $\mathbf{E}(\mathbf{r}, t)$  and  $\mathbf{B}(\mathbf{r}, t)$  under the operation of time reversal?

# 6

---

## The quark model

---

Young man, if I could remember the names of all these particles, I would have been a botanist.

*Enrico Fermi<sup>1</sup>*

The existence and properties of quarks were first inferred from the close correspondence between the observed hadrons and those predicted by the simple quark model, where the prefix ‘simple’ means that only  $q\bar{q}$ ,  $qqq$  and  $\bar{q}\bar{q}\bar{q}$  states are considered. In the main part of this chapter, we will explore this correspondence in more detail and use it to infer some fundamental properties of the quarks and the strong interactions between them. In particular, we will present evidence that the strong force is ‘flavour independent’, by which we mean that its properties do not depend on the flavour  $u, d, s\dots$  of the quark on which it is acting, so that, for example, the strong forces between  $qq = uu, ud, us, sd, \dots$  pairs are identical.<sup>2</sup> We will also introduce a new property of quarks, called *colour*, together with the hypothesis of *colour confinement*. This lies at the heart of our understanding of strong interactions and explains, among other things, why hadrons have integer electric charges, while fractionally charged combinations like  $qq$  and  $qq\bar{q}$  are forbidden.

In what follows, we shall not attempt a comprehensive discussion of the quark model,<sup>3</sup> but will concentrate on just five particularly important topics. To begin, we introduce an approximate symmetry of hadrons, called *isospin symmetry*, which has its origins in

---

<sup>1</sup> Remark made to Leon Lederman, reported in Lederman and Teresi (1993).

<sup>2</sup> The same principle applies to quark–antiquark forces, but these are not identical to quark–quark forces because in the former case annihilations can occur.

<sup>3</sup> Fuller descriptions may be found in Close (1979), Godfrey and Napolitano (1999), Klempert and Richard (2010) and Amsler *et al.* (2013).

flavour independence and the almost equal masses of the  $u$  and  $d$  quarks. In the next two sections we consider hadrons made of the light quarks  $u$ ,  $d$ ,  $s$  and their antiquarks, and then hadrons containing the heavy quarks  $b$  and  $c$ , restricting ourselves for simplicity to the lightest particles, in which the quarks have zero orbital angular momentum. In both cases, the simple quark model is strikingly successful in accounting for the observed states, but apparently violates the Pauli principle. This paradox is removed in the next section, where we introduce colour, with emphasis on its role in quark confinement. Finally, we consider the spectra of the quark–antiquark systems  $c\bar{c}$  and  $b\bar{b}$ , called *charmonium* and *bottomonium*, respectively. These are of considerable interest because the lower-lying states are approximately non-relativistic, and a simple discussion based on the Schrödinger equation leads both to a detailed understanding of the observed states and a determination of the quark–antiquark potential.

## 6.1 Isospin symmetry

Perhaps the most striking fact about hadrons is that they occur in families of particles with approximately equal masses. Within a given family, all particles have the same spin, parity, baryon number, and strangeness, charm and bottom quantum numbers, but differ in their electric charges. Many examples of this were seen in Chapter 3, including the nucleons

$$p(938) = uud, \quad n(940) = udd \quad (6.1)$$

and the  $K$ -mesons

$$K^+(494) = u\bar{s}, \quad K^0(498) = d\bar{s}. \quad (6.2)$$

This behaviour reflects a symmetry between  $u$  and  $d$  quarks: if

$$m_u = m_d \quad (6.3)$$

and if the forces acting on  $u$  and  $d$  quarks were exactly equal, then replacing the  $u$  quark in the  $K^+$  by a  $d$  quark would produce another particle – presumably the  $K^0$  – with exactly the same mass. The observed masses are slightly different, indicating that this symmetry in practice is not exact. The strong forces on  $u$  and  $d$  quarks are the same, as far as we know, but the electromagnetic forces are different because the quarks have different electric charges. In addition, the  $d$  quark is a few  $\text{MeV}/c^2$  heavier than the  $u$  quark, as we shall show below. However, this quark mass difference is small compared with typical hadron masses and electromagnetic forces are weak compared with strong forces. Consequently, *isospin symmetry*, as it is called, is a good approximation and greatly simplifies the interpretation of

hadron physics. Families of particles like (6.1) and (6.2) are called *isospin multiplets*.

### 6.1.1 Isospin quantum numbers

In order to formulate isospin symmetry more precisely, we introduce three more quantum numbers that are conserved in strong interactions. Two of these are not really new, but are combinations of ones previously introduced. The first is the *hypercharge*  $Y$  defined by

$$Y \equiv B + S + C + \tilde{B} + T, \quad (6.4)$$

where, as usual,  $B$ ,  $S$ ,  $C$ ,  $\tilde{B}$  and  $T$  are the baryon number, strangeness, charm, bottom and top, respectively. As these quantum numbers have the same values for all members of a given isospin multiplet, so does the hypercharge. The second combination is defined by

$$I_3 \equiv Q - Y/2, \quad (6.5)$$

where  $Q$  is the electric charge. Alternatively,  $I_3$  can be expressed in terms of the quark numbers

$$N_u = N(u) - N(\bar{u}) \quad \text{and} \quad N_d = N(d) - N(\bar{d}),$$

and we leave it as an exercise for the reader to show that (6.5) is equivalent to

$$I_3 = \frac{1}{2}(N_u - N_d) \quad (6.6)$$

using Equations (3.2) to (3.5).

Whereas the hypercharge  $Y$  takes the same value for each member of a multiplet,  $I_3$  takes different values, and if we call the maximum value within a multiplet  $I$ , that is

$$I \equiv (I_3)_{\max}, \quad (6.7)$$

then we find that all observed multiplets contain precisely  $(2I + 1)$  members, with  $I_3$  values

$$I_3 = I, I-1, \dots, -I. \quad (6.8)$$

This is very reminiscent of the formalism for the spin quantum numbers  $S$  and  $S_3$ , where

$$S_3 = S, S-1, \dots, -S, \quad (6.9)$$

and, by analogy,  $I$  is called the *isospin* quantum number and  $I_3$  the third component of isospin. This is illustrated in Table 6.1, where the values of these various quantum numbers are listed for four multiplets that we met in Chapter 3, including the ‘isosinglet’  $\Lambda$  with  $I = I_3 = 0$ . The theory of isospin symmetry is formally very similar to the theory of angular momentum. We shall not pursue this in detail here, but argue by analogy with the latter. (For the interested

**Table 6.1** Values of the baryon number  $B$ , hypercharge  $Y$ , charge  $Q$  and isospin quantum numbers  $I$  and  $I_3$  for some typical hadron isospin multiplets.

Particle	$B$	$Y$	$Q$	$I_3$	$I$
$\Lambda(1116)$	1	0	0	0	0
$p(938)$	1	1	1	$1/2$	$1/2$
$n(940)$	1	1	0	$-1/2$	$1/2$
$K^+(494)$	0	1	1	$1/2$	$1/2$
$K^0(498)$	0	1	0	$-1/2$	$1/2$
$\pi^+(140)$	0	0	1	1	1
$\pi^0(135)$	0	0	0	0	1
$\pi^-(140)$	0	0	-1	-1	1

reader, a more formal treatment of isospin is given in Appendix C.) In particular, the rules for addition of isospins are identical to those for the addition of ordinary spins or orbital angular momenta. Thus the sum of two isospins  $I^a$  and  $I^b$  can give the values

$$I^a + I^b, I^a + I^b - 1, \dots, |I^a - I^b|, \quad (6.10a)$$

while the sum of their third components is

$$I_3 = I_3^a + I_3^b. \quad (6.10b)$$

The latter result follows directly from the definition (6.5), since both  $Q$  and  $Y$  are additive quantum numbers.

### 6.1.2 Allowed quantum numbers

The isospin quantum numbers  $I$  and  $I_3$  could in principle take on any integer or half-integer values subject to the conditions (6.7) and (6.8). However, in the simple quark model, the only possible values are those that can arise from the quark combinations  $q\bar{q}$  and  $qqq$ , to which we now turn.

The values of  $Y$  and  $I_3$  for the quarks themselves can be deduced from (6.4) and (6.6) using the values given in Table 3.1, and are given in Table 6.2. In this table, the  $u$  and  $d$  quarks have been assigned to an isodoublet (i.e. an isospin doublet with  $I = 1/2$ ), since they have  $I_3 = 1/2$  and  $-1/2$ , respectively, and approximately equal masses.

**Table 6.2** Values of the baryon number  $B$ , hypercharge  $Y$ , charge  $Q$ , and isospin quantum numbers  $I$  and  $I_3$  for quarks. For the corresponding antiquarks,  $B$ ,  $Y$ ,  $Q$  and  $I_3$  are reversed in sign, while  $I$  remains unchanged.

Quark	$B$	$Y$	$Q$	$I_3$	$I$
$d$	$1/3$	$1/3$	$-1/3$	$-1/2$	$1/2$
$u$	$1/3$	$1/3$	$2/3$	$1/2$	$1/2$
$s$	$1/3$	$-2/3$	$-1/3$	0	0
$c$	$1/3$	$4/3$	$2/3$	0	0
$b$	$1/3$	$-2/3$	$-1/3$	0	0
$t$	$1/3$	$4/3$	$2/3$	0	0

**Table 6.3** Values of the isospin  $I$  and strangeness  $S$  allowed in the quark model for hadrons with  $C = \tilde{B} = T = 0$ .

(a) Baryons		(b) Mesons	
$S$	$I$	$S$	$I$
0	$3/2, 1/2$	1	$1/2$
-1	$1, 0$	0	$1, 0$
-2	$1/2$	-1	$1/2$
-3	0		

For antiquarks, the sign of  $I_3$  is reversed, so that the  $\bar{d}$  and  $\bar{u}$  antiquarks have  $I_3 = 1/2$  and  $-1/2$ , respectively, and are also assigned to an isodoublet. The remaining quarks and antiquarks all have  $I_3 = 0$ , with no partners of similar mass, and are assigned to isosinglets (i.e.  $I = 0$ ). With these assignments, the isospins of hadrons may be deduced using the rules for the addition of isospin given in (6.10).

The allowed isospins in the quark model are correlated to the values of the strangeness, charm and bottom quantum numbers. For simplicity, we will confine ourselves to zero values of charm, bottom and top since these are zero for the majority of known hadrons. For baryons the possibilities are

$$sss, \quad ssq_i, \quad sq_iq_j, \quad q_iq_jq_k, \quad (6.11)$$

where  $s$  is the strange quark with  $I = 0$  and  $q_i = u$  or  $d$  is a quark with  $I = 1/2$ . The  $s$  quark contributes nothing to the total isospin, so that the states  $sss$  with  $S = -3$  and  $ssq_i$  with  $S = -2$  must have  $I = 0$  and  $I = 1/2$ , respectively. The isospin of the states with  $S = -1$  is just the sum of the isospins of the quarks  $q_i$  and  $q_j$ , both with  $I = 1/2$ . By (6.10a) this is  $I = 0$  or 1. Finally, for states  $q_iq_jq_k$  with  $S = 0$ , we have to combine a third  $I = 1/2$  isospin with the latter results, giving  $I = 1/2$  or  $I = 3/2$ , respectively. These values are given in Table 6.3(a). The corresponding values for mesons with  $C = \tilde{B} = T = 0$ , summarised in Table 6.3(b), are left as an exercise for the reader. These values are the only ones allowed if mesons are  $q\bar{q}$  and baryons  $qqq$  states, provided  $C = \tilde{B} = T = 0$ , and so far they are the only values observed in nature.

### 6.1.3 An example: the sigma ( $\Sigma$ ) baryons

We will illustrate the use of internal quantum numbers, particularly isospin, by considering the  $\Sigma$  baryons.<sup>4</sup> The  $\Sigma^+(1189)$  is a particle that can be produced in the strong interaction



<sup>4</sup> Baryons with non-zero strangeness are also called *hyperons*.

and which decays to pions and nucleons,

$$\Sigma^+ \rightarrow \pi^+ + n, \quad \pi^0 + p, \quad (6.13)$$

with a lifetime of  $0.8 \times 10^{-10}$  s, typical of weak interactions. By applying the conservation laws of baryon number and strangeness, charm, bottom and top quantum numbers to (6.12), and using the known values for  $K^-(B=0, S=-1, C=\tilde{B}=T=0)$ ,  $p(B=1, S=C=\tilde{B}=T=0)$  and  $\pi^-(B=0, S=C=\tilde{B}=T=0)$ , it follows that, for the  $\Sigma^+$ ,

$$B = 1, \quad S = -1, \quad C = \tilde{B} = T = 0, \quad (6.14)$$

and hence, from (6.4) and (6.5),  $Y = 0$  and  $I_3 = 1$ .

If a particle is found with a non-zero value of  $I_3$ , it follows as a simple consequence of isospin symmetry that other particles belonging to the same isospin multiplet must exist. Thus the  $\Sigma^+$  is a member of a multiplet in which all the particles have the quantum numbers (6.14), and hence, by (6.7),  $Q = I_3$ . Within this multiplet it follows that there must be states with charges

$$Q = I_3 = I, I-1, \dots, -I,$$

where  $I$  must be an integer greater than or equal to one, since we already know that the  $\Sigma^+$  has  $I_3 = 1$ . The smallest multiplet possible is therefore an isotriplet ( $\Sigma^+, \Sigma^0, \Sigma^-$ ). Other partners would exist in addition to these if  $I \geq 2$ , but this would contradict the quark model (see Table 6.3(a)).

The above argument shows that isospin symmetry unambiguously predicts the existence of  $\Sigma^0$  and  $\Sigma^-$  particles with approximately the same mass as the  $\Sigma^+$  and the quantum numbers (6.14). These particles do indeed exist and can be produced in the reactions

$$K^- + p \rightarrow \pi^0 + \Sigma^0 \quad (6.15a)$$

and

$$K^- + p \rightarrow \pi^+ + \Sigma^-, \quad (6.15b)$$

which are similar to (6.12). The  $\Sigma^-(1197)$  decays by the strangeness-violating weak reaction

$$\Sigma^- \rightarrow \pi^- + n, \quad (6.16)$$

analogous to (6.13), while the  $\Sigma^0(1193)$  decays by the strangeness-conserving electromagnetic reaction

$$\Sigma^0 \rightarrow \Lambda + \gamma, \quad (6.17)$$

because the  $\Lambda(1116)$  has the same internal quantum numbers (6.14) as the  $\Sigma^0$  and is lighter. No doubly-charged partners have been found, confirming that the  $\Sigma^+, \Sigma^-$  and  $\Sigma^0$  are an isotriplet, as expected from the quark model.

### 6.1.4 The $u, d$ quark mass splitting

We next deduce the quark composition of the  $\Sigma$  baryons in order to discuss the origins of the small mass differences between them and obtain an estimate of the mass difference between the  $u$  and  $d$  quarks. From (6.14) we see that each  $\Sigma$  particle contains one  $s$  quark but no  $c$ ,  $b$  or  $t$  quarks. The identity of the remaining quarks as  $u$  or  $d$  can be inferred from the observed electric charges or from (6.6). Either way, we find the compositions

$$\Sigma^+(1189) = uus, \quad \Sigma^0(1193) = uds, \quad \Sigma^-(1197) = dds. \quad (6.18)$$

We now assume that the mass differences arise from quark mass differences and from electromagnetic interactions between pairs of quarks, and to obtain a rough estimate of the former we assume that the latter are proportional to the products of the quark charges. This requires that pairs of quarks are on average separated by the same distance and that their magnetic moments are proportional to their charges. This is a reasonable first approximation.<sup>5</sup> We then have

$$\begin{aligned} M(\Sigma^-) &= M_0 + m_s + 2m_d + \delta(e_d^2 + e_d e_s + e_d e_s) \\ &= M_0 + m_s + 2m_d + \delta/3, \end{aligned} \quad (6.19a)$$

$$\begin{aligned} M(\Sigma^0) &= M_0 + m_s + m_d + m_u + \delta(e_u e_d + e_u e_s + e_d e_s) \\ &= M_0 + m_s + m_d + m_u - \delta/3 \end{aligned} \quad (6.19b)$$

and

$$M(\Sigma^+) = M_0 + m_s + 2m_u + \delta(e_u^2 + e_u e_s + e_u e_s) = M_0 + m_s + 2m_u, \quad (6.19c)$$

where  $e_q$  are the quark charges,  $\delta$  is a constant and  $M_0$  is the contribution to the  $\Sigma$  masses arising from the strong interactions between the quarks. This latter contribution will be the same in all three cases if the strong interactions of  $u$  and  $d$  quarks are identical, as we have assumed. From these equations, we directly obtain

$$m_d - m_u = \frac{1}{3}[M(\Sigma^-) + M(\Sigma^0) - 2M(\Sigma^+)] = 4.0 \text{ MeV}/c^2$$

as a rough estimate of the quark mass difference from this simple model. This estimate agrees with the range of values

$$2 \lesssim m_d - m_u \lesssim 4 \text{ MeV}/c^2, \quad (6.20)$$

found in other more sophisticated determinations, and justifies our assertion that this quark mass difference is small compared with typical hadron masses.

---

<sup>5</sup> Magnetic moments of quarks are discussed in Section \*6.2.3.

## 6.2 The lightest hadrons

The lowest-lying states of any quark system have zero orbital angular momentum and the lightest quarks are the  $u$ ,  $d$  and  $s$  quarks. We therefore confine our attention to states composed of only these three quarks and/or their antiquarks, and having zero orbital angular momentum. As we shall see, there are 36 such states – 18 mesons and 18 baryons – whose quantum numbers match exactly those of the lightest experimentally observed hadrons.

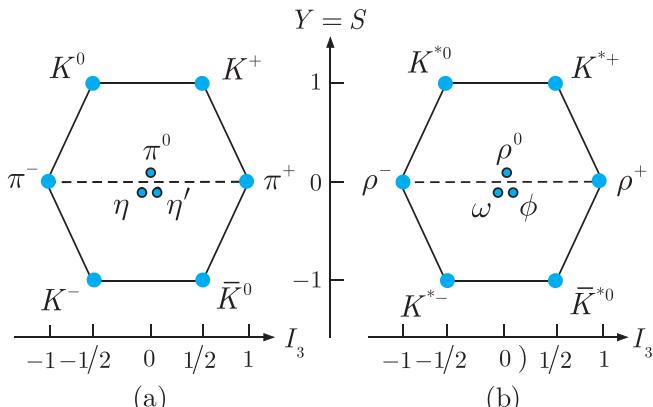
Since all the particles we discuss have zero charm, bottom and top, the hypercharge (6.4), which will be used repeatedly in this section, reduces to

$$Y = B + S. \quad (6.21)$$

### 6.2.1 The light mesons

The lightest mesons observed experimentally are a family of nine particles with spin-parity  $J^P = 0^-$ , called the *pseudoscalar meson nonet*, and a family of nine particles with spin-parity  $1^-$ , called the *vector meson nonet*. The members of each nonet are conveniently summarised on the so-called *weight diagrams* of Figures 6.1(a) and (b). In these diagrams, each particle is associated with a dot at a position corresponding to its quantum numbers  $(Y, I_3)$  and if two or more dots coincide they are displaced slightly for clarity. The pseudoscalar nonet of Figure 6.1(a) includes the familiar pions and kaons, together with the  $\eta(549)$  meson, which was discussed in Section 5.4.1, and a heavier meson called the  $\eta'(958)$ . The vector mesons (Figure 6.1(b)) are all resonances and include the  $K^*$  mesons discussed in Section 3.5.

Different members of the same isospin multiplet – like the pions or the kaons in Figure 6.1(a) – lie along horizontal lines in these weight diagrams, and on comparing Figures 6.1(a) and Figure 6.1(b) we see that precisely the same set of internal quantum number values



**Figure 6.1** Weight diagrams for (a) the  $0^-$  meson nonet and (b) the  $1^-$  meson nonet.

**Table 6.4** The states of the light  $L = 0$  meson nonets.

Quark content	$0^-$ state	$1^-$ state	$I_3$	$I$	$Y = S$
$u\bar{s}$	$K^+(494)$	$K^{*+}(892)$	$1/2$	$1/2$	$1$
$d\bar{s}$	$K^0(498)$	$K^{*0}(896)$	$-1/2$	$1/2$	$1$
$u\bar{d}$	$\pi^+(140)$	$\rho^+(768)^{(*)}$	$1$	$1$	$0$
$\frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d})$	$\pi^0(135)$	$\rho^0(768)^{(*)}$	$0$	$1$	$0$
$d\bar{u}$	$\pi^-(140)$	$\rho^-(768)^{(*)}$	$-1$	$1$	$0$
$s\bar{d}$	$\bar{K}^0(498)$	$\bar{K}^{*0}(896)$	$1/2$	$1/2$	$-1$
$s\bar{u}$	$K^-(494)$	$K^{*-}(892)$	$-1/2$	$1/2$	$-1$
See text	$\eta(549)$	$\omega(782)$	$0$	$0$	$0$
See text	$\eta'(958)$	$\phi(1019)$	$0$	$0$	$0$

\*The measured mass difference between the neutral and charged  $\rho$  mesons is  $m(\rho^0) - m(\rho^+) = 0.3 \pm 2.2 \text{ MeV}/c^2$ .

occurs for both nonets. These striking results are exactly what is expected if we are dealing with mesons  $M = a\bar{b}$  made of the lightest quarks  $a = u, d, s$  and antiquarks  $\bar{b} = \bar{u}, \bar{d}, \bar{s}$  with zero orbital angular momentum  $L = 0$ . For  $L = 0$ , the parity of the meson is

$$P = P_a P_{\bar{b}} (-1)^L = (-1)^{L+1} = -1$$

by (5.39a), while the spin of the mesons is just the sum of the quark and antiquark spins, and can be 0 or 1. Since there are nine possible combinations  $a\bar{b}$ , we arrive at two nonets of mesons with spin-parities  $0^-$  and  $1^-$ , as is experimentally observed. The  $Y$  and  $I_3$  values for each quark combination  $a\bar{b}$  are obtained by adding the values for  $a$  and  $\bar{b}$  listed in Table 6.2. The resulting values correspond to those for the observed mesons with the quark assignments listed in Table 6.4. In most cases these are completely unambiguous: for example, there is only one quark–antiquark pair  $u\bar{s}$  with the values  $I_3 = 1/2$ ,  $Y = S = 1$  corresponding to the  $K^+$  meson in the  $0^-$  nonet and the  $K^{*+}$  meson in the  $1^-$  nonet. However, there are three particles in each nonet with

$$I_3 = Y = S = 0$$

which could correspond to any of the particle–antiparticle pairs

$$u\bar{u}, \quad d\bar{d} \quad \text{or} \quad s\bar{s}, \quad (6.22)$$

or to any linear combination of them. It can be shown that the linear combinations of (6.22) corresponding to definite isospins  $I = 0$  or  $I = 1$  are<sup>6</sup>

$$\frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d}) \quad (I = 1, \quad I_3 = 0), \quad (6.23)$$

$$\frac{1}{\sqrt{2}}(u\bar{u} + d\bar{d}) \quad (I = 0, \quad I_3 = 0) \quad (6.24a)$$

<sup>6</sup> These results are derived as Equations (C.24b) and (C.26) in Appendix C.

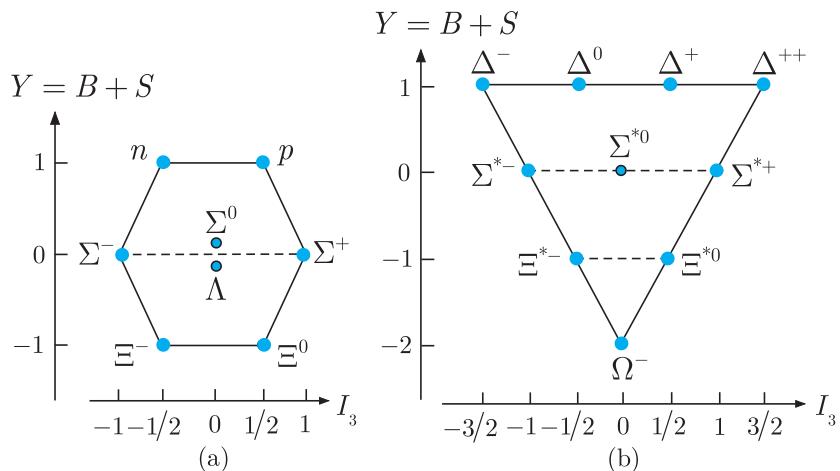
and

$$s\bar{s} \quad (I = 0, I_3 = 0) \quad (6.24b)$$

so that the quark assignments of the  $I = 1, I_3 = S = 0$  states, corresponding to the  $\pi^0$  and  $\rho^0$  mesons in Table 6.4 are unique. In contrast, the assignment of the  $I = I_3 = S = 0$  particles to the quark states (6.24a) and (6.24b), or to linear combinations of them, is more complicated, as we will see in Section \*6.2.4.

## 6.2.2 The light baryons

Families of hadrons with the same baryon number, spin and parity, like the  $0^-$  and  $1^-$  meson nonets, are called *supermultiplets*, and all hadrons belong to such families. For mesons, these supermultiplets are all nonets, whereas baryon supermultiplets can have 1, 8 or 10 members, and are called *singlets*, *octets* and *decuplets*, respectively. Here we shall consider only the lightest baryon supermultiplets that are experimentally observed, which are an octet of  $J^P = 1/2^+$  particles and a decuplet of  $J^P = 3/2^+$  particles. These are summarised in the weight diagrams of Figure 6.2, where for baryons the hypercharge is given by  $Y = S + 1$  by (6.21). The  $1/2^+$  octet includes the nucleons, the  $\Lambda(1116)$  and the  $\Sigma$  particles discussed in Section 6.1.3, together with the xi or ‘cascade’ particles  $\Xi^0$  and  $\Xi^-$ , which decay by weak interactions with lifetimes of order  $10^{-10}$  s. The  $3/2^+$  states are all resonances, except for the  $\Omega^-(1672)$  which decays by the weak interaction<sup>7</sup> with a lifetime of order  $10^{-10}$  s.



**Figure 6.2** Weight diagrams for (a) the  $J^P = 1/2^+$  octet of light baryons and (b) the  $J^P = 3/2^+$  baryon decuplet.

<sup>7</sup> This particle has strangeness  $S = -3$ . It cannot decay by strong interactions, unlike the other members of the decuplet, because there is no lighter state with  $S = -3$  to which it can decay.

We shall now show that the remarkable patterns of experimental states shown in Figures 6.2(a) and (b) are predicted by the quark model if we assume that the combined space and spin wavefunctions are *symmetric* under the interchange of like quarks. At first sight this seems impossible, since quarks are spin-1/2 fermions and hence obey the Pauli exclusion principle. The explanation will emerge in Section 6.4, where we introduce the fundamental concept of ‘colour’. Here we shall simply assume that only overall symmetric space–spin wavefunctions are allowed and show that this assumption leads to the octet and decuplet states that are observed.

The baryons of Figure 6.2 correspond to three-quark states  $B = abc$  containing only the light quarks  $u, d, s$  with zero orbital angular momentum. Hence their parity

$$P = P_a P_b P_c = 1, \quad (6.25)$$

by (5.39b), while the baryon spin is equal to the sum of the quark spins. In addition, for simple systems the lowest-lying states almost invariably have spatial wavefunctions that are symmetric under the interchange of like particles, and we will assume this to be the case here. Since we are assuming symmetric space–spin wavefunctions, this implies that the spin wavefunction must also be symmetric under the exchange of like quarks. That is, any like pair of quarks  $aa$  must have spin 1, corresponding to parallel quark spins and the symmetric spin wavefunction (5.43) and may not have spin 0, corresponding to antiparallel spins and the antisymmetric spin wavefunction (5.44). Hence in the six combinations

$$uud, \quad uus, \quad ddu, \quad dds, \quad ssu, \quad ssd \quad (6.26a)$$

of the form  $B = aab$ , where  $a \neq b$ , the  $aa$  pair must be in a spin-1 state. Adding the spin of the third quark gives

$$J = S = 1/2, \quad 3/2, \quad (6.26b)$$

as the possible values for the baryon spin, since the orbital angular momentum is zero. For the three combinations

$$uuu, \quad ddd, \quad sss \quad (6.27a)$$

the only way to have parallel spins for all pairs is for all three quarks to have their spins parallel, giving

$$J = S = 3/2 \quad (6.27b)$$

as the only possibility for the baryon spin. Finally, for the remaining case

$$uds, \quad (6.28a)$$

**Table 6.5** The states of the  $L = 0, J^P = 1/2^+$  octet of light baryons.

Quark composition	Observed state	$I_3$	$I$	$S$
$uud$	$p(938)$	$1/2$	$1/2$	$0$
$udd$	$n(940)$	$-1/2$	$1/2$	$0$
$uds$	$\Lambda(1116)^{(*)}$	$0$	$0$	$-1$
$uus$	$\Sigma^+(1189)$	$1$	$1$	$-1$
$uds$	$\Sigma^0(1193)^{(*)}$	$0$	$1$	$-1$
$dds$	$\Sigma^-(1197)$	$-1$	$1$	$-1$
$uss$	$\Xi^0(1315)$	$1/2$	$1/2$	$-2$
$dss$	$\Xi^-(1321)$	$-1/2$	$1/2$	$-2$

\*The  $\Lambda(1116)$  and the  $\Sigma^0(1193)$  differ in that in the former the  $ud$  pair is in a spin-0 state, while in the latter it is in a spin-1 state, as discussed at the end of Section 6.2.2.

in which all the quarks are different, there are two spin-1/2 baryons and one spin-3/2 baryon, that is

$$J = S = 1/2, \quad 1/2, \quad 3/2. \quad (6.28b)$$

These arise because the  $ud$  pair can have spin-0, leading to  $S = 1/2$  on adding the spin of the  $s$  quark, or spin-1, leading to either  $S = 1/2$  or  $S = 3/2$  on adding the spin of the  $s$  quark.

The rest is bookkeeping. All the above states have positive parity by (6.25), and on counting the various spin states in (6.26), (6.27) and (6.28) we find that there are eight with  $J^P = 1/2^+$  and ten with  $J^P = 3/2^+$ , as observed experimentally.<sup>8</sup> Furthermore, the  $Y$  and  $I_3$  values of the various states, obtained by adding the values for the individual quarks given in Table 6.2, correspond exactly to those of the observed states with the assignments given in Tables 6.5 and 6.6.

We finally comment on one detailed point that will be required in our discussion of baryon magnetic moments in Section \*6.2.3 below.<sup>9</sup> In the argument following (6.28a), we saw that there were two  $uds$  states with spin 1/2, corresponding to  $S = 0$  and  $S = 1$  for the spin of the  $ud$  pair alone. These states have  $Y = I_3 = 0$ , corresponding to the observed particles  $\Lambda(1116)$  and  $\Sigma^0(1193)$  in Table 6.5, and the question is: ‘Which is which?’ To resolve this ambiguity, we recall that the  $\Sigma^0$  is a member of the isotriplet

$$\Sigma^+(1189) = uus, \quad \Sigma^0(1193) = uds, \quad \Sigma^-(1197) = dds \quad (6.29)$$

discussed in Section 6.1.3, and that the members of isospin multiplets differ in the replacement of  $u$  quarks by  $d$  quarks (or vice versa), with the wavefunctions unchanged. Since in the  $\Sigma^+$  the two like quarks

<sup>8</sup> If we had assumed that the space–spin wavefunction was antisymmetric, then we would instead have obtained a quite different spectrum (see Problem 6.6).

<sup>9</sup> Readers who plan to omit this starred section may go directly to Section 6.3.

**Table 6.6** The states of the  $L = 0, J^P = 3/2^+$  decuplet of light baryons.

Quark composition	Observed state	$I_3$	$I$	$S$
$uuu$	$\Delta^{++}(1232)^{(*)}$	$3/2$	$3/2$	$0$
$uud$	$\Delta^+(1232)^{(*)}$	$1/2$	$3/2$	$0$
$udd$	$\Delta^0(1232)^{(*)}$	$-1/2$	$3/2$	$0$
$ddd$	$\Delta^-(1232)^{(*)}$	$-3/2$	$3/2$	$0$
$uus$	$\Sigma^+(1383)$	$1$	$1$	$-1$
$uds$	$\Sigma^0(1384)$	$0$	$1$	$-1$
$dds$	$\Sigma^-(1387)$	$-1$	$1$	$-1$
$uss$	$\Xi^0(1532)$	$1/2$	$1/2$	$-2$
$dss$	$\Xi^-(1535)$	$-1/2$	$1/2$	$-2$
$sss$	$\Omega^-(1672)$	$0$	$0$	$-3$

\*We do not show the small mass differences between the various charged and neutral states of the  $\Delta$  as these are not all precisely determined.

$uu$  must be in the symmetric spin-1 state, it follows that in the  $\Sigma^0$  the  $ud$  pair must also have spin 1. The remaining spin-1/2  $uds$  state, in which the  $ud$  pair is in a spin singlet state, is then assigned to the  $\Lambda(1116)$ .

### ★6.2.3 Baryon magnetic moments

We now consider the magnetic moments of the  $1/2^+$  octet of baryons, which have all been measured except for that of the very short-lived  $\Sigma^0(1193)$ . If the quarks have zero orbital angular momenta, then these magnetic moments are just the sums of contributions from the quark magnetic moments, which we assume to be of the Dirac form (1.14). That is, we assume

$$\mu_q \equiv \left\langle q, S_z = \frac{1}{2} \middle| \hat{\mu}_z \middle| q, S_z = \frac{1}{2} \right\rangle = e_q e / 2m_q = (e_q M_p / m_q) \mu_N, \quad (6.30a)$$

where  $e_q$  is the quark charge in units of  $e$  and

$$\mu_N = e / 2M_p \quad (6.30b)$$

is the nuclear magneton. In the case of the  $\Lambda(1116) = uds$ , the  $ud$  pair is in a spin-0 state as discussed in Section 6.2.2 and hence makes no contribution to the  $\Lambda$  spin or magnetic moment, so that

$$\mu_\Lambda = \mu_s = -\frac{1}{3} \frac{M_p}{m_s} \mu_N, \quad (6.31a)$$

where we have used (6.30a). For  $1/2^+$  baryons  $B$  with the quark configuration  $aab$ , the  $aa$  pair is in the symmetric spin-1 state, as

shown in Section 6.2.2, with parallel spins and magnetic moment  $2\mu_a$ . The spin-up baryon state is given by<sup>10</sup>

$$\begin{aligned} \left|B; S = \frac{1}{2}, S_z = \frac{1}{2}\right\rangle &= \sqrt{\frac{2}{3}} \left|b; S = \frac{1}{2}, S_z = -\frac{1}{2}\right\rangle |aa; S = 1, S_z = 1\rangle \\ &\quad - \sqrt{\frac{1}{3}} \left|b; S = \frac{1}{2}, S_z = \frac{1}{2}\right\rangle |aa; S = 1, S_z = 0\rangle. \end{aligned}$$

The first term corresponds to a state with magnetic moment  $2\mu_a - \mu_b$ , since the  $b$  quark has  $S_z = -1/2$ , and the second term to a state with magnetic moment  $\mu_b$ , since the  $aa$  pair has  $S_z = 0$  and does not contribute. Hence the baryon magnetic moment is given by

$$\mu_B = \frac{2}{3}(2\mu_a - \mu_b) + \frac{1}{3}\mu_b = \frac{4}{3}\mu_a - \frac{1}{3}\mu_b. \quad (6.31b)$$

In particular, the proton magnetic moment is given by

$$\mu_p = \frac{4}{3}\mu_u - \frac{1}{3}\mu_d = \frac{M_p}{m_{u,d}}\mu_N, \quad (6.31c)$$

where we have neglected the small mass difference between the  $u$  and  $d$  quarks. The predictions (6.31a) and (6.31c) agree with the measured values of the  $p$  and  $\Lambda$  magnetic moments for the constituent quark masses

$$m_s = 510 \text{ MeV}/c^2, \quad m_{u,d} = 336 \text{ MeV}/c^2. \quad (6.32)$$

Substituting these values into (6.30) and (6.31b), and using the quark assignments of Table 6.5, leads to predictions for the other magnetic moments, which are compared with experiment in Table 6.7.

The agreement is reasonable, but not exact. This suggests that while the assumption that the observed baryons are three-quark

**Table 6.7** A comparison of the observed magnetic moments of the  $1/2^+$  baryon octet and the predictions of the simple quark model, Equations (6.31a) and (6.31b), for  $m_u = m_d = 336 \text{ MeV}/c^2$ ,  $m_s = 510 \text{ MeV}/c^2$ .

Particle	Prediction ( $\mu_N$ )	Experiment ( $\mu_N$ )
$p(938)$	2.79	$2.793^{(*)}$
$n(940)$	-1.86	$-1.913^{(*)}$
$\Lambda(1116)$	-0.61	$-0.613 \pm 0.004$
$\Sigma^+(1189)$	2.69	$2.458 \pm 0.010$
$\Sigma^-(1197)$	-1.04	$-1.160 \pm 0.025$
$\Xi^0(1315)$	-1.44	$-1.250 \pm 0.014$
$\Xi^-(1322)$	-0.51	$-0.6507 \pm 0.0025$

\*The errors on the proton and neutron magnetic moments are of the order  $6 \times 10^{-8}$  and  $5 \times 10^{-7}$ , respectively.

<sup>10</sup>This result is derived in Appendix C, Equation (C.36c).

states with zero orbital angular momentum is a good approximation, there could be a small admixture of other states with the same quantum numbers. One possibility is suggested by our discussion of the deuteron magnetic moment in Section 5.2.2. This showed that the orbital angular momentum quantum number is only approximately conserved in strong interactions. Thus the ground states of strongly interacting systems are not in general pure  $S$ -waves, but could contain small admixtures of states with non-zero orbital angular momenta. Such admixtures are neglected in the simple quark model, but may well occur. A second possibility is that in addition to the dominant  $qqq$  state, there may be small admixtures of the more complicated  $qqqq\bar{q}$  states. We return to this possibility briefly in Sections 6.4.1 and 7.1.3.

#### \*6.2.4 Hadron mass splittings

The mass differences between different members of a given supermultiplet are conveniently separated into the small mass differences between members of the same isospin multiplet and the large mass differences between members of different isospin multiplets. Here we will concentrate on the latter, ignoring the small mass differences within isospin multiplets.<sup>11</sup> That is, we will work in the approximation that all members of an isospin multiplet have the same mass, so that, for example,

$$m_p = m_n \equiv m_N$$

for the nucleons.

From Tables 6.5 and 6.6 we see that within each supermultiplet the heaviest isospin multiplets are those that contain the most  $s$  quarks, suggesting that the main contribution to the mass differences may be the mass difference between the  $s$  quark and the lighter  $u$  and  $d$  quarks. If this were the only contribution, the mass differences could be calculated directly from the sums of the masses of the constituent quarks. For the  $3/2^+$  decuplet this gives

$$M_\Omega - M_\Xi = M_\Xi - M_\Sigma = M_\Sigma - M_\Delta = m_s - m_{u,d}$$

and for the  $1/2^+$  octet

$$M_\Xi - M_\Sigma = M_\Xi - M_\Lambda = M_\Lambda - M_N = m_s - m_{u,d},$$

where  $m_s - m_{u,d} \approx 170 \text{ MeV}/c^2$  from (6.32). Experimentally, the above mass differences are  $(142, 145, 153) \text{ MeV}/c^2$  for the  $3/2^+$  decuplet, and  $(123, 202, 177) \text{ MeV}/c^2$  for the  $1/2^+$  octet, where the mass of each isospin multiplet has been taken to be the mean mass of its members.

---

<sup>11</sup>These have been discussed in Section 6.1.4.

These results support the suggestion that baryon mass differences (and by analogy meson mass differences) are dominantly due to the mass differences of their constituent quarks. However, this cannot be the complete explanation, because if it were then the  $1/2^+$  nucleon would have the same mass as the  $3/2^+ \Delta(1232)$ , as they have the same quark constituents, and similarly for other related particles in the  $1/2^+$  octet and  $3/2^+$  decuplet. The absence of orbital angular momentum in these states means that there is nothing equivalent to the ‘fine structure’ of atomic physics. The difference lies in the spin structures of these states.

In atomic physics, there are very small splittings in the energy levels resulting from the interaction between the magnetic moments of the constituent particles. This is called the *hyperfine interaction*, and for two particles in a relative S-state it can be shown that the interaction energy is given by<sup>12</sup>

$$\Delta E = -\frac{2}{3}\mu_0 |\psi(0)|^2 \boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j, \quad (6.33a)$$

where  $\boldsymbol{\mu}_i$  and  $\boldsymbol{\mu}_j$  are the magnetic moments of the particles,  $\psi(0)$  is the wavefunction at the origin and  $\mu_0$  is the magnetic permeability of free space ( $\mu_0 = 1/\epsilon_0$  in natural units). In the case of point-like particles, the magnetic moments are of the Dirac form  $\mu_i = (e_i/m_i) \mathbf{S}_i$  and so (6.33a) can be written

$$\Delta E = -\frac{2}{3}\mu_0 \frac{e_i e_j}{m_i m_j} |\psi(0)|^2 \mathbf{S}_i \cdot \mathbf{S}_j. \quad (6.33b)$$

In the hadron case, the constituent particles are quarks and the electric charges in (6.33b) must be replaced by their strong interaction equivalents, with appropriate changes to the overall numerical factor. The resulting interaction, called the *chromomagnetic interaction*, is different for  $qq$  and  $q\bar{q}$  pairs, but is independent of the quark flavours, for reasons that will become clear in Chapter 7. As we cannot calculate the equivalent quark–quark wavefunction, as a first approximation we will write the contribution to a hadron mass as

$$\Delta M = a \frac{\mathbf{S}_1 \cdot \mathbf{S}_2}{m_1 m_2},$$

where  $a$  is a constant to be found from experiment. This assumes that the wave function is the same for all states, which will not be exactly true.

Consider firstly the case of mesons. By writing the total spin squared as

$$\mathbf{S}^2 \equiv (\mathbf{S}_1 + \mathbf{S}_2)^2 = \mathbf{S}_1^2 + \mathbf{S}_2^2 + 2\mathbf{S}_1 \cdot \mathbf{S}_2,$$

---

<sup>12</sup>See, for example, p. 307 of Gasiorowicz (1979).

we easily find that the expectation values of  $\mathbf{S}_1 \cdot \mathbf{S}_2$  in natural units are  $-3/4$  for the  $\mathbf{S} = \mathbf{0}$  mesons and  $1/4$  for the  $\mathbf{S} = \mathbf{1}$  mesons. The masses may then be written

$$M(\text{meson}) = m_1 + m_2 + \Delta M, \quad (6.34\text{a})$$

where  $m_{1,2}$  are the masses of the constituent quarks and

$$\begin{aligned} \Delta M(J^P = 0^- \text{ meson}) &= -\frac{3a}{4} \frac{1}{m_1 m_2}, \\ \Delta M(J^P = 1^- \text{ meson}) &= \frac{a}{4} \frac{1}{m_1 m_2}. \end{aligned} \quad (6.34\text{b})$$

The masses of the members of the  $0^-$  and  $1^-$  meson supermultiplets then follow from a knowledge of their quark compositions. For example, the  $K$ -mesons have one  $u$  or  $d$  quark and one  $s$  quark and so

$$M_K = m + m_s - \frac{3a}{4mm_s}.$$

The corresponding equations for the other masses are given in Table 6.8. The quark masses can then be inferred from the two relations

$$m_\pi + 3m_\rho = 8m, \quad m_K + 3m_{K^*} = 4(m + m_s)$$

which follow from these equations, and the parameter  $a$  from the resulting masses, together with the mass of the  $\rho$  meson. The results are

$$m = 0.306 \text{ GeV}/c^2, \quad m_s = 0.487 \text{ GeV}/c^2, \quad a = 0.0592 (\text{GeV}/c^2)^3, \quad (6.35)$$

and the predicted masses for other mesons are in good agreement with experiment, as is also shown in Table 6.8. Note that the quark mass values are similar to, but somewhat smaller than, those

**Table 6.8** Predicted meson masses (in  $\text{GeV}/c^2$ ) in the approximation (6.34) compared with experimental values. The predictions have been obtained using  $m = 0.306 \text{ GeV}/c^2$ ,  $m_s = 0.487 \text{ GeV}/c^2$  and  $a = 0.0592 (\text{GeV}/c^2)^3$ .

Particle	Mass	Prediction	Experiment
$\pi$	$2m - \frac{3a}{4m^2}$	0.14	0.137
$K$	$m + m_s - \frac{3a}{4mm_s}$	0.50	0.496
$\rho$	$2m + \frac{a}{4m^2}$	0.77	0.770
$\omega$	$2m + \frac{a}{4m^2}$	0.77	0.782
$K^*$	$m + m_s + \frac{a}{4mm_s}$	0.89	0.892
$\phi$	$2m_s + \frac{a}{4m_s^2}$	1.04	1.020

obtained previously from the baryon magnetic moments, reflecting the different approximations made. In addition, in Table 6.8 we have assumed that the  $\omega$  meson corresponds to the light quark combination (6.24a) while the  $\phi$  is a pure  $s\bar{s}$  state (6.24b). As can be seen from Table 6.8, this automatically leads to the result  $m_\rho = m_\omega$ , which is in good agreement with experiment. On the other hand, the same analogous assumption for the pseudoscalar nonet leads to the result  $m_\pi = m_\eta$ . This is obviously a very bad approximation, indicating that the eta and eta-prime are mixtures of the quark states (6.24a) and (6.24b), and we will not pursue the matter further.<sup>13</sup>

For baryons we have three quarks and the mass formula becomes

$$M = m_1 + m_2 + m_3 + b \sum_{i < j} \frac{\mathbf{S}_i \cdot \mathbf{S}_j}{m_i m_j}, \quad i, j = 1, 3, \quad (6.36)$$

where  $b$  is a constant to be determined from experiment. (There is no reason for  $b$  to be equal to the constant  $a$  used in the meson case because the quark wavefunctions and numerical factors in the baryonic equivalent of (6.33) will be different in the two cases.) In the case of the  $3/2^+$  decuplet, all three quarks have their spins aligned and every pair therefore combines to make spin 1. Thus, for example,

$$(\mathbf{S}_1 + \mathbf{S}_2)^2 = \mathbf{S}_1^2 + \mathbf{S}_2^2 + 2\mathbf{S}_1 \cdot \mathbf{S}_2 = 2,$$

giving  $\mathbf{S}_1 \cdot \mathbf{S}_2 = 1/4$  and in general

$$\mathbf{S}_1 \cdot \mathbf{S}_2 = \mathbf{S}_1 \cdot \mathbf{S}_3 = \mathbf{S}_2 \cdot \mathbf{S}_3 = 1/4.$$

Using this result, the mass of the  $\Sigma^*(1385)$ , for example, may be written

$$M_{\Sigma^*} = 2m + m_s + \frac{b}{4} \left( \frac{1}{m^2} + \frac{2}{mm_s} \right).$$

In the case of the  $1/2^+$  octet, we have

$$\begin{aligned} (\mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_3)^2 &= \mathbf{S}_1^2 + \mathbf{S}_2^2 + \mathbf{S}_3^2 + 2(\mathbf{S}_1 \cdot \mathbf{S}_2 + \mathbf{S}_1 \cdot \mathbf{S}_3 + \mathbf{S}_2 \cdot \mathbf{S}_3) \\ &= 3/4 \end{aligned}$$

and hence

$$\mathbf{S}_1 \cdot \mathbf{S}_2 + \mathbf{S}_1 \cdot \mathbf{S}_3 + \mathbf{S}_2 \cdot \mathbf{S}_3 = -3/4. \quad (6.37)$$

In addition, we have to consider the symmetry of the spin wave functions of individual hadrons. For example, the spins of the  $u$  and  $d$  pair in the  $\Lambda$  must combine to give  $\mathbf{S} = \mathbf{0}$ , as discussed at the end

---

<sup>13</sup>See, for example, pp. 155–157 of Perkins (1987).

of Section 6.2.2. In this case,  $(\mathbf{S}_u + \mathbf{S}_d)^2 = 0$ , so that  $\mathbf{S}_u \cdot \mathbf{S}_d = -3/4$ . Then,

$$M_\Lambda = m_u + m_d + m_s + b \left[ \frac{\mathbf{S}_u \cdot \mathbf{S}_d}{m_u m_d} + \frac{\mathbf{S}_u \cdot \mathbf{S}_s}{m_u m_s} + \frac{\mathbf{S}_d \cdot \mathbf{S}_s}{m_d m_s} \right].$$

Finally, setting  $m_u = m_d = m$  gives

$$\begin{aligned} M_\Lambda &= 2m + m_s + b \left[ \frac{\mathbf{S}_u \cdot \mathbf{S}_d}{m^2} + \frac{(\mathbf{S}_u \cdot \mathbf{S}_d + \mathbf{S}_u \cdot \mathbf{S}_s + \mathbf{S}_d \cdot \mathbf{S}_s - \mathbf{S}_u \cdot \mathbf{S}_d)}{m m_s} \right] \\ &= 2m + m_s - \frac{3b}{4m^2}, \end{aligned} \quad (6.38)$$

where we have used (6.37). The resulting formulas for all the  $1/2^+$  octet and  $3/2^+$  decuplet masses are shown in Table 6.9, and imply

$$m_N + m_\Delta = 6m, \quad m_\Delta - m_N = 6b/4m^2, \quad m_N - m_\Lambda = m - m_s$$

Substituting the experimental masses gives

$$m = 0.362 \text{ GeV}/c^2, \quad m_s = 0.539 \text{ GeV}/c^2, \quad b = 0.0256 (\text{GeV}/c^2)^3 \quad (6.39)$$

and the resulting predictions for the other masses agree well with experiment, as shown in Table 6.9. Again, the agreement between these quark masses and those previously obtained is reasonable, but not perfect. This is to be expected. Since free quarks are not observed, their masses cannot be measured directly and, in

**Table 6.9** Baryon masses (in  $\text{GeV}/c^2$ ) in the approximation (6.36) compared with experimental values. These have been obtained using  $m = 0.362 \text{ GeV}/c^2$ ,  $m_s = 0.539 \text{ GeV}/c^2$  and  $b = 0.0256 (\text{GeV}/c^2)^3$ .

Particle	Mass	Prediction	Experiment
N	$3m - \frac{3b}{4m^2}$	0.94	0.939
$\Lambda$	$2m + m_s - \frac{3b}{4m^2}$	1.12	1.116
$\Sigma$	$2m + m_s + \frac{b}{4} \left( \frac{1}{m^2} - \frac{4}{mm_s} \right)$	1.18	1.193
$\Xi$	$m + 2m_s + \frac{b}{4} \left( \frac{1}{m_s^2} - \frac{4}{mm_s} \right)$	1.32	1.318
$\Delta$	$3m + \frac{3b}{4m^2}$	1.23	1.232
$\Sigma^*$	$2m + m_s + \frac{b}{4} \left( \frac{1}{m^2} + \frac{2}{mm_s} \right)$	1.38	1.385
$\Xi^*$	$m + 2m_s + \frac{b}{4} \left( \frac{2}{mm_s} + \frac{1}{m_s^2} \right)$	1.53	1.533
$\Omega$	$3m_s + \frac{3b}{4m_s^2}$	1.68	1.673

**Table 6.10** The lightest observed mesons containing  $c$  and  $b$  quarks. Apart from the  $c\bar{c}$  and  $b\bar{b}$  states, they all have distinct antiparticles, which are not listed. For the  $c\bar{b}$  case, with  $C = B = 1$ , only the lightest state has so far been observed.

Quark content	$J^P = 0^-$	$J^P = 1^-$	$I_3$	$I$	$S$	$C$	$\tilde{B}$
$c\bar{u}$	$D^0(1865)$	$D^{*0}(2007)$	$-1/2$	$1/2$	0	1	0
$c\bar{d}$	$D^+(1870)$	$D^{*+}(2010)$	$1/2$	$1/2$	0	1	0
$c\bar{s}$	$D_s^+(1968)$	$D_s^{*+}(2112)$	0	0	1	1	0
$c\bar{c}$	$\eta_c(2984)$	$J/\psi(3097)$	0	0	0	0	0
$b\bar{u}$	$B^-(5279)$	$B^{*-}(5325)$	$-1/2$	$1/2$	0	0	-1
$b\bar{d}$	$B^0(5280)$	$B^{*0}(5325)$	$1/2$	$1/2$	0	0	-1
$b\bar{s}$	$B_s^0(5367)$	$B_s^{*0}(5415)$	0	0	1	0	-1
$b\bar{b}$	$\eta_b(9398)$	$\Upsilon(9460)$	0	0	0	0	0
$c\bar{b}$	$B_c^+(6276)$		0	0	0	1	1

the absence of a precise theory, the constituent masses are best regarded as effective masses inferred indirectly in the context of particular models and environments; while reasonable models of the light hadrons always give approximately the same values, there will inevitably be some variation reflecting the different approximations made.<sup>14</sup>

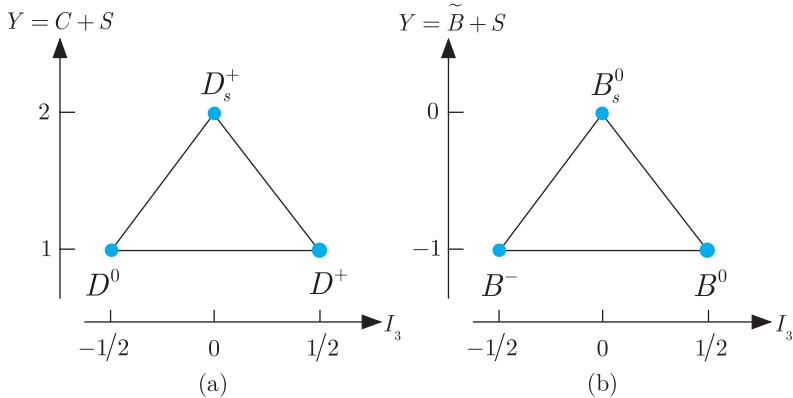
### 6.3 The $L = 0$ heavy quark states

Although in the above discussion we have only considered states containing the ‘light quarks’  $u$ ,  $d$  and  $s$ , it is straightforward to extend these ideas to include the ‘heavy quarks’  $b$  and  $c$ .<sup>15</sup>

We start with the mesons. According to the simple quark model, the lightest states of any  $q\bar{q}$  pair are predicted to have  $L = 0$ , with  $J^P = 0^-, 1^-$ , irrespective of the quark flavours. In Table 6.10 we list the predicted states for all combinations of a  $c$  or  $b$  quark with an antiquark, together with their internal quantum numbers. The experimentally observed states, which are also listed, are completely compatible with these predictions. The  $c\bar{c}$  and  $b\bar{b}$  states have zero charm and beauty and will be discussed in some detail in Section 6.5. Apart from these, the  $J^P = 0^-$  states are the lightest mesons with their particular values of the quantum numbers  $Q$ ,  $S$ ,  $C$  and  $\tilde{B}$  and so cannot decay by strong or electromagnetic interactions. They each decay by the weak interaction to a very large number of possible final states, with lifetimes of order  $10^{-12}$  s. The  $J^P = 1^-$  states decay to the  $0^-$  states together with a photon or a pion.

<sup>14</sup>We will return to this point in Section 8.5.

<sup>15</sup>The third heavy quark  $t$  decays too rapidly to form hadron bound states, as noted in Section 3.1.



**Figure 6.3** Weight diagrams for the  $L = 0, J^P = 0^-$  triplets of mesons with  
 (a)  $C = 1, \tilde{B} = 0$  and  
 (b)  $C = 0, \tilde{B} = -1$ .

Because the  $b$  and  $c$  quarks are so heavy, the masses of the states in Table 6.10 vary greatly depending on which of these quarks are present. However, if we consider the  $0^-$  particles with  $C = 1$  and  $\tilde{B} = 0$ , we find a *triplet* of states

$$D^0(1865) = c\bar{u}, \quad D^+(1870) = c\bar{d}, \quad D_s^+(1968) = c\bar{s},$$

of relatively similar masses, consisting of an isodoublet ( $D^0, D^+$ ) and an isosinglet  $D_s^+$ . These states form a supermultiplet of states that differ by the interchange of light antiquarks  $\bar{u}, \bar{d}, \bar{s}$  and are conveniently displayed on the weight diagram of Figure 6.3(a). The weight diagram for the triplet of  $0^-$  states

$$B^-(5279) = b\bar{u}, \quad B^0(5280) = b\bar{d}, \quad B_s^0(5367) = b\bar{s},$$

with  $\tilde{B} = -1, C = 0$  is shown in Figure 6.3(b), and in each case there are similar weight diagrams for the corresponding  $1^-$  states. The remaining states in Table 6.10 do not contain any light quarks or antiquarks and can be regarded as *singlet* supermultiplets.

Turning to baryons, extensive data exist only for baryons containing a single heavy quark. We shall therefore concentrate primarily on the predictions for these, starting with charmed baryons.

As in Section 6.2, we shall assume that the lightest particles are  $3q$  states with zero angular momentum and hence positive parity, and that quarks of the same flavour must have parallel spins. Two cases need to be considered. For the combinations  $cuu, cdd, css$ , the light quark pairs have the same flavour and so must have spin 1. Adding the spin of the  $c$ -quark then gives states with  $J^P = 1/2^+$  and  $J^P = 3/2^+$  for all three combinations. On the other hand, for combinations  $cud, cus, cds$ , the light quark pair can have spin 0 or spin 1, leading to two  $1/2^+$  states and one  $3/2^+$  state in each case. On gathering together all the possibilities, we have nine states with  $J^P = 1/2^+$  and six with  $J^P = 3/2^+$ , all having  $C = 1$  and  $\tilde{B} = 0$ . These predictions, together with the experimentally observed particles, are listed in Tables 6.11 and 6.12, and the corresponding weight diagrams are shown in Figures 6.4(a) and (b). The measured

**Table 6.11** The states of the  $L = 0, J^P = 1/2^+$  nonet of baryons with  $C = 1, \tilde{B} = 0$ .

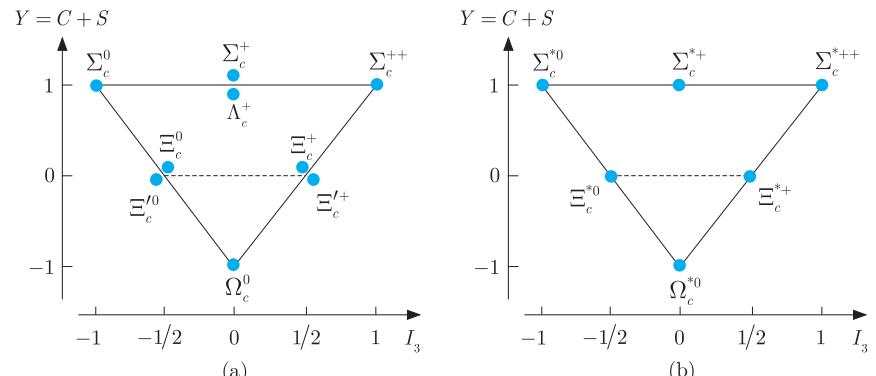
Quark composition	Observed state	$I_3$	$I$	$S$
$udc$	$\Lambda_c^+(2286)$	0	0	0
$uuc$	$\Sigma_c^{++}(2454)$	1	1	0
$udc$	$\Sigma_c^+(2453)$	0	1	0
$ddc$	$\Sigma_c^0(2454)$	-1	1	0
$usc$	$\Xi_c^+(2468)$	$1/2$	$1/2$	-1
$dsc$	$\Xi_c^0(2471)$	$1/2$	$-1/2$	-1
$usc$	$\Xi_c^{*+}(2577)$	$1/2$	$1/2$	-1
$dsc$	$\Xi_c^{*0}(2577)$	$1/2$	$-1/2$	-1
$ssc$	$\Omega_c^0(2695)$	0	0	-2

**Table 6.12** The states of the  $L = 0, J^P = 3/2^+$  sextet of baryons with  $C = 1, \tilde{B} = 0$ .

Quark composition	Observed state	$I_3$	$I$	$S$
$uuc$	$\Sigma_c^{*++}(2518)$	1	1	0
$udc$	$\Sigma_c^{*+}(2518)$	0	1	0
$ddc$	$\Sigma_c^0(2519)$	-1	1	0
$usc$	$\Xi_c^{*+}(2646)$	$1/2$	$1/2$	-1
$dsc$	$\Xi_c^{*0}(2646)$	$-1/2$	$1/2$	-1
$ssc$	$\Omega_c^{*0}(2695)$	0	0	-2

quantum numbers  $Q$ ,  $I$ ,  $I_3$  and  $S$  coincide precisely with those of the 15 states predicted. However, the spins and parities of most of the observed states have not yet been measured. The exceptions are the  $\Lambda_c(2286)$  and the isotriplet  $\Sigma_c(2454)$ , which have  $J^P = 1/2^+$  as predicted.

The experimental observation of other baryons with non-zero charm and/or bottom is considerably less complete. The quark model predictions for baryons with  $C = 0$  and  $\tilde{B} = -1$  are simply obtained by replacing the  $c$ -quarks by  $b$ -quarks in the states with  $C = 1$  and



**Figure 6.4** Weight diagrams for (a) the  $1/2^+$  nonet and (b) the  $3/2^+$  sextet of baryons with  $C = 1, \tilde{B} = 0$ .

$\tilde{B} = 0$  considered above. Hence 15 states are predicted, but only 9 have been observed to date, and none of their spins or parities has been measured. The predictions for baryons containing two or more heavy quarks are left as an exercise for the reader but none have yet been detected. The spectroscopy of hadrons with non-zero charm and bottom is still very much a developing area of experimental research.

Finally, in this and the previous section, we have grouped the hadrons into supermultiplets of particles with the same spin-parity and roughly similar masses, whose members differ by the interchange of light quarks of different flavours. These supermultiplets arise because the forces between different quarks are independent of their flavour and the mass splittings between the three light quarks are relatively small on the scale of hadron masses. These ideas are sometimes generalised to form extended supermultiplets that include states in which the light quarks are replaced by  $b$  or  $c$  quarks. However, although the forces are still flavour-independent, the symmetry between these different states is very badly broken by the large mass differences between the light and heavy quarks and we will not consider such extended supermultiplets further.

## 6.4 Colour

The quark theory of hadrons is conspicuously successful, but appears to contradict the Pauli principle. This states that the wavefunction of any state must be antisymmetric under the interchange of any two identical spin-1/2 fermions, and is one of the fundamental assumptions of quantum mechanics. Despite this, in discussing the lightest baryons we assumed that their combined space and spin wavefunctions were symmetric under the interchange of any two quarks of the same flavour contained within them. This assumption is crucial to the success of the model, and, as remarked earlier, the assumption of antisymmetry leads to a spectrum that is incompatible with experiment (cf. Problem 6.6). The same remark applies to the excited states of baryons, which are also well accounted for by the quark model provided the combined space and spin wavefunctions are assumed to be symmetric.

The apparent contradiction between the quark model and the Pauli principle was resolved in 1964 by Greenberg, who argued that in addition to space and spin degrees of freedom, quarks must possess another attribute, which he called *colour*. The total wavefunction is then written as the product of a spatial part  $\psi(\mathbf{r})$ , a spin part  $\chi$  and a colour wavefunction  $\chi_C$ , that is

$$\Psi = \psi(\mathbf{r})\chi\chi_C, \quad (6.40)$$

where the colour wavefunction will be described shortly. The Pauli principle is now interpreted as applying to the total wavefunction,

including the colour part  $\chi_C$ . The combined space and spin wavefunctions can then be symmetric under the interchange of quarks of the same flavour (to accord with experiment) provided the colour wavefunction is antisymmetric. Associated with the colour wavefunction are several conserved quantum numbers, called *colour charges*, which play a similar role in strong interactions to that played by the electric charge in electromagnetic interactions. This will be discussed in the next chapter. Here we will introduce the colour charges and wavefunctions, and investigate the hypothesis of *colour confinement*. This is the hypothesis that hadrons can only exist in states, called *colour singlets*, which have zero values for all colour charges, while quarks, which have non-zero colour charges, can only exist confined within them. As we shall see shortly, this explains why hadrons have integer electric charges, while fractionally charged combinations like  $qq$  and  $qq\bar{q}$  are forbidden, in accordance with experimental observation.

#### 6.4.1 Colour charges and confinement

The basic assumption of the colour theory is that any quark  $q = u, d, s, \dots$  can exist in three different colour states  $\chi_C = r, g, b$ , standing for ‘red’, ‘green’ and ‘blue’, respectively.<sup>16</sup> Just as the spin states  $\chi = \alpha, \beta$  correspond to different values of the spin component  $S_3$ , the colour states  $\chi_C$  correspond to different values of two of the colour charges called the *colour hypercharge* and the *colour isospin charge*.<sup>17</sup> They are denoted  $Y^C$  and  $I_3^C$ , and their values for the single quark states  $\chi_C = r, g, b$  are listed in Table 6.13(a). Their values for other states composed of quarks and antiquarks then follow by using the

**Table 6.13** Values of the colour charges  $I_3^C$  and  $Y^C$  for the colour states of quarks and antiquarks.

(a) Quarks		(b) Antiquarks			
	$I_3^C$	$Y^C$			
$r$	1/2	1/3	$\bar{r}$	-1/2	-1/3
$g$	1/2	1/3	$\bar{g}$	1/2	-1/3
$b$	0	-2/3	$\bar{b}$	0	2/3

<sup>16</sup> Explicit forms for these colour wavefunctions, analogous to the Pauli representation (5.22) for the spin wavefunctions  $\chi = \alpha, \beta$ , will be given in Section \*6.4.2.

<sup>17</sup> The origin of the colour charges is discussed in more detail in Section \*6.4.2, but is not required here, and the values of Table 6.13 may be regarded as definitions. The names colour hypercharge and colour isospin arise because of formal similarities with the ‘ordinary’ hypercharge and isospin discussed previously. There is, however, no physical connection between the latter quantum numbers and the two colour charges introduced here.

fact that they are additive quantum numbers, like electric charge, whose values for particles and antiparticles are equal in magnitude but opposite in sign. Thus under charge conjugation a quark  $q$  in the  $r$ -state is transformed into a quark  $\bar{q}$  in the colour state  $\chi_C = \bar{r}$  with reversed values of  $Y^C$  and  $I_3^C$ , as listed in Table 6.13(b).

Direct evidence for the three colour states of quarks will be presented in Chapter 7, where we will also discuss the physical interpretation of the colour charges as sources of the strong interaction. Here we will discuss the consequences of assuming that any hadron that can be observed as an isolated particle in free space must have colour charges

$$I_3^C = Y^C = 0, \quad (6.41)$$

in accordance with the hypothesis of colour confinement.

We first consider baryons, assumed to be composed of three quarks. From Table 6.13(a) it is clear that the colour charges of the three quarks can only add up to give  $I_3^C = Y^C = 0$  if baryons are composed of one  $r$ , one  $g$  and one  $b$  quark. It is this observation that gave rise to the term ‘colour’, since in the theory of colour vision white (i.e. colourless) light is formed by combining three primary colours, which can be chosen to be red, green and blue.<sup>18</sup> The most general colour wavefunction for a baryon is then a linear superposition of six possible combinations, that is

$$\begin{aligned} \chi_B^C = & \alpha_1 r_1 g_2 b_3 + \alpha_2 g_1 r_2 b_3 + \alpha_3 b_1 r_2 g_3 \\ & + \alpha_4 b_1 g_2 r_3 + \alpha_5 g_1 b_2 r_3 + \alpha_6 r_1 b_2 g_3, \end{aligned} \quad (6.42)$$

where, for example,  $r_3$  means that the third quark is in an  $r$ -state and the  $\alpha_i (i = 1, 2, \dots, 6)$  are constants. In particular, if we choose the totally antisymmetric combination

$$\chi_B^C = \frac{1}{\sqrt{6}}(r_1 g_2 b_3 - g_1 r_2 b_3 + b_1 r_2 g_3 - b_1 g_2 r_3 + g_1 b_2 r_3 - r_1 b_2 g_3), \quad (6.43)$$

then the Pauli principle will indeed require the combined space and spin wavefunction of the baryon to be symmetric under the interchange of quarks of the same flavour, as assumed in our earlier discussion of baryons. In fact it can be shown that the choice (6.43) is not only allowed by colour confinement but is required by it. However, this needs a fuller treatment than the one given here.<sup>19</sup>

We next consider the combinations  $q^m \bar{q}^n$  of  $m$  quarks and  $n$  anti-quarks that are allowed by the confinement condition (6.41). In doing this, it will be sufficient to consider only combinations with baryon

<sup>18</sup>This is of course only an analogy, and the quark colours have nothing whatsoever to do with ‘ordinary’ colours!

<sup>19</sup>It is given, for the interested reader, in Section \*6.4.2, but is not required elsewhere in the book.

number  $B \geq 0$ , and hence  $m \geq n$ , since the particles with  $B < 0$  are just antiparticles of those with  $B > 0$ . The corresponding colour wavefunctions can be written

$$r^\alpha g^\beta b^\gamma \bar{r}^{\bar{\alpha}} \bar{g}^{\bar{\beta}} \bar{b}^{\bar{\gamma}}, \quad (6.44)$$

where  $r^\alpha$  means that there are  $\alpha$  quarks in the  $r$ -state, etc.,<sup>20</sup> and

$$m = \alpha + \beta + \gamma \geq n = \bar{\alpha} + \bar{\beta} + \bar{\gamma}.$$

We obtain the colour hypercharge  $Y^C$  and colour isospin  $I_3^C$  for the colour wavefunction (6.44) by adding the colour charges listed in Table 6.13. Imposing the confinement condition (6.41) then gives

$$I_3^C = (\alpha - \bar{\alpha})/2 - (\beta - \bar{\beta})/2 = 0$$

and

$$Y^C = (\alpha - \bar{\alpha})/3 + (\beta - \bar{\beta})/3 - 2(\gamma - \bar{\gamma})/3 = 0,$$

implying

$$\alpha - \bar{\alpha} = \beta - \bar{\beta} = \gamma - \bar{\gamma} \equiv p$$

and hence

$$m - n = 3p,$$

where  $p$  is a non-negative integer. Thus the only combinations  $q^m \bar{q}^n$  allowed by colour confinement are of the form

$$(3q)^p (q\bar{q})^n \quad (p, n \geq 0), \quad (6.45)$$

where we recall that we have restricted ourselves to states with baryon number  $B \geq 0$ . It follows from this that hadrons with fractional electric charges such as

$$qq, \quad qq\bar{q}, \quad qqqq, \dots$$

are forbidden by colour confinement, while the combinations  $q\bar{q}$  and  $3q$  used in the simple quark model are allowed.

Nuclei composed of  $A$  nucleons and hence  $3A$  quarks are also examples of (6.45), as are *hypernuclei*, in which a hyperon, for example  $\Lambda(1115)$ , is bound to a nucleus to produce unstable states that decay by weak interactions with lifetimes of order  $10^{-10}$  s. These states can, to a very good approximation, be regarded as bound states of nucleons, or in the case of hypernuclei, nucleons and a hyperon. In 1964, Dyson and Xuong suggested that in addition to the deuteron, which was the only known nucleon–nucleon bound state,  $N\Delta$  and  $\Delta\Delta$  bound states might exist, where  $N$  is a nucleon and  $\Delta$  is the  $\Delta(1232)$  resonance. This led to many investigations of nucleon–nucleon scattering, in the hope, among other things, of discovering these or other short-lived *dibaryon* states. In general, nucleon–nucleon scattering is dominated by non-resonant scattering and, for

---

<sup>20</sup>There is no need to consider the order of the factors here, since, for example,  $rrb$  has the same values of  $I_3^C$  and  $Y^C$  as  $rbr$  or  $brr$ .

many years, signs of dibaryons were conspicuous by their absence. However, in 2011, in an experiment at the COSY proton synchrotron in Jülich in Germany, a resonance denoted  $d^*(2380)$  was observed in the reaction

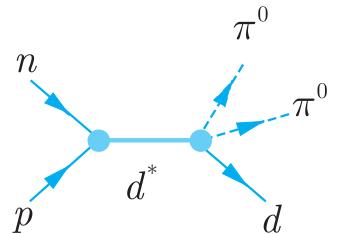
$$n + p \rightarrow d^* \rightarrow d + \pi^0 + \pi^0, \quad (6.46)$$

as shown in Figure 6.5, where  $d$  is a deuteron, and subsequently confirmed in 2014 by a detailed analysis of polarised neutron–proton scattering. The  $d^*(2380)$  is an isospin zero state (cf. Problem 6.12) with spin-parity  $J^P = 3^+$ , whose properties are consistent with those expected of a  $\Delta\Delta$  state, although other configurations cannot be ruled out.

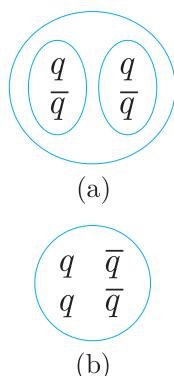
We next turn to combinations such as

$$qq\bar{q}\bar{q} \quad \text{and} \quad qqqq\bar{q}, \quad (6.47)$$

called *tetraquarks* and *pentaquarks*, respectively, which could give rise to mesons and baryons that are not included in the simple quark model. One possibility is that, like the dibaryons discussed above, these could be loosely bound states of a pair of ‘normal’ hadrons that retain, to a significant degree, their individual identities. Such states, whose masses are approximately equal to the sum of the masses of their constituents, are referred to as *molecular states* and are illustrated schematically, for the case of a di-meson, in Figure 6.6(a). Alternatively, they could be more tightly bound states, which are not well approximated as bound states of two hadrons, as illustrated for the tetraquark case in Figure 6.6(b). Such states would represent a more radical departure from the simple quark model than molecular states, but there is no obvious theoretical reason why they should not exist. Assuming that they do, one would expect them to be heavier than conventional  $q\bar{q}$  and  $qqq$  states with the same internal quantum numbers, since they contain more particles and more degrees of freedom. In addition, like any other hadrons, one would expect them to occur in supermultiplets of particles with similar masses, which differ by the interchange of light quarks  $u, d, s$ ; these supermultiplets will usually include particles with exotic quantum numbers. For example, if a pentaquark  $uuuu\bar{u}$  existed with allowed quantum numbers  $Q = 2, S = 0$ , then one would also expect to observe, among other particles, an isospin partner  $uuuud$  with approximately the same mass and exotic quantum numbers  $Q = 3, S = 0$ , and a somewhat heavier state  $uuuu\bar{s}$  with exotic quantum numbers  $Q = 2, S = -1$ . Similarly, one would expect a tetraquark  $u\bar{u}u\bar{u}$  with allowed quantum numbers  $Q = 0, S = 0$  to belong to a supermultiplet that included, among others, states  $u\bar{d}d\bar{d}$  and  $u\bar{d}u\bar{s}$  with exotic quantum numbers  $Q = 2, S = 0$  and  $Q = 2, S = 1$ , respectively. Hence the absence of well-established states with exotic quantum numbers, despite extensive attempts to detect them, implies at best a limited role for such



**Figure 6.5** Production of a  $d^*$  resonance via reaction (6.46).



**Figure 6.6** Schematic diagram of possible tetraquark structures: (a) a molecular quark state; (b) a more tightly bound four-quark state.

states (6.47) composed of light quarks. On the other hand, supermultiplets involving heavy quark–antiquark pairs do not necessarily contain particles with exotic quantum numbers. For example, the members of a *tetraquark–charmonium* supermultiplet with the composition  $c\bar{c}q\bar{q}$ , where  $q$  again indicates a light quark  $u$ ,  $d$  or  $s$ , would have exactly the same internal quantum numbers as the members of a normal  $q\bar{q}$  supermultiplet, and there is evidence that at least some such states exist, as we shall see in Chapter 7.

### \*6.4.2 Colour wavefunctions and the Pauli principle

In this section we give explicit expressions for the colour wavefunctions and operators, and show that for baryons the totally antisymmetric colour wavefunction (6.43) is not only allowed by colour confinement but is required by it.

The three independent colour wavefunctions  $\chi^C = r, g, b$  of a quark are conveniently represented by the ‘colour spinors’

$$r = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad g = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad (6.48)$$

in the same way that the two spin wavefunctions  $\chi = \alpha, \beta$  were represented by the spinors

$$\alpha = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

in Section 5.2.1. Just as the spin wavefunctions are acted on by spin operators, the colour wavefunctions are acted on by ‘colour operators’. The latter are represented by three-dimensional matrices in the same way that the spin operators are represented by the two-dimensional matrices (5.21). There are eight such independent colour operators

$$\hat{F}_i = \frac{1}{2}\lambda_i \quad (i = 1, 2, \dots, 8), \quad (6.49a)$$

where the matrices

$$\begin{aligned} \lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda_2 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ \lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, & \lambda_5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, & \lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \\ \lambda_7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, & \lambda_8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \end{aligned} \quad (6.49b)$$

can be regarded as the three-dimensional analogues of the two-dimensional Pauli matrices (5.21).<sup>21</sup> The colour wavefunctions (6.48) are eigenfunctions of the operators  $\hat{F}_3$  and  $\hat{F}_8$  with, for example,

$$\hat{F}_3 r = \frac{1}{2}r, \quad \hat{F}_8 r = \frac{1}{2\sqrt{3}}r \quad (6.50)$$

for the  $r$ -state. The colour charges  $I_3^C$  and  $Y^C$  discussed in the previous section are the eigenvalues of the operators

$$\hat{I}_3^C \equiv \hat{F}_3 \quad \text{and} \quad \hat{Y}^C \equiv \frac{2}{\sqrt{3}}\hat{F}_8, \quad (6.51)$$

where the factor  $2/\sqrt{3}$  was introduced for historical reasons. From (6.48) and (6.49) it follows that  $I_3^C = \frac{1}{2}$  and  $Y^C = \frac{1}{3}$  for the  $r$ -state, and the corresponding values for the  $g$  and  $b$  states listed in Table 6.13(a) are obtained in a similar fashion. The remaining operators  $\hat{F}_1, \hat{F}_4, \hat{F}_5, \hat{F}_6$  and  $\hat{F}_7$  mix the colour states (6.41), and it is easily shown using (6.49) that, for example,

$$\hat{F}_1 r = \frac{1}{2}g, \quad \hat{F}_1 g = \frac{1}{2}r, \quad \hat{F}_1 b = 0. \quad (6.52)$$

The observables associated with the operators  $\hat{F}_i$  are all believed to be exactly conserved in nature, that is

$$[\hat{F}_i, H] = 0 \quad (i = 1, 2, \dots, 8),$$

and for this reason they are all called colour charges. Colour confinement is the requirement that all eight colour charges vanish for any observed hadron  $h$ , implying

$$\hat{F}_i \chi_h^C = 0 \quad (i = 1, 2, \dots, 8) \quad (6.53)$$

for the corresponding colour wavefunction  $\chi_h^C$ . This is reminiscent of a spin singlet state, in which all three spin components vanish, so that the spin wavefunctions satisfy

$$\hat{S}_x \chi = 0, \quad \hat{S}_y \chi = 0, \quad \hat{S}_z \chi = 0. \quad (6.54)$$

It is for this reason that colour states that satisfy (6.53) are called colour singlets. Equation (6.53) implies, but is more restrictive than, the conditions (6.41) that were exploited extensively in the previous section.

Finally, we shall use colour confinement to fix completely the colour wavefunction for any baryon. We have already seen that the most general colour wavefunction compatible with (6.41) is

$$\begin{aligned} \chi_B^C &= \alpha_1 r_1 g_2 b_3 + \alpha_2 g_1 r_2 b_3 + \alpha_3 b_1 r_2 g_3 \\ &\quad + \alpha_4 b_1 g_2 r_3 + \alpha_5 g_1 b_2 r_3 + \alpha_6 r_1 b_2 g_3, \end{aligned} \quad (6.55)$$

---

<sup>21</sup>In both cases, they are traceless Hermitian matrices, and the most general Hermitian matrix in two or three dimensions can be written as a linear combination of the unit matrix and the  $\sigma_i$  or  $\lambda_i$ , respectively.

where the  $\alpha_i (i = 1, 2, \dots, 6)$  are constants. The values of the latter follow from applying (6.50), which we will illustrate for the case of  $\hat{F}_1$ . It follows from (6.52) that acting with  $\hat{F}_1$  on the first term of (6.55) gives

$$\begin{aligned}\alpha_1 \hat{F}_1(r_1 g_2 b_3) &= \alpha_1 (\hat{F}_1 r_1) g_2 b_3 + \alpha_1 r_1 (\hat{F}_1 g_2) b_3 + \alpha_1 r_1 g_2 (\hat{F}_1 b_3) \\ &= \frac{\alpha_1}{2} (g_1 g_2 b_3 + r_1 r_2 b_3).\end{aligned}$$

Similar contributions are obtained by acting with  $\hat{F}_1$  on the other terms in (6.48), and collecting these together we obtain

$$\begin{aligned}\hat{F}_1 \chi_B^C &= \frac{(\alpha_1 + \alpha_2)}{2} (g_1 g_2 b_3 + r_1 r_2 b_3) + \frac{(\alpha_3 + \alpha_4)}{2} (b_1 g_2 g_3 + b_1 r_2 r_3) \\ &\quad + \frac{(\alpha_5 + \alpha_6)}{2} (g_1 b_2 g_3 + r_1 b_2 r_3).\end{aligned}$$

This is only compatible with (6.53) for  $i = 1$  if the conditions

$$\alpha_1 = -\alpha_2, \quad \alpha_3 = -\alpha_4, \quad \alpha_5 = -\alpha_6 \quad (6.56)$$

are satisfied. Analogous relations follow from considering the other operators  $\hat{F}_2, \hat{F}_4$ , etc., in (6.53), and together they uniquely determine the wavefunction (see Problem 6.11) to be the totally antisymmetric combination (6.43)

$$\chi_B^C = \frac{1}{\sqrt{6}} (r_1 g_2 b_3 - g_1 r_2 b_3 + b_1 r_2 g_3 - b_1 g_2 r_3 + g_1 b_2 r_3 - r_1 b_2 g_3),$$

up to an arbitrary normalisation constant. Since the Pauli principle requires the total wavefunction to be antisymmetric under the interchange of identical quarks, the combined space and spin part must therefore be symmetric, as assumed in our discussion of baryons in Section 6.2.2.

## 6.5 Charmonium and bottomonium

We next consider the spectra of charmonium ( $c\bar{c}$ ) and bottomonium ( $b\bar{b}$ ). The  $c$  and  $b$  quarks are much heavier than the  $u$ ,  $d$  and  $s$  quarks (see Table 3.1) and the lower-lying  $c\bar{c}$  and  $b\bar{b}$  states are approximately non-relativistic systems, in which the quark kinetic energies are much smaller than their rest masses. There is thus a strong analogy with the  $e^+e^-$  bound system positronium, except that the potential energy is different. As in positronium, states can be labelled by their angular momentum quantum numbers,  $J$ ,  $L$  and  $S$ , where the spin  $S$  can be 0 or 1, and by their principal quantum number  $n = \nu + L + 1$ , where  $\nu$  is the number of radial nodes, not counting nodes at the origin. Alternatively, instead of  $n$ , the radial quantum number  $n_r = \nu + 1 = n - L = 1, 2, 3, \dots$  is sometimes used to label

**Table 6.14** Predicted  $c\bar{c}$  and  $b\bar{b}$  states with principal quantum numbers  $n = 1$  and 2, and radial quantum number  $n_r = n - L$ , compared with experimentally observed states. Masses are given in  $\text{MeV}/c^2$ .

$2S+1 L_J$	$n$	$n_r$	$J^{PC}$	$c\bar{c}$ state	$b\bar{b}$ state
$^1 S_0$	1	1	$0^{-+}$	$\eta_c(2984)$	$\eta_b(9398)^{(*)}$
$^3 S_1$	1	1	$1^{--}$	$J/\psi(3097)$	$\Upsilon(9460)$
$^3 P_0$	2	1	$0^{++}$	$\chi_{c0}(3415)$	$\chi_{b0}(9859)$
$^3 P_1$	2	1	$1^{++}$	$\chi_{c1}(3511)$	$\chi_{b1}(9893)$
$^3 P_2$	2	1	$2^{++}$	$\chi_{c2}(3556)$	$\chi_{b2}(9912)$
$^1 P_1$	2	1	$1^{+-}$	$h_c(3525)$	$h_b(9899)$
$^1 S_0$	2	2	$0^{-+}$	$\eta_c(3639)$	$\eta_b(9999)$
$^3 S_1$	2	2	$1^{--}$	$\psi(3686)$	$\Upsilon(10023)$

\*State is not well established and its quantum number assignments are unknown.

the states. The resulting states for positronium with principal quantum number  $n = 1$  or  $n = 2$  have been listed in Table 5.1, together with the values of their parity

$$P = P_{e^+} P_{e^-} (-1)^L = (-1)^{L+1} \quad (6.57)$$

and  $C$  parity

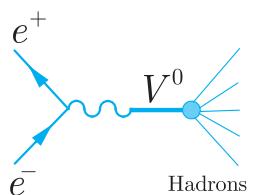
$$C = (-1)^{L+S}, \quad (6.58)$$

which follow from (5.35) and (5.53) respectively. The same results apply to charmonium and bottomonium, and are listed in Table 6.14, together with the observed mesons that correspond to them. As we shall see, the measured energies of these states can be used to partially determine the potential energies in the two systems, leading to the important conclusion that the forces acting in the  $c\bar{c}$  and  $b\bar{b}$  systems are the same within errors.

### 6.5.1 Charmonium

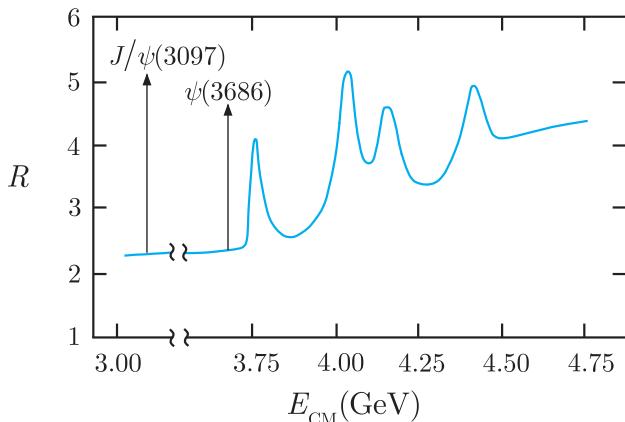
The charmonium system is most conveniently studied by first forming those states  $V^0 = J/\psi(3097), \psi(3686), \dots$  that have the same  $J^{PC}$  values  $1^{--}$  as the photon (cf. Table 6.14) and then observing their radiative decays. The basic formation process is shown in Figure 6.7: the electron–positron pair annihilates to a virtual photon, which converts to the  $V^0$ , which then decays, mainly into hadrons. This is only possible when the total centre-of-mass energy is equal to the  $V^0$  mass, leading to a peak in the total cross-section for electron–positron annihilation into hadrons. This cross-section is usually plotted in terms of the ratio

$$R \equiv \frac{\sigma(e^+ e^- \rightarrow \text{hadrons})}{\sigma(e^+ e^- \rightarrow \mu^+ \mu^-)}, \quad (6.59a)$$



**Figure 6.7** Mechanism for the formation of mesons  $V^0$  with quantum numbers  $J^{PC} = 1^{--}$  in electron–positron annihilation, and their subsequent decay to hadrons.

**Figure 6.8** Measured cross-section ratio (6.59a) in the vicinity of the threshold for producing pairs of charmed particles. The extremely narrow peaks associated with the  $J/\psi(3097)$  and  $\psi(3686)$  are indicated by arrows because they extend far above the scale shown. In the partially suppressed region between these peaks, the ratio is essentially constant. (Data from Brandelik *et al.* 1978 and Schindler *et al.* 1980.)

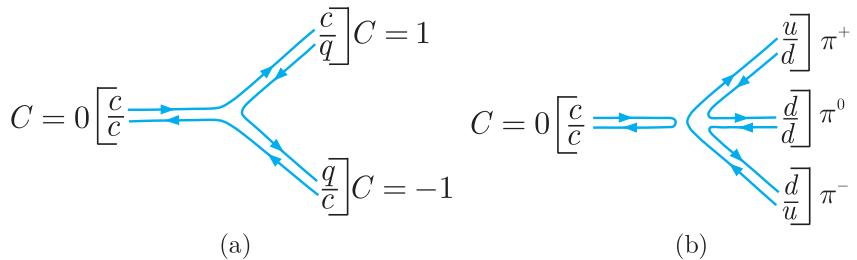


where the measured cross-section for muon pair production,<sup>22</sup>

$$\sigma(e^+e^- \rightarrow \mu^+\mu^-) = \frac{4\pi\alpha^2}{3E_{CM}^2}, \quad (6.59b)$$

and is a smoothly varying function of the centre-of-mass energy  $E_{CM}$ , as shown in Figure 6.8.

As can be seen in this figure, there are two sharp peaks in the ratio, corresponding to the masses of the  $J/\psi(3097)$  and  $\psi(3686)$ , followed by several relatively low, broad peaks at energies above 3730 MeV. This latter energy is called the *charm threshold*, and is twice the mass of the lightest charmed particles, which are the  $D$  mesons (3.18). Charmonium states that lie above this threshold are resonances that decay readily into pairs of charmed particles by the mechanism of Figure 6.9(a), involving the creation of a light quark–antiquark pair, and the broad peaks in Figure 6.8 are evidence for the existence of several such states. In contrast, charmonium states that lie below this threshold are all very narrow and cannot decay by the mechanism of Figure 6.9(a) because of energy conservation; instead they decay much more slowly to non-charmed hadrons by mechanisms like that of Figure 6.9(b), involving the annihilation of a charmed quark–antiquark pair.



**Figure 6.9** Quark diagrams for (a) the decay of a charmonium state to a pair of charmed mesons and (b) an example of a decay to non-charmed mesons.

<sup>22</sup>This cross-section can be calculated theoretically from the mechanism of Figure 1.3 (see, for example, pp. 135–139 of Mandl and Shaw, 2010). However, it is sufficient for our purposes to take this result from experiment.

In general, in strong interactions, processes involving the creation or annihilation of  $c\bar{c}$  and  $b\bar{b}$  pairs are always very heavily suppressed compared to processes involving the creation or annihilation of light quark–antiquark pairs ( $u\bar{u}, d\bar{d}, s\bar{s}$ ). This characteristic, which does not apply to weak or electromagnetic interactions, is called the OZI rule, after Okubo, Zweig and Iizuka.<sup>23</sup> In this context it explains why the states below the charm threshold are much narrower than those above it. Consequently, as narrow states are much easier to identify experimentally than broad ones, the spectrum is much better known below the charmed threshold than above it. We will therefore confine ourselves to this part of the spectrum in the rest of this section.

The  $J/\psi(3097)$  and  $\psi(3686)$  are the only well-established charmonium states listed in Table 6.14 that can be produced by the mechanism of Figure 6.7, because they are the only ones with the same quantum numbers  $J^{PC} = 1^{--}$  as the photon. Some of the other states shown in Table 6.14, however, can be found by studying the radiative decays of the  $J/\psi(3097)$  and  $\psi(3686)$ . In particular, the  $\psi(3686)$  decays mostly into hadrons, but about 25% of the time it decays by the radiative transitions

$$\psi(3686) \rightarrow \chi_{ci} + \gamma \quad (i = 0, 1, 2), \quad (6.60)$$

where  $\chi_{c0}(3415)$ ,  $\chi_{c1}(3511)$  and  $\chi_{c2}(3556)$  are three new charmonium states, called chi-states. These particles themselves decay, mostly to give hadrons in the final state, and an analysis of these decays shows that they have  $J^{PC}$  values  $0^{++}$ ,  $1^{++}$  and  $2^{++}$ , respectively. We therefore identify them with the  $^3P_0$ ,  $^3P_1$  and  $^3P_2$  states of charmonium, which have these predicted  $J^{PC}$  values, as can be seen from Table 6.14. In addition, the  $n = 1$ ,  $^1S_0$  ground state  $\eta_c(2980)$  has been observed in the radiative transitions

$$\psi(3686) \rightarrow \eta_c(2984) + \gamma \quad (6.61a)$$

and

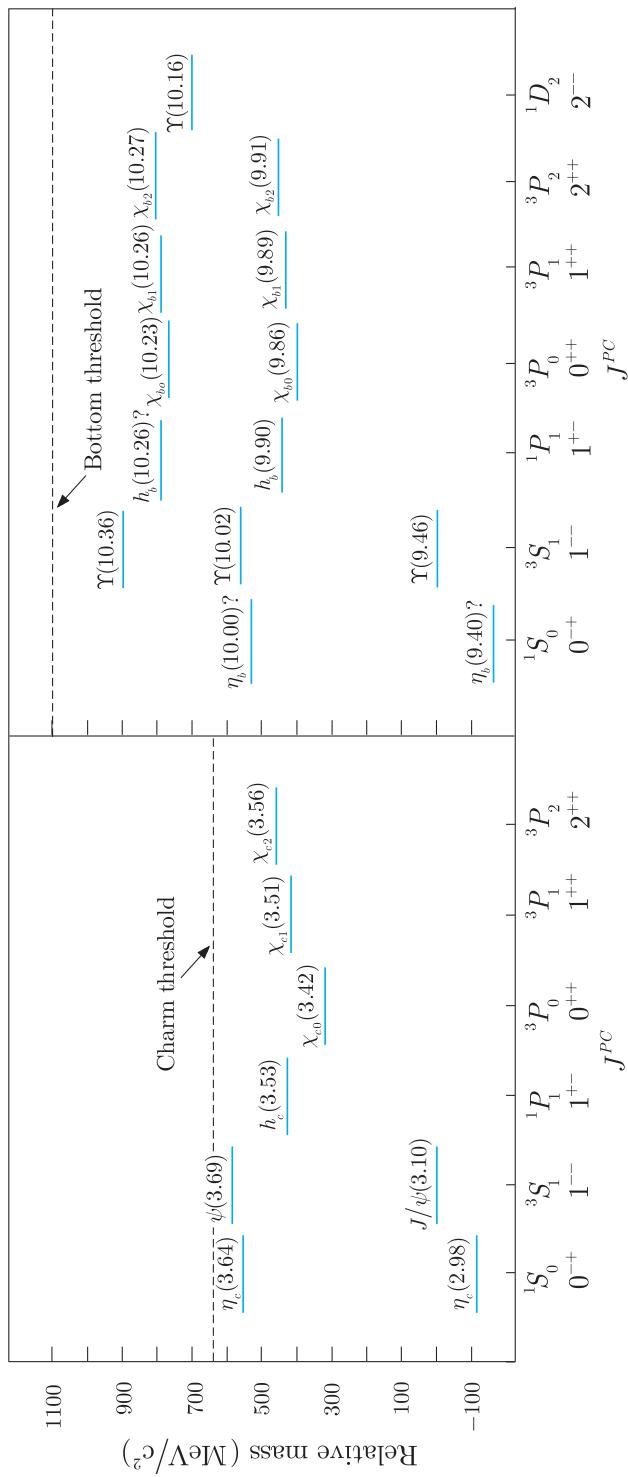
$$J/\psi(3097) \rightarrow \eta_c(2984) + \gamma, \quad (6.61b)$$

despite the small branching ratios of about  $3 \times 10^{-3}$  and  $10^{-2}$ , respectively. The remaining charmonium states in Table 6.14 with  $n \leq 2$  are the  $^1S_0$  state  $\eta_c(3639)$  and the  $^1P_1$  state  $h_c(3526)$ . The former has been seen in  $\gamma\gamma$  interactions and the latter in  $p\bar{p}$  annihilations.

To summarise, all eight predicted states of charmonium with principal quantum number  $n = 1$  or  $n = 2$  have been identified and are the only states observed below the charm threshold. The resulting experimental spectrum is shown in Figure 6.10. Despite the difference in energy scales, it bears a striking resemblance to the spectrum of

---

<sup>23</sup>The dynamical origin of the OZI rule will be discussed in Section 7.1.1.



**Figure 6.10** The states of charmonium ( $c\bar{c}$ ) and bottomonium ( $b\bar{b}$ ) observed below the charm and bottom thresholds respectively, together with their quark model assignments. The masses are plotted relative to that of the  ${}^3S_1$  ground state, and the absolute mass values in  $\text{GeV}/c^2$  are shown in brackets after the particle symbols.

positronium shown in Figure 5.2, except that the  $2S$  levels are somewhat higher in energy than the  $2P$  levels, instead of being degenerate with them.

### 6.5.2 Bottomonium

The bottomonium spectrum is observed in much the same way as the charmonium spectrum, and we confine ourselves here to stating the results, which are summarised in Figure 6.10. In this case there are three bands of narrow states, corresponding to principal quantum number  $n = 1, 2$  and  $3$ , lying below the bottom threshold of  $10.56 \text{ GeV}/c^2$ . This is just twice the mass of the  $B$  mesons (3.19), which are the lightest particles with non-zero bottom, and bottomonium states above this threshold are broad resonances that decay to pairs of particles with  $\tilde{B} = 0$  by a mechanism analogous to that of Figure 6.9(a). As can be seen in Figure 6.10, the ordering of the levels and the magnitudes of their splittings are similar in charmonium and bottomonium, indicating an underlying similarity between the forces acting in the two systems.

### \*6.5.3 The quark–antiquark potential

The experimental spectra of charmonion and bottomonium shown in Figure 6.10 bear a striking resemblance to the spectrum of positronium shown in Figure 5.2, and like it can be understood in terms of a simple non-relativistic treatment. In the centre-of-mass frame of the  $q\bar{q} = c\bar{c}$  or  $b\bar{b}$  pair, the Schrödinger equation is

$$-\frac{1}{2\mu} \nabla^2 \psi(\mathbf{r}) + V(r) \psi(\mathbf{r}) = E \psi(\mathbf{r}), \quad (6.62)$$

where  $r = |\mathbf{r}|$  is the distance between the quarks,  $\mu = m_q/2$  is their reduced mass and we have neglected spin-dependent effects. In this approximation, the energies  $E$  and the particle masses

$$M(q\bar{q}) = 2m_q + E \quad (6.63)$$

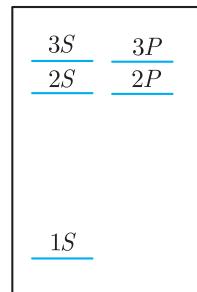
depend only on the principal quantum number  $n$  and the orbital angular momentum quantum number  $L$ , for a given potential  $V(r)$ . In the particular case of a Coulomb-like potential

$$V(r) \propto r^{-1}, \quad (6.64)$$

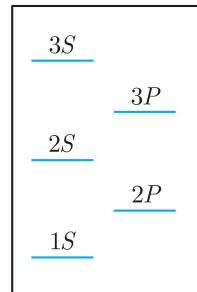
the energies depend on  $n$  only, but for other  $r$ -dependencies this degeneracy is broken.

This is illustrated in Figure 6.11, where the spectrum of  $S$ - and  $P$ -wave states resulting from a Coulomb potential (6.64) is compared with that resulting from a simple harmonic oscillator potential

$$V(r) \propto r^2. \quad (6.65)$$

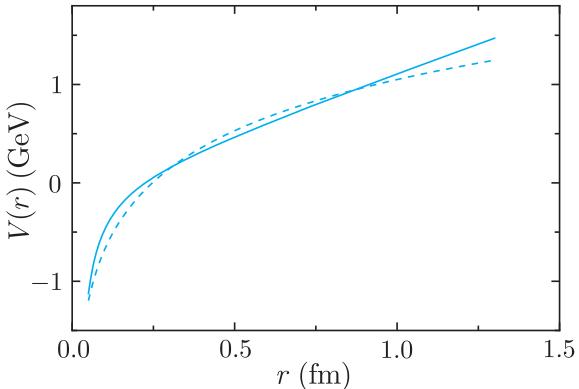


(a) Coulomb



(b) Oscillator

**Figure 6.11** Patterns of  $S$ - and  $P$ -wave energy levels arising from (a) Coulomb-like ( $r^{-1}$ ) and (b) simple harmonic oscillator ( $r^2$ ) potentials for principal quantum numbers  $n = 1, 2, 3$ . The strengths of the potentials are assumed to be such that the  $3s$  states have the same energy, measured relative to the  $1s$  state, in each case.



**Figure 6.12** Heavy quark-antiquark potentials obtained from fitting the energy levels of charmonium and bottomonium. The solid and dashed lines show the results obtained from the forms (6.66) and (6.67), respectively.

The heavy quarkonia spectra of Figure 6.10 are intermediate between these possibilities, and can be fitted using a potential of the form

$$V(r) = -\frac{a}{r} + b r, \quad (6.66a)$$

which behaves like a Coulomb potential at small  $r$  but rises linearly at large  $r$  values. The constants  $a$  and  $b$  are determined by solving the Schrödinger equation numerically and adjusting their values until good fits to the experimental spectra are obtained.<sup>24</sup> In this way one finds that the values

$$a = 0.30 \quad \text{and} \quad b = 0.23 \text{ GeV}^2 \quad (6.66b)$$

can be used to fit both the charmonium and bottomonium spectra, and the resulting potential is shown in Figure 6.12. The potential is not completely determined by this analysis, because of uncertainties in the quark masses and because equally satisfactory results can be obtained using other forms, for example

$$V(r) = a \ln(br), \quad (6.67a)$$

where

$$a = 0.75 \text{ GeV} \quad \text{and} \quad b = 0.80 \text{ GeV}. \quad (6.67b)$$

However, all the potentials that successfully explain the spectrum for given quark masses are in good agreement with each other in the range

$$0.2 \text{ fm} \lesssim r \lesssim 1.0 \text{ fm}, \quad (6.68)$$

as illustrated in Figure 6.12 for the cases (6.66) and (6.67). We therefore conclude that the potential is well-determined in the range (6.68) and that, within this range at least, the potentials for charmonium

<sup>24</sup>The agreement is not perfect, because relativistic effects are not completely negligible. These include a spin-dependent interaction to account for the fine structure seen in Figure 6.10. This interaction is similar to that used to describe the positronium fine structure (cf. Section 5.5.1).

$(c\bar{c})$  and bottomonium ( $b\bar{b}$ ) are, within small experimental uncertainties, the same, as expected from flavour independence.<sup>25</sup>

## Problems 6

- 6.1 Show that the two definitions (6.5) and (6.7) of the third component of isospin are equivalent.
- 6.2 The hadron  $\Sigma_c^+$  (2455) is observed to decay by  $\Sigma_c^+ \rightarrow \Lambda_c^+ + \pi^0$  with a rate typical of strong interactions, where  $\Lambda_c^+$  is the isosinglet hadron (3.21). Deduce the values of the quantum numbers (3.35) and hence the quark content of the  $\Sigma_c^+$ . Does it possess any isospin partners, and if so what is their quark content?
- 6.3 In the simple model of the  $\Sigma$  baryon mass splittings discussed in Section 6.1.4, the electromagnetic interaction energy between two quarks  $a$  and  $b$  was assumed to be of order  $\delta e_a e_b$ , where  $e_a$  and  $e_b$  are the quark charges in units of  $e$ . Deduce the value of  $\delta$  obtained from the observed  $\Sigma$  baryon masses in this approximation, and use it to make a rough estimate of the typical distance  $r$  between the quarks by assuming that the Coulomb interaction dominates.
- 6.4 Show that a meson that decays to  $\pi^+\pi^-$  pairs by the strong interaction must have  $C = P = (-1)^J$ , where  $J$  is the spin of the meson. The  $\rho^0(775)$  and  $f_2^0(1275)$  mesons decay by strong interactions to give  $\pi^+\pi^-$  pairs and have spin-1 and spin-2, respectively. Which of the decays  $\rho^0 \rightarrow \pi^0\gamma$  and  $f_2^0 \rightarrow \pi^0\gamma$  is forbidden in electromagnetic interactions? Which of the decays  $\rho^0 \rightarrow \pi^0\pi^0$  and  $f_2^0 \rightarrow \pi^0\pi^0$  is forbidden in any interaction?
- 6.5 The  $K^+$  and its excited states with masses below  $1.5 \text{ GeV}/c^2$  are shown in Figure 3.12. Identify these mesons with states  $^{2S+1}L_J$  of the appropriate quark–antiquark system, specifying the value of the principal quantum number  $n$  in each case.
- 6.6 In Section 6.2.2 we assumed that the combined space–spin wavefunction of a baryon was symmetric under the interchange of any pair of like quarks. What spectrum of low-lying baryon states composed of light quarks with zero orbital angular momenta  $L_{12} = L_3 = 0$  would be predicted if, instead, we had assumed that the wavefunctions were antisymmetric under the exchanges of like quarks?
- 6.7 In Section 6.2.4 we used (6.36) to derive the masses of the  $\Lambda$  and  $\Sigma^*$  baryons in Table 6.9. Derive the mass formulas for the remaining baryons listed in this table.

<sup>25</sup>These analyses also determine the mean square radii

$$r_0^2 \equiv \overline{r^2} = \int r^2 |\psi(r)|^2 d^3\mathbf{r} \quad (6.69)$$

of the various states, where the wavefunctions are determined by the solutions of the Schrödinger equation for successful potentials like (6.66) and (6.67). The resulting values of  $r_0$  for the various states shown in Figure 6.10 span just the range (6.68) in which the potential is well-determined.

- 6.8** The branching ratios for the decays of the  $D^{*0}(2007)$  (listed in Table 6.10) to  $D^0\pi^0$  and  $D^0\gamma$  are 0.62 and 0.38, respectively, while the branching ratios of the  $D_s^{*+}(2112)$  to  $D_s^+\pi^0$  and  $D_s^+\gamma$  are 0.06 and 0.93, respectively. Suggest a qualitative explanation for these results.
- 6.9** In the text we considered the simple quark model predictions for the baryons with angular momenta  $L_{12} = L_3 = 0$  containing 0 or 1 of the heavy quarks  $b,c$ . Extend this to baryons with (a)  $C = 2$ ,  $\tilde{B} = 0$  and (b)  $C = 1$ ,  $\tilde{B} = -1$ , and draw the weight diagram for the supermultiplet in each case.
- 6.10** Calculate the quark model prediction for the magnetic moment of the  $\Sigma^0(1193)$ .
- 6.11** In Section 6.4.2 we showed that the confinement condition  $\hat{F}_1 \chi_B^C = 0$  led to the conditions  $\alpha_1 = -\alpha_2, \alpha_3 = -\alpha_4, \alpha_5 = -\alpha_6$  for the coefficients  $\alpha_i$  in the general colour wavefunction for baryons (6.55). Find the conditions resulting from the other confinement conditions (6.53). Show that they are consistent with each other, and uniquely determine the colour wavefunction for baryons to be the totally antisymmetric form (6.43), up to an arbitrary overall normalisation constant.
- 6.12** Use the methods of Appendix C to show that a state with  $I = 1, I_3 = 0$  cannot decay to  $\pi^0\pi^0$  by strong interactions. Hence show that the  $d^*$  resonance produced in reaction (6.46) has isospin zero, given that the deuteron has  $I = 0$ .
- 6.13** Draw the lowest-order quark diagrams for the processes:
- $e^+ + e^- \rightarrow K^+ + K^-$ ,
  - $J/\psi \rightarrow \mu^+ + \mu^-$ ,
  - $J/\psi \rightarrow \eta_c + \gamma$ .
- 6.14** A hypothetical spin-1 bottomonium resonance  $\Upsilon$  with mass  $9\text{ GeV}/c^2$  is observed in the reaction  $e^+e^- \rightarrow \mu^+\mu^-$  and in the total electron–positron annihilation cross-section. The resonance is too narrow for its width to be resolved directly, but the cross-sections integrated over the vicinity of the resonance are

$$\int \sigma_\mu(E) dE = 10 \text{ nb MeV} \quad \text{and} \quad \int \sigma_t(E) dE = 300 \text{ nb MeV},$$

where  $E$  is the total centre-of-mass energy,  $\sigma_\mu = \sigma(e^+e^- \rightarrow \mu\mu)$  and  $\sigma_t$  is the total cross-section. Assuming the cross-sections are dominated by a Breit–Wigner resonance of the form discussed in Section B5.2, evaluate the partial widths  $\Gamma(\Upsilon \rightarrow e^+e^-), \Gamma(\Upsilon \rightarrow \mu^+\mu^-)$  and the total decay width.

# 7

---

## QCD, jets and gluons

---

In Chapter 6, we have seen how the static properties of hadrons can be understood in terms of their quark structure. In this chapter we will consider the dynamical effects of this substructure as seen in scattering experiments. Such effects are observed in many reactions, but here we will concentrate on electron–positron annihilation, in which quarks and gluons are closely associated with jets of hadrons observed in the final state, and where the quark interpretation of the experiments is relatively direct. In the following chapter we will extend the discussion to high-energy scattering experiments in which beams of leptons, which are themselves structureless, are used to probe the internal structures of the proton and neutron. Firstly, however, we introduce the theory of strong interactions, whose qualitative features will underpin almost our whole account.

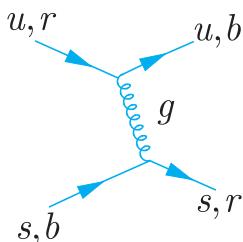
### 7.1 Quantum chromodynamics

The theory that describes strong interactions in the standard model is called *quantum chromodynamics*, or QCD for short. Although this is not tested to the same extent or precision as quantum electrodynamics (QED), it is nevertheless in impressive agreement with a large body of experimental data, and is not contradicted by any experiment. QCD is similar to QED in that both describe interactions that are mediated by massless spin-1 bosons coupling to conserved charges. Theories of this type are called *gauge theories* because they have a characteristic symmetry, called *gauge invariance*. Gauge invariance plays a fundamental role in theoretical treatments of QED and QCD, where it can be used to infer the detailed forms of the interactions.<sup>1</sup> Here we shall adopt a more

phenomenological approach. The spin-1 bosons are called gauge bosons; in QED they are photons and in QCD they are called gluons and denoted by the symbol  $g$ . Gluons have zero electric charge, like photons, but couple to the colour charges that were discussed in Section 6.4.1, rather than to the electric charge. This leads immediately to the so-called *flavour independence* of strong interactions: that is, the different quark flavours  $u$ ,  $d$ ,  $s$ ,  $c$ ,  $b$  and  $t$  must have identical strong interactions, because they exist in the same three colour states  $r$ ,  $g$ ,  $b$ , with the same possible values of the colour charges. This has its most striking consequences for  $u$  and  $d$  quarks, which have almost equal masses, where it leads to the phenomenon of isospin symmetry described in Section 6.1. In the same way, because the mass difference between the  $s$  quark and the  $u$  and  $d$  quarks is small on the scale of hadron masses, it leads to the supermultiplets discussed in Sections 6.2 and 6.3. Flavour independence also implies the equality of the potentials in the charmonium and bottomonium systems, which emerged from experiment as shown in Section \*6.5.3.

A second property of strong interactions that follows from the above picture without detailed argument is that the forces between the quarks must be of long range, because the gluons have zero mass. This does not imply that the forces between hadrons are also long range, because hadrons have zero colour charges overall, as we saw in the last chapter. The forces between the ‘colourless’ hadrons are the residues of the forces between their quark constituents, and cancel when the hadrons are far apart.

We have noted that QED and QCD both describe interactions, albeit of very different strengths, which are mediated by massless spin-1 bosons that couple to conserved charges. However, there is a crucial difference between them that profoundly affects the character of the resulting forces. It is that while the photons that couple to the electric charge are themselves neutral, gluons have non-zero values of the colour charges to which they couple. This is illustrated in Figure 7.1, which shows a particular example of a quark–quark interaction by gluon exchange, where the gluon is represented by a ‘corkscrew’ line to distinguish it from a photon. In this diagram, the colour states of the two quarks are interchanged, and the gluon has the colour quantum numbers



**Figure 7.1** Example of quark–quark scattering by gluon exchange. The gluon is represented by a ‘corkscrew’ line to distinguish it from a photon. In this diagram the quark flavour  $u$  or  $s$  is unchanged on gluon emission, but the colour state can change, as shown.

$$I_3^C = I_3^C(r) - I_3^C(b) = 1/2 \quad (7.1a)$$

and

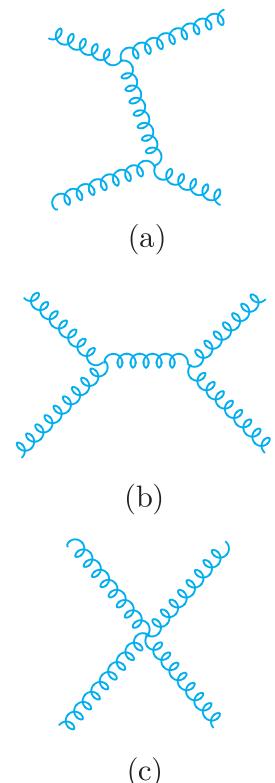
$$Y^C = Y^C(r) - Y^C(b) = 1, \quad (7.1b)$$

<sup>1</sup> This approach is called the *principle of minimal gauge interactions*, or the *gauge principle* for short. It is discussed explicitly in Appendix D, but is not required in what follows.

which follow from applying colour conservation at, for example, the lower vertex, and using the values for the quark colour states  $r$ ,  $b$  listed in Table 6.13. Just as quarks exist in three different colour states, gluons can exist in eight different colour states,<sup>2</sup> and the  $I_3^C, Y^C$  values (7.1) correspond to just one of them. We do not need to consider these states in detail for what follows, since we shall not attempt any detailed calculations in QCD. Rather, we note that if gluons couple to particles with non-zero colour charges, and if gluons themselves also have non-zero charges, then it follows that gluons couple to other gluons. The two types of gluon self-coupling that occur in QCD are illustrated in Figure 7.2, which shows the three lowest-order contributions to gluon–gluon scattering. The strength of these interactions is characterised by the strong coupling constant  $\alpha_s$ , which is analogous to the fine structure constant  $\alpha$  in QED. Just as each electron–photon vertex in QED corresponds to a basic process with a transition probability of order  $\alpha$ , each quark–gluon vertex and three-gluon vertex in Figures 7.1 and 7.2, respectively, corresponds to a basic process of order  $\alpha_s$ . However, the zero-range ‘contact’ interaction corresponding to the four-gluon vertex in Figure 7.2(c) is of order  $\alpha_s^2$ , so that all three diagrams in Figure 7.2 are second-order diagrams, even though Figure 7.2(c) has only one vertex.

The gluon–gluon interactions of Figure 7.2 have no analogue in QED, and it can be shown that they lead to properties of the strong interaction that differ markedly from those of the electromagnetic interaction. These properties are *colour confinement* and *asymptotic freedom*. Colour confinement is the requirement that observed states have zero colour charges. This was discussed in Section 6.4.1, where, for example, it was shown to imply that all quark bound states (hadrons) must have integral electric charges. It also implies that gluons, like quarks, cannot be observed as isolated free particles, because they have non-zero values of the colour charges. Bound states of two or more gluons with zero colour charges overall can be formed in principle, due to the strong interaction between gluons themselves. Such states are called *glueballs*. Similarly, gluons may bind to, for example,  $q\bar{q}$  pairs, to form colourless states, called *hybrids*, composed of both quarks and gluons. Both glueballs and hybrids will be discussed further in Section 7.1.3.

Asymptotic freedom means that the interaction gets weaker at short distances, and at distances less than about 0.1fm the lowest order diagrams dominate. At these distances, quark–quark scattering, for example, is given approximately by one-gluon exchange diagrams like Figure 7.1. However, as the distance between the



**Figure 7.2** The three lowest-order contributions to gluon–gluon scattering in QCD.

<sup>2</sup> This can be shown to follow as a consequence of the assumed symmetry properties of QCD, but we will not discuss this further.

quarks increases, the interaction gets stronger, and many higher-order diagrams become important. In this strong interaction regime perturbation theory is no longer applicable, and it has not yet been possible to evaluate the theory precisely. We therefore have to rely on approximate results obtained by numerical simulations of the theory on very large computers, and the demonstration of confinement in QCD rests largely on such simulations. They are done using an approach called *lattice gauge theory* in which space (and sometimes time) is approximated by a finite lattice of discrete points. The exact theory can then in principle be recovered by letting the lattice spacing go to zero, and the number of lattice points become infinite.<sup>3</sup> In practice, the number of lattice points that can be handled is limited by the computing power available, but nonetheless good results have been obtained for several static properties, for example the masses and decay constants of the lower-lying hadrons.<sup>4</sup>

The above features are conveniently illustrated by considering the static potential between a heavy quark and its antiquark in a colour singlet state. This is the appropriate potential for a non-relativistic discussion of charmonium and bottomonium, and was determined empirically for the limited range  $0.2 \leq r \leq 0.8$  fm in \*Section 6.5.3. Here we are concerned with the behaviour predicted on the basis of QCD. At short interquark distances  $r \leq 0.1$  fm, the interaction is dominated by one-gluon exchange and we might expect a Coulomb-like potential analogous to that arising from one-photon exchange in QED, so that the potential is given by

$$V(r) = -\frac{4}{3} \frac{\alpha_s}{r} \quad (r \leq 0.1 \text{ fm}), \quad (7.2a)$$

where  $\alpha_s$  is the strong coupling constant discussed above.<sup>5</sup> Because of asymptotic freedom, the strength of the interaction, and hence  $\alpha_s$ , decreases with decreasing  $r$ , but for  $r \leq 0.1$  fm this variation is slight and can in many applications be neglected, as we shall see below. At distances beyond 0.1 fm, however, the strength of the interaction increases more rapidly and one-gluon exchange no longer dominates. In this region, we have to rely mainly on lattice gauge theory calculations of limited precision. These are consistent with the empirical potential of Figure 6.12 in the range where this is determined and show that at large distances the potential increases approximately linearly

$$V(r) \approx \lambda r \quad (r \geq 1 \text{ fm}), \quad (7.2b)$$

---

<sup>3</sup> A simple account of such calculations is given in Weingarten (1996).

<sup>4</sup> See, for example, Section 15.6 of Amsler *et al.* (2013).

<sup>5</sup> The numerical factor multiplying  $\alpha_s$  (i.e.  $-4/3$  in this case) depends on the colour state chosen, and we will not discuss it further.

where the constant  $\lambda$  is of order 1 GeV/fm. This is an example of a confining potential in that it does not die away with increasing separation and the force between the quark and antiquark cannot be neglected, even when they are very far apart.

### 7.1.1 The strong coupling constant

The strong interaction derives its name from the strong forces acting at distances of order 1 fm that, among other things, bind quarks in hadrons. However, the remarkable phenomena discussed in this chapter depend on the fact that the interaction gets weaker at short distances, that is, on asymptotic freedom. Such short-distance interactions are associated with large momentum transfers  $|q|$  between the particles, with

$$|\mathbf{q}| = \mathcal{O}(r^{-1}), \quad (7.3)$$

where  $r = |\mathbf{r}|$  is the distance at which the interaction occurs. For example, the amplitude (1.32) for scattering from a spherically symmetric potential  $V(r)$  becomes

$$\mathcal{M}(q) = 4\pi \int_0^\infty V(r) \left( \frac{\sin(qr)}{qr} \right) r^2 dr \quad (7.4)$$

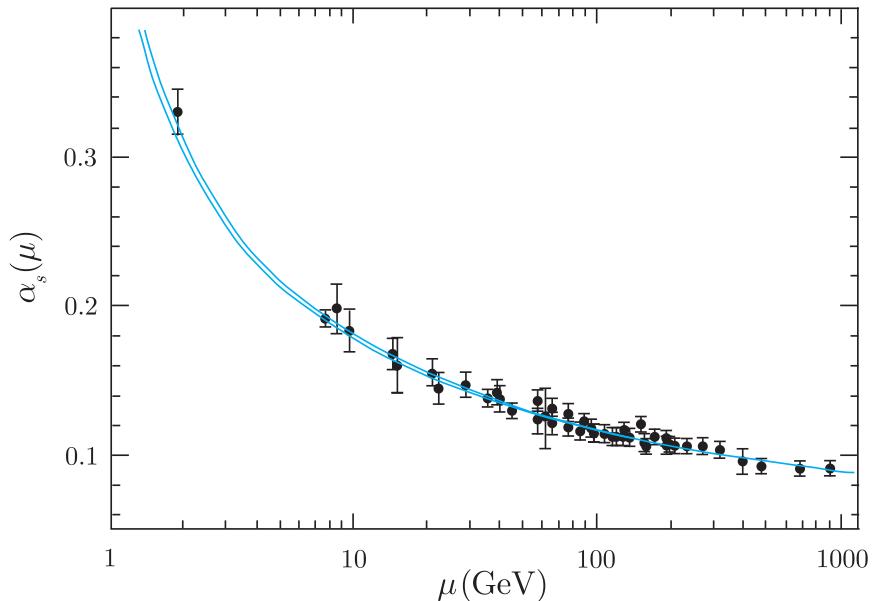
on integrating over all angular directions. The dominant contributions arise from  $r$  values of order  $q^{-1}$  as asserted, since for smaller  $r$  the integrand is suppressed by the factor  $r^2$ , while for large  $r$  it is suppressed by the average over the rapidly oscillating sine factor. Hence in discussing scattering from the static potential (7.2a), in which the strong coupling constant  $\alpha_s$  decreases with decreasing  $r$ , we can equally well regard it as decreasing with increasing momentum transfer, since the one implies the other by (7.3).

In this discussion, we have considered scattering from a static potential leaving the energy of the particle unchanged. Particle energies are also unchanged in elastic scattering in the centre-of-mass frame, but in other reference frames, and in other processes, energy as well as momentum can be exchanged between the particles. In such cases, the strength of the interaction can be shown to depend on the Lorentz-invariant quantity

$$\mu^2 \equiv |\mathbf{q}^2 - E_q^2|, \quad (7.5)$$

which reduces to  $\mathbf{q}^2$  when the energy exchanged  $E_q$  is zero. More specifically, it can be shown that, for  $\mu^2 \gg 1 \text{ GeV}^2$ , the QCD coupling constant  $\alpha_s$  is given to a good first approximation by

$$\alpha_s(\mu) = \alpha_s(\mu_0) \left[ 1 + \frac{(33 - 2N_f)}{6\pi} \alpha_s(\mu_0) \ln(\mu/\mu_0) \right]^{-1}, \quad (\mu^2 \gg 1 \text{ GeV}^2). \quad (7.6)$$



**Figure 7.3** Values of the running coupling constant  $\alpha_s$  obtained from the following sources: hadronic  $\tau$  decays; lattice QCD; jets in deep inelastic lepton–nucleon scattering; heavy quarkonia decays; hadronic final states of  $e^+ e^-$  annihilations; hadron collider jets; and electroweak precision fits to  $Z^0$  decay. The solid curves show the evolution of  $\alpha_s$  with  $\mu$ , as predicted by QCD, assuming the value (7.7) for  $\alpha_s(M_Z)$ . (Adapted from Bethke *et al.* 2015. Reproduced with permission from Elsevier.)

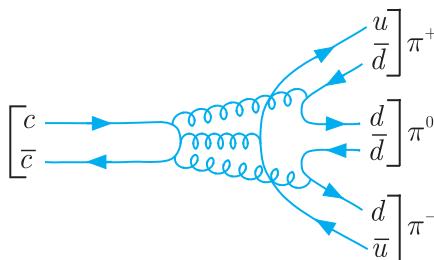
Here  $N_f$  is the number of quark flavours  $u, d, s, \dots$ ,<sup>6</sup> with  $4m_q^2 < \mu^2$  and  $\alpha_s(\mu_0)$  is the value of the coupling at a chosen reference value  $\mu_0$ , which is usually taken to be  $\mu_0 = M_Z$ , the mass of the  $Z$  boson. Measured values of  $\alpha_s(\mu)$  obtained from a variety of different processes are shown in Figure 7.3, where the curves show the predicted behaviours corresponding to the ‘best-fit’ value

$$\alpha_s(M_Z) = 0.118 \pm 0.002 \quad (7.7)$$

at the reference value  $\mu_0 = M_Z$ . The decrease in  $\alpha_s(\mu)$  as  $\mu$  increases, corresponding to shorter distances, is clearly seen, and because of this variation,  $\alpha_s(\mu)$  is often referred to as the *running coupling constant*. Conversely,  $\alpha_s(\mu)$  increases more and more rapidly as  $\mu$  decreases towards 1 GeV, until it is no longer small compared to unity, when the interaction becomes strong and perturbative ideas are no longer applicable.

We are now able to discuss the origin of the OZI rule, which was used in Section 6.5 to explain the narrowness of the charmonium and bottomonium states below the charm and bottom thresholds, respectively. When a quark–antiquark pair annihilates to gluons, the appropriate mass scale is the invariant mass of the pair. Hence light quark pairs are easily produced in the region where the interaction is strong, and typical hadron decays, like  $K^*(890)$  decay (cf. Figure 3.11) or the decay of charmonium states above the charmed threshold

<sup>6</sup> The change in  $\alpha_s(\mu)$  at  $\mu^2 = 4m_c^2, 4m_b^2$  is, of course, not really discontinuous as implied by the approximation (7.6), but is ‘smoothed out’ over a threshold region.



**Figure 7.4** OZI-suppressed decay of a  $C = -1$  charmonium state below the  $D\bar{D}$  threshold. The corresponding quark diagram, in which all virtual gluon lines are by convention suppressed, has already been given in Figure 6.9(b).

(cf. Figure 6.9(a)), proceed via the creation of one or more such pairs, with typical widths of order tens of MeV. On the other hand, the annihilation or creation of  $c\bar{c}$  or  $b\bar{b}$  pairs is relatively suppressed because of the running of the coupling constant, which becomes small enough for perturbative arguments to be used.

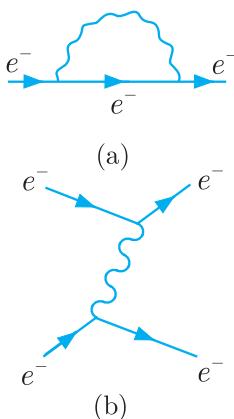
Now let us consider the decays of charmonium or bottomonium states below the charm or bottom threshold, by the annihilation mechanism illustrated in Figure 6.9(b), and let us suppose that the decaying state has charge parity  $C = -1$ , like the  $J/\psi$  or the lightest  $\Upsilon$  states. Then, because the initial heavy quark–antiquark pair is colourless, it cannot annihilate to a single gluon without violating colour conservation, and because the initial state has  $C = -1$  and  $n$  gluons, like  $n$  photons, have  $C$  parity  $(-1)^n$ , it must annihilate to at least three gluons to conserve  $C$  parity. This is illustrated for a  $c\bar{c}$  state decaying to a three-pion final state in Figure 7.4, and because of the additional powers of  $\alpha_s(\mu)$  required to produce the three gluons, the corresponding decay rates are heavily suppressed. Of course more quantitative results require more detailed arguments. In the case of the corresponding  $\Upsilon$  states, as they are heavier, more accurate results are available because perturbation theory converges more rapidly than for charmonium states. The corresponding value of  $\alpha_s(\mu)$  is smaller, and this leads to one of the empirical values of  $\alpha_s$  shown in Figure 7.3.

### 7.1.2 Screening, antiscreening and asymptotic freedom

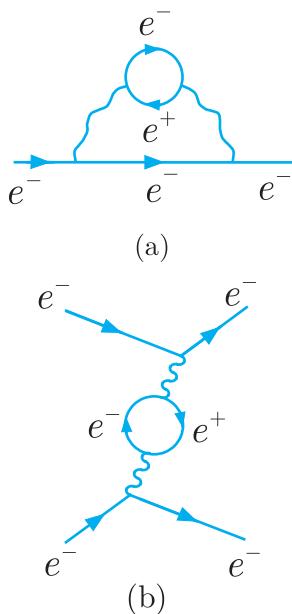
A derivation of asymptotic freedom needs quantum field theory<sup>7</sup> and we will content ourselves here with a qualitative discussion, making plausible its connection to gluon self-interactions. Firstly, however, it will be helpful to take a closer look at QED, where the analogous photon self-interactions are absent.

Students studying QED for the first time, when meeting the one-photon exchange contribution to electron scattering in Figure 1.7,

<sup>7</sup> Such a derivation can be found in, for example, Chapter 15 of Mandl and Shaw (2010).



**Figure 7.5** (a) The lowest-order quantum fluctuation of an electron and (b) the associated exchange process.



**Figure 7.6** (a) A higher-order quantum fluctuation of the electron, together with (b) the associated exchange process.

often ask the question: ‘When the first electron emits the photon, how can it know that there is a second electron close enough to absorb it in the short time allowed by the energy–time uncertainty principle?’ Of course the electron knows no such thing, and the question suggests itself because our earlier account is incomplete. According to QED, what happens is that the electron emits and reabsorbs virtual photons all the time, as shown in Figure 7.5(a), whether another electron is nearby or not. If, however, a second electron is nearby, it will sometimes absorb the photon before it is reabsorbed by the first electron, giving rise to the familiar exchange diagram of Figure 7.5(b).<sup>8</sup>

Processes in which one particle is converted for a short time to two or more particles are called *quantum fluctuations*. Figure 7.5(a) is only the simplest quantum fluctuation of an electron, and others are possible, such as that shown in Figure 7.6(a). This is similar to Figure 7.5(a), except that in this case the emitted photon itself fluctuates to an electron–positron pair before it is re-absorbed by the initial electron. Thus an electron continuously emits and re-absorbs not only photons but also, indirectly, electron–positron pairs. These pairs have small but measurable consequences called *vacuum polarisation effects*<sup>9</sup> by analogy with those observed when a charge is immersed in a dielectric medium. If the charge is positive, then the molecules of the dielectric will tend to align themselves as shown in Figure 7.7, and if a small test charge is introduced at a distance that is large compared with the size of the molecules, it will experience a force that is the same as if the charge  $Q$  had been replaced by a smaller ‘effective charge’  $Q_{\text{eff}}$ . However, at distances between the charges that are of the same order as the molecular sizes, this screening becomes less effective, so that  $Q_{\text{eff}}$  increases as the charges get closer together. There is a somewhat similar effect in QED, even in a vacuum, associated with the electron–positron pairs produced by quantum fluctuations. For example, if a second electron is close by, it will sometimes absorb the second photon in Figure 7.6(a) before it is re-absorbed by the initial electron, giving rise to Figure 7.6(b). This

<sup>8</sup> In interpreting these diagrams, it is important to remember that the particles do not travel on well-defined classical trajectories. For example, an argument analogous to that given in Section 1.4.1 shows that the de Broglie wavelength of the exchanged photon in Figure 7.5(b) is greater than, or about equal to, the distance between the two vertices, and the uncertainty in the position of the photon must be of at least the same order. It is possible, therefore, for it to be absorbed by either electron, despite the distance between them.

<sup>9</sup> The calculation of these and other effects due to quantum fluctuations presents considerable difficulties, which took many years to overcome. Here we will simply state the results of such calculations without further comment. They are described in, for example, Chapter 9 of Mandl and Shaw (2010).

diagram contributes to electron–electron scattering, and a detailed calculation shows that, at short distances,

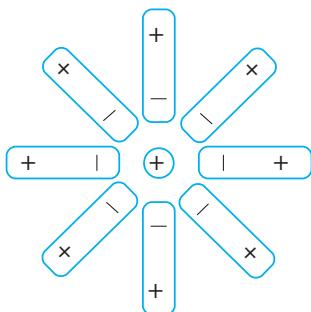
$$r \leq r_c \equiv \hbar/mc = 3.9 \times 10^{-13} \text{ m},$$

it leads to corrections that are of order  $\alpha$  smaller than the familiar Coulomb potential arising from the one-photon exchange diagram in Figure 7.5(b). Specifically, if we write the effective potential in the form  $\phi_{\text{eff}} = \alpha_{\text{eff}}/r$ , then

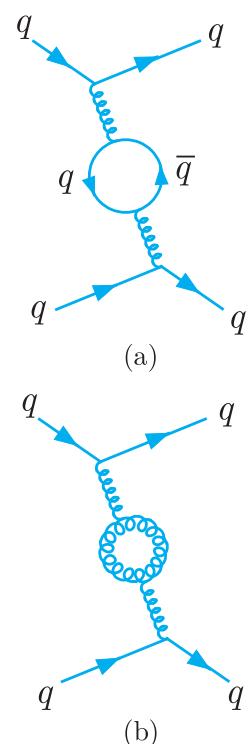
$$\alpha_{\text{eff}} = \alpha \approx 1/137 \quad (r \gg r_c),$$

but for  $r < r_c$  it is somewhat larger and increases as  $r$  becomes smaller. In other words, the interaction strength increases at very short distances, and because of this one can show, for example, that the  $2s$  state in hydrogen is more tightly bound (by  $2.2 \times 10^{-17}$  eV) than it would be in a pure Coulomb potential. This effect is tiny, but nonetheless is present and must be taken into account if the very precise experimental data on the hydrogen spectrum are to be fully understood.<sup>10</sup>

Quantum fluctuations also exist in QCD, and as in QED they also lead to a variation in the interaction strength with distance. Specifically, if we consider quark–quark scattering there are two lowest-order vacuum polarisation corrections to the one-gluon exchange diagram of Figure 7.1. These are shown in Figure 7.8. The first of these is a direct analogue of Figure 7.6(b) in electron–electron scattering, and like it leads to a screening correction. If this were the only contribution, the interaction would grow stronger at short distances as in QED. However, there is also the second diagram of Figure 7.8(b), involving a gluon–gluon pair produced in a gluon self-interaction of the type shown in Figure 7.2(a). This diagram has no counterpart in QED and the nature of its contribution is far from obvious. In fact, detailed calculation shows that it leads to an *antiscreening* effect, meaning that it causes the interaction to grow weaker at short distances. This effect is bigger than the screening correction from Figure 7.8(a), and the net result is that the interaction grows weaker at short distances; that is, there is asymptotic freedom.



**Figure 7.7** Schematic diagram representing the polarisation of the molecules of a dielectric medium by a positive charge placed within it.



**Figure 7.8** The two lowest-order vacuum polarisation corrections to one-gluon exchange in quark–quark scattering.

### \*7.1.3 Exotic hadrons

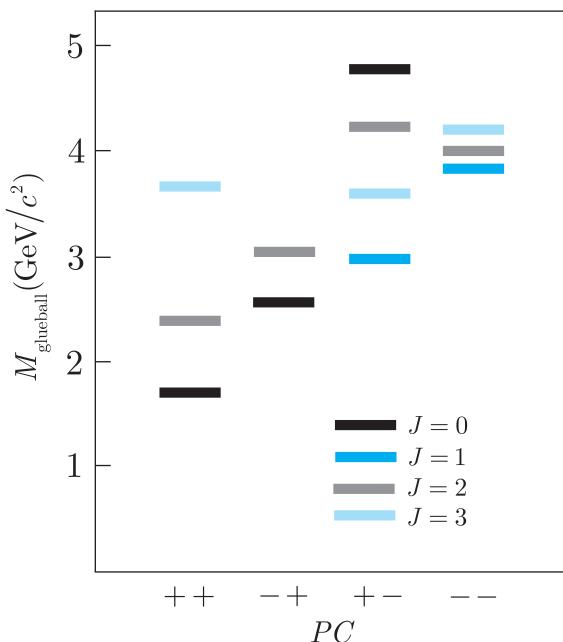
Any hadrons that lie outside the simple quark model (so not having the structure  $q\bar{q}$ ,  $qqq$  or  $\bar{q}\bar{q}\bar{q}$ ) are termed *exotic*, whether or not their quantum numbers are exotic, that is whether or not their quantum numbers are forbidden to occur in the simple quark model. Such states are allowed in QCD, and lattice gauge theory calculations, for

<sup>10</sup> This is discussed in pp. 207–208 of Mandl and Shaw (2010).

example, suggest that while conventional quark model states dominate at lower masses, exotic states do play a role in hadron spectroscopy. In this section, we discuss the evidence for such states.<sup>11</sup> We will start by considering glueballs and hybrids involving the light quarks  $u$ ,  $d$ ,  $s$  and then go on to consider exotic mesons and baryons containing heavy quark–antiquark pairs.<sup>12</sup>

### 7.1.3(a) Glueballs and hybrids

We start with glueballs. Since they are composed of gluons alone, they are predicted to be strongly interacting neutral bosons with  $I = S = C = \tilde{B} = 0$  and, if they exist, one would expect their electromagnetic interactions to be much weaker than those of ‘ordinary’ mesons composed of charged quarks and antiquarks. Unfortunately, precise theoretical calculations of glueball properties are not yet possible because of our limited understanding of confinement. However, approximate lattice gauge theory calculations that ignore the existence of quarks yield glueball spectra similar to that shown in Figure 7.9, where the lightest glueball is a scalar meson with  $J^{PC} = 0^{++}$  and a mass of around  $1.5 - 1.7 \text{ GeV}/c^2$ , depending on the details of the approximations made. However, when quarks are included in the theory, such states can usually mix with  $q\bar{q}$  mesons of the same



**Figure 7.9** Predicted glueball mass spectrum from lattice QCD calculations in the absence of quarks. (Adapted from Chen *et al.* 2006. © American Physical Society 2006. Reproduced with permission.)

<sup>11</sup>The possible evidence for one such state has already been considered in Section 3.6.

<sup>12</sup>For a comprehensive review of exotic hadrons see, for example, Olsen (2015).

quantum numbers and similar masses, so that the observed states are unlikely to be pure glueballs, but states with both glueball and  $q\bar{q}$  components. These mixed states are difficult to distinguish from ordinary  $q\bar{q}$  mesons, but there is some evidence for them, as we shall see shortly.

Firstly, however, we note that the problem of glueballs mixing with conventional  $q\bar{q}$  mesons would be avoided if the  $J^{PC}$  values of the observed state could not occur in the simple quark model. By (5.39a) and (5.53) we have  $P = (-1)^{L+1}$  and  $C = (-1)^{L+S}$  for a bound state of a quark with its own antiquark. Hence for  $J = 0$  there are just two possibilities:  $L = S = 0$  with  $J^{PC} = 0^{-+}$  and  $L = S = 1$  with  $J^{PC} = 0^{++}$ . The remaining possibilities,  $0^{--}$  and  $0^{+-}$ , do not occur in the simple quark model. Similarly, for arbitrary  $J > 0$  one finds that combination  $P = (-1)^J$ ,  $C = (-1)^{J+1}$  does not occur (cf. Problem 7.6). Hence the quantum number combinations

$$B = Q = S = C = \tilde{B} = 0 \quad (7.8a)$$

and

$$J^{PC} = 0^{--} \quad \text{and} \quad J^{PC} = 0^{+-}, 1^{-+}, 2^{+-}, 3^{-+}, \dots \quad (7.8b)$$

are ‘exotic’, that is they are forbidden in the simple quark model, and a glueball with these quantum numbers cannot mix with a  $q\bar{q}$  state. Unfortunately, as can be seen from Figure 7.9, the lightest glueball with exotic quantum numbers is predicted to be a  $2^{+-}$  state with a mass of about  $4\text{GeV}/c^2$ . Because it is so heavy, such a state is likely to be very unstable and difficult to identify, and has not been observed.

However, while there are no established examples of pure glueballs, there is evidence for both hybrid states, which contain one or more gluons in addition to a  $q\bar{q}$  pair, and for mixed states, which are superpositions of  $q\bar{q}$  and glueball components. In particular, the 2014 edition of the Particle Data Group compilation lists two isotriplets, the  $\pi_1(1400)$  and the  $\pi_1(1600)$ , as established states, with  $S = C = \tilde{B} = 0$  and members with  $Q = I_3 = +1, 0, -1$  in both cases. These states cannot be glueballs, because they have non-zero isospin and cannot be  $q\bar{q}$  states because, in both cases, the neutral member has the exotic quantum numbers  $J^{PC} = 1^{-+}$ . They could be four-quark states or hybrids. However, four-quark states formed from light quarks are expected to occur in supermultiplets containing members with exotic internal quantum numbers, and no such states have been identified.<sup>13</sup> Hence the simplest explanation is that they are hybrids. In addition, there is a ‘surplus’ of observed states in the range  $(1 - 2)\text{GeV}/c^2$  with the quantum numbers  $I = S = C = \tilde{B} = 0$  and  $J^{PC} = 0^{++}$ , when compared to the

---

<sup>13</sup>See the discussion in Section 6.4.1.

predictions of the simple quark model. The latter predicts just two such states with  $L = S = 1$ , corresponding to the quark combinations  $(u\bar{u} + d\bar{d})/\sqrt{2}$  and  $s\bar{s}$ , while, as we have seen, the lightest glueball is expected to lie in this region. Experimentally four states are observed – called  $f^0(980)$ ,  $f^0(1370)$ ,  $f^0(1500)$  and  $f^0(1710)$  – which are plausibly interpreted as mixtures of the two predicted quark states, the predicted glueball and a hybrid state. Clearly, states do occur that cannot be interpreted as  $q\bar{q}$  pairs, but when they do, unless they have exotic quantum numbers, they are likely to be quantum mechanical superpositions of allowed and exotic states.

### 7.1.3(b) Heavy quarkonia

The narrow, low-lying states of charmonium, below the charm threshold, are well described by the simple quark model, as we have seen in the Section 6.5. However, there are also states above the charm threshold, and while many of these exhibit the behaviour expected in the simple quark model, others do not. In particular, in 2005 the BaBar Collaboration at SLAC identified a neutral vector meson, conventionally called  $X(4260)$ , whose existence was subsequently confirmed by other experiments. It was first observed as a resonance peak in the  $(\pi^+\pi^-J/\psi)$  mass spectrum seen in the reaction

$$e^+ + e^- \rightarrow \gamma + \pi^+ + \pi^- + J/\psi \quad (7.9a)$$

and has also been observed in the reaction

$$e^+ + e^- \rightarrow X(4260) \rightarrow \pi^+ + \pi^- + J/\psi, \quad (7.9a)$$

corresponding to the production mechanism of Figure 6.7, characteristic of a vector with  $J^{PC} = 1^{--}$ . The observed mass and width are

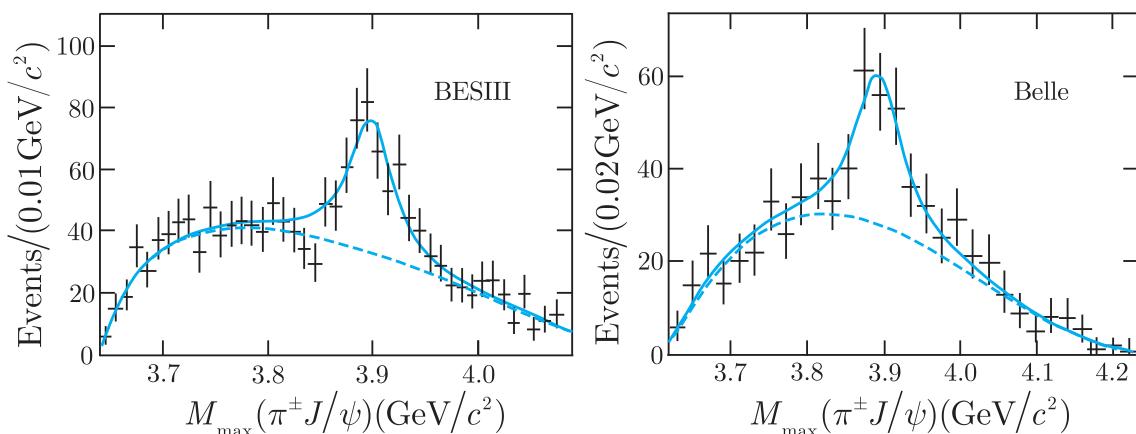
$$M = 4251 \pm 9 \text{ MeV}/c^2, \quad \Gamma = 120 \pm 12 \text{ MeV}.$$

The  $X(4260)$  has been observed to decay to  $\pi^0\pi^0J/\psi$  and to  $K^+K^-J/\psi$ , as well as to  $\pi^+\pi^-J/\psi$ , and since the  $\pi^0\pi^0J/\psi$  can only have isospin 0 or 2 (cf. Problem 6.11) and  $K^+K^-J/\psi$  can only have isospin 0 or 1, the  $X(4260)$  can only have isospin 0.

In the simple quark model, the  $X(4260)$  would be interpreted as a  $c\bar{c}$  state lying above the charm threshold. However, such states are expected to decay predominantly by the mechanism of Figure 6.9(a) into final states that contain particles with a non-zero charm quantum number, and decay only rarely to give a  $J/\psi$  particle in the final state. This is precisely the behaviour observed for the vector charmonium states  $\psi(3770)$ ,  $\psi(4040)$ ,  $\psi(4160)$  and  $\psi(4415)$ , corresponding to the four broad peaks in the total cross-section seen in Figure 6.8. In contrast, the  $X(4260)$  couples very weakly to  $e^+e^-$ , and open

charm decays, far from dominating, have so far not been detected. This behaviour is very difficult to understand if the  $X(4260)$  is a  $c\bar{c}$  state, but can be understood if it is a hybrid state  $c\bar{c}g$  or a *tetraquark-charmonium* state  $c\bar{c}q\bar{q}$ , where  $q = u, d, s$  or a mixture of the two. The first of these possibilities – that it is a hybrid – has the advantage that such states obviously have isospin zero. On the other hand, if it is a four-quark state, one would expect it to belong to a supermultiplet together with so far undetected states with non-zero isospin or strangeness, such as like  $c\bar{c}u\bar{d}$  or  $c\bar{c}s\bar{u}$ , although it is possible this might not be the case if it were a very loosely bound molecular state.

The ambiguity in the interpretation of the  $X(4260)$  only arises because it is electrically neutral; a charged state containing a  $c\bar{c}$  pair could not be a simple  $c\bar{c}$  or a hybrid, but must necessarily contain extra quarks. That such states do exist was first established in 2013 independently by two experiments, which studied the  $J/\psi\pi^+$  and  $J/\psi\pi^-$  pairs produced in the decay  $X(4260) \rightarrow \pi^+\pi^-J/\psi$ . The BESIII group, working at the Beijing Electron Positron Collider (BEPC) studied the reaction  $e^+e^- \rightarrow \pi^+\pi^-J/\psi$  at the centre-of-mass energy corresponding to the mass of the  $X(4260)$ . For each event, the invariant mass of the  $J/\psi\pi^+$  and  $J/\psi\pi^-$  pairs were calculated and the invariant mass distribution of the heaviest pairs found. It is shown in Figure 7.10 and exhibits a clear resonance peak above a background (shown by the dotted line) arising from uncorrelated pairs produced by some other, non-resonant, mechanism. The mass and width of the peak obtained from the fit shown are  $M = 3899.0 \pm 3.6 \pm 4.9 \text{ MeV}/c^2$  and  $\Gamma = 46 \pm 10 \pm 20 \text{ MeV}$ , respectively, where the first errors are statistical and the second are



**Figure 7.10** Invariant mass distribution of the heaviest  $J/\psi\pi^-$  and  $J/\psi\pi^+$  pairs observed in the reaction  $e^+e^- \rightarrow \pi^+\pi^-J/\psi$  by the BESIII Collaboration (adapted from Ablikim *et al.*, 2013) and the Belle Collaboration (adapted from Liu *et al.*, 2013). (Copyright, American Physical Society, reproduced with permission.)

systematic. The Belle Collaboration, working at the KEK laboratory near Tokyo, studied the reaction (7.9a) at  $e^+e^-$  centre-of-mass energies between 10 and 11 GeV. Events were selected with  $J/\psi\pi^\pm$  masses between 4.15 and 4.45 GeV/ $c^2$  to obtain a sample of 689  $X(4260) \rightarrow \pi^\pm\pi^-J/\psi$  events, with an estimated background of 139 events. The invariant mass distribution of the heavier  $J/\psi\pi^\pm$  pair is shown in Figure 7.10 and again there is a clear resonance peak with parameters  $M = 3894 \pm 6.6 \pm 4.5$  MeV/ $c^2$ ,  $\Gamma = 63 \pm 24 \pm 26$  MeV, in agreement with those from the BESIII Collaboration. These experiments clearly establish the existence of particles, now called  $X^\pm(3900)$ , decaying to  $J/\psi\pi^\pm$ . Subsequently their spin-parity has been determined to be  $J^P = 1^+$  and a neutral isospin partner  $X^0(3900)$  decaying to  $J/\psi\pi^0$  has also been identified.<sup>14</sup> These particles must contain a  $c\bar{c}$  pair, since otherwise the observed decays would be heavily suppressed by the OZI rule, and would make a tiny contribution to the decay rate compared to the very many possible OZI allowed final states that do not contain charmed quarks.<sup>15</sup> Their minimal quark compositions must be  $c\bar{c}ud\bar{d}$  and  $c\bar{c}d\bar{u}$  for the positive and negatively charged states, respectively.

In the above, we focused on the  $X(4260)$  because its study led to the discovery of the first four-quark states  $X^\pm(3900)$  to be unambiguously identified. However, several other neutral meson states that do not fit into the simple quark model have also been observed. These include the  $X(4360)$  and  $X(4660)$ , which are  $J^{PC} = 1^{--}$  states, somewhat analogous to the  $X(4260)$ , and the  $X(3872)$ , which is a  $1^{++}$  state with a decay width measured to be less than 1.2 MeV, even though it lies well above the charm threshold.<sup>16</sup> The  $X(3872)$  sits very close to the  $D\bar{D}^*$  threshold and is known to decay to  $D^0\bar{D}^{*0}$  with a branching ratio of at least 24%. It is therefore a natural candidate for a molecular  $D\bar{D}^*$  state, although other explanations are undoubtedly possible at present. In addition, several other anomalous particles have been reported in both the charmonium and bottomonium spectra, but need further verification before they can be regarded as firmly established states. In short, while some exotic states clearly exist, much more experimental data are needed to clarify their nature and to establish the existence of others. Hopefully, the much higher statistics available at the recently

---

<sup>14</sup>The notation for these various states is not completely settled and the letters  $X$ ,  $Y$  and  $Z$  are used in the literature for suspected exotic  $c\bar{c}$  states in a way that is not always mutually consistent. Here we follow the Particle Data Group in using the letter  $X$  for all of them.

<sup>15</sup>Indeed, the decay width of a  $q\bar{q}$  state of this mass composed of light quarks would be expected to be so large it is very unlikely such a state could ever be detected.

<sup>16</sup>Historically, this was the first exotic candidate to be discovered in the charmonium sector, by the Belle Collaboration in 2003.

upgraded facilities at KEK(SuperBelle) and BEPC, as well as data from other facilities, including LHCb, will provide these in the next few years.

### 7.1.3(c) Exotic baryons

From time to time, experiments have claimed evidence for the existence of baryons with exotic quantum numbers, but as more experiments have been performed and more data accumulated, the evidence has receded. The present consensus is that there is no convincing evidence for such baryons. However, following the discovery of the tetraquark–charmonium states described above, it was natural to search for the analogous *pentaquark–charmonium* states  $c\bar{c}qqq$ , where  $q = u, d, s$ , which would not have exotic quantum numbers. The first evidence for such states was published in 2015 by the LHCb Collaboration, who used the detector described in Section 4.5.2. This detector was designed specifically to study the decays of short-lived particles containing heavy quarks, which are produced in large numbers at the LHC. In this case, it was used to measure about 26 000 decays of the form

$$\Lambda_b^0 \rightarrow K^- + p + J/\psi. \quad (7.10a)$$

Here  $\Lambda_b^0$  denotes  $\Lambda_b^0(5624) = ubd$ , the lightest baryon with a non-zero bottom quantum number. The main contribution to the decay (7.10a) is expected to arise from the sequence

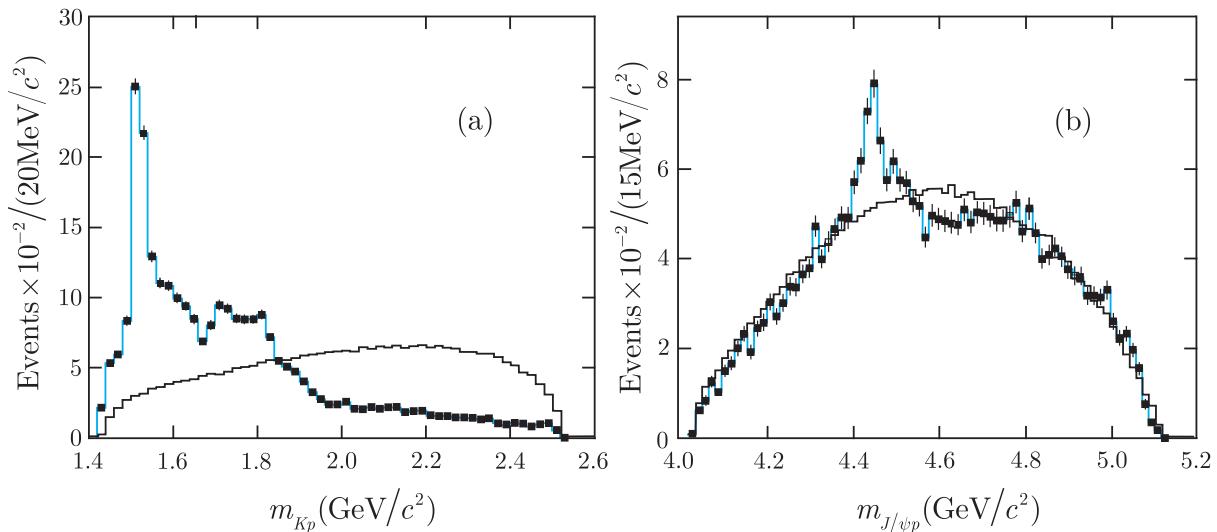
$$\Lambda_b^0 \rightarrow \Lambda^* + J/\psi; \quad \Lambda^* \rightarrow K^- + p, \quad (7.10b)$$

where  $\Lambda^*$  is any resonant excitation of the  $\Lambda(1116) = uds$  state discussed in Section 3.4. This is confirmed by the observed invariant mass distribution of the  $K^-p$  pairs shown in Figure 7.11(a), which shows a prominent peak corresponding to the  $\Lambda(1520) = uds$  resonance and a broader enhancement at higher masses to which several known  $\Lambda^*$  states can contribute. In addition, the  $J/\psi p$  mass distribution, shown in Figure 7.11(b), exhibits a peak at about  $4.5 \text{ GeV}/c^2$ , which is naturally interpreted as a  $J/\psi p$  resonance  $P_c^+$  occurring in the decay sequence

$$\Lambda_b^0 \rightarrow K^- + P_c^+; \quad P_c^+ \rightarrow J/\psi + p \quad (7.10c)$$

However, the peak sits upon a large background arising mostly from reaction (7.10b), and in order to clarify the situation, the authors carried out a very detailed analysis including contributions to the reaction (7.10b) from all fourteen known  $\Lambda^*$  resonances. This confirmed that the data could not be understood without contributions of the form (7.10c) and showed that to understand the detailed mass and angular distributions required not one, but two,  $P_c^+$  resonances, which were found to have the masses and widths

$$M = (4380 \pm 8 \pm 29) \text{ MeV}/c^2, \quad \Gamma = (205 \pm 18 \pm 86) \text{ MeV}$$



**Figure 7.11** Invariant mass of (a)  $K^-p$  and (b)  $J/\psi p$  combinations from  $\Lambda_b^0 \rightarrow J/\psi K^- p$  decays. The black line is the expectation from phase space; the blue line is the data. (Adapted from Aaij *et al.* 2015, with permission from the American Physical Society.)

and

$$M = (4449.8 \pm 1.7 \pm 2.5) \text{ MeV}/c^2, \quad \Gamma = (35 \pm 5 \pm 19) \text{ MeV},$$

respectively. They have the same internal quantum numbers as the proton, but, as in the case of the tetraquark–charmonium states discussed above, the presence of the  $J/\psi$  in their decay products is a strong indication of a  $c\bar{c}$  pair, so the obvious interpretation is that they are pentaquark–charmonium states with the quantum composition  $P_c^+ = uudc\bar{c}$ .

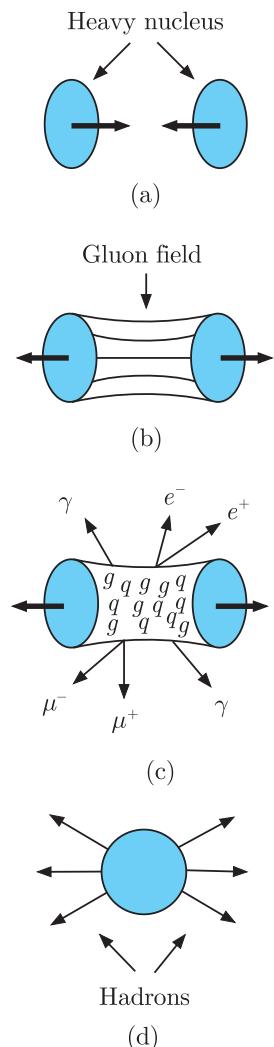
#### \*7.1.4 The quark–gluon plasma

In ordinary matter, quarks are confined within hadrons. However, as the density and/or temperature is increased, a phase transition can in principle occur to a state in which individual hadrons lose their identities, and quarks and gluons become free to move across a volume that is large compared to a hadron. It has not yet been possible to evaluate QCD precisely to see if and when this occurs, because perturbation theory breaks down in the region of interest, so again we have to rely on lattice gauge theory and other approximate methods. These suggest that in the low-temperature limit, the transition should occur at densities about five times higher than that at the centre of a heavy nucleus, while in the limit of low nucleon density, it should occur at temperatures of order 150–200 MeV, which translates to a conventional temperature of order  $10^{12}\text{K}$ , about five orders

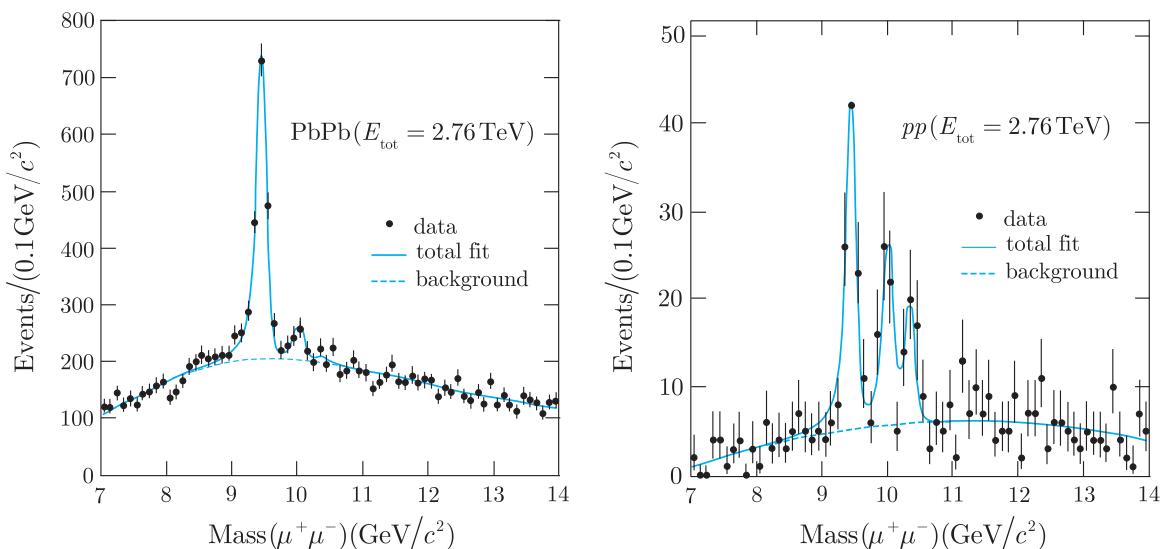
of magnitude greater than the temperature at the centre of the Sun. The resulting new state of matter is called a *quark-gluon plasma* (*QGP*). A quark-gluon plasma is believed to have existed in the first few microseconds after the big bang, and it may exist today at the centre of neutron stars. More prosaically, a quark-gluon plasma may be created briefly in collisions between heavy ions, if the collision energy is large enough. The collisions in the relativistic ion colliders described below are characterised by length and time scales typically of order 10 fm and  $10^{-23}$ s, and if any QGP is created in the collision, the initial ‘fireball’ rapidly cools and expands under its own pressure, first forming a plasma emitting lepton and photon radiation. Then, as it cools further, it will pass through the transition temperature and a gas of hadrons will be formed, resulting in large numbers of hadrons being observed in the final states. This is illustrated schematically in Figure 7.12. The existence of a new state of matter then can be inferred from the properties of the emitted particles and, particularly, the differences from what would be expected if they were simply produced by collisions between the individual nucleons contained in the heavy ions.

High-energy collisions between heavy ions have been studied in several experiments, at the Relativistic Heavy Ion Collider (RHIC) at the Brookhaven National Laboratory, starting in 2000, and more recently using ALICE and other detectors at the LHC, starting in 2010. The RHIC typically collides two counter-circulating beams of fully stripped gold ions at a maximum energy of 200 GeV per nucleon, that is a total collision energy of about 32 TeV. At the LHC, the collisions are between lead ions at a much higher total energy of 574 TeV. If the ions collide centrally (i.e. head-on) several thousand final-state particles are produced. An example of an event seen in the STAR detector (which was shown in Figure 4.26) is shown in Plate 6.

The key question is whether the energy-density in these collisions is sufficient to have created a quark-gluon plasma and its subsequent cooling phases. There are many signatures for this, including the relative abundances of different final-state particle types. For example, the large numbers of gluons in the plasma would lead to copious production of  $s\bar{s}$  pairs via gluon fusion  $gg \rightarrow s\bar{s}$ , and hence production of strange particles in excess of that expected if the colliding nuclei behaved simply as a source of nucleon–nucleon collisions. On the other hand, the production of  $J/\psi$  and upsilon states would be suppressed because the  $c$  and  $\bar{c}$ , or  $b$  and  $\bar{b}$ , quarks produced (also from gluon fusion) would be separated by many quarks of other flavours, leading instead to the production of mesons with non-zero charm and bottom quantum numbers. Data on the  $\mu^+\mu^-$  pairs produced in both heavy ion and proton–proton collisions are shown in Figure 7.13, where the  $\mu^+\mu^-$  invariant mass region covers the mass range



**Figure 7.12** Stages in the formation of a quark-gluon plasma and subsequent hadron emission: (a) two heavy nuclei collide at high energies and (b) interact via the gluon field; (c) the very high energy-density produced causes the quarks and gluons to de-confine and form a plasma that can radiate photons and lepton pairs; (d) finally, as the plasma cools, hadrons condense and are emitted in large numbers. (Adapted from NRC99, with permission from the National Academic Press.)



**Figure 7.13** Dimuon invariant-mass distributions in PbPb (left) and  $pp$  (right) data at a total centre-of-mass energy of 2.76 TeV. (Adapted from S. Chatrchyan *et al.* 2012, with permission from the American Physical Society.) The same reconstruction and analysis selection criteria are applied to both datasets. The solid (signal + background) and dashed (background) curves show the results of the simultaneous fit to the two datasets.

of the lowest three upsilon states  $\Upsilon(9.46)$ ,  $\Upsilon(10.02)$  and  $\Upsilon(10.36)$ . The three states are significantly suppressed in the Pb–Pb ion collisions, with the  $\Upsilon(9.46)$  rate about a factor 2 below what would be expected if the ions behaved as just a collection of nucleons. This is in accord with theoretical expectations, and, in general, these and many other measurements support the existence of a new phase of matter and confirm approximate predictions for the transition temperature. However, whether the transition is a true phase transition, or a more gradual transition as the temperature decreases, is not yet clear. The experiments also suggest that, presumably because of the strength of the colour force between quarks, the plasma observed at present temperatures behaves more like a dense liquid than as a gas of quasi-free particles, as in a true plasma formed of electrons and positive ions. Future experiments at the RHIC and at the LHC will play a crucial role in resolving these and other detailed questions.

## 7.2 Electron–positron annihilation

In this section we shall consider reactions of the type

$$e^+ + e^- \rightarrow \text{hadrons} \quad (7.11)$$

in the centre-of-mass energy range 15–40 GeV. Such reactions have been extensively studied in electron–positron colliding-beam experiments, and give clear evidence for the existence of colour and gluons.

Their study will also introduce us to jets of hadrons, like those shown in Figure 4.12, which are produced in a wide variety of high-energy processes. These jets are closely related to the underlying quark and gluon interactions, and are the closest thing to quark and gluon ‘tracks’ that we are ever likely to see.

### 7.2.1 Two-jet events

In the centre-of-mass energy range 15–40 GeV, electron–positron annihilation into hadrons (7.11) is dominated by processes like that shown in Figure 7.14.<sup>17</sup> These can be regarded as occurring in two stages: a primary electromagnetic process

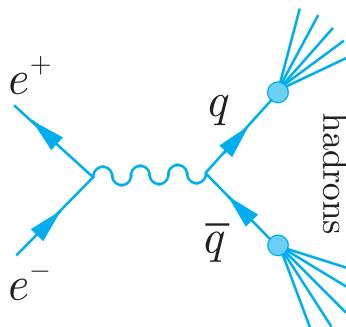
$$e^+ + e^- \rightarrow q + \bar{q}, \quad (7.12)$$

leading to the production of a quark–antiquark pair, followed by a strong interaction process, called *fragmentation*, which converts the high-energy  $q\bar{q}$  pair into two jets of hadrons. These jets are emitted in opposite directions in the centre-of-mass frame in order to conserve momentum, and a typical example of such an event observed in an electron–positron colliding-beam experiment is shown in Figure 4.12.

The fragmentation process that converts the quarks into hadrons is complicated, and the composition of the jets – that is the numbers and types of particles in the jet and their momenta – varies from event to event. However, the direction of a jet, defined by the total momentum

$$\mathbf{P} = \sum_i \mathbf{p}_i, \quad (7.13)$$

where the sum extends over all the particles within the jet, reflects closely the parent quark or antiquark direction. This is because the QCD interaction is relatively weak at very short distances, and the



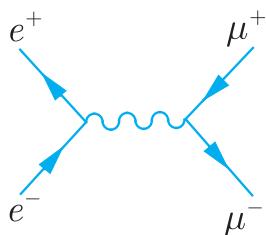
**Figure 7.14** Basic mechanism of two-jet production in electron–positron annihilation.

<sup>17</sup> At lower energies, resonance formation processes like Figure 6.7 contribute, as we saw in Section 6.5, while at higher energies, effects due to weak interactions become important, as we shall see in Chapter 10.

quark and antiquark do not interact strongly until they are separated by a distance  $r$  of order 1 fm, which according to (7.3) gives rise to a typical momentum transfer of order  $200 \text{ MeV}/c$  between them. This is small compared with the momentum of the quark and antiquark produced in the initial reaction (7.12), and the total momenta (7.13) of the jets that subsequently develop point almost exactly in the initial quark and antiquark directions. That is, the jet angular distributions reflect the angular distributions of the quark and antiquark in the basic reaction (7.12).

The above interpretation of the two-jet events is confirmed by comparing their angular distributions with those of muons produced in the reaction

$$e^+ + e^- \rightarrow \mu^+ + \mu^- \quad (7.14)$$



**Figure 7.15** Dominant mechanism for electron–positron annihilation to muon pairs.

by the mechanism of Figure 7.15. This mechanism is identical to that assumed for quark pair production in Figure 7.14, so that the two reactions are very closely related. At high beam energies  $E \gg m_\mu$  the differential cross-section for (7.14) is<sup>18</sup>

$$\frac{d\sigma}{d\cos\theta} (e^+ + e^- \rightarrow \mu^+ + \mu^-) = \frac{\pi\alpha^2}{8E^2} (1 + \cos^2\theta), \quad (7.15)$$

where  $\theta$  is the production angle of either muon with respect to the initial electron direction in the centre-of-mass frame.<sup>19</sup> The corresponding cross-section for producing quark–antiquark pairs  $a\bar{a}$  of a given flavour  $a = u, d, s, \dots$  is then obtained by replacing the muon charge  $e$  with the quark charge  $ee_a$ , and since this charge is the same for all three colour states  $r, g$  and  $b$ , we must multiply by three to account for the three colour states  $r\bar{r}, g\bar{g}$  and  $b\bar{b}$ , giving

$$\frac{d\sigma}{d\cos\theta} = \frac{3\pi e_a^2 \alpha^2}{8E^2} (1 + \cos^2\theta) \quad (7.16)$$

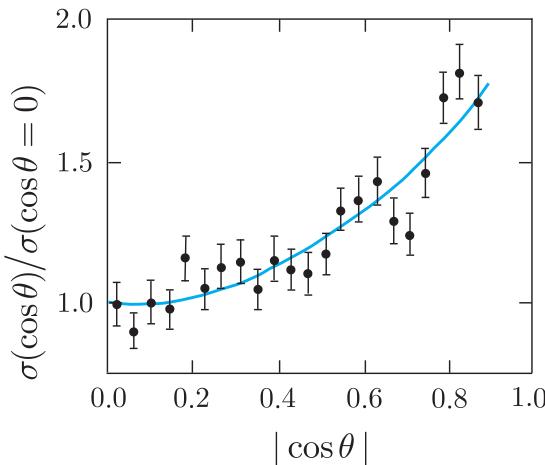
overall. In particular, the angular distributions of the jets should be proportional to

$$(1 + \cos^2\theta), \quad (7.17)$$

irrespective of which quark is produced, if the jet directions reflect the initial quark and antiquark directions, as we have assumed. This is indeed the case, as is shown in Figure 7.16.

<sup>18</sup>This cross-section can be calculated theoretically from the mechanism of Figure 7.15 (see, for example, pp. 146–150 of Mandl and Shaw, 2010). However, it is sufficient for our purposes to take (7.15) from experiment.

<sup>19</sup>The  $\mu^+$  and  $\mu^-$  emerge in opposite directions in the centre-of-mass frame, so that  $\theta_+ = \pi - \theta_-$ , in an obvious notation, implying  $\cos^2\theta_+ = \cos^2\theta_- = \cos^2\theta$ .



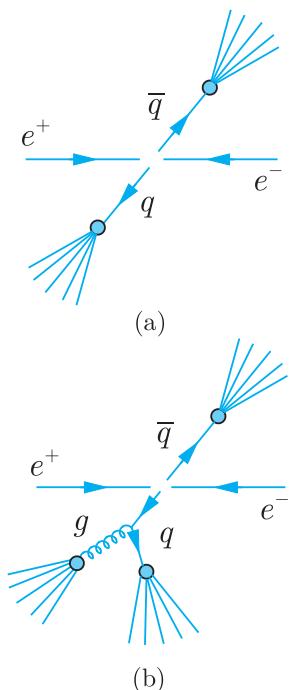
**Figure 7.16** Comparison of the angular distribution of two-jet events observed in an electron–positron colliding-beam experiment by the CELLO Collaboration at DESY, and the theoretically predicted  $(1 + \cos^2 \theta)$  behaviour (7.17). (Behrend *et al.* 1987. Reproduced with permission from Elsevier.)

## 7.2.2 Three-jet events

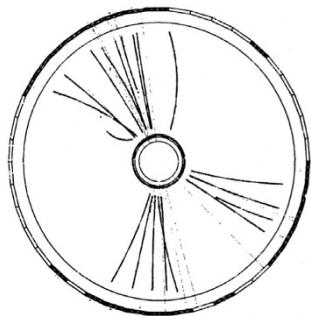
The dominant process in  $e^+e^-$  annihilation to hadrons at high energies is the formation of ‘back-to-back’ jets in the centre-of-mass frame, as discussed above and illustrated in Figure 7.17(a). However, sometimes we might expect a high-momentum gluon to be emitted at a wide angle by the quark or antiquark before fragmentation occurs, leading to the formation of a three-jet event as illustrated in Figure 7.17(b).

Such events are indeed observed, as shown in Figure 7.18, and provide strong evidence for the spin-1 gluons whose existence is assumed in QCD. The analysis is not straightforward because it is not obvious which of the jets in Figure 7.18 is due to the gluon. Instead one focuses on the total energy of each jet. The jets are classified in the order  $E_1 > E_2 > E_3$  in the overall centre-of-mass frame, before transforming to the mutual centre-of-mass frame of jets 2 and 3. The angular distribution of jet 1 is then plotted with respect to the angle  $\phi$  between its direction and the mutual line-of-flight of jets 2 and 3. The latter is almost certainly coincident with the line-of-flight of the gluon, since from Figure 7.17(b) we would expect the highest-energy jet to be associated with the quark or antiquark that does not emit the gluon, and detailed calculations show that the distribution described is sensitive to the gluon spin. The measured distribution is compared with the theoretical expectations for spin-0 and spin-1 gluons in Figure 7.19, and clearly favours the latter.

The probability that a quark will emit a gluon is determined by the strong coupling constant  $\alpha_s$  in the same way that the probability that an electron will emit a photon is determined by the fine structure constant  $\alpha$ . Hence the observed rates of production of three or more jets in electron–positron annihilation can be used to determine the strong coupling constant  $\alpha_s$  and some of the experimental values shown in Figure 7.3 were obtained in this way.



**Figure 7.17** Schematic diagrams representing (a) two-jet and (b) three-jet formations in electron–positron annihilation in the centre-of-mass frame.



**Figure 7.18** Computer reconstruction of a three-jet event observed by the JADE Collaboration at DESY. The curved tracks correspond to charged particles and the dotted lines to neutral particles whose trajectories are unaffected by the magnetic field. (Wu 1984. Reproduced with permission from Elsevier.)

### 7.2.3 The total cross-section

The measured values of the total cross-section for electron–positron annihilation to hadrons are shown in Figure 7.20 in terms of the ratio

$$R \equiv \frac{\sigma(e^+e^- \rightarrow \text{hadrons})}{\sigma(e^+e^- \rightarrow \mu^+\mu^-)}, \quad (7.18)$$

where the total cross-section for muon production

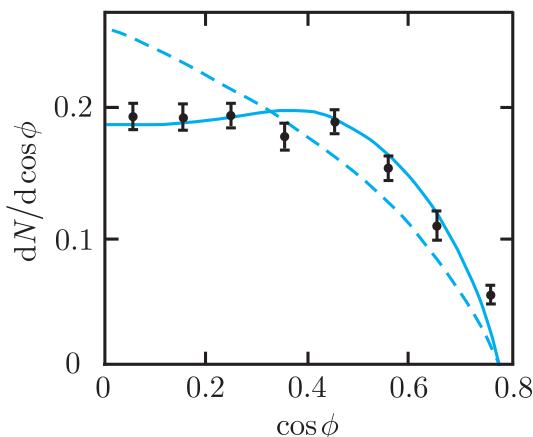
$$\sigma(e^+e^- \rightarrow \mu^+\mu^-) = \frac{\pi\alpha^2}{3E^2} \quad (7.19)$$

follows from integrating the differential cross-section (7.15) over all angles  $\theta$ . The near constancy of this ratio follows from the dominance of the two-step mechanism of Figure 7.14, with the total annihilation rate being determined by that of the initial reaction (7.12), while the value of the ratio directly confirms the existence of three colour states, each with the same electric charge, for a given quark flavour.

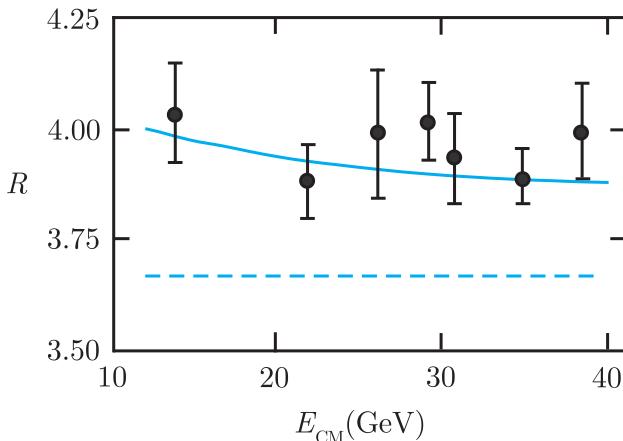
To understand this, let us suppose that each quark flavour  $a = u, d, s, \dots$  exists in  $N_C$  colour states so that  $N_C = 3$  according to QCD, while  $N_C = 1$  if the colour degree of freedom does not exist. Since the different colour states all have the same electric charge, they will all be produced equally readily by the mechanism of Figure 7.14, and the cross-section for producing quark pairs of any given flavour  $a = u, d, s, \dots$  will be proportional to the number of colours  $N_C$ . The differential cross-section for three colours is given by (7.16), and on multiplying by  $N_C/3$  and integrating over all angles, we obtain

$$\sigma(e^+e^- \rightarrow q\bar{q}) = N_C e_a^2 \sigma(e^+e^- \rightarrow \mu^+\mu^-) \quad (7.20)$$

for the total cross-section, where the muon pair production cross-section is given by (7.19). Hence if hadron production were



**Figure 7.19** Angular distribution of three-jet events measured in the TASSO Collaboration, where  $\phi$  is the angle defined in the text. The dashed and solid lines show the theoretical expectations for spin-0 and spin-1 gluons respectively. (Adapted from Brandelik 1980 and Wu 1984. Reproduced with permission from Elsevier.)



**Figure 7.20** Comparison between the measured values of the cross-section ratio  $R$ , defined in (7.18), and the theoretical prediction (7.22) for three colours,  $N_C = 3$ . The dashed line shows the corresponding prediction (7.21), omitting small contributions of order  $\alpha_s$ . (Data from the compilation of Wu (1984) and Behrend (1987).)

completely dominated by the two-step process of Figure 7.14, we would have

$$R = R_0 \equiv N_C(e_u^2 + e_d^2 + e_s^2 + e_c^2 + e_b^2) = 11N_C/9, \quad (7.21)$$

because quark–antiquark pairs of five flavours can be produced at the energies we are considering. The small contribution from the three-jet events in Figure 7.17(b) is more difficult to calculate, but when it and other corrections of order  $\alpha_s$  are taken into account, (7.21) is modified to

$$R = R_0(1 + \alpha_s/\pi), \quad (7.22)$$

where  $\alpha_s$  is given by (7.6) evaluated at  $\mu^2$  equal to the centre-of-mass energy squared. Although these corrections of order  $\alpha_s$  are very small compared with the dominant contribution (7.21), they must be included if the most precise experimental data on  $R$ , which have errors of order 2–3% are to be accounted for. As can be seen from Figure 7.19, these data are in excellent agreement with the theoretical prediction (7.22) for the value  $N_C = 3$ , in accord with the basic assumption of QCD.

## Problems 7

- 7.1 Draw the lowest-order Feynman diagrams that contribute to the processes:
  - (a)  $g + q \rightarrow g + q$ ,
  - (b)  $q + \bar{q} \rightarrow g + g$ ,
  - (c)  $q + \bar{q} \rightarrow g + \gamma$ .
- 7.2 Make a rough estimate of the relative magnitude of the static potential between two quarks separated by distances of 0.1 fm and 0.001 fm.
- 7.3 Estimate the relative magnitudes of the static potential between (a) a quark–antiquark pair and (b) two protons, both separated by a distance of 3 fm, given that the value of the dimensionless Yukawa

coupling (1.31) for the  $p\bar{p}\pi^0$  vertex is about 14. Comment briefly on your answer in the light of QCD.

- 7.4** Would you expect the widths of three chi states listed in Table 6.14 to be bigger, smaller or about the same as the widths of the  $J/\psi(3097)$  and  $\psi(3686)$ ? Check your answer with the experimental widths given in the PDG tables.
- 7.5** Two identical particles can scatter elastically by both single photon and single  $Z^0$  exchange. Calculate the typical squared energy-momentum transfer  $q^2$  between them for a distance of closest approach of (a) 1 fm and (b)  $10^{-3}$  fm in their centre-of-mass frame. Assuming that the intrinsic strengths of the fundamental weak and electromagnetic interactions are approximately equal, compare the relative sizes of the invariant (scattering) amplitudes for  $Z^0$  and photon exchange processes at these two energy-momentum transfers.
- 7.6** Show that mesons with the quantum numbers (7.8) are forbidden in the simple quark model.
- 7.7** Table 5.3 lists the allowed combinations of strangeness  $S$  and isospin  $I$  for mesons composed of light quarks  $u, d, s$  in the simple quark model. What are the allowed combinations for tetraquarks composed of light quarks? What are the allowed values of the electric charge  $Q$  for tetraquarks with strangeness  $S = 0, 1, 2$ ?
- 7.8** Draw two quark diagrams corresponding to the decay sequence (7.10b), given that the  $b$  quark couples to lighter quarks almost exclusively via the vertex  $bcW$  analogous to the  $udW$  vertex of Figure 3.1.
- 7.9** Draw quark diagrams for the decay sequence (7.10c), assuming  $P_c^+$  is the charmonium–pentaquark state  $P_c^+ = uudcc\bar{c}$ .
- 7.10** Estimate the cross-section ratio  $R$ , defined by (7.18), at total centre-of-mass energies  $E_{CM} = 2.8, 5$  and  $15$  GeV, assuming that there are no resonance peaks at these energies and that (7.22) remains approximately valid even at energies as low as 2.8 GeV. How would you expect  $R$  to change when the energy becomes large enough to produce  $t\bar{t}$  pairs, where  $t$  is the top quark?

# 8

---

## Quarks and partons

---

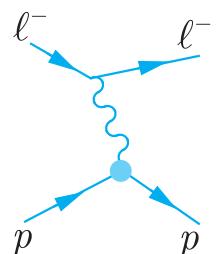
In Chapter 6, we have seen how the static properties of hadrons can be understood in terms of their quark structures. In this chapter we will consider the dynamical effects of this substructure as seen in scattering experiments. Such effects are most directly seen in high-energy inelastic scattering experiments in which beams of leptons, which are themselves structureless, are used to probe the structures of the proton and neutron. This leads to a detailed picture of the internal structure of nucleons at high resolution, called the quark–parton model, which is an essential tool in understanding many other reactions. However, before studying inelastic lepton scattering it is useful to first understand the corresponding elastic process.

### 8.1 Elastic electron scattering: the size of the proton

Electron and muon scattering have played a crucial role in exploring the structure of hadrons. Elastic scattering led to the first measurements of the size of the proton, while inelastic scattering experiments gave the first clear evidence for scattering from individual quarks within the nucleon. Here we shall discuss the elastic scattering reactions

$$\ell^- + p \rightarrow \ell^- + p \quad (\ell = e, \mu), \quad (8.1)$$

where the same arguments apply to electrons and muons, since they have identical electromagnetic interactions. In particular, both reactions (8.1) are dominated by the single-photon exchange mechanisms of Figure 8.1, where the ‘blob’ indicates an interaction whose nature we seek to explore. In this section we shall consider only these single-photon exchange contributions, neglecting completely the small corrections of higher order in  $\alpha$  arising from multiphoton exchange.



**Figure 8.1** Dominant one-photon exchange mechanism for elastic lepton–proton scattering, where  $\ell = e$  or  $\mu$ .

### 8.1.1 Static charge distributions

Before considering high-energy electron–proton scattering, it is instructive to first consider non-relativistic scattering from static charge distributions. The starting point for this is the *Rutherford cross-section* formula

$$\left(\frac{d\sigma}{d\Omega}\right)_R = \frac{m^2\alpha^2}{4p^4 \sin^4(\theta/2)} \quad (8.2)$$

for scattering of an electron of mass  $m$  and momentum  $p$  through an angle  $\theta$  by a static point charge  $e$ . This formula is derived in Section \*8.1.3, where we also show how it is modified if the same total charge is spread out in a spherically symmetric density distribution  $e\rho(r)$ , where

$$\int \rho(r) d^3r = 1 \quad (8.3)$$

and  $r = |\mathbf{r}|$  as usual. In this case the cross-section (8.2) is replaced by

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega}\right)_R G_E^2(q^2), \quad (8.4)$$

where the *electric form factor*

$$G_E(q^2) = \int \rho(r) e^{i\mathbf{q}\cdot\mathbf{r}} d^3r \quad (8.5)$$

is the Fourier transform of the charge distribution with respect to the momentum transfer

$$\mathbf{q} = \mathbf{p} - \mathbf{p}', \quad (8.6)$$

and  $\mathbf{p}$  and  $\mathbf{p}'$  are the initial and final electron momenta.<sup>1</sup> For  $\mathbf{q} = \mathbf{0}$ , (8.5) reduces to the integral of the density  $\rho(r)$  so that

$$G_E(0) = 1 \quad (8.7)$$

with our normalisation (8.3). On the other hand, for very large momentum transfers

$$G_E(q^2) \rightarrow 0 \quad \text{as} \quad q^2 \rightarrow \infty, \quad (8.8)$$

because of the increasingly rapid oscillation of the exponential factor in (8.5).

Clearly, measurements of the cross-section (8.4) determine the form factor, and hence the Fourier transform of the charge distribution that has caused the scattering. In particular, it follows from

---

<sup>1</sup> Despite appearances, the right-hand side of (8.5) is a function of  $q^2 = |\mathbf{q}^2|$  only, as can be seen by explicitly integrating over the angles (cf. Problem 8.1(a)).

(8.5) that the root-mean-square (rms) charge radius  $r_E$  is given by (see Problem 8.1(b))

$$r_E^2 \equiv \overline{r^2} = \int r^2 \rho(r) d^3 \mathbf{r} = -6 \frac{dG_E(q^2)}{dq^2} \Big|_{q^2=0}. \quad (8.9)$$

### 8.1.2 Proton form factors

In discussing electron scattering from the proton, several other factors must be taken into account. In addition to the electric form factor  $G_E$  associated with the charge distribution, there is also a *magnetic form factor*  $G_M$  associated with the magnetic moment distribution within the proton. For momentum transfers that are much smaller than the proton mass  $M$ , the recoil energy of the proton is negligible, and the electric form factor is again given by (8.5), with an analogous expression for the magnetic form factor  $G_M$ . However, at high momentum transfers the recoil energy of the proton cannot be neglected, so that the initial and final electron energies are no longer equal. In this regime, it can be shown that the form factors are functions of the Lorentz-invariant generalisation of the momentum transfer

$$Q^2 \equiv (\mathbf{p} - \mathbf{p}')^2 = (E - E')^2, \quad (8.10)$$

where the initial and final electrons have momenta and energies  $(E, \mathbf{p})$  and  $(E', \mathbf{p}')$ , respectively, and the interpretation of the form factor in terms of static charge and magnetic moment distributions breaks down. It is, however, regained in the limit of low momentum transfers, when  $Q^2$  reduces to our previous variable  $q^2$  and, in particular, the expression (8.9) for the rms radius of the proton remains valid. Finally, in addition to recoil effects and the magnetic form factor of the proton, we must also take account of the magnetic moment of the electron, which is assumed to be a point particle, without internal structure, throughout the analysis. When all these factors are taken into account, the cross-section formula becomes quite complicated, and we shall simply quote it in the limit that the electron mass can be neglected compared with its energy. The differential cross-section is then given by

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E^2 \sin^4(\theta/2)} \left( \frac{E'}{E} \right) [G_1(Q^2) \cos^2(\theta/2) + 2\tau G_2(Q^2) \sin^2(\theta/2)]. \quad (8.11)$$

Here

$$G_1(Q^2) = \frac{G_E^2 + \tau G_M^2}{1 + \tau}, \quad G_2(Q^2) = G_M^2 \quad (8.12a)$$

with

$$\tau = Q^2/4M^2, \quad (8.12b)$$

and the form factors are normalised so that

$$G_E(0) = 1 \quad \text{and} \quad G_M(0) = \mu_p, \quad (8.13)$$

where  $\mu_p = 2.79$  is the proton magnetic moment in nuclear magnetons. The same formulas hold for muon elastic scattering in the limit that the muon mass can be neglected compared with its energy. The values of the form factors  $G_E$  and  $G_M$  have been measured in many muon and electron scattering experiments, starting with the pioneering work of Hofstadter and co-workers in 1956, which gave the first measurement of the finite size of the proton. The results are conveniently divided into three  $Q^2$  regions.

At low  $Q^2$ , the contributions of the electric form factor  $G_E$  dominate the cross-section (8.11), since the contributions from the magnetic form factor are suppressed by a factor  $\tau \ll 1$ . Hence  $G_E$  can be precisely measured, enabling the size of the proton – or more correctly its rms charge radius – to be determined by (8.9). The average value from different analyses is

$$r_E = 0.8775 \pm 0.0051 \text{ fm}. \quad (8.14)$$

In the intermediate range  $0.02 \geq Q^2 \leq 3 \text{ GeV}^2$ , both form factors make measurable contributions, which can be separated by varying the scattering angle  $\theta$ , as can be seen from (8.11). The results in this region can be summarised to within a few percent by the approximate ‘dipole fit’

$$G_E(Q^2) \simeq \frac{G_M(Q^2)}{\mu_p} \simeq G_D \equiv \left( \frac{0.71}{0.71 + Q^2} \right)^2, \quad (8.15)$$

where  $Q^2$  is measured in  $\text{GeV}^2$ . Finally, for  $Q^2 > 3 \text{ GeV}^2$ , the contributions of the electric form factor become negligible, and only the magnetic form factor  $G_M$  can be measured with reasonable accuracy.

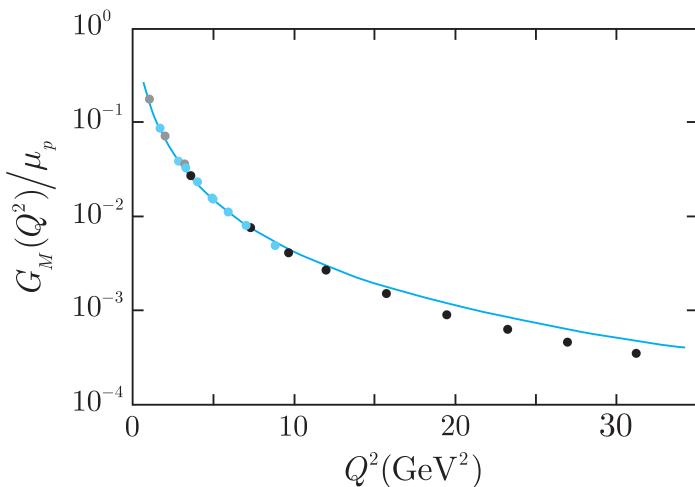
The resulting values for  $G_M(Q^2)$  are compared with the predictions of the dipole fit (8.15) in Figure 8.2. As can be seen, the form factor continues to fall rapidly with  $Q^2$ , decreasing by three orders of magnitude over the  $Q^2$  range for which it has been measured. This is in marked contrast to the behaviour of a hypothetical point proton, with a Dirac magnetic moment, for which the form factors reduce to<sup>2</sup>

$$G_E(Q^2) = 1 \quad \text{and} \quad G_M(Q^2) = 1 \quad (8.16)$$

independent of  $Q^2$ .

---

<sup>2</sup> These results follow directly from (8.5) and the analogous relation for  $G_M$  in the non-relativistic limit. They continue to hold relativistically.



**Figure 8.2** Measured values of the proton magnetic form factor and the ‘dipole fit’  $G_D$  from (8.15). The data points are as follows: grey, Walker *et al.* (1994); blue, Bosted *et al.* (1992); black, Sill *et al.* (1993). The errors are too small to be shown.

### \*8.1.3 The basic cross-section formulas

It remains to give the derivations for the non-relativistic cross-sections that were the starting point for our discussion. These are the Rutherford cross-section (8.2) for scattering from a Coulomb potential

$$V_C(\mathbf{r}) = -\frac{e^2}{4\pi\varepsilon_0} \frac{1}{|\mathbf{r}|} = -\frac{\alpha}{r} \quad (8.17)$$

and the modified cross-section (8.4) for scattering from the electrostatic potential

$$V_\rho(\mathbf{r}) = -\frac{e^2}{4\pi\varepsilon_0} \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r}' \quad (8.18)$$

due to a spherically symmetric charge distribution

$$\rho(\mathbf{r}) \equiv \rho(r). \quad (8.19)$$

These formulas follow directly from the Born approximation<sup>3</sup>

$$\frac{d\sigma}{d\Omega} = \frac{m^2}{4\pi^2} |\mathcal{M}(\mathbf{q})|^2 \quad (8.20a)$$

for scattering from an arbitrary potential  $V(\mathbf{r})$  in lowest-order perturbation theory,<sup>4</sup> where

$$\mathcal{M}(\mathbf{q}) = \int V(\mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r}) d^3\mathbf{r} \quad (8.20b)$$

<sup>3</sup> For a discussion see, for example, pp. 397–399 of Mandl (1992).

<sup>4</sup> This gives cross-sections of order  $\alpha^2$ , corresponding to the lowest-order diagram of Figure 8.1. Higher orders in perturbation theory give additional powers of  $\alpha$ , corresponding to higher-order diagrams, which we are neglecting (cf. the discussion for scattering from a Yukawa potential in Section 1.4.2).

and  $\mathbf{q}$  is the momentum transfer. Unfortunately the integral (8.20b) is ill-defined for a pure Coulomb potential, because it diverges at large  $r$ . However, charges in nature are in practice always screened at large distances by intervening matter, and it is reasonable to regard the pure Coulomb potential as a screened potential in the limit that the screening is removed. That is, we interpret

$$\mathcal{M}_C(\mathbf{q}) = \int \left( -\frac{\alpha}{r} \right) \exp(i\mathbf{q} \cdot \mathbf{r}) d^3\mathbf{r}$$

as

$$\mathcal{M}_C(\mathbf{q}) = \lim_{\mu \rightarrow 0} \int \left( -\frac{\alpha e^{-\mu r}}{r} \right) \exp(i\mathbf{q} \cdot \mathbf{r}) d^3\mathbf{r}, \quad (8.21a)$$

which gives

$$\mathcal{M}_C(\mathbf{q}) = -\frac{4\pi\alpha}{q^2} \quad (8.21b)$$

on evaluating the integral (cf. Problem 8.2). Finally, we note that the magnitude of the electron momentum is unchanged by scattering from a static potential, so that  $|\mathbf{p}| = |\mathbf{p}'| = p$ ,  $\mathbf{p} \cdot \mathbf{p}' = p^2 \cos \theta$  and

$$q^2 = (\mathbf{p} - \mathbf{p}')^2 = 4p^2 \sin^2(\theta/2). \quad (8.22)$$

The Rutherford cross-section formula (8.2) then follows on substituting (8.21) and (8.22) into (8.20a).

The corresponding cross-section for scattering from a static charge distribution is obtained by substituting the potential (8.18) into (8.20b). This gives

$$\begin{aligned} \mathcal{M}_\rho(\mathbf{q}) &= \int V_\rho(\mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r}) d^3\mathbf{r} \\ &= -\frac{e^2}{4\pi\epsilon_0} \int \rho(\mathbf{r}') \left\{ \int \frac{e^{i\mathbf{q} \cdot \mathbf{r}}}{|\mathbf{x} - \mathbf{x}'|} d^3\mathbf{r} \right\} d^3\mathbf{r}', \end{aligned} \quad (8.23)$$

where we have changed the order of integration in the double integral. Changing variables from  $\mathbf{x}$  to  $\mathbf{x} - \mathbf{x}'$  and using (8.21) then gives

$$\mathcal{M}_\rho(\mathbf{q}) = \mathcal{M}_C(\mathbf{q}) G_E(\mathbf{q}), \quad (8.24)$$

where

$$G_E(\mathbf{q}) \equiv \int \rho(\mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r}) d^3\mathbf{r}. \quad (8.25)$$

For a spherically symmetric charge distribution,  $G_E(\mathbf{q})$  reduces to a function of  $q^2$  (see Problem 8.1(a)), and the desired cross-section formula (8.4) follows on substituting (8.24) into (8.20a) and using (8.21) and (8.22).

## 8.2 Inelastic electron and muon scattering

The study of inelastic electron scattering at high energies is of great historical importance because it led to the first clear evidence for

scattering from individual quark constituents confined within the proton. Here we shall consider both electron and muon scattering reactions

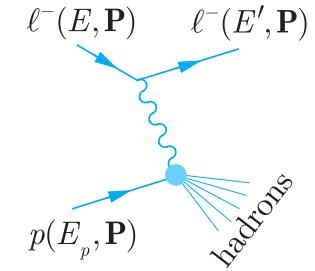
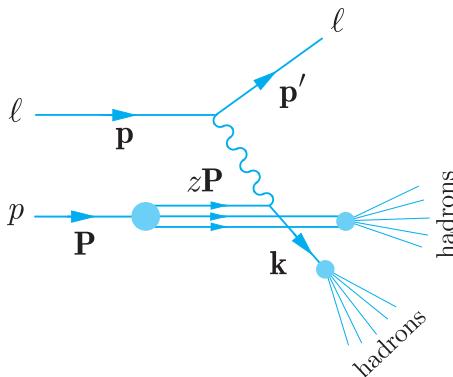
$$\ell^- + p \rightarrow \ell^- + X \quad (\ell = e, \mu), \quad (8.26)$$

where  $X$  as usual denotes any set of hadrons allowed by the conservation laws.<sup>5</sup> These reactions are dominated by the one-photon exchange mechanism of Figure 8.3, where the energies and momenta of the leptons and proton are also defined. This is a direct generalisation of Figure 8.1 for elastic scattering  $X \equiv p$ , and as in our discussion of elastic scattering, we shall consider only these contributions, neglecting completely the small corrections arising from multi-photon exchanges.

In Figure 8.3 the exchanged photon acts as a probe of the proton structure and to resolve individual quarks within it requires large momentum transfers  $\mathbf{q} = \mathbf{p} - \mathbf{p}'$ , corresponding to photon wavelengths that are very small compared with the size of the proton. In this region, the process is referred to as *deep inelastic scattering* and is dominated by processes like that shown in Figure 8.4, that is by two-step processes in which the first step is elastic lepton scattering

$$\ell^- + q \rightarrow \ell^- + q \quad (\ell = e, \mu) \quad (8.27)$$

from one of the quarks contained in the proton, while the second is a fragmentation process that converts the recoil quark and the remaining constituents of the proton into hadrons. In the experiments we shall discuss, the final state hadrons are not observed, and we shall not need to consider this fragmentation process further. Instead we shall concentrate on the energy and angular distributions of the recoil leptons, which directly reflect the properties of the quarks from which they scatter.



**Figure 8.3** Dominant one-photon exchange mechanism for inelastic lepton–proton scattering, where  $\ell = e$  or  $\mu$ .

**Figure 8.4** Dominant contribution to deep inelastic lepton–proton scattering in the quark model, where  $\ell = e$  or  $\mu$ .

<sup>5</sup> The corresponding neutrino scattering reactions are also of considerable importance and will be discussed in Section 8.3.

In what follows, we firstly describe the striking features exhibited by the data and then show how these are explained by the mechanism of Figure 8.4. Before doing that, however, it is convenient to introduce two variables that will prove useful in the discussion. These are the variable  $\nu$  defined by

$$2M\nu \equiv W^2 + Q^2 - M^2 \quad (8.28)$$

and the dimensionless variable

$$x \equiv Q^2 / 2M\nu. \quad (8.29)$$

Here,  $M$  is the proton mass,  $W$  is the invariant mass of the hadron state  $X$  and  $Q^2$  is again the squared energy-momentum transfer (8.10). In the rest frame of the proton, the variable  $\nu$  reduces to (see Problem 8.3)

$$\nu = E - E' \quad (8.30)$$

so that it can be regarded as a Lorentz-invariant generalisation for the energy transferred from the lepton to the proton. The physical interpretation of the variable  $x$ , which is called the *scaling variable*, will become clear later.

### 8.2.1 Bjorken scaling

In a typical high-energy lepton scattering experiment, the hadron state is left unidentified and the scattered lepton momentum, and hence its energy, are measured in a spectrometer. The measured quantities are therefore the initial and final lepton energies  $E$  and  $E'$  and the lepton scattering angle  $\theta$ . For elastic scattering the invariant mass  $W$  of the hadrons is just the proton mass, and for a given beam energy  $E$ , the final energy  $E'$  is determined by the scattering angle  $\theta$  (see Problem 8.3). However, in inelastic scattering  $W$  can vary continuously, depending on the energies and momenta of the produced hadrons, so that both  $E'$  and  $\theta$  are independent variables for a given initial energy  $E$ . We therefore have to consider a differential cross-section for scattering to an energy  $E'$  and angle  $\theta$ , and the appropriate generalisation of the elastic scattering formula (8.11) is

$$\frac{d\sigma}{dE'd\Omega'} = \frac{\alpha^2}{4E^2 \sin^4(\theta/2)} \frac{1}{\nu} \left[ \cos^2(\theta/2) F_2(x, Q^2) + \sin^2(\theta/2) \frac{Q^2}{xM^2} F_1(x, Q^2) \right]. \quad (8.31)$$

Here the dimensionless *structure functions*  $F_1(x, Q^2)$  and  $F_2(x, Q^2)$  parameterise the interaction at the lower vertex in Figure 8.3, in the same way that the dimensionless functions  $G_1(Q^2)$  and  $G_2(Q^2)$  in (8.11) parameterise the interaction at the lower vertex in elastic scattering in Figure 8.1. We have written them as functions of the

scaling variable  $x$  and  $Q^2$ , rather than the energy transfer  $\nu$  and  $Q^2$ , because for fixed  $x$  and large  $Q^2$  the measured values of  $F_1$  and  $F_2$  become approximately independent of  $Q^2$ , that is

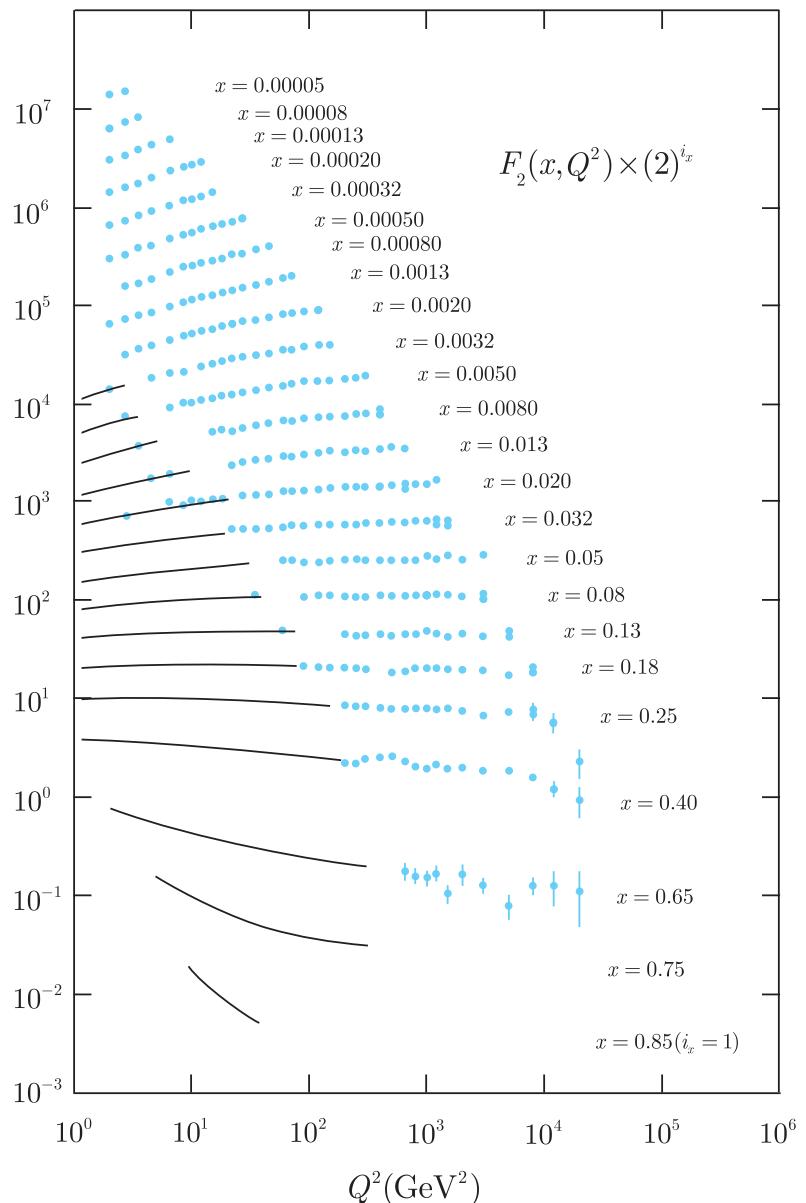
$$F_{1,2}(x, Q^2) = F_{1,2}(x) \quad (Q^2 \gg M^2) \quad (8.32)$$

is a good approximation over quite large ranges of  $Q^2$ . This behaviour is called *Bjorken scaling* or *scale invariance* because (8.32) implies that the structure functions are left unchanged by a scale transformation, that is by a transformation in which all particle masses, energies and momenta are multiplied by a scale factor  $\kappa$ , so that  $Q^2 \rightarrow \kappa^2 Q^2$  but  $x$  remains unchanged. Scale invariance is not exact as we have said; deviations from it are called *scaling violations*.

These features can be clearly seen in Figure 8.5, which shows the measured  $Q^2$  dependence of the proton structure function  $F_2$  for  $Q^2 > 1 \text{ GeV}^2$  at various fixed values of  $x$ . This figure should be treated with care, because to fit all the data shown on one graph, log scales have been used and at each fixed  $x$  the data have been multiplied by a scale factor that varies from 2 at  $x = 0.85$  to  $2^{24} \approx 1.7 \times 10^7$  at the smallest value of  $x$ . In contrast to the analogous function  $G_2 \equiv G_M^2$  for elastic scattering, which varies by about five orders of magnitude over the much smaller range  $1 < Q^2 < 30 \text{ GeV}^2$ , the curves in Figure 8.5 are slowly varying, that is there is approximate scale invariance. On the other hand, deviations from exact scaling are clearly seen and are largest at very small and very large  $x$ .

Historically, the approximate scaling behaviour (8.32) was first observed in inelastic scattering experiments at the Stanford Linear Accelerator Center (SLAC) in 1969, and caused considerable excitement, not least because of the contrast with the elastic form factors. However, if the proton were a point particle, the elastic form factors would also be independent of  $Q^2$ , as can be seen from (8.16), and it was soon realised that the observed scaling behaviour of inelastic scattering could be successfully accounted for if the scattering was from charged point-like constituents within the proton, rather than from the proton as a whole, as we shall discuss below. This was the first dynamical evidence for quarks, whose existence had previously been inferred solely on the basis of static quantities, like the masses and quantum numbers of the hadrons discussed in the previous two chapters. Somewhat later, the development of QCD led to the prediction of small deviations from scaling. These scale violations had not been observed at the time because of the limited accuracy and range of the data that were then available, and their subsequent detection with the predicted form was in turn one of the earliest triumphs of QCD.

In what follows we shall concentrate initially on understanding the phenomenon of scale invariance, deferring the discussion of the deviations from it to Section 8.2.3.



**Figure 8.5** Measured values  $2^{i_x} \times F_2(x, Q^2)$  for  $Q^2 > 1 \text{ GeV}^2$ , where  $i_x = 1, 2, \dots, 24$  is an integer labelling the various  $x$ -values in turn starting from  $x = 0.85$ . The blue points are data obtained using the HERA electron–proton collider at DESY, Hamburg, while the curves indicate the behaviour obtained in several earlier experiments using lepton scattering from a fixed proton target. (Adapted from the review of Foster *et al.*, 2013).

### 8.2.2 The parton model

The approximate scaling behaviour (8.32) was quickly understood, as we have said, in terms of scattering from individual charged point-like constituents within the proton. These were called *partons*, because the existence of quarks, with which they are now identified, was still controversial at the time. Correspondingly, the simplest model that explains the main features of the data is called the *parton model*. In discussing it we shall not initially assume that the partons are

quarks, in order to see how far this identification can be justified by experimental data.

The parton model is defined by assuming the mechanism of Figure 8.4 and completely neglecting the corrections to it that were subsequently predicted by QCD. It takes its simplest form in a reference frame in which the target proton has a very large momentum. In such a frame, the momenta of the constituents will be almost collinear with the proton momentum, so that to a good approximation the target can be viewed as a stream of partons, each carrying a fraction of the proton momentum  $P$ . Furthermore, for  $Q^2 \gg M^2$  it can be shown (see Problem 8.5) that the fraction  $z$  of the proton momentum carried by the struck parton in Figure 8.4 is given by

$$z = x = Q^2/2M\nu. \quad (8.33)$$

This is a key result in interpreting the data. It implies that the measured cross-section at a given value of  $x = Q^2/2M\nu$  is proportional to the probability of finding a parton with a fraction  $z = x$  of the proton momentum. The full calculation is more complicated and we will simply quote below the results obtained for the structure functions in the parton model.<sup>6</sup> We will do this for both spin-0 and spin-1/2 partons with arbitrary electric charges in order to see which spin and charge values, if any, are selected by the data.

We start with the structure function  $F_2$ , which is given for both spin-0 and spin-1/2 partons by

$$F_2(x, Q^2) = \sum_a e_a^2 x f_a(x), \quad (8.34)$$

where  $f_a(x) dx$  is the probability of finding a parton  $a$  of charge  $e_a$  with fractional momentum between  $x$  and  $x + dx$ , and the sum extends over all parton types  $a$ . The parton distributions  $f_a(x)$  are not determined by the model, so that the  $x$ -dependence of the structure functions must be determined by experiment. However, they are predicted to be the same at all  $Q^2$ ; that is one has exact scaling for point-like partons, irrespective of their spin. In contrast, the predictions obtained for the structure function  $F_1$  do depend on the spin of the partons. They are

$$F_1(x, Q^2) = 0 \quad (\text{spin 0}) \quad (8.35a)$$

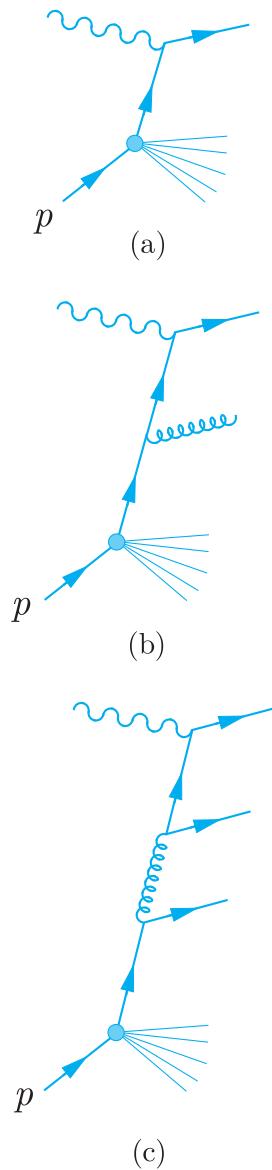
and

$$2xF_1(x, Q^2) = F_2(x, Q^2) \quad (\text{spin 1/2}), \quad (8.35b)$$

where the latter result is called the *Callan–Gross relation*. The data rule decisively in favour of relation (8.35b), as one would expect if the partons were spin-1/2 quarks.

---

<sup>6</sup> The derivation is given, for example, in pp. 191–192 of Halzen and Martin (1984).



**Figure 8.6** (a) The interaction of the exchanged photon with the struck quark in the parton model, together with (b) and (c), two of the additional processes that occur when quark–gluon interactions are taken into account.

Finally, if the partons are quarks, their squared charges  $e_a^2$  must correspond to the squared quark charges  $1/9$  and  $4/9$ . Unfortunately, this cannot be confirmed from measurements of the proton structure functions alone, because of the occurrence of the unknown probability distributions  $f_a(x)$  in (8.34). However, by combining measurements of the proton structure functions with those for neutron structure functions, and with similar results obtained in neutrino scattering, it is possible to check that the squared parton charges are consistent with those expected for quarks, as we shall see in Section 8.3.1 below. For the moment, we will simply accept this to be the case.

In summary: the scaling behaviour of  $F_2$  is evidence for point-like constituents within the proton, the Callan–Gross relation shows that they are spin-1/2 particles and the checks on their squared charges confirm their identification with quarks beyond all reasonable doubt.

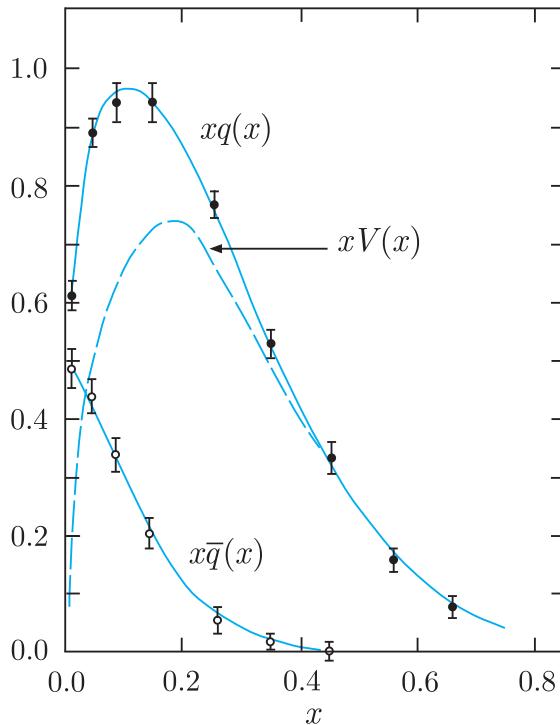
### 8.2.3 Parton distributions and scaling violations

The parton model described in Section 8.2.2 was developed before the advent of QCD and completely ignores the quark–gluon interaction. This interaction is, of course, what holds the proton together, and it manifests itself in deep inelastic scattering in three important ways. The first is that the leptons not only scatter from the three so-called *valence* quarks  $uud$ , which give rise to the internal quantum numbers

$$B = Q = 1, \quad S = C = \tilde{B} = T = 0$$

of the proton, but also from a so-called *sea* of quark–antiquark pairs  $a\bar{a}$ , which are created by the interaction. The second is that these quarks and antiquarks together carry a little less than one-half of the momentum of a high-energy proton, the rest being accounted for by gluons. Finally, the resulting structure functions become weakly dependent on  $Q^2$ , as seen experimentally in Figure 8.5. Here we shall discuss in turn how each of these features arises.

In the absence of quark–gluon interactions, the interaction of the exchanged photon with the struck quark in the nucleon may be depicted by Figure 8.6(a). When the quark–gluon interactions are taken into account, other processes besides those of Figure 8.6(a) can contribute, like those shown in Figures 8.6(b) and (c). In Figure 8.6(b) the quark emits a gluon before interacting with the exchanged photon. In Figure 8.6(c) the emitted gluon is converted to a quark–antiquark pair, which forms part of the sea of quark–antiquark pairs mentioned above, and it is one of these particles that absorbs the exchanged photon. Detailed calculations show that the parton model remains a good first approximation when such processes are taken into account, provided the parton



**Figure 8.7** Quark and antiquark distributions (8.37a), together with the valence quark distribution (8.37b), measured at a  $Q^2$  value of about  $10 \text{ GeV}^2$ , from neutrino experiments at CERN and Fermilab.

distribution functions  $f_a(x)$  are extended to include contributions from antiquarks as well as quarks. That is, (8.34) becomes

$$F_2(x, Q^2) = \sum_{a, \bar{a}} \left[ e_a^2 x f_a(x) + e_{\bar{a}}^2 x f_{\bar{a}}(x) \right], \quad (8.36)$$

where the sum extends over all quark flavours  $a = u, d, s, \dots$ . The Callan–Gross relation (8.35b) remains unchanged, since both quarks and antiquarks have spin-1/2.

Electron and muon scattering cannot separate the quark and antiquark contributions because the structure functions depend on the squared charges of the constituents, which are the same for quarks and antiquarks. However, neutrino scattering experiments can distinguish between quarks and antiquarks, and are sensitive to the same distribution functions  $f_a(x)$  and  $f_{\bar{a}}(x)$  as electron and muon scattering.<sup>7</sup> The resulting quark and antiquark distributions

$$x q(x) \equiv \sum_a x f_a(x), \quad x \bar{q}(x) \equiv \sum_{\bar{a}} x f_{\bar{a}}(x), \quad (8.37a)$$

where the sum extends over all quark flavours  $a = u, d, s, \dots$ , are shown in Figure 8.7. Also shown is the difference

$$x V(x) \equiv x q(x) - x \bar{q}(x), \quad (8.37b)$$

<sup>7</sup> Neutrino scattering is discussed briefly in the next section.

which corresponds to the valence quark distribution, if we assume that the quarks belonging to the sea of  $q\bar{q}$  pairs produced by the gluons have the same momentum distribution as the antiquarks.<sup>8</sup> As can be seen, the valence quarks dominate the distributions for  $x \geq 0.3$ , while the sea of  $q\bar{q}$  pairs contributes mainly at small  $x$  values.

The second important result follows directly from the curves in Figure 8.7. In a reference frame in which the proton has a very large momentum,  $x$  is the fraction of the proton's momentum carried by the struck parton, as we saw in the previous section. Hence the integral

$$\int_0^1 [xq(x) + x\bar{q}(x)]dx$$

is the fraction of the proton's momentum carried by the quarks and antiquarks, and a quick glance at the area under the relevant curves in Figure 8.7 is sufficient to see that this fraction is considerably less than one. The remaining momentum is carried away by gluons, like that emitted in Figure 8.6(b) and other similar diagrams. These gluons, like the quarks and antiquarks, are also referred to as partons and account for about half the proton's momentum. Since they only have strong interactions, their  $x$  distribution cannot be measured directly in lepton scattering experiments, but can be determined by other means, as we shall see.

The third important result of the quark–gluon interaction is that the effective parton distributions, which are shown in Figure 8.7 for  $Q^2$  values of about  $10 \text{ GeV}^2$ , depend on  $Q^2$ , albeit weakly. At larger  $x$  values, where the valence quarks dominate, the most important correction to the parton model diagram in Figure 8.6(a) is the process shown in Figure 8.6(b), in which the struck quark emits a gluon before interacting with the exchanged photon. Detailed calculations show that as  $Q^2$  increases, this process increases in importance,<sup>9</sup> and that the emitted gluon carries off some of the initial quark momentum.<sup>10</sup> Thus the effective quark distribution shrinks towards smaller values of the fractional momentum  $x$  as  $Q^2$  increases. This is the reason for the decrease in the structure functions with increasing  $Q^2$  seen for fixed  $x > 0.2$  in Figure 8.5. In addition, because gluons are emitted more readily as  $Q^2$  increases, the sea distribution of

<sup>8</sup> This follows from  $C$  invariance if we assume  $q\bar{q}$  pairs are produced in the basic process  $g \rightarrow q\bar{q}$ .

<sup>9</sup> This is analogous to the situation in electromagnetism. In QED one finds that the more an electron is accelerated, the more likely it is that the process will be accompanied by photon emission. Similarly in QCD, the more a quark is accelerated, the more likely it is that the process will be accompanied by gluon emission.

<sup>10</sup> See also Problem 8.6.

quark–antiquark pairs is also found to increase with  $Q^2$ , leading to a rise in the structure functions with  $Q^2$  at small  $x$  values, where the sea is important. Finally, it is clear from Figures 8.6(b) and 8.6(c) that the magnitude of these deviations from exact scaling depends on the magnitude of the strong coupling constant  $\alpha_s(Q^2)$ , and their measurement is one of the classic means of determining its value.

### 8.3 Inelastic neutrino scattering

Next we turn to inelastic neutrino reactions such as

$$\nu_\mu + p \rightarrow \mu^- + X^{++} \quad (8.38a)$$

and

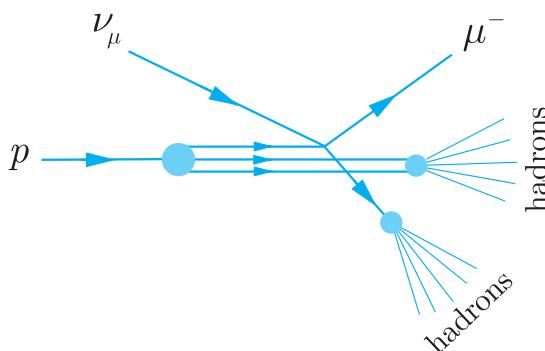
$$\nu_\mu + n \rightarrow \mu^- + X^+, \quad (8.38b)$$

where  $X^+$  and  $X^{++}$  are as usual any hadronic states allowed by the conservation laws. At high energies these reactions are well described by the parton model, and they enable the contributions of different types of partons to be separated, as we noted in the previous section. Firstly, however, we shall use the parton model to predict the energy dependence of the cross-sections for reactions (8.38) at centre-of-mass energies that are much larger than the proton mass. This leads to one of the most striking experimental tests of the parton model, as we shall immediately see.

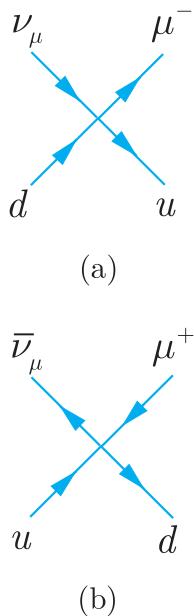
In the parton model, the reactions (8.38) are assumed to proceed via two-step processes like that shown in Figure 8.8. In the first step, the neutrino scatters from one of the quark constituents in a weak interaction like

$$\nu_\mu + d \rightarrow \mu^- + u, \quad (8.39)$$

while the second step is a fragmentation process that converts the recoil quark and the remaining constituents of the nucleons into hadrons. This mechanism is similar to that shown in Figure 8.4 for inelastic electron and muon scattering. As in that case, we shall only



**Figure 8.8** Dominant contribution to deep inelastic neutrino scattering in the parton model.



**Figure 8.9** Allowed neutrino quark scattering processes (8.39) and (8.46).

discuss experiments in which the hadrons are not identified, and we will not need to consider the fragmentation process further.

In Section 2.2.1 we saw that at low energies the weak interaction could be approximated by a zero-range interaction whose strength is characterised by the Fermi coupling constant  $G_F$  given in (2.16). In this approximation, the neutrino reaction (8.39) is represented by the point interaction of Figure 8.9(a), in the same way that inverse muon decay was represented in Figure 2.7. This zero-range approximation has been used for the neutrino scattering process (8.39) in Figure 8.8, and it remains a good approximation for the experiments we shall discuss.<sup>11</sup> The energy dependence of the cross-sections for reactions (8.38) then follows from a simple dimensional argument that rests on three points. The first is that the cross-sections are proportional to  $G_F^2$ , since the amplitudes for weak interactions like (8.39) are proportional to  $G_F$ . The second is that in the parton model these basic scattering processes only involve point-like particles (i.e. leptons and quarks) with no length scales associated with them. The third point is that, if we work in the centre-of-mass frame, it is reasonable to neglect the proton mass as well as the lepton masses, since the proton will also have a high energy in this frame. Hence the only energy scale available is the centre-of-mass energy  $E_{\text{CM}}$ , and a simple dimensional argument gives

$$\sigma_\nu = K G_F^2 E_{\text{CM}}^2 \quad (8.40)$$

for the total cross-section. Here  $K$  is a dimensionless constant and the power of energy is determined by the fact that both  $\sigma_\nu$  and  $G_F$  have dimensions  $[E^{-2}]$  in natural units (see Table 1.1). This result is conventionally rewritten in terms of the neutrino energy  $E_\nu$  in the laboratory frame, which is given by<sup>12</sup>

$$4 E_{\text{CM}}^2 = 2 M E_L + M^2 \approx 2 M E_L$$

at high energies  $E_L \gg M$ . Hence we finally obtain

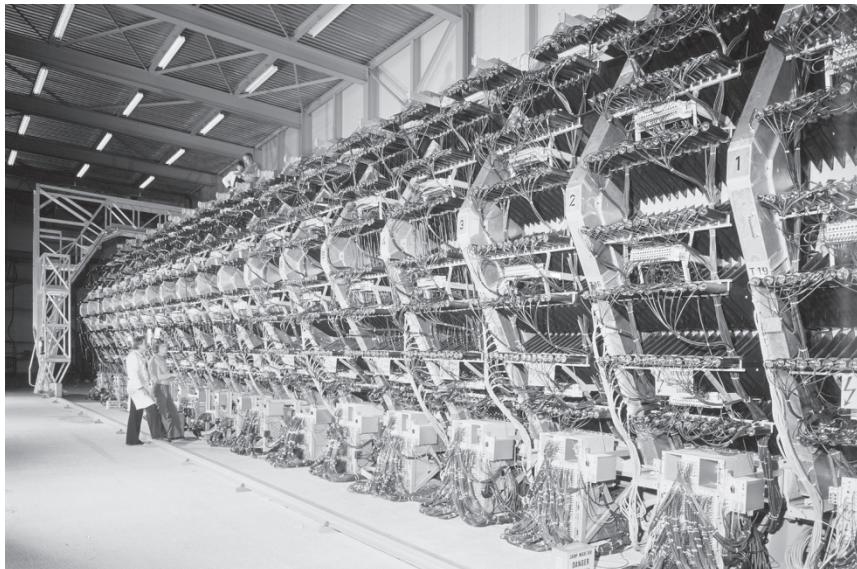
$$\sigma \propto E_L \quad (8.41)$$

for the energy dependence of the total cross-section in the parton model, and a similar result holds for antineutrinos.

In order to verify the predicted behaviour (8.41), accurate measurements of the neutrino cross-sections are required. In spite of their small size, this has been done by exploiting the intense neutrino

<sup>11</sup>It is expected to break down for total centre-of-mass energies of order 80 GeV, for reasons that will become clear in the next chapter. Here we restrict ourselves to centre-of-mass energies  $E_{\text{CM}} < 25$  GeV, which are smaller than this, but still large compared with the proton mass.

<sup>12</sup>This follows directly from Equation (A.12) of Appendix A.



**Figure 8.10** The electronic detector used in the WA1 neutrino experiment at CERN from 1976 to 1984. (Reproduced by permission of CERN.)

beams available at a number of accelerators<sup>13</sup> together with very large target/detector systems. An example of such a detector is shown in Figure 8.10.

This 20 m long, 1400 ton device was a modular assembly consisting of a succession of magnetised iron toroids interspersed with drift chambers and scintillation counters. The iron toroids served as the target and also produced the magnetic field in the detector. The muons produced in the reactions (8.38) penetrated many modules of the detector and their momenta were determined from the curvatures of their trajectories in the magnetic field. Finally, the whole detector acted as a hadron calorimeter, enabling the total energy of the unidentified hadron state  $X$  to be measured. Because the targets were iron nuclei, the experiment essentially measured the sum of the cross-sections for the reactions (8.38a) and (8.38b) for incident neutrinos, and of the cross-sections for the corresponding reactions

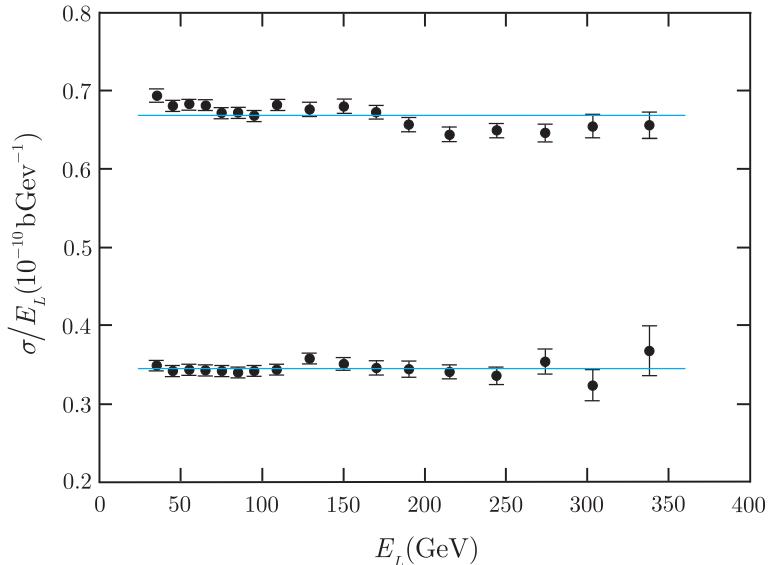
$$\bar{\nu}_\mu + p \rightarrow \mu^+ + X^0 \quad (8.42a)$$

and

$$\bar{\nu}_\mu + n \rightarrow \mu^+ + X^- \quad (8.42b)$$

for incident antineutrinos. The results for the total cross-sections are shown in Figure 8.11. As can be seen, they are in excellent agreement with the prediction of the parton model (8.41).

<sup>13</sup>The production of such beams was discussed briefly in Section 4.2.4.



**Figure 8.11** Neutrino and antineutrino total cross-sections. (Data from Seligman, 1997)

### 8.3.1 Quark identification and quark charges

The results shown in Figure 8.11 strikingly confirm the idea of scattering from point-like constituents within the nucleons. However, to extract the different quark and antiquark distributions from neutrino data, it is necessary to measure differential cross-sections analogous to those given in (8.31) for charged lepton scattering. The corresponding formulas for neutrino and antineutrino scattering are somewhat similar to (8.31), but in addition to structure functions  $F_1^{\nu N}, F_2^{\bar{\nu} N}$  analogous to  $F_1$  and  $F_2$  in (8.31), there is a third structure function  $F_3^{\nu N}$  that occurs because of the existence of parity violation in weak interactions. We will not discuss this here, but simply cite without proof the results obtained for the neutrino and antineutrino structure functions analogous to  $F_2$ . In addition, we shall confine ourselves for simplicity to the region  $x \geq 0.3$ , where the sea of quark–antiquark pairs can to a good approximation be neglected. In this region, the only non-vanishing quark distributions are

$$u(x) \equiv f_u(x) \quad \text{and} \quad d(x) \equiv f_d(x) \quad (8.43)$$

arising from the valence quarks  $uud$  within the proton, and the proton structure function (8.36) measured in electron or muon scattering reduces to

$$F_2(x) = e_u^2 x u(x) + e_d^2 x d(x) \quad (x \geq 0.3). \quad (8.44)$$

The corresponding results for neutrino and antineutrino scattering are

$$F_2^{\nu p}(x) = 2x d(x) \quad \text{and} \quad F_2^{\bar{\nu} p}(x) = 2x u(x) \quad (x \geq 0.3) \quad (8.45)$$

and can be understood, apart from their normalisation, from the conservation of electric charge, baryon number and muon number. The reaction (8.39), shown in Figure 8.9(a), and the corresponding reaction

$$\bar{\nu}_\mu + u \rightarrow \mu^+ + d, \quad (8.46)$$

shown in Figure 8.9(b), satisfy all these conservation laws. However, other hypothetical reactions like

$$\nu_\mu + u \rightarrow \mu^+ + d \quad \text{and} \quad \bar{\nu}_\mu + d \rightarrow \mu^- + u$$

that satisfy charge and baryon number conservation, violate the conservation of muon number (2.3b), and there are no  $\nu_\mu u$  or  $\bar{\nu}_\mu d$  scattering reactions that satisfy all three conservation laws. Hence neutrinos can scatter from  $d$  quarks but not from  $u$  quarks, while antineutrinos can scatter from  $u$  quarks but not from  $d$  quarks, as implied by (8.45). It is for this reason that neutrino and/or antineutrino beams are ideal tools for separating  $u$  and  $d$  quark contributions, among others.

The above discussion can be extended to small  $x$  by including the sea contributions, but we shall not pursue it further.<sup>14</sup> Instead, we shall continue to work at  $x > 0.3$  and show how the comparison of neutrino with electron or muon scattering on nuclear targets leads to a determination of the sum of quark charges ( $e_u^2 + e_d^2$ ). To do this, we first note that, according to isospin symmetry, a neutron  $n(940) = ddu$  differs from a proton  $p(938) = uud$  in that  $u$  and  $d$  quarks are interchanged. Hence the quark distributions  $u^n(x)$  and  $d^n(x)$  of the neutron are related to those of the proton by

$$u^n(x) = d^p(x) \equiv d(x) \quad \text{and} \quad d^n(x) = u^p(x) \equiv u(x)$$

to a very good approximation. The neutron structure function  $F_n^2$  measured in electron or muon scattering is then given by

$$F_2^n(x) = e_u^2 x u^n(x) + e_d^2 x d^n(x) = e_u^2 x d(x) + e_d^2 x u(x) \quad (x > 0.3) \quad (8.47)$$

in analogy to (8.44), and the corresponding neutrino structure functions are given by

$$F_2^{\nu n}(x) = 2x d^n(x) = 2x u(x) \quad \text{and} \quad F_2^{\bar{\nu} n}(x) = 2x u^n(x) = 2x d(x) \quad (8.48)$$

in analogy to (8.45). From these results, one can easily show that

$$\frac{F_2^p(x) + F_2^n(x)}{F_2^{\nu p}(x) + F_2^{\bar{\nu} n}(x)} = \frac{e_u^2 + e_d^2}{2} = \frac{5}{18} \quad (8.49)$$

---

<sup>14</sup>See, for example, Section 9.3 of Halzen and Martin (1984).

where we have substituted the values of the quark charges. This prediction can be tested by taking data on nuclei that contain equal numbers of protons and neutrons. It is verified within experimental errors of order 10%, and was historically important in confirming that the charges of the ‘charged partons’ found in deep inelastic scattering were indeed the same as the charges of the quarks with which they are now identified.

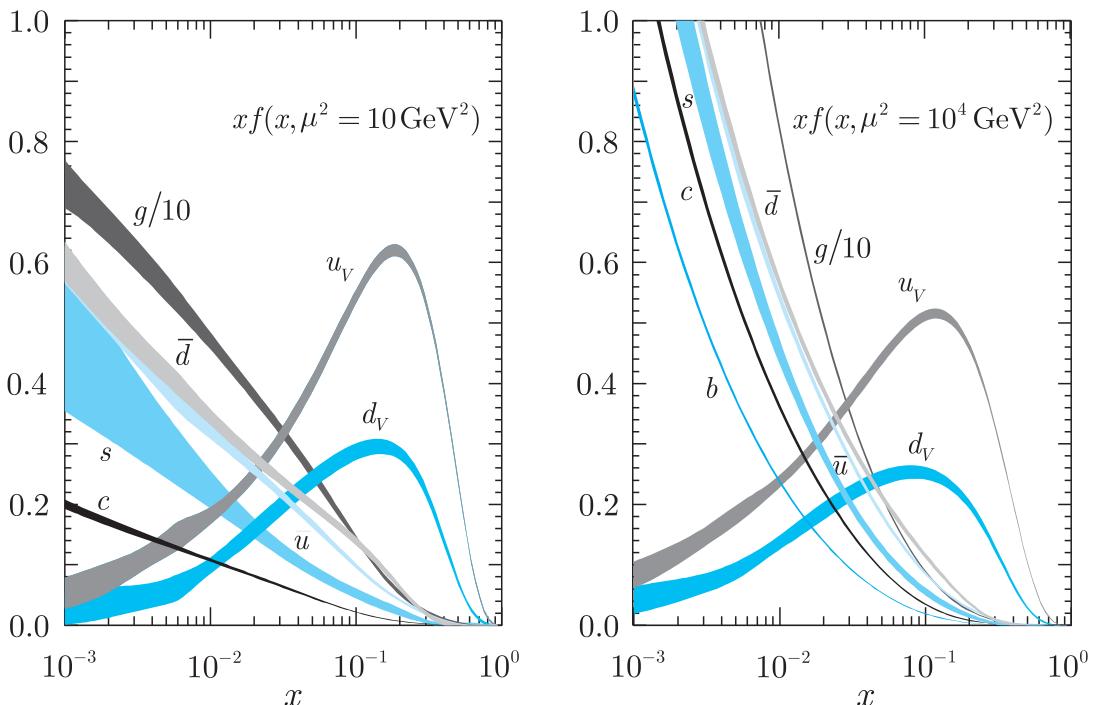
## \*8.4 Other processes

In many experiments at high-energy colliders, in which proton and/or antiproton beams are accelerated, the main interest lies in the interactions between the fundamental constituents within the protons and/or antiprotons, and in the particles or jets these interactions produce. For example, we saw in Section 4.5.1 that the  $W^\pm$  and  $Z^0$  bosons were discovered in proton–antiproton collisions via the mechanism of Figure 4.19, in which a quark from the proton annihilates with an antiquark from the antiproton to produce the bosons, which were then identified from their decay products. In order to understand the rate and other characteristics of this and many other reactions, in addition to calculating the fundamental interactions between the partons, it is necessary to know the momentum distributions of the partons within the initial hadrons. Conversely, the data from such reactions can be used to extend our knowledge of the parton distributions, and our most precise knowledge of these is based on the simultaneous analysis of many reactions,<sup>15</sup> including the deep inelastic lepton scattering processes already discussed. Recent results from one such analysis are shown in Figure 8.12 at two different mass scales  $\mu^2 = Q^2$ , using a log scale for  $x$  in order to be able to show the detailed information at very small  $x$ . Several points should be noted:

- (i) The valence distributions, which dominate at large  $x$ , vary relatively slowly as  $\mu^2$  increases.
- (ii) At small  $x$ , the antiquark distributions arising from the sea depend on the quark flavour and increase rapidly as  $\mu^2$  increases.
- (iii) The gluon distribution, which also increases rapidly as  $\mu^2$  increases, is very large at small  $x$ . Because of this, gluons play a very important, and often dominant, role in many reactions at the LHC where, for example, a parton with  $x$  as small as  $10^{-2}$  has an energy of 80 GeV.

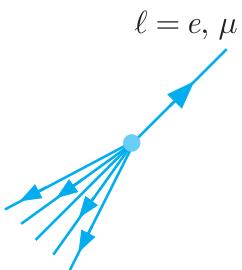
---

<sup>15</sup>The Particle Data Group lists about 25 reactions as important sources of information on parton distributions. See Table 19.1 in Olive *et al.* (2014).



**Figure 8.12** Parton distribution functions from a global analysis of data at two different mass scales  $\mu^2 = Q^2$ . (Adapted from Foster *et al.*, 2013.)

In what follows, we shall illustrate the use of parton distribution functions by considering a few important topics in proton–proton scattering. However, before doing so it will be useful to comment briefly on the jet structure in deep inelastic electron and muon scattering. As we have seen, the dominant term is the parton model diagram in Figure 8.4, in which the final state comprises two parts. When a quark is knocked out of the proton, the remaining partons continue on their way and are converted by a fragmentation process into hadrons with relatively small momenta transverse to the beam direction. This is referred to as *beam fragmentation* or sometimes *target fragmentation*. On the other hand, the quark recoils against the lepton and fragments into a jet of hadrons in the same way as the quarks and gluons produced in electron–positron annihilation, and described in Section 7.2. The global properties of this or any other jet can then be summarised by the total energy  $E$  and momentum  $\mathbf{p}$  of its constituents. In addition, it will be useful to set up Cartesian axes with associated unit vectors  $\mathbf{i}, \mathbf{j}, \mathbf{k}$  with the  $z$ -axis in the beam direction and to separate the momentum into the *longitudinal momentum*  $p_z \mathbf{k}$  and the transverse momentum  $\mathbf{p}_T = p_x \mathbf{i} + p_y \mathbf{j}$ . For example, in deep inelastic electron and muon scattering, when an event is viewed along the beam line, so that one sees the projection of the event in the  $xy$ -plane, the charged lepton and the quark jet



**Figure 8.13** Sketch of a final-state charged lepton and the associated jet produced by the mechanism of Figure 8.4, viewed along the initial lepton beam direction. The blue circle indicates the collision point.

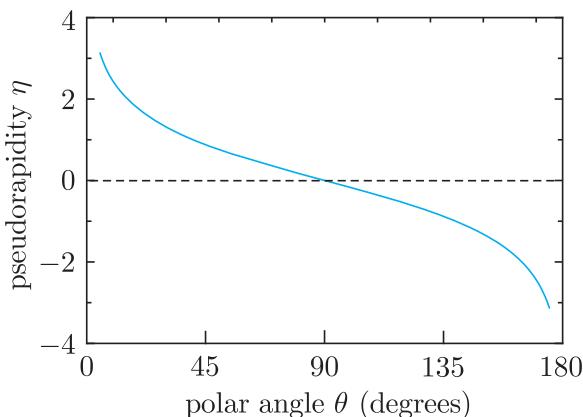
appear ‘back to back’, as illustrated in Figure 8.13. This is because the internal transverse momenta of the quarks within the proton can be neglected,<sup>16</sup> so that, by momentum conservation, the lepton and the struck quark in Figure 8.4 recoil against each other with equal and opposite transverse momenta. They do not in general emerge back-to-back in three dimensions, because they do not have equal and opposite longitudinal momenta  $p_z$ , and as we will see, it is often convenient to replace the longitudinal momentum by the *rapidity*, defined by

$$y \equiv \frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right). \quad (8.50)$$

If the jet has zero invariant mass  $M^2 = E^2 - p^2 = 0$ , then  $p_z = E \cos \theta$ , where  $\theta$  is the polar angle of the jet momentum, and the rapidity reduces to the *pseudorapidity*

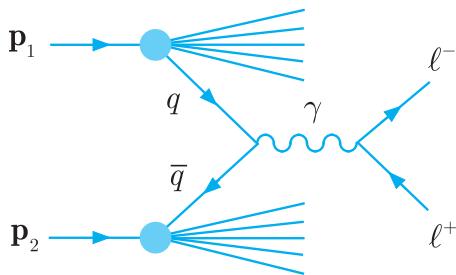
$$\eta = -\ln(\tan \theta/2). \quad (8.51)$$

For high-energy jets, it is usually a good approximation to neglect the invariant mass of the jet compared to its energy, so that the pseudorapidity is a good approximation to the rapidity itself. It is often used as it depends only on one angle, and is therefore relatively easy to measure experimentally. A plot of the pseudorapidity against the polar angle  $\theta$  is shown in Figure 8.14. As can be seen, it tends rapidly to  $\pm\infty$  for jets emitted close to, or opposite to, the beam direction, and is less than, or of order, unity for jets emitted at a large angle to both these directions.



**Figure 8.14** Pseudorapidity  $\eta$  as a function of the polar angle  $\theta$  relative to the beam direction.

<sup>16</sup>The internal momenta of the quarks within a proton at rest are of order 300 MeV (because their de Broglie wavelengths are of the order of the proton diameter), which is very small compared to the typical values of the transverse momenta of the recoil electron in deep inelastic scattering. The same result applies in the electron–proton centre-of-mass, or any frame related by a Lorentz transformation along the  $z$ -axis, since such a transformation leaves transverse momenta invariant.



**Figure 8.15** The Drell–Yan mechanism for charged lepton-pair production in proton–proton collisions.

### 8.4.1 Lepton pair production

In high-energy proton–proton collisions, charged lepton pairs with large invariant mass are produced in the so-called *Drell–Yan* reaction

$$p + p \rightarrow \ell^+ + \ell^- + X, \quad (8.52)$$

where  $\ell = e, \mu, \tau$  and  $X$  represents unobserved hadrons, as usual. For lepton-pair invariant masses  $M$  less than about 50 GeV, the dominant mechanism for this reaction is shown in Figure 8.15, where we assume for the moment that the annihilating quarks are a  $u\bar{u}$  pair. The probabilities of finding a  $u$ -quark in proton 1 with fractional momentum  $x_1$  to  $x_1 + dx_1$ , and a  $\bar{u}$ -antiquark in proton 2 with fractional momentum  $x_2$  to  $x_2 + dx_2$ , are  $u(x_1) dx_1$  and  $\bar{u}(x_2) dx_2$ , respectively. Hence the contribution to the cross-section for the reaction (8.52) from Figure 8.15 is

$$d^2\sigma = u(x_1) dx_1 \bar{u}(x_2) dx_2 \sigma(u\bar{u} \rightarrow \ell^+ \ell^-), \quad (8.53)$$

where the momenta of the  $u$  and  $\bar{u}$  quarks are  $x_1 \mathbf{p}_1$  and  $x_2 \mathbf{p}_2$ , respectively. At high energies, when lepton and quark masses can be neglected, the cross-section for  $u\bar{u} \rightarrow \ell^+ \ell^-$  can be deduced from the cross-section (cf. (7.19))

$$\sigma(e^+ e^- \rightarrow \mu^+ \mu^-) = \frac{4\pi\alpha^2}{3M^2}, \quad (8.54)$$

which arises from the essentially identical mechanism of Figure 7.14. For  $u\bar{u} \rightarrow \ell^+ \ell^-$ , we must replace  $e^2$  by  $e_u^2 e^2$  and divide by a factor three, because the quark and antiquark must be in the same colour state to annihilate.<sup>17</sup> Hence (8.53) becomes

$$d\sigma^2 = \frac{4\pi\alpha^2}{9M^2} e_u^2 u(x_1) \bar{u}(x_2) dx_1 dx_2.$$

<sup>17</sup>This should be contrasted to the discussion in Section 7.2.1, where in obtaining the cross-section for  $e^+ e^- \rightarrow q\bar{q}$ , we had to multiply by three, because the quark pairs could be produced in all three colour states.

Finally, summing over all quark flavours  $a = u, d, s, c, b$  in protons and taking account of the fact that the quark in Figure 8.15 could originate equally well in the second proton, we obtain

$$d^2\sigma = \frac{4\pi\alpha^2}{9M^2} \sum_a e_a^2 [f_a(x_1)f_{\bar{a}}(x_2) + f_{\bar{a}}(x_1)f_a(x_2)] dx_1 dx_2. \quad (8.55)$$

It remains to relate the variables  $x_1$  and  $x_2$  to variables that are directly measured in experiments. To do this, we shall assume that the proton and quark energies are large enough for their masses to be neglected and that we have colliding beams such that the energy–momenta  $(E, \mathbf{p})$  of protons 1 and 2 are

$$(E, 0, 0, E) \quad \text{and} \quad (E, 0, 0, -E), \quad (8.56a)$$

respectively. The quark and antiquark energy–momenta are then

$$E(x_1, 0, 0, x_1) \quad \text{and} \quad E(x_2, 0, 0, -x_2), \quad (8.56b)$$

respectively, and by energy–momentum conservation, the energy–momentum of the lepton pair is

$$(\tilde{E}, \tilde{\mathbf{p}}) = E(x_1 + x_2, 0, 0, x_1 - x_2). \quad (8.56c)$$

For a fixed proton energy  $E$ , the squared centre-of-mass energy of the proton–proton system is

$$s = 4E^2, \quad (8.57)$$

while the squared invariant mass of the lepton pair is

$$M^2 = \tilde{E}^2 - \tilde{\mathbf{p}}^2 = 4x_1 x_2 E^2 = x_1 x_2 s. \quad (8.58)$$

The rapidity (8.50) of the lepton pair follows from (8.56c) and is

$$y = \frac{1}{2} \ln \left( \frac{\tilde{E} + \tilde{p}_z}{\tilde{E} - \tilde{p}_z} \right) = \frac{1}{2} \ln \left( \frac{x_1}{x_2} \right), \quad (8.59)$$

while the inverse relations

$$x_1 = \frac{M}{\sqrt{s}} e^y, \quad x_2 = \frac{M}{\sqrt{s}} e^{-y} \quad (8.60)$$

follow from (8.58) and (8.59), leading to the Jacobian

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial M} & \frac{\partial x_1}{\partial y} \\ \frac{\partial x_2}{\partial M} & \frac{\partial x_2}{\partial y} \end{vmatrix} = -\frac{2M}{s}.$$

Hence

$$dx_1 dx_2 = |J| dM dy = \frac{2M}{s} dM dy, \quad (8.61)$$

and (8.55) becomes

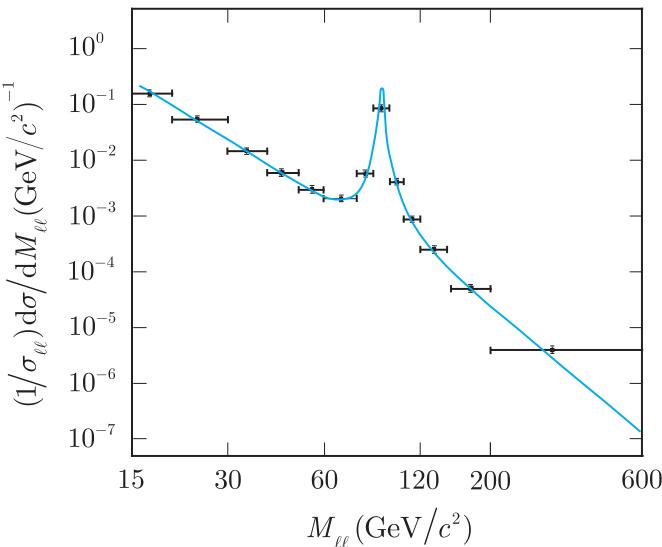
$$\frac{d^2\sigma}{dMdy} = \frac{8\pi\alpha^2}{9Ms} F(x_1, x_2), \quad (8.62a)$$

where

$$F(x_1, x_2) = e_a^2 \sum_a [f_a(x_1)f_{\bar{a}}(x_2) + f_{\bar{a}}(x_1)f_a(x_2)], \quad (8.62b)$$

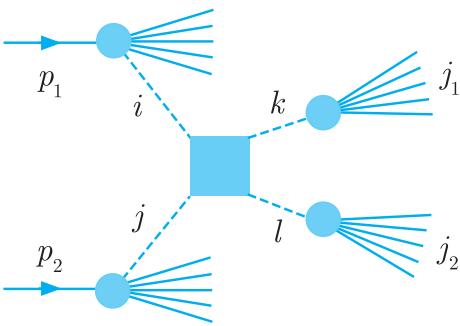
and  $x_1, x_2$  are given by (8.60).

The Drell–Yan mechanism shown in Figure 8.15, with the corresponding cross-section formula (8.62), is the dominant contribution to the production of lepton pairs with invariant mass between about 15 GeV and 50 GeV.<sup>18</sup> Above the latter value, one must also take into account a second Drell–Yan production mechanism, in which the intermediate photon in Figure 8.15 is replaced by an intermediate  $Z^0$  boson. This is closely analogous to the mechanism of Figure 4.19, which leads to a sharp peak in the  $\ell^+\ell^-$  mass spectrum at the mass of the  $Z^0$  in proton–antiproton collisions, as discussed in Section 4.5.1(a). In addition, while these are the dominant mechanisms, if very precise comparisons with data are to be made it is necessary to include corrections from more complicated higher-order diagrams. The resulting predictions are compared with data from the CMS experiment at the LHC for the dilepton mass range  $15 < M < 600$  GeV in Figure 8.16. As can be seen, there is excellent agreement between theory and experiment, including a sharp peak at the  $Z^0$  mass, as expected.



**Figure 8.16** Dilepton invariant mass spectrum in the mass region  $15 < M < 600$  GeV/ $c^2$ . The vertical bars indicate the small experimental uncertainties (statistical and systematic); the horizontal bars indicate bin sizes. The blue line is the prediction of the Drell–Yan mechanism, normalised to the  $Z$  resonance region.  
(Chatrchyan *et al.*, 2011a.)

<sup>18</sup>Below this value, there are peaks corresponding to the production and decay of other particles, such as  $\Upsilon$  and  $J/\psi$  mesons.



**Figure 8.17** The dominant mechanism for di-jet production in proton–proton collisions. The dashed lines represent partons of type  $i$ ,  $j$ ,  $k$  and  $l$ , and the square box represents a short-range interaction between them.

#### 8.4.2 Jets in pp collisions

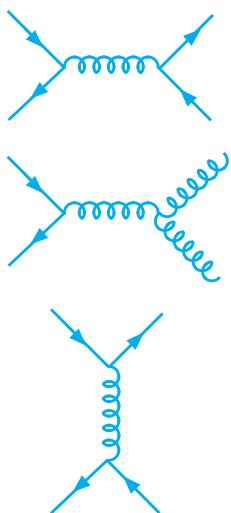
In the preceding section, we saw that dilepton production could be understood by the Drell–Yan mechanism. Dijet production arises from a large number of analogous mechanisms, summarised generically in Figure 8.17. In this figure, a parton  $i$  from a proton interacts with a parton  $j$  from another proton to produce two partons labelled by  $k$  and  $l$ , which are observed as jets in the final state. Since the partons  $i$  and  $j$  have negligible transverse momentum, as in the Drell–Yan case (cf. (8.56b)), the two jets must have equal and opposite transverse momenta. Hence the jets appear ‘back-to-back’ when observed along the beam axis, in analogy to the lepton and the jet in Figure 8.13. This is confirmed by experimental observations. However, this is not in general true when the event is viewed in other projections, since their longitudinal momenta need not be equal and opposite.

To progress further, it is necessary to evaluate the various possibilities in Figure 8.17. To illustrate this, consider the case in where  $i = u$  and  $j = \bar{u}$ . Then, in analogy to (8.53) for di-lepton production, the contribution to the cross-section for di-jet production is

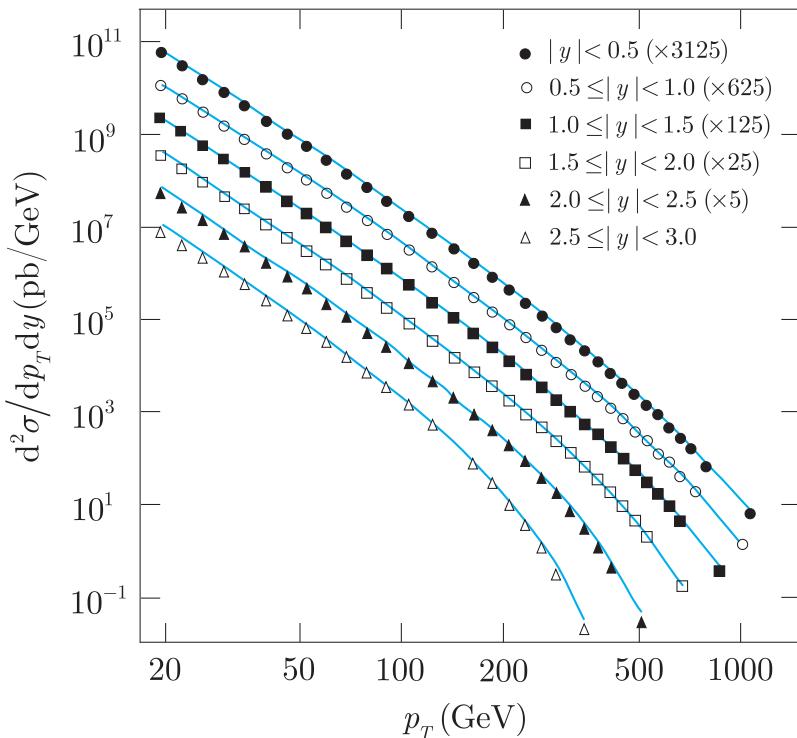
$$d^2\sigma = u(x_1)dx_1\bar{u}(x_2)dx_2\sigma(u\bar{u} \rightarrow j_1j_2), \quad (8.63)$$

where the cross-section  $\sigma(u\bar{u} \rightarrow j_1j_2)$  for producing two jets is given in lowest order by the Feynman diagrams of Figure 8.18. The other combinations of  $i$ ,  $j$  can be treated in similar ways, although the relevant Feynman diagrams will be different in each case. Because the partons  $i$ ,  $j$  can each be either gluons or any of the five quarks  $u$ ,  $d$ ,  $s$ ,  $c$ ,  $b$  or their antiquarks, there are 121 such terms to be taken into account. In practice, this is done using very large computer programs, which also include the higher-order corrections in  $\alpha_s$  that are required for precise comparisons with experimental data, as we shall illustrate below.

At this point, we note that the lowest-order contributions to two-jet production are of order  $\alpha_s^2$ , as can be seen from Figure 8.18, and in general, the production of  $n$  jets is of order  $\alpha_s^n$ . However, if only a single jet is produced, the relevant parton cross-section in (8.63)



**Figure 8.18** Feynman diagrams contributing to the reaction  $u\bar{u} \rightarrow j_1 j_2$  in lowest order.



**Figure 8.19** The inclusive differential cross-sections for jet production in  $pp$  collisions at a total energy of 7 TeV, as a function of the transverse momentum  $p_T$  for a range of values of rapidity  $y$ . Experimental uncertainties are smaller than the size of the data points. The blue lines are QCD predictions. (Adapted from S. Chatrchyan *et al.* 2011b, with permission from the American Physical Society.)

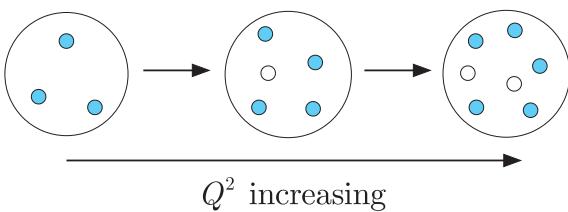
becomes  $\sigma(u\bar{u} \rightarrow j)$  and the jet associated with the parton  $j$  will have negligible transverse momentum, like the  $u$  and  $\bar{u}$ , and hence would not be easy to distinguish from beam fragments. In practice, it is usual to consider only jets with reasonably large  $p_T$ , which ensures that they are clearly separated from beam fragments.<sup>19</sup> Hence, in practice we are only concerned with  $n \geq 2$  jets. Events have been observed with up to at least eight jets, although their frequency decreases with  $n$ , because, as noted above, they are of order  $\alpha_s^n$ . The data on such jets is in excellent agreement with the theoretical predictions, as illustrated by the example in Figure 8.19, which shows the cross-section for producing jets as a function of transverse momentum at several values of rapidity, irrespective of how many other jets are produced. This is also called the *inclusive* production cross-section.

## 8.5 Current and constituent quarks

In previous sections, we have established that when the proton is probed with high spatial resolution, corresponding to large  $Q^2$ , it is

<sup>19</sup>It also ensures that the parton–parton interaction must be short-range in order to generate large  $p_T$  values, and so can be treated perturbatively.

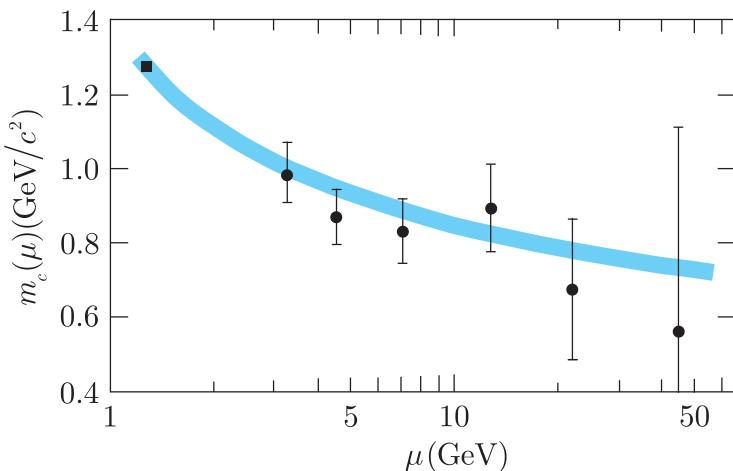
**Figure 8.20** Schematic representation of the evolution of the quark content of the proton with increasing resolution. Quarks and antiquarks are denoted by blue and white circles, respectively; gluons are not shown.



found to be composed of three valence quarks  $uud$ , plus gluons and a sea of quark–antiquark pairs. In this section we will briefly digress to consider the relation between this description and the simple quark model of Chapters 3 and 6.

From Figure 8.7 we saw that at  $Q^2 = 10 \text{ GeV}^2$ , gluons account for about half the proton’s momentum and most of the rest is carried by the valence quarks. The contribution of the sea of quark–antiquark pairs is far smaller and increases/decreases in importance as  $Q^2$  gets larger/smaller. This trend is indicated schematically in Figure 8.20, where we have assumed that for low  $Q^2 \ll M^2$ , corresponding to low resolution, the sea contribution can be neglected and the proton can be approximated by the three valence quarks moving in a gluon field. The success of the simple quark model in this regime implies that this can in turn be approximated by three constituent quarks moving a simple potential. This is a very useful approximation, but not exact, as shown by the fact that the masses of the constituent quarks depends on their environment, as seen in Section 6.2.4, where we found that they had heavier masses in baryons than in mesons.

The quarks and antiquarks of the QCD-improved parton model are, together with gluons, the fundamental constituents of hadrons and are often called *current quarks* to distinguish them from the constituent quarks of the simple quark model. Just like constituent quarks, current quarks are confined within hadrons, so their masses cannot be defined as a measure of the inertia of a free quark. However, unlike the masses of constituent quarks, the masses of current quarks appear as parameters in the fundamental Hamiltonian that defines QCD. Their values can then be estimated by relating them to observables, like hadron masses, using various approximations, including lattice gauge theory. However, two complicating features must be noted before doing this. Firstly, because quarks are confined in QCD, their masses cannot be directly measured. Rather, they are parameters that occur in the QCD Hamiltonian. This is in contrast to QED, where we usually define the electron mass as the inertial mass of a free electron, which can be directly measured. Hence its value is independent of any theoretical considerations. Secondly, QCD can be formulated in various different ways, using slightly different definitions of the quark masses. These are just different accounts of the same theory and the different mass definitions



**Figure 8.21** Measured charm mass  $m_c(\mu)$  in the  $\overline{\text{MS}}$  scheme (black circles). (Behnke *et al.* 2015. Reproduced with permission from Elsevier Ltd.) The black square at scale  $m_c$  is the PDG world average and the blue band is its expected running.

are related to one another, but a choice has to be made when quoting mass values. Here we will use the most popular scheme, called the  $\overline{\text{MS}}$  scheme.<sup>20</sup> In addition, like the running coupling constant  $\alpha_s(\mu)$ , the quark masses depend on the momentum scale  $\mu$  (cf. Equations (7.5) and (7.6)) at which they are measured. These *running masses*<sup>21</sup>  $m_a(\mu)$  ( $a = u, d, s, \dots$ ) decrease with increasing  $\mu$  and their values are quoted at a chosen reference value. For the  $u, d, s$  quarks this is taken to be  $\mu = 2$  GeV<sup>2</sup>; for the  $c$  and  $b$  quarks it is chosen to be  $m_c$  and  $m_b$ , respectively, that is the value for which  $m_c(\mu) = \mu$ , for example. The resulting values are

$$m_u = 2.3^{+0.7}_{-0.5}, \quad m_d = 4.8^{+0.5}_{-0.3}, \quad m_s = 95 \pm 5, \quad (8.64\text{a})$$

in units of  $\text{MeV}/c^2$ , and

$$m_c = 1.275 \pm 0.023, \quad m_b = 4.18 \pm 0.03, \quad (8.64\text{b})$$

in units of  $\text{GeV}/c^2$ , while the predicted behaviour of  $m_c(\mu)$  as a function of  $\mu$  is shown in Figure 8.21.

The most striking feature of these results is that the mass of the light quarks,  $u, d, s$ , are small compared to the typical hadron mass scale of  $1 \text{ GeV}/c^2$ . In the limit that their masses vanish, it can be shown that QCD possesses a new symmetry, called *chiral symmetry*, and the pions and kaons would also be massless. The symmetry is broken by the non-zero quark masses in a well-defined way, yielding relations between the pion, kaon and quark masses. These relations, together with the complementary results from lattice gauge theories

<sup>20</sup>The  $\overline{\text{MS}}$  scheme and an alternative are defined in Section 15.3.1 of Mandl and Shaw (2010).

<sup>21</sup>The running mass is defined formally in Section 15.3.2 of Mandl and Shaw (2010).

and the measured meson masses, lead directly to the quark masses (8.64a).

Finally, we note that (8.64a) implies that

$$2m_u + m_d \simeq 9 \text{ MeV}/c^2 \ll m_{\text{proton}}.$$

Even when allowing for the increase in the quark masses with decreasing  $\mu$ , one can conclude that three valence quarks in the proton make only a small contribution to the mass of the proton. The bulk of the proton mass arises from the interaction energy of the quarks and the gluon field that holds the proton together. This is in marked contrast to more familiar examples of bound states, like atoms and nuclei, whose masses are to a good approximation equal to the sum of the masses of their constituent particles.

## Problems 8

- 8.1** Perform the angular integration in (8.5) and use the result (a) to show that  $G_E(q^2)$  is a function of  $q^2$  only and (b) to derive (8.9) for the root-mean-square charge radius  $r_E$ .
- 8.2** Verify (8.21b) by explicitly evaluating the integral (8.21a).
- 8.3** Show that the scattering variable  $\nu$  defined by (8.28) reduces to (8.30) in the rest frame of the proton. Hence show that the scaling variable (8.29) lies in the range  $0 \leq x \leq 1$  if the lepton masses are neglected.
- 8.4** Elastic lepton scattering (8.1) can be regarded as a special case of inelastic scattering (8.26) in which the final hadronic state  $X$  is a single proton. Find the corresponding value of the scaling variable (8.29). Hence show that in the rest frame of the proton, the initial ( $E$ ) and final ( $E'$ ) lepton energies are related by

$$M(E - E') = EE'(1 - \cos \theta),$$

where  $\theta$  is the lepton scattering angle,  $M$  is the proton mass and lepton masses are assumed to be negligible compared with their energies.

- 8.5** According to (8.33), the fraction  $z$  of the proton momentum carried by the struck parton in Figure 8.4 is equal to the scaling variable  $x$  if certain approximations are made. These are: the proton momentum  $\mathbf{P}$  is much larger than its mass, that is  $|\mathbf{P}| \gg M$ ; the masses and transverse momenta of the partons are negligible compared with  $|\mathbf{P}|$  and  $Q^2$ ; and  $Q^2 \gg M^2$ . Verify this starting from the relation

$$\omega^2 - \mathbf{k}^2 = m^2 \simeq 0,$$

where  $\mathbf{k}$  is the final momentum of the struck parton, as shown in Figure 8.4, and  $\omega$  and  $\tilde{m}$  are its energy and mass, respectively. (*This problem should not be attempted before Problem 8.3.*)

- 8.6** Use the quark distributions of Figure 8.7 to make a rough estimate of the fraction of the proton momentum in a high-momentum frame that is carried by its quark and antiquark constituents. To what would you ascribe the ‘missing momentum’?

- 8.7** Derive the Gottfried sum rule,

$$\int_0^1 [F_2^p(x) - F_2^n(x)] \frac{dx}{x} = \frac{1}{3} + \frac{2}{3} \int_0^1 [\bar{u}(x) - \bar{d}(x)] dx,$$

where the quark distributions refer to the proton.

- 8.8** Show that: (a) the difference between two rapidities  $y_1$  and  $y_2$  is invariant under a Lorentz transformation along the beam direction; (b) the rapidity reduces to the pseudorapidity if the mass of the jet or particle referred to can be neglected compared to its energy.
- 8.9** Generalise the Drell–Yan production formula (8.62) to apply to the reaction  $\pi^- p \rightarrow \mu^+ \mu^- X$ . Then show that, under reasonable assumptions, the ratio

$$R = \frac{d^2\sigma(\pi^- + C \rightarrow \mu^+ + \mu^- + X)}{d^2\sigma(\pi^+ + C \rightarrow \mu^+ + \mu^- + X)} \rightarrow 4$$

as  $M^2 \rightarrow s$ , where  $C$  is a carbon-12 nucleus and  $M$  is the invariant mass of the lepton pair.

- 8.10** (a) The cross-section  $\sigma(u\bar{d} \rightarrow W^+)$  near the mass of the  $W^+$  is given by a Breit–Wigner formula (3.34b). For  $\sigma(u\bar{d} \rightarrow W^+)$  this may be written as

$$\sigma_{u\bar{d}} = \frac{12\pi\Gamma\Gamma_{u\bar{d}}}{M^2[4(E-M)^2 + \Gamma^2]},$$

where  $(M, \Gamma)$  are the mass and total width of the  $W^+$ ,  $\Gamma_{u\bar{d}}$  is the partial width for  $W^+ \rightarrow u\bar{d}$  and  $E$  is the total centre-of-mass energy of the  $u\bar{d}$  pair. Find  $\sigma_{\max}$ , the maximum value of  $\sigma$ , in practical units, given that the branching ratio for  $W^+ \rightarrow u\bar{d}$  is  $1/3$ .

(b) Find an expression for the cross-section  $\sigma(p\bar{p} \rightarrow W^+ + \dots)$  via the mechanism of Figure 4.19 in terms of the quark distributions in the proton, assuming that the  $p\bar{p}$  total centre-of-mass energy  $\sqrt{s}$  is sufficiently small for the contributions of sea quarks to be neglected. Use the narrow width approximation (cf. Equation (B.47b) of Appendix B), which can be rewritten in the form

$$\sigma_{u\bar{d}}(E) = \frac{\pi\Gamma}{M} \sigma_{\max} \delta\left(1 - \frac{E^2}{M^2}\right),$$

where  $\delta$  is the Dirac delta function, to find an expression for the cross-section  $\sigma(p\bar{p} \rightarrow W^+ + \dots)$  in terms of  $\sigma_{\max}$ , the  $p\bar{p}$  total centre-of-mass energy squared  $s$ , and a single integral over the quark distribution functions of the proton.

---

## Weak interactions: quarks and leptons

---

We turn now to the third fundamental force of nature, the weak interaction. Like the strong and electromagnetic interactions, the weak interaction is also associated with elementary spin-1 bosons that act as ‘force carriers’ between quarks and/or leptons. However, in contrast to photons and gluons, these are very massive particles and the resulting interactions are consequently of very short range. There are three such ‘intermediate vector bosons’: the charged bosons  $W^+$  and  $W^-$ , and the neutral  $Z^0$ . All were discovered at CERN in 1983.<sup>1</sup> Their masses are measured to be

$$M_W = 80.38 \text{ GeV}/c^2, \quad M_Z = 91.19 \text{ GeV}/c^2, \quad (9.1)$$

which from (1.26) give ranges

$$R_W \approx R_Z \approx 2 \times 10^{-3} \text{ fm} \quad (9.2)$$

for the weak interaction. These are very small distances, even when compared with the size of the nucleon, and at low energies the weak interaction can be treated as a zero-range interaction, as discussed in Section 2.2.1 (cf. Figures 2.7 and 2.8). In this and the following chapter, we shall extend the boson-exchange theory to include quarks and also consider high-energy weak interactions, where the zero-range approximation is no longer appropriate. Firstly, however, it will be useful to make a few remarks on the historical development of the modern theory of weak interactions.

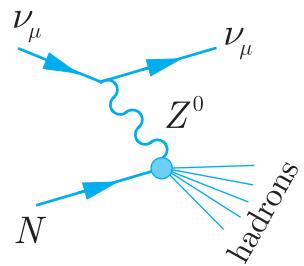
---

<sup>1</sup> One of the experiments (UA1) is described in Section 4.5.1.

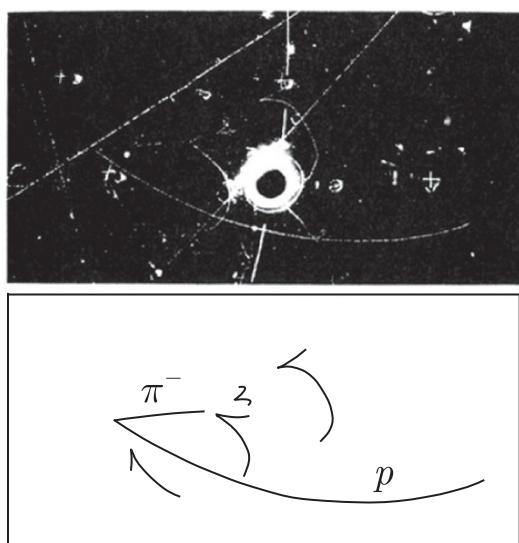
The idea that weak interactions are due to the exchange of massive charged bosons seems to have been first proposed by Klein in 1938.<sup>2</sup> These particles were later called  $W^\pm$  bosons, and until 1973 all observed weak interactions were consistent with the hypothesis that they were mediated by the exchange of heavy charged bosons  $W^\pm$  only. However, in the 1960s Glashow, Salam and Weinberg had developed a theory that unified electromagnetic and weak interactions in a way that is often compared to the unification of electric and magnetic interactions by Faraday and Maxwell a century earlier. This new theory made several remarkable predictions, including the existence of a neutral vector boson  $Z^0$  and of weak reactions arising from its exchange. These processes are called *neutral current* reactions to distinguish them from the so-called *charged current* reactions arising from charged  $W^\pm$  boson exchange. In particular, neutral current reactions of the type

$$\nu_\mu + N \rightarrow \nu_\mu + X \quad (9.3)$$

were predicted to occur via the mechanism of Figure 9.1, where  $N$  is a nucleon and  $X$ , as usual, is any set of hadrons allowed by the conservation laws. Although difficult to detect, such reactions were first observed in an experiment at CERN in 1973 using a heavy-liquid bubble chamber, and one of the events obtained is shown in Figure 9.2. In this event, an incoming muon neutrino interacts with a nucleon inside the chamber, producing several hadrons but no muon. The presence of a neutrino in the final state is inferred from the law of conservation of muon number (2.3b) discussed in Chapter 2.



**Figure 9.1** Mechanism for neutral current reactions of the type (9.3).



**Figure 9.2** Neutral current reaction of the type (9.3) observed in a heavy-liquid bubble chamber at CERN in 1973. The neutrino is incident from the left and interacts with an atomic nucleus inside the chamber to produce a negative pion and a proton. Three photons are also detected via their conversion to  $e^+ e^-$  pairs. These photons almost certainly arise from the decays of neutral pions produced in the initial neutrino interaction and in the subsequent interaction of the negative pion with another atomic nucleus. (Reproduced by permission of CERN.)

<sup>2</sup> Klein (1938). See also Klein (1948).

The prediction of the existence and properties of neutral currents, prior to their discovery, is only one of many spectacular successes of the unified theory of electromagnetic and weak interactions. Others include the prediction of the existence of the charmed quark prior to its discovery in 1974; the prediction of the masses of the  $W^\pm$  and  $Z^0$  bosons prior to their detection in 1983; and the prediction of the Higgs boson long before its existence was announced in 2012. In general, the theory is in complete agreement with all the data on both weak and electromagnetic interactions, which are now usually referred to collectively as the *electroweak interaction*, in the same way that electric and magnetic interactions are referred to collectively as electromagnetic interactions. However, the new unification only becomes manifest at very high energies, and at lower energies weak and electromagnetic interactions can still be clearly separated, as we shall see.

In this chapter we will discuss only the charged-current reactions, i.e. those that involve the emission, absorption or exchange of  $W^\pm$  bosons, leaving neutral current reactions, which involve  $Z^0$  bosons, and the unified theory to Chapter 10.

## 9.1 Charged current reactions

The simplest charged current reactions are purely leptonic processes like muon decay

$$\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu \quad (9.4)$$

and we discuss these first. We then discuss purely hadronic processes, like lambda decay

$$\Lambda \rightarrow p + \pi^- \quad (9.5)$$

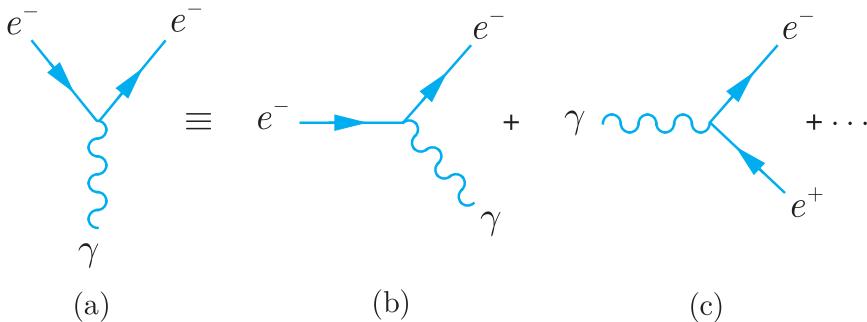
and *semileptonic* reactions, such as neutron decay

$$n \rightarrow p + e^- + \bar{\nu}_e, \quad (9.6)$$

which involve both hadrons and leptons. As usual, the interactions of the hadrons are interpreted in terms of their quark constituents, and we will see that there is a remarkable symmetry between the weak interactions of leptons and quarks.

### 9.1.1 $W^\pm$ -lepton interactions

In Chapter 1 we saw that all the electromagnetic interactions of electrons and positrons could be built from eight basic interactions in which a photon is either emitted or absorbed. These processes were derived from hole theory in Section 1.3.1, and they are conveniently summarised by the vertex of Figure 9.3(a), from which they can be obtained by associating different particle lines with the initial



**Figure 9.3** The basic vertex for electron–photon interactions, together with two of the eight basic processes derived from it.

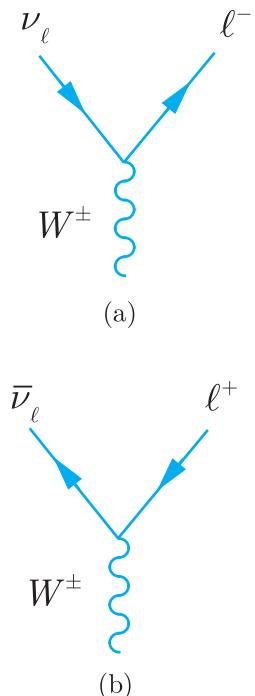
and final states. This is illustrated in Figure 9.3(b) and (c), while the full set of eight processes corresponding to the vertex is shown in Figure 1.4. Leptonic weak interaction processes like muon decay (9.4) and inverse muon decay

$$\nu_\mu + e^- \rightarrow \mu^- + \nu_e \quad (9.7)$$

can also be built from a limited number of basic vertices in the same way. In the case of the charged current reactions, it can be shown that for each lepton type  $\ell = e, \mu, \tau$  there are 16 such basic reactions corresponding to the two vertices of Figure 9.4. The eight processes corresponding to Figure 9.4(a) are shown explicitly in Figure 9.5. The eight corresponding to Figure 9.4(b) are not shown explicitly since they can be obtained from those of Figure 9.5 by replacing all the particles by their antiparticles, as illustrated in Figure 9.6.

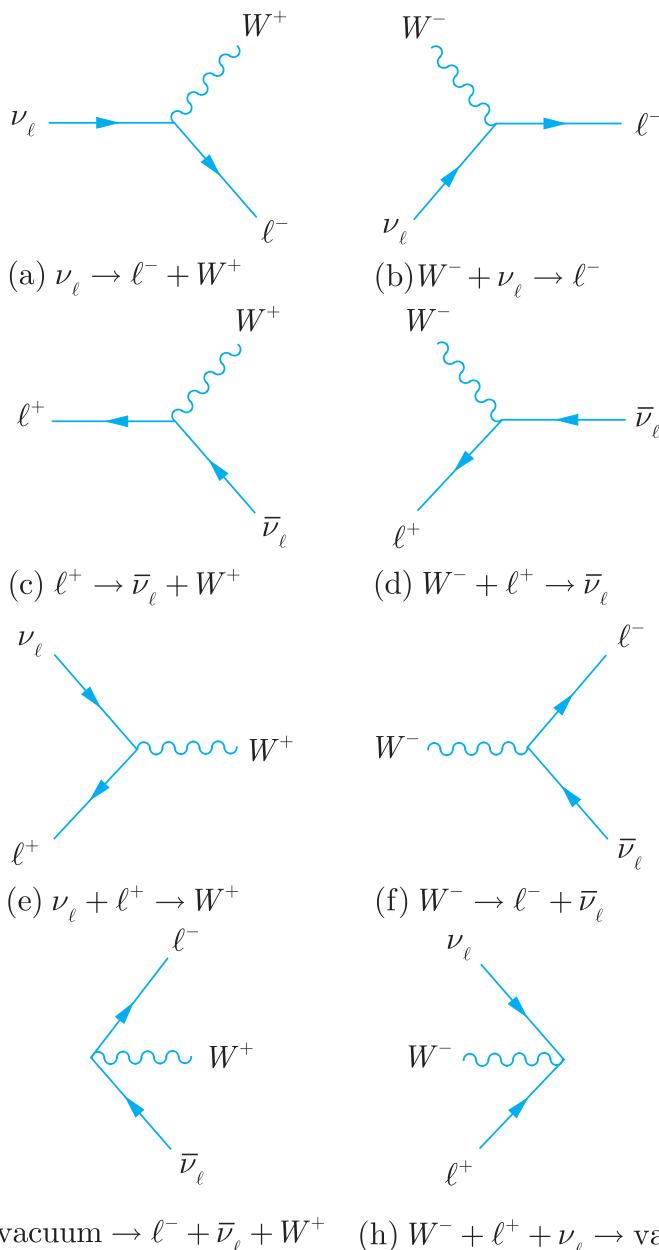
One of the most important properties of the weak interaction is that it conserves the lepton numbers  $L_e, L_\mu, L_\tau$  defined in Equations (2.3a), (2.3b) and (2.3c). This is guaranteed by the fact that in each of the vertices of Figure 9.4 there is one arrow pointing into the vertex and one pointing out of it, and the lepton label  $\ell$  is the same on both lines. In contrast, vertices like those in Figure 9.7(a), which gives rise to processes that violate lepton number, like those shown in Figures 9.7(b) and (c), are excluded from the scheme.<sup>3</sup>

Although the processes of Figure 9.5 conserve lepton numbers, this does not mean that they can occur as isolated reactions in free space. As we saw in Section 1.3.2, the basic electromagnetic processes of Figure 1.4 violate energy conservation if momentum conservation is assumed, and they can only contribute to physical processes if two or more are combined in such a way that energy is conserved overall. The same applies to most of the processes of Figure 9.5, but in this case there are some important exceptions. Specifically, by working in the rest frame of the  $W$ -boson one can easily show



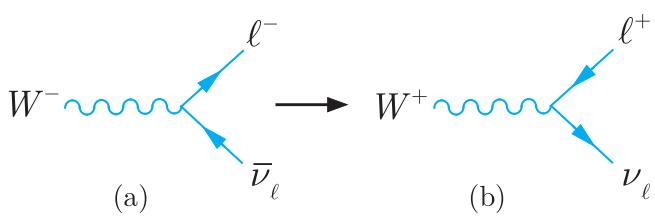
**Figure 9.4** The two basic vertices for  $W^\pm$ –lepton interactions.

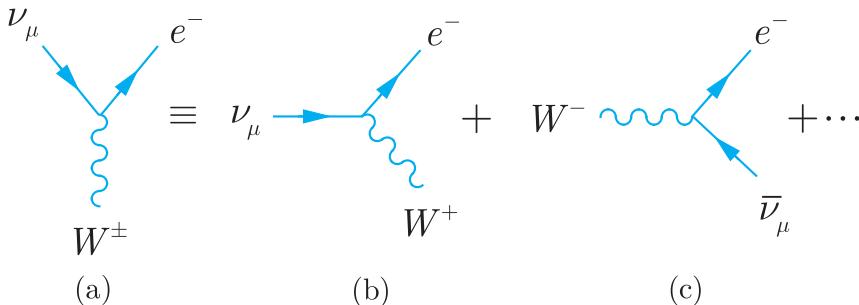
<sup>3</sup> In fact the diagrams of Figure 9.4 are the only possible  $W^\pm$ –lepton vertices that satisfy both charge and lepton number conservation (see Problem 9.1).



**Figure 9.5** The eight basic vertices derived from the vertex of Figure 9.4(a).

**Figure 9.6** A pair of weak interaction processes that are related by replacing all particles by their antiparticles. The two diagrams (a) and (b) arise from the vertices of Figures 9.4(a) and (b), respectively.





**Figure 9.7** Example of a vertex that violates lepton number conservation, together with two of the forbidden processes to which it would give rise.

that Figures 9.5(e) and (f) are compatible with both energy and momentum conservation provided that

$$M_W > M_\ell + M_{\nu_\ell} \quad (\ell = e, \mu, \tau),$$

which is obviously the case. In particular, Figure 9.5(f) and its ‘antiparticle’ reaction, which are shown together in Figure 9.6, are the dominant mechanisms for the leptonic decays (4.28a) and (4.28b), which were used to detect the  $W^\pm$  bosons in the experiments discussed in Section 4.5.1.

We turn now to the strength of the interaction. In Section 2.2.2 we showed that the weak and electromagnetic interactions of electrons, muons and tauons are identical, provided their mass differences are taken into account. Correspondingly, each of the vertices of Figure 9.4 is characterised by the same dimensionless strength parameter  $\alpha_W$ , analogous to the fine structure constant  $\alpha$  in QED, independent of which lepton type is involved. We can obtain an order-of-magnitude estimate for  $\alpha_W$  by applying the method of dimensions to the rate for the leptonic decays shown in Figures 9.5(a) and (b) for  $\ell = e$ . This decay rate has a natural dimension [E], since in natural units time has a dimension [E<sup>1</sup>] (see Table 1.1) and is measured to be

$$\Gamma(W \rightarrow e\nu) \simeq 0.223 \pm 0.007 \text{ GeV}. \quad (9.8)$$

Since it involves only a single basic process, Figure 9.6(a) or (b), this decay rate is proportional to  $\alpha_W$ , and if we neglect the lepton masses, the only other parameter is the  $W^-$  boson mass. Hence a simple dimensional estimate gives

$$\Gamma(W \rightarrow e\nu) \simeq \alpha_W M_W \simeq 80 \alpha_W \text{ GeV},$$

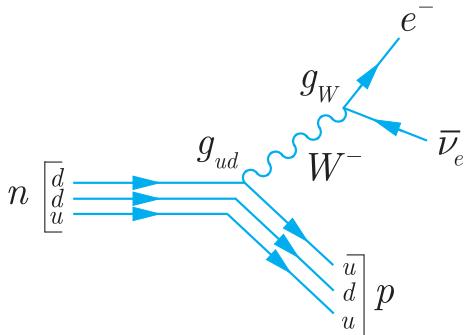
which combined with (9.8) gives

$$\alpha_W \simeq 1/350.$$

Alternatively, a more detailed calculation<sup>4</sup> gives

$$\Gamma(W \rightarrow e\nu) = 2\alpha_W M_W/3,$$

<sup>4</sup> See, for example, Section 16.6.3 of Mandl and Shaw (2010).



**Figure 9.8** Dominant diagram for neutron decay (9.6).

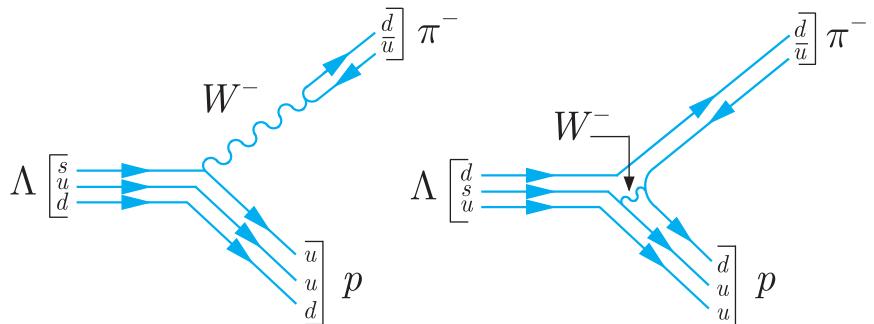
resulting in the value

$$\alpha_W = 0.0042 \pm 0.0002, \quad (9.9)$$

in good agreement with the value (2.19) obtained from the precise values of the Fermi coupling constant and the W-mass using (2.17). As already noted in Section 2.2.1, the weak and electromagnetic interactions are actually of comparable intrinsic strength, and weak interaction rates are only small at low energies because of the large mass of the W boson, which enters the low-energy effective coupling constant (2.17) as the inverse square.

### 9.1.2 Lepton–quark symmetry and mixing

The weak interactions of hadrons are understood in terms of basic processes in which  $W^\pm$  bosons are emitted or absorbed by their constituent quarks. These can give rise to semileptonic processes like neutron decay (9.6), for which the dominant diagram is shown in Figure 9.8, and purely hadronic decays such as  $\Lambda$  decay (9.5), for which the dominant diagrams are shown in Figure 9.9. The purely hadronic decays are not nearly so well understood as the semileptonic ones, because the final-state particles interact strongly with each other, leading to effects that are very difficult to calculate. We shall therefore concentrate mainly on the semileptonic interactions



**Figure 9.9** Dominant diagrams for  $\Lambda$  decay (9.5).

in what follows, and we shall initially restrict ourselves to the first two generations of quarks

$$\begin{pmatrix} u \\ d \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} c \\ s \end{pmatrix}, \quad (9.10)$$

of which the majority of known hadrons are composed.

The weak interactions of quarks are best understood in terms of two ideas: *lepton–quark symmetry* and *quark mixing*. Here we shall introduce these ideas and use them to deduce the form of the  $W^\pm$  boson–quark interactions. They may also be used to deduce the form of the  $Z^0$ -boson–quark interaction,<sup>5</sup> and their justification lies in the fact that the interactions so deduced successfully explain the observed weak interactions of hadrons in both cases.

We start with lepton–quark symmetry. In its simplest form this asserts that the two generations of quarks (9.10) and the two generations of leptons

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix} \quad (9.11)$$

have identical weak interactions. That is, one obtains the basic  $W^\pm$ –quark vertices by making the replacements  $\nu_e \rightarrow u$ ,  $e^- \rightarrow d$ ,  $\nu_\mu \rightarrow c$ ,  $\mu^- \rightarrow s$  in the basic  $W^\pm$ –lepton vertices in Figure 9.4, leaving the coupling constant  $g_W$  unchanged. In this way one obtains the vertices of Figure 9.10, where the coupling constants

$$g_{ud} = g_{cs} = g_W. \quad (9.12)$$

This works quite well for many reactions, like the pion decay

$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu, \quad (9.13a)$$

which corresponds to

$$d\bar{u} \rightarrow \mu^- + \bar{\nu}_\mu \quad (9.13b)$$

at the quark level, and is interpreted by the Feynman diagram of Figure 9.11(a).

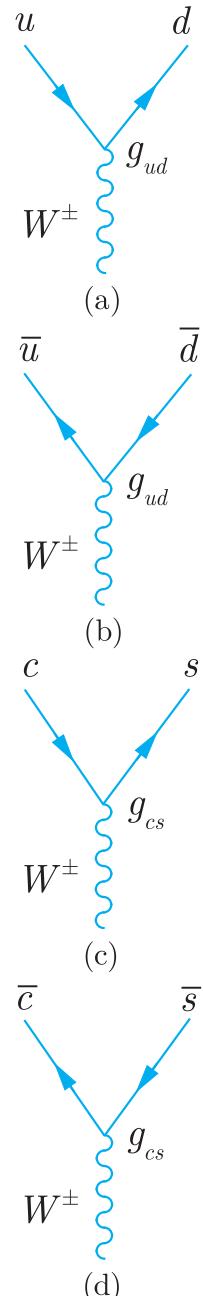
However, many other decays that are experimentally observed are forbidden in this simple scheme. An example is the kaon decay

$$K^- \rightarrow \mu^- + \bar{\nu}_\mu, \quad (9.14a)$$

which corresponds to

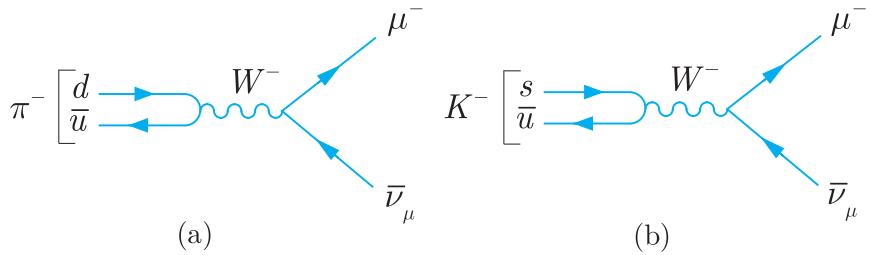
$$s\bar{u} \rightarrow \mu^- + \bar{\nu}_\mu \quad (9.14b)$$

at the quark level. This decay would be naturally explained by the Feynman diagram of Figure 9.11(b), which includes a  $usW$  vertex, but the latter is not included in the vertices of Figure 9.10. It can be



**Figure 9.10** The  $W^\pm$  quark vertices obtained from lepton–quark symmetry when quark mixing is ignored.

<sup>5</sup> This is done in Section 10.1.1.



**Figure 9.11** Feynman diagrams for the semileptonic decays: (a)  $\pi^- \rightarrow \mu^- + \bar{\nu}_\mu$  and (b)  $K^- \rightarrow \mu^- + \bar{\nu}_\mu$ .

incorporated, however, by introducing an hypothesis due to Cabibbo called quark mixing, mentioned above. According to this idea, the  $d$  and  $s$  quarks participate in the weak interactions via the linear combinations

$$d' = d \cos \theta_C + s \sin \theta_C \quad (9.15a)$$

and

$$s' = -d \sin \theta_C + s \cos \theta_C, \quad (9.15b)$$

where the parameter  $\theta_C$  is called the *Cabibbo angle*. That is, lepton–quark symmetry is assumed to apply to the doublets

$$\begin{pmatrix} u \\ d' \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} c \\ s' \end{pmatrix}. \quad (9.16)$$

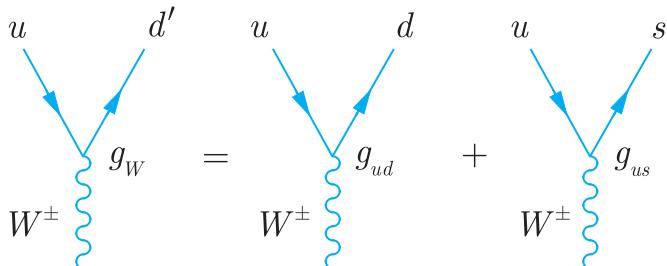
This is illustrated, for the first of these, in Figure 9.12. As can be seen, the coupling of the previously allowed  $udW$  vertex is suppressed by a factor  $\cos \theta_C$  relative to (9.12), while the previously forbidden  $usW$  vertex is now allowed with a coupling  $g_W \sin \theta_C$ . The same argument applies to the other three vertices in Figure 9.10, so that, in addition to the four vertices of Figure 9.10 with the couplings

$$g_{ud} = g_{cs} = g_W \cos \theta_C, \quad (9.17a)$$

one has the vertices of Figure 9.13 with the couplings

$$g_{us} = -g_{cd} = g_W \sin \theta_C. \quad (9.17b)$$

The set of  $W^\pm$ –quark couplings of (9.17a) and (9.17b) successfully accounts for the charged current interactions of hadrons. Since the



**Figure 9.12** The  $ud'W$  vertex and its interpretation in terms of  $udW$  and  $usW$  vertices.

$$d' = d \cos \theta_C + s \sin \theta_C \quad g_{ud} = g_W \cos \theta_C \quad g_{us} = g_W \sin \theta_C$$

lepton coupling  $g_W$  is already fixed, these couplings are specified in terms of a single free parameter, the Cabibbo angle  $\theta_C$ . This is determined by deducing the values of the couplings  $g_{ud}$  and  $g_{us}$  from the measured rates of various hadron decays. One way of doing this is to compare the rates of decays like those shown in Figure 9.11. As can be seen, they differ in that a  $d$  quark has been replaced by an  $s$  quark in the initial state, and a coupling  $g_{ud}$  replaced by a coupling  $g_{us}$ , so that the ratio of their rates

$$\frac{\Gamma(K^- \rightarrow \mu^-\bar{\nu}_\mu)}{\Gamma(\pi^- \rightarrow \mu^-\bar{\nu}_\mu)} \propto \frac{g_{us}^2}{g_{ud}^2} = \tan^2 \theta_C,$$

by (9.17). Of course, the difference in the  $d$  and  $s$  quark masses will also have an effect on the rates, and this must be taken into account. We omit the details and merely quote the mean value obtained from this and other similar determinations, which is

$$g_{us}/g_{ud} = \tan \theta_C = 0.2313 \pm 0.0007, \quad (9.18)$$

corresponding to a Cabibbo angle of

$$\theta_C = 13.02 \pm 0.04 \text{ degrees}. \quad (9.19)$$

A similar value is obtained by comparing the rates for neutron and muon decay, which depends on the ratio

$$(g_{ud}/g_W)^2 = \cos^2 \theta_C,$$

as can be seen by comparing the quark diagrams of Figures 9.8 and 2.5.

It remains to consider the charmed quark couplings  $g_{cd}$  and  $g_{cs}$ . These are measured most accurately in neutrino scattering experiments and yield a value consistent with (9.19), but with a larger error. However, the most striking result is for charmed particle decays, which almost always yield a strange particle in the final state. In order to understand this we first note that decays that involve the couplings (9.17b) are called *Cabibbo-suppressed* because their rates are typically reduced by a factor of order

$$g_{us}^2/g_{ud}^2 = g_{cd}^2/g_{cs}^2 = \tan^2 \theta_C \approx 1/20 \quad (9.20)$$

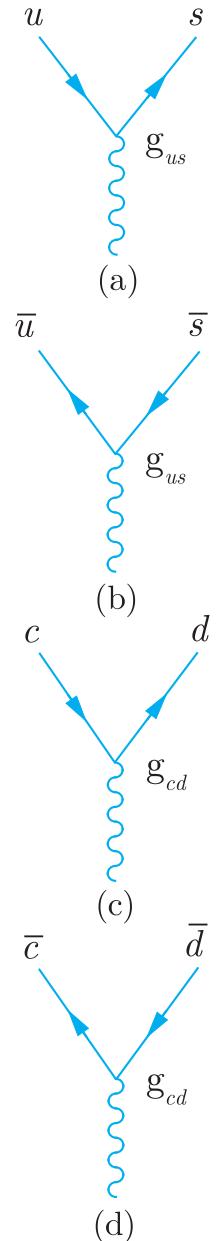
compared with similar *Cabibbo-allowed* decays, which involve the couplings (9.17a). The Cabibbo-allowed decays

$$c \rightarrow s + \ell^+ + \nu_\ell \quad (\ell = e, \mu) \quad (9.21a)$$

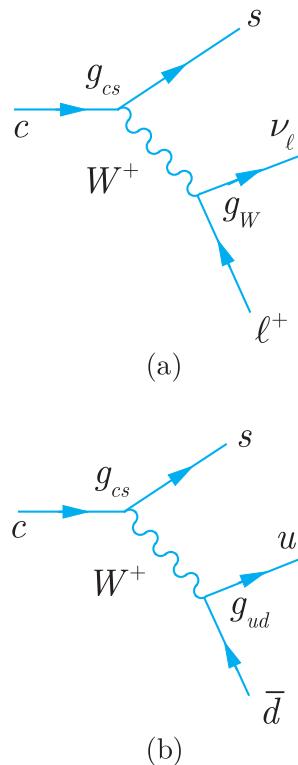
and

$$c \rightarrow s + u + \bar{d} \quad (9.21b)$$

of a charmed quark to lighter quarks and leptons are shown in Figure 9.14 and they necessarily produce an  $s$  quark in the final state, as can be seen. Of course, both these quarks must be bound into



**Figure 9.13** The additional vertices arising from lepton-quark symmetry when quark mixing is taken into account.



**Figure 9.14** Cabibbo-allowed decays (9.21) of a charmed quark.

hadrons, and the implication of this result is that the decay of a charmed hadron will almost always lead to a strange hadron in the final state. Experimentally, this is indeed the case.

### 9.1.3 $W$ boson decays

In Section 9.1.1, we saw that the  $W^+$  boson can decay to lepton pairs  $\ell^+\nu_\ell$  ( $\ell = e, \mu, \tau$ ) by the mechanism of Figure 9.6(a). Lepton–quark symmetry implies that it can also decay to quark–antiquark pairs

$$W^+ \rightarrow u\bar{d}', c\bar{s}', \quad (9.22)$$

by the corresponding mechanism, shown in Figure 9.15 with the same coupling strength  $\alpha_W$ . Here  $d'$  and  $s'$  are given by (9.15). Thus (9.22) implies, for example,

$$u\bar{d}' = u\bar{d} \cos \theta_C + u\bar{s} \sin \theta_C$$

so that  $\bar{u}d$  and  $\bar{u}s$  will occur with relative probabilities  $\cos^2 \theta_C$  and  $\sin^2 \theta_C$ , respectively. In the rest frame of the decaying  $W^+$ , the quark and antiquark are emitted back-to-back with high energies  $E \approx M_W/2 \approx 40$  GeV. They fragment into hadron jets, like those shown in Figure 4.12, so that the observed decay reaction

$$W^+ \rightarrow \text{hadrons} \quad (9.23)$$

is dominated by the two-step mechanism of Figure 9.16. However, the rate for hadron decays is determined by the initial decays (9.22), in the same way that the rate for electron–positron annihilation into hadrons is determined by the rate for the initial quark reaction  $e^+e^- \rightarrow q\bar{q}$  (cf. Section 7.2.1).

The relative rates for the quark decays (9.22) and the leptonic decays

$$W^+ \rightarrow \ell^+ + \nu_\ell \quad (\ell = e, \mu, \tau) \quad (9.24)$$

are easily estimated in the good approximation where the final-state lepton and quark masses are neglected. We then have

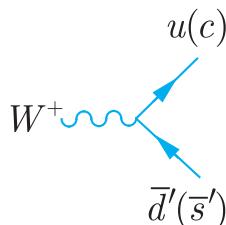
$$\Gamma(W^+ \rightarrow u\bar{d}') = \Gamma(W^+ \rightarrow c\bar{s}') = 3\Gamma(W^+ \rightarrow e^+\nu_e), \quad (9.25a)$$

since the mechanisms of these reactions are identical, but the  $q\bar{q}$  pairs can be produced in three colour states with equal probabilities, while universality gives

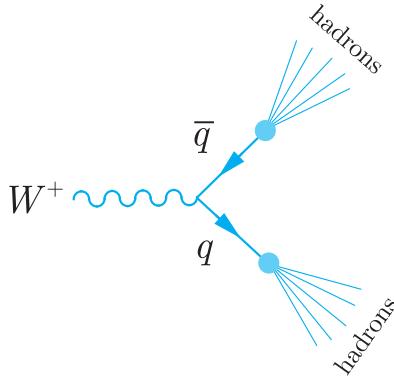
$$\Gamma(W^+ \rightarrow e^+\nu_e) = \Gamma(W^+ \rightarrow \mu^+\nu_\mu) = \Gamma(W^+ \rightarrow \tau^+\nu_\tau). \quad (9.25b)$$

Since these are the only first-order weak decays possible and there are two quark combinations (9.22) contributing to the hadron decays (9.23), we immediately arrive at the predicted branching ratios

$$B(W^+ \rightarrow \text{hadrons}) \approx 2/3, \quad (9.26a)$$



**Figure 9.15** Feynman diagram for the decays  $W^+ \rightarrow u\bar{d}'$ ,  $c\bar{s}'$ .



**Figure 9.16** The dominant mechanism for the decay of  $W^\pm$  bosons into hadrons.

together with

$$B(W^+ \rightarrow \ell^+ \nu_\ell) \approx 1/9 \quad (\ell = e, \mu, \tau), \quad (9.26b)$$

for each of the three leptonic decay modes. This simple approximation is in good agreement with experiment.

#### \*9.1.4 Selection rules in weak decays

Many observations about the weak decays of hadrons are explained by  $W^\pm$  exchange without the need for detailed calculation. For example, the decays

$$\Sigma^- \rightarrow n + e^- + \bar{\nu}_e \quad (9.27)$$

and

$$\Sigma^+ \rightarrow n + e^+ + \nu_e \quad (9.28)$$

seem very similar, where  $\Sigma^+(1189) = uus$  and  $\Sigma^-(1197) = dds$  are the charged  $\Sigma$  baryons discussed in Section 6.1.3. However, while reaction (9.27) is observed, reaction (9.28) is not, and the experimental upper limit on its rate relative to the observed decay is

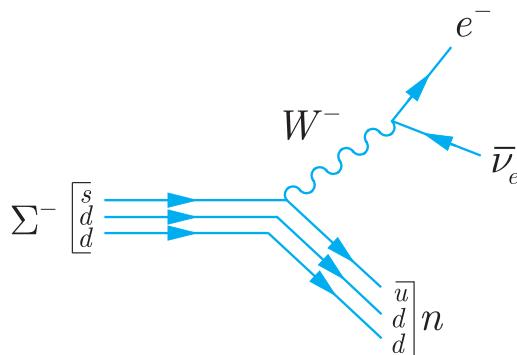
$$\frac{\Gamma(\Sigma^+ \rightarrow n + e^+ + \nu_e)}{\Gamma(\Sigma^- \rightarrow n + e^- + \bar{\nu}_e)} < 5 \times 10^{-3}.$$

The reason for this is that reaction (9.27) is allowed via the mechanism of Figure 9.17, whereas no diagram with single  $W$ -exchange can be drawn for reaction (9.28), which at the quark level is

$$uus \rightarrow udd + e^+ + \nu_e$$

and would require two separate quark transitions. It can go via a mechanism involving the emission and absorption of two  $W$  bosons, but this contribution is higher order in the weak interaction and is negligibly small.

Reaction (9.28) is just one of many that cannot proceed via single  $W^\pm$  exchange, and are therefore not observed, despite the fact that they satisfy all the appropriate conservation laws for weak



**Figure 9.17** Mechanism for the decay  $\Sigma^- \rightarrow n + e^- + \bar{\nu}_e$ .

interactions. These ‘forbidden’ reactions can be identified systematically using a number of selection rules for single  $W^\pm$  exchange processes, which can be deduced from the vertices of Figures 9.10 and 9.13. We shall illustrate this by considering the allowed changes of strangeness.

We consider firstly semileptonic decays, like those of Figures 9.11 and 9.17. Since these involve a single  $W^\pm$  quark vertex, the changes in the strangeness and electric charge of the hadrons are given by the possible changes in  $S$  and  $Q$  at this vertex. There are just two possibilities. If no strange quarks are involved at the vertex as in Figures 9.10(a) and (b) and Figures 9.13(c) and (d), there is obviously no change in strangeness, while the quark charge changes by  $\pm 1$  depending on the charge of the  $W$  boson. Hence the changes  $\Delta S$  and  $\Delta Q$  in the strangeness and the electric charge of the hadrons satisfy

$$\Delta S = 0, \Delta Q = \pm 1. \quad (9.29)$$

On the other hand, those vertices like Figures 9.10(c) and (d) and Figures 9.13(a) and (b) which do involve a strange quark give rise to processes like

$$u \rightarrow s + W^+ \quad \text{or} \quad W^- \rightarrow s + \bar{c},$$

in which the total quark charge and strangeness both decrease, giving  $\Delta S = \Delta Q = -1$ , or processes like

$$s \rightarrow u + W^- \quad \text{or} \quad W^+ \rightarrow \bar{s} + c,$$

in which the total quark charge and strangeness both increase, giving  $\Delta S = \Delta Q = 1$ . Thus the allowed semileptonic decays are characterised by the selection rules (9.29) and

$$\Delta S = \Delta Q = \pm 1, \quad (9.30)$$

where  $\Delta Q$  is the change in the charge of the hadrons only. The latter is called the  $\Delta S = \Delta Q$  rule for strangeness-changing decays, and decays with

$$\Delta S = -\Delta Q = \pm 1 \quad (9.31)$$

are forbidden. Reaction (9.28) is a typical example of a forbidden  $\Delta S = -\Delta Q$  reaction requiring changes (9.31) since the  $\Sigma^+$  has strangeness  $S = -1$  and  $Q = +1$  while the neutron has  $S = 0$  and  $Q = 0$ . Other examples are discussed in Problem 9.7.

Interesting results are also obtained for purely hadronic decays. In such decays, the exchanged  $W$  boson must be both emitted and absorbed at  $W^\pm$ -quark vertices, as illustrated in Figure 9.9. Hence the selection rules for strangeness can be inferred by applying the selection rules (9.29) and (9.30) to each individual vertex, subject to the constraint that the change in the hadron charge must now be  $\Delta Q = 0$  overall since no leptons are involved and the total charge must of course be conserved. If two vertices satisfying (9.29) are involved, strangeness is conserved and  $\Delta S = 0$ , while if one satisfies (9.29) and the other (9.30) we obviously have  $\Delta S = \pm 1$ . Finally, if two vertices satisfying (9.30) are involved, we again obtain  $\Delta S = 0$  overall because of the charge conservation condition  $\Delta Q = 0$ . We thus obtain the selection rule

$$\Delta S = 0, \pm 1 \quad (9.32)$$

for hadronic weak decays and the same selection rule also holds for semileptonic decays since it is guaranteed by the semileptonic selection rules (9.29) and (9.30).

The selection rule (9.32), which holds for all weak decays, has its most important implications for the decays of the so-called *cascade* particles

$$\Xi^0(1532) = ssu, \quad \Xi^-(1535) = ssd \quad (S = -2) \quad (9.33)$$

and the omega-minus baryon

$$\Omega^-(1672) = sss \quad (S = -3), \quad (9.34)$$

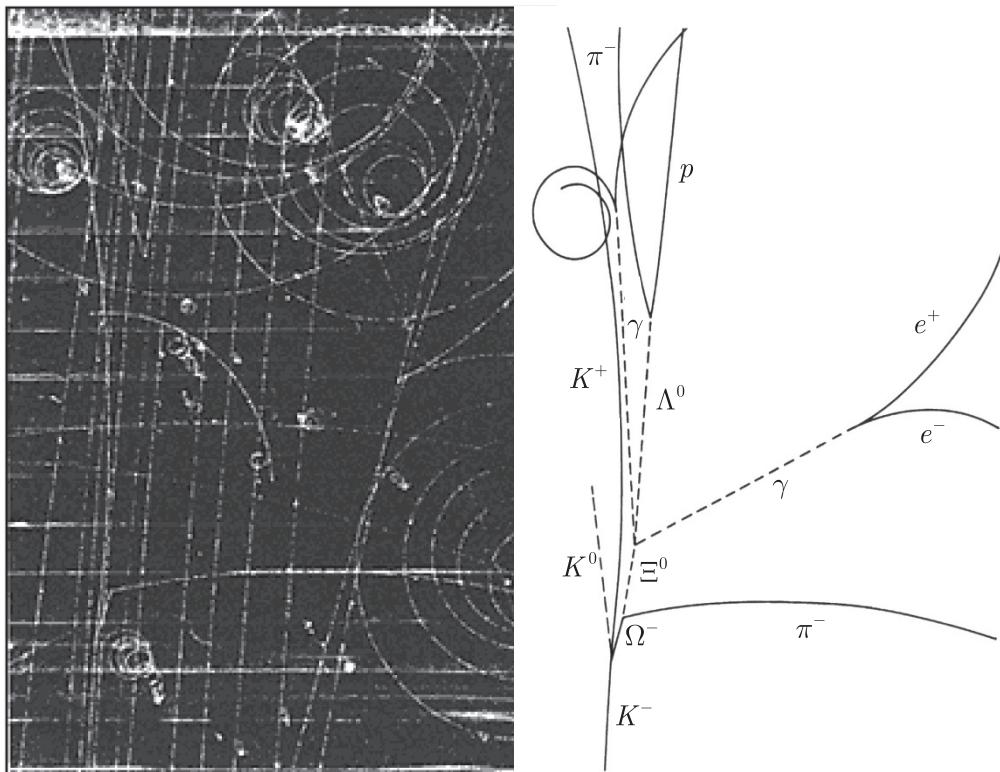
which we met briefly in Section 6.2.2 (see Tables 6.5 and 6.6). Because the baryon number is conserved, these must ultimately decay to give a proton in the final state, since this is the only stable baryon. However, this cannot occur directly because of the selection rule (9.32), but must proceed via a series of successive decays in which strangeness changes by at most one unit in each step. Thus the  $\Xi^-$  particle frequently decays by the sequence

$$\begin{aligned} \Xi^- &\rightarrow \Lambda^0 + \pi^- \\ &\quad \swarrow p + \pi^- \end{aligned} \quad (9.35)$$

giving rise to a characteristic pattern of charged particle tracks. In contrast, decay processes like

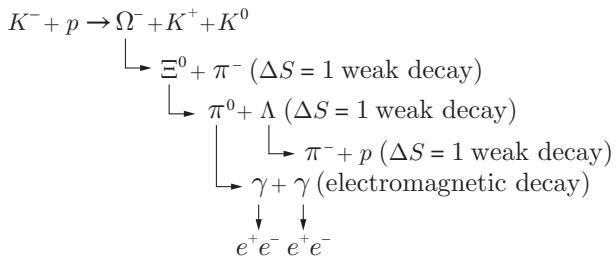
$$\Xi^- \rightarrow n + \pi^-$$

that satisfy all known conservation laws, but violate the selection rule (9.32), are never observed. Similar phenomena occur in  $\Omega^-$



**Figure 9.18** Characteristic pattern of tracks produced by the production and decay of the  $\Omega^-$ . (From Barnes *et al.* 1964, photo courtesy of Brookhaven National Laboratory.)

decays, and it is a remarkable fact that when the  $\Omega^-$  particle was first observed both its production and the whole decay sequence



was captured in a single bubble chamber picture, as shown in Figure 9.18.

## 9.2 The third generation

In Section 9.1.2 we deduced the weak couplings of the first two generations of quarks ( $u, d$ ), ( $c, s$ ) using the ideas of lepton–quark

symmetry and quark mixing. Historically, the most remarkable thing about these ideas is that they were formulated before the discovery of the charmed quark. In 1971 only seven fundamental fermions were known: the four leptons  $\nu_e, e^-, \nu_\mu, \mu^-$  and the three quarks  $u, d, s$ . This led Glashow, Iliopoulos and Maiani to propose the existence of a fourth quark  $c$  to complete the lepton–quark symmetry and to solve problems associated with neutral currents.<sup>6</sup> The charmed quark was subsequently discovered in 1974 and its measured weak couplings are consistent with the predictions of lepton–quark symmetry, as we have seen.

Since 1974, events have moved on. By 1975 there were six known leptons

$$\begin{pmatrix} e^- \\ \nu_e \end{pmatrix}, \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix}, \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix}, \quad (9.36)$$

and by 1977 five known quarks

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} b \\ t \end{pmatrix},$$

so that once again an extra quark – the top quark – was needed to restore lepton–quark symmetry. By 1994, the mass of this quark had been predicted to be

$$m_t = 170 \pm 30 \text{ GeV}/c^2, \quad (9.37)$$

by arguments based on small effects in the unified theory of electroweak interactions (see Section 10.1.2). The top quark was finally detected at Fermilab in 1995 with a mass

$$m_t \simeq 180 \text{ GeV}/c^2 \quad (9.38)$$

compatible with the prediction of (9.37).

The rest of this section is devoted to the third generation of quarks  $b$  and  $t$ , and to the experiments that first established the existence of the latter.

### 9.2.1 More quark mixing

Lepton–quark symmetry was applied to the first two generations (9.10) and (9.11) in Section 9.1.2. In doing so, we allowed for mixing between the  $d$  and  $s$  quarks according to (9.15), which is conveniently rewritten in matrix form<sup>7</sup>

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos \theta_C & \sin \theta_C \\ -\sin \theta_C & \cos \theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}. \quad (9.39)$$

<sup>6</sup> Neutral currents are discussed in Section 10.1, and these problems in particular in Section 10.1.1.

<sup>7</sup> This is the most general form of mixing possible between two generations (see Problem 9.9).

When the third generation of quarks ( $t$ ,  $b$ ) is included, we must allow for the possibility of mixing between all three ‘lower’ quarks  $d$ ,  $s$  and  $b$  instead of just the first two. This is done by generalising (9.39) to

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \quad (9.40)$$

where the so-called *CKM matrix*<sup>8</sup>  $V_{\alpha\beta}$  ( $\alpha = u, c, t$ ;  $\beta = d, s, b$ ) must be unitary to ensure that  $d'$ ,  $s'$  and  $b'$  are orthonormal single-quark states, like  $d$ ,  $s$  and  $b$ . Lepton–quark symmetry is then applied to the doublets

$$\begin{pmatrix} u \\ d' \end{pmatrix}, \begin{pmatrix} c \\ s' \end{pmatrix}, \begin{pmatrix} t \\ b' \end{pmatrix} \quad (9.41)$$

in order to obtain values for the  $\alpha\beta W$  couplings  $g_{\alpha\beta}$ , where again  $\alpha = u, c, t$  and  $\beta = d, s, b$ . The argument is analogous to that used earlier following Equation (9.16) and gives

$$g_{\alpha\beta} = g_W V_{\alpha\beta} \quad (\alpha = u, c, t; \beta = d, s, b). \quad (9.42)$$

To progress further, we must consider the values of the CKM matrix elements  $V_{\alpha\beta}$ . The general form of this matrix, analogous to (9.39) for two generations, is quite complicated and will be considered later, in Section 11.3. However, in the limit that the mixing between the  $b$  quark and the  $(d, s)$  quarks can be neglected, it reduces to

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \approx \begin{pmatrix} \cos \theta_C & \sin \theta_C & 0 \\ -\sin \theta_C & \cos \theta_C & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (9.43)$$

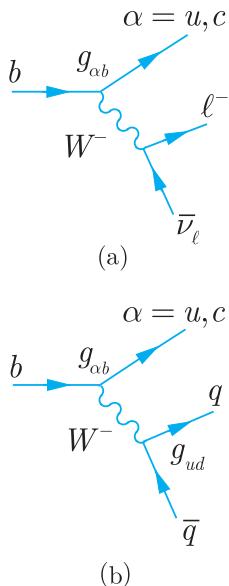
which is equivalent to our previous result (9.39) together with

$$b' = b. \quad (9.44)$$

This must be at least a good first approximation, since its predictions (9.22) for the couplings of the  $u$ ,  $d$ ,  $c$  and  $s$  quarks agree with the experimental data within errors,<sup>9</sup> as we saw in Section 9.1.2.

The accuracy of the approximation (9.43) can be very sensitively tested by considering decays of the  $b$  quark and of hadrons containing it. The  $b$  quark has a mass of approximately  $4.5 \text{ GeV}/c^2$ . It can decay to lighter quarks by the mechanisms of Figure 9.19. As can be seen, all of these decay modes have rates that are proportional to one of the squared couplings

$$|g_{ub}|^2 = |V_{ub}|^2 g_W^2 \quad \text{or} \quad |g_{cb}|^2 = |V_{cb}|^2 g_W^2. \quad (9.45)$$



**Figure 9.19** The dominant decays of the  $b$  quark to lighter quarks and leptons. Here  $\ell = e, \mu$  or  $\tau$  and  $q\bar{q} = d\bar{u}, d\bar{c}, s\bar{u}$  or  $s\bar{c}$ .

<sup>8</sup> The initials stand for Cabibbo, Kobayashi and Maskawa. The last two authors extended the original Cabibbo scheme to include the third generation of quarks.

<sup>9</sup> If the predictions (9.17) were exact, then (9.43) would follow exactly since the CKM matrix  $V_{\alpha\beta}$  is a unitary matrix (cf. Problem 9.10).

Hence the  $b$  quark is stable in the approximation (9.43), since both of these couplings vanish.

Experimentally,  $b$  decays occur at rates that can be inferred from the decays of hadrons containing  $b$  quarks and that imply

$$\tau_b \approx 10^{-12} \text{ s} \quad (9.46)$$

for the  $b$  quark lifetime. While this is not infinite, it is very long compared with what one would expect if the couplings (9.45) were not suppressed, that is if

$$g_{ub} = g_{cb} = g_W. \quad (9.47)$$

To see this, we compare the decays of the  $b$  quark shown in Figure 9.19 with the decays of the lighter  $\tau$  lepton shown in Figure 9.20. As can be seen, all the mechanisms are identical in the approximation (9.47) and the difference in the total decay rates of the  $\tau$  and  $b$  can only arise from the different masses of the particles involved. This contributes in two ways. Firstly, by counting the numbers of different final states in Figures 9.19 and 9.20, it can be seen that for each  $\tau$  lepton decay there are  $N$  decays of the  $b$  quark possible, where  $N = 3$  and 4 for diagrams (a) and (b), respectively. Secondly, and more importantly, weak decay rates increase rapidly with the energy released in the decay. If the masses of the final-state particles can be neglected, the decay rates increase as the fifth power of the mass of the decaying particle, as shown for leptonic  $\tau$  decay in Section 2.2.2. In this simple but rough approximation<sup>10</sup> we obtain

$$\tau_b \approx \frac{1}{N} \left( \frac{m_\tau}{m_b} \right)^5 \tau_\tau \approx 10^{-15} \text{ s}$$

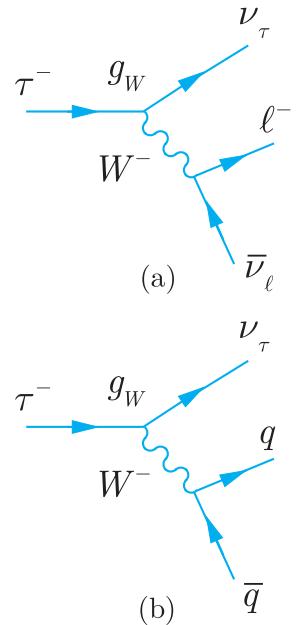
on substituting the experimental  $\tau$  lifetime  $\tau_\tau \approx 3 \times 10^{-13} \text{ s}$ . This is much shorter than the observed lifetime (9.46), implying that the couplings (9.45) are heavily suppressed relative to (9.47), as stated. Furthermore, this qualitative conclusion is confirmed by a more thorough analysis of the various decay modes, which yields

$$|V_{ub}|^2 = (1.7 \pm 0.4) \times 10^{-5} \quad \text{and} \quad |V_{cb}|^2 = (1.7 \pm 0.1) \times 10^{-3}. \quad (9.48)$$

In other words, while the neglect of mixing between the  $b$  quark and the ( $d, s$ ) quarks is not exact, it is a very good first approximation.

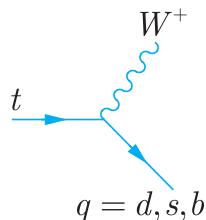
## 9.2.2 Properties of the top quark

The properties of the top quark differ markedly from those of the other quarks because it is so much heavier. In particular it is much



**Figure 9.20** The dominant decays of the  $\tau$  lepton to quarks and leptons. Here  $\ell = e$  or  $\mu$  and  $q\bar{q} = d\bar{u}$  or  $s\bar{u}$ .

<sup>10</sup> A more accurate treatment of the leptonic decay modes is given in Problem 9.11.



**Figure 9.21** The mechanism for  $t$  quark decays. The decays that lead to  $b$  quarks are overwhelmingly the most important.

heavier than the  $W^\pm$  bosons. Hence it can decay by the first-order weak interaction

$$t \rightarrow q + W^+ \quad (q = d, s, b)$$

with rates proportional to the squared couplings  $|g_{tq}|^2$ , as shown in Figure 9.21. In the approximation (9.43), these couplings are predicted by (9.47) to be

$$g_{td} = 0, \quad g_{ts} = 0, \quad g_{tb} = g_W.$$

Hence the only significant decay mode is<sup>11</sup>

$$t \rightarrow b + W^+, \quad (9.49)$$

with a rate proportional to  $\alpha_W$ , given by (9.9). A crude dimensional estimate of the decay width  $\Gamma \sim \alpha_W m_t \sim 1 \text{ GeV}$  is enough to establish that the top quark is highly unstable. A full calculation of Figure 9.21 for  $m_t = 180 \text{ GeV}/c^2$  leads to the prediction  $\Gamma \approx 1.7 \text{ GeV}$ , with a corresponding lifetime

$$\tau = \Gamma^{-1} \approx 4 \times 10^{-25} \text{ s}. \quad (9.50)$$

The predicted lifetime (9.50) is at the crux of top quark physics. By relativity, a hadron state of diameter  $d \approx 1 \text{ fm}$ <sup>12</sup> cannot form in a time less than  $t \approx d/c = \mathcal{O}(10^{-23} \text{ s})$ . The other five quarks  $u$ ,  $d$ ,  $s$ ,  $c$  and  $b$  have lifetimes of order  $10^{-12} \text{ s}$  or more, and there is plenty of time for them to form hadrons that can be observed in the laboratory. In contrast, when top quarks are created they decay too rapidly to form observable hadrons. Instead they decay by (9.49) to give a  $b$  quark and a  $W$  boson, which in turn decays predominantly to either light quarks<sup>13</sup>

$$W^+ \rightarrow q_1 + \bar{q}_2 \quad (q_1 \bar{q}_2 = u\bar{d}, u\bar{s}, c\bar{d} \text{ or } c\bar{s}) \quad (9.51a)$$

or leptons

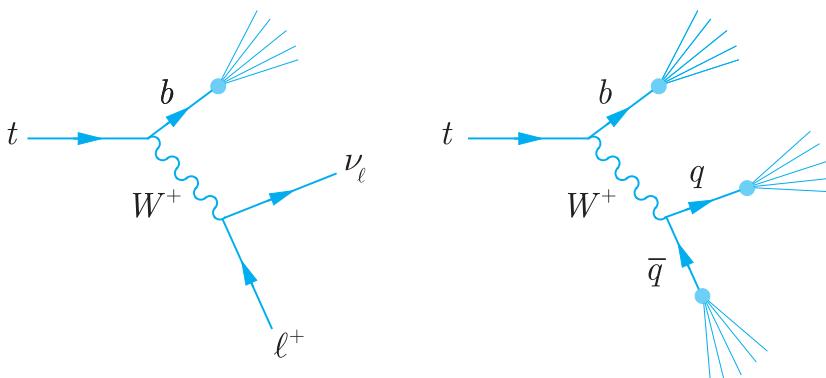
$$W^+ \rightarrow \ell^+ + \nu_\ell \quad (\ell = e, \mu, \tau). \quad (9.51b)$$

Furthermore, the quarks released in these decays are not seen directly, but ‘fragment’ into jets of hadrons. This is illustrated in Figure 9.22, which shows the observable final states resulting from top quark decay.

<sup>11</sup>When deviations from the approximation (9.43) are taken into account, about 2 decays in 1000 lead to a  $d$  or  $s$  quark instead.

<sup>12</sup>A state formed of top quarks only (e.g.  $t\bar{t}$ ) would be much smaller than this. This case is dealt with separately in Problem 9.12.

<sup>13</sup>In identifying the dominant decays, we neglect mixing between the  $b$  and  $(d, s)$  quarks.



**Figure 9.22** Production of hadron jets from the decay  $t \rightarrow b + W^+$ , where the  $W$  boson decays to give either leptons or hadrons.

### \*9.2.3 Discovery of the top quark

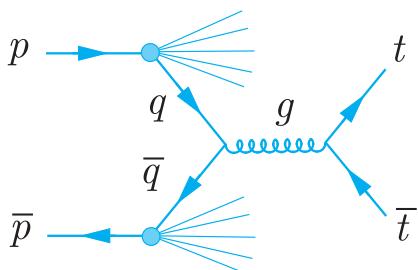
Top quarks were first produced in pairs in the reaction

$$p + \bar{p} \rightarrow t + \bar{t} + X^0, \quad (9.52)$$

where  $X^0$  is an arbitrary hadronic state allowed by the conservation laws. The  $t\bar{t}$  pairs were identified by their subsequent decay products. The dominant decays of the  $t$  quark are shown in Figure 9.22, while the  $\bar{t}$  quark decays by the corresponding antiparticle reactions. Clearly the final state resulting from the initial  $t\bar{t}$  pair can be very complicated and difficult to identify in the presence of backgrounds from other processes. In addition, very high energies are required if such pairs are to be produced at a reasonable rate in the reaction (9.52). The dominant mechanism for this is shown in Figure 9.23 and involves the quark–antiquark annihilation process<sup>14</sup>

$$q + \bar{q} \rightarrow t + \bar{t}. \quad (9.53)$$

This can only occur if the total energy of the  $q\bar{q}$  pair is at least  $2m_t \approx 360$  GeV, corresponding to the top quark and antiquark being produced at rest. Since each initial quark (or antiquark) carries only

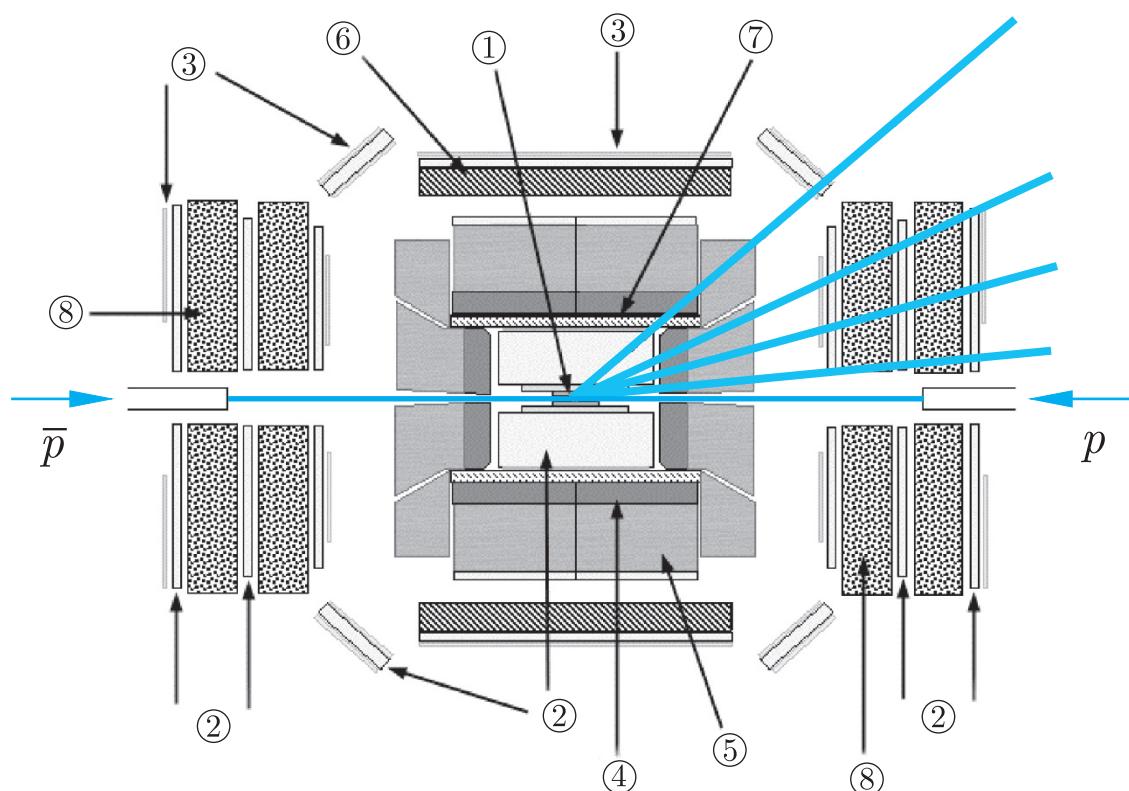


**Figure 9.23** The dominant mechanism for top quark production in proton–antiproton collisions at 1.8 TeV, the energy of the  $p\bar{p}$  collider at Fermilab.

<sup>14</sup>This mechanism is very similar to the mechanism for  $W^\pm$  and  $Z^0$  production shown in Figure 4.19. Consequently, there are several similarities between the present discussion and that of the discovery of the  $W^\pm$  and  $Z^0$  bosons in Section 4.5.1.

a fraction of the initial proton (or antiproton) energy, the energy of the  $p\bar{p}$  system must be much higher if a reasonable reaction rate is to be obtained.

These problems were first overcome by two experiments at Fermilab in 1995. In both cases, proton and antiproton beams each had an energy of 900 GeV, corresponding to a total centre-of-mass energy of 1.8 TeV overall. Like the UA1 detector described in Section 4.5.1, both detectors could reconstruct complex multiparticle events and both could detect and identify all long-lived particles except neutrinos. One of these experiments used the so-called CDF detector. This is shown schematically in Figure 9.24. Its dimensions were approximately 12 m in all three directions and it weighed about 5000 tonnes. The beams of protons and antiprotons, shown as thin coloured lines, enter from each end through focusing quadrupole magnets and interact in the central region. The thicker blue lines indicate some



**Figure 9.24** The CDF detector used at the Tevatron collider at Fermilab. The various subdetectors shown are: (1) a silicon vertex detector; (2) drift chambers; (3) scintillation counters; (4) electromagnetic calorimeters; (5) hadron calorimeters. Steel shielding (6) is also shown, while magnetic fields are produced by a solenoid coil (7) and magnetic toroids (8). (Adapted from a Fermilab graphic, reproduced with permission.)

particles produced in the collision. In the central region there is a silicon vertex detector (1) to detect the decay of very short-lived particles. The intersection point is surrounded by a 2000 tonne detector system, consisting of a tracking detector composed of inner drift chambers (2), electromagnetic calorimeters (4), hadron calorimeters (5) time-of-flight detectors (not indicated) and further drift chambers (2) on the outside to detect muons. The whole system is in a magnetic field with the solenoid coil shown at (7) and steel shielding at (6). The rest of the detector consists of two symmetrical sets of drift chambers (2) sandwiched between scintillation counters (3) and magnetic toroids (8) to provide momentum measurements, primarily for muons.

We will discuss how the CDF detector was used to identify a particular class of events, where top quarks decay by (9.49) to give

$$t + \bar{t} \rightarrow b + W^+ + \bar{b} + W^- . \quad (9.54)$$

One of the  $W$  bosons then decays by (9.51a) to give light quarks, while the other decays by (9.51b) to give either an electron or a muon.<sup>15</sup> The net result is therefore either

$$t + \bar{t} \rightarrow \ell^+ + \nu_\ell + q + \bar{q} + b + \bar{b} \quad (\ell = e, \mu) \quad (9.55a)$$

or

$$t + \bar{t} \rightarrow \ell^- + \bar{\nu}_\ell + q + \bar{q} + b + \bar{b} \quad (\ell = e, \mu) \quad (9.55b)$$

where the quarks manifest themselves as hadron jets like those in Figure 7.14. The reconstruction of a typical event seen in the CDF detector is shown in Plate 7.

We shall proceed as follows: firstly, we identify an initial experimental signal corresponding to the desired events; then we consider what background processes could give rise to a similar signal and how they can be eliminated; finally, we present the result.

### \*9.2.3(a) Initial event selection<sup>16</sup>

For each top quark event of the type (9.52) there are more than  $10^9$  events in which hadrons alone are produced. The extraction of a signal in the presence of this background is only possible because the

<sup>15</sup>A top quark signal was also identified corresponding to events in which both  $W$ -bosons decayed leptonically. However, there were fewer events in this case and they did not lead to a good determination for the top quark mass. Events in which both  $W$ -bosons decay to give light quarks are much more difficult to separate from ‘background’ processes which give rise to a similar experimental signal.

<sup>16</sup>The method of selection described here is very similar to that used to select events in which a  $W^\pm$  boson is produced in reaction (4.27) and decays to leptons by (4.28) (cf. Section 4.5.1). The essential difference is that we now also require the presence of hadron jets.

top quarks are heavy and are produced with relatively low momenta. Because of this, their decay products are often emitted with large momenta at large angles to the initial beam direction. It is extremely rare for the hadrons produced in proton–antiproton collisions to be emitted with these characteristics. Hence the overwhelming majority of background events can be eliminated if events are selected that contain the combinations

$$\ell^+ + \nu_l + N \text{ jets} \quad (\ell = e, \mu) \quad (9.56\text{a})$$

and

$$\ell^- + \bar{\nu}_l + N \text{ jets} \quad (\ell = e, \mu), \quad (9.56\text{b})$$

where  $N \geq 3$  and the leptons and jets are all required to have large momenta transverse to the incoming beams.<sup>17</sup> Of course neutrinos cannot be observed directly. However, since they are the only long-lived particles that are not detected, their presence can be revealed by summing the transverse momenta  $p_T$  of all the observed particles. If this sum is not zero within errors, as required by momentum conservation, the ‘missing transverse momentum’ is ascribed to a neutrino.

Two further comments on the initial trigger (9.56) are required before we go on to see whether it can be produced in other ways. The first is that while (9.55) gives rise to four quarks, they will not always give rise to four distinct jets with high  $p_T$ . Sometimes one or more jets will be emitted close to the beam direction, where there are many hadrons from other sources. Sometimes two jets will be emitted in more or less the same direction and mimic a single jet in the detector. The trigger specifies events in which at least three distinct jets emerge at high transverse momentum.<sup>18</sup> The second point is that a high-momentum lepton could arise from the decay  $Z^0 \rightarrow \ell^+ \ell^-$  of a produced  $Z^0$  boson, rather than a  $W$ -boson decay. In this case, the lepton would be part of an  $\ell^+ \ell^-$  pair with invariant mass equal to the  $Z^0$  mass. Such events were removed from the sample.

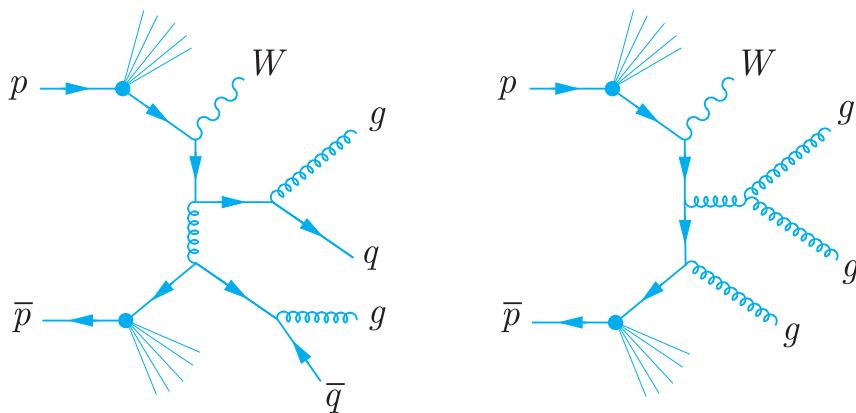
### \*9.2.3(b) Background events

So far we have identified a distinctive class of events that can arise from the production and decay of top quarks. As in all experiments,

---

<sup>17</sup>In the CDF experiment, the leptons were required to have transverse energies greater than 20 GeV. The jets were required to have transverse momenta greater than 15 GeV/c and their longitudinal momenta in the beam direction were restricted to ensure that they were emitted at a reasonably large angle to the beam direction.

<sup>18</sup>In complex reactions there is always a balance to be struck between a very precise trigger (e.g. four jets observed) with a smaller number of events and a looser trigger (e.g. at least two jets observed) giving rise to more severe background problems.



**Figure 9.25** Reactions involving the subprocesses (9.57) that contribute to the background for top quark production. The quarks and gluons are observed as jets.

it is necessary to consider whether such events could arise from other ‘background’ processes. In the present case, the most important backgrounds arise when the subprocess (9.53)

$$q + \bar{q} \rightarrow t + \bar{t}$$

is replaced by a subprocess of the type

$$q + \bar{q} \rightarrow W^\pm + (N \geq 3) \text{ jets.} \quad (9.57)$$

Examples of such processes, corresponding to

$$q + \bar{q} \rightarrow W + q + \bar{q} + g + g \quad (9.58a)$$

and

$$q + \bar{q} \rightarrow W + g + g + g, \quad (9.58b)$$

are shown in Figure 9.25. If the  $W$  boson decays to leptons, such processes can give rise to events that satisfy the trigger (9.56). Theoretical calculations and experimental observations both indicate that the ratio of events corresponding to this background to those for the ‘signal’ of top quarks is about 4:1.

### \*9.2.3(c) *b*-Jet tagging

Background events of the type (9.57) do not usually contain any jets associated with  $b$  quarks. This is illustrated by the examples of Figure 9.24, where the jets arise either from gluons or from quarks or antiquarks that originate from the proton or antiproton. In contrast, the top quark reactions (9.55) do invariably give rise to  $b$  quarks. Hence the signal can be considerably enhanced relative to the background if ‘ $b$  quark jets’ can be distinguished from jets arising from other quarks and gluons. This is called ‘ $b$ -jet tagging’.

One of the most successful methods of  $b$ -jet tagging relies on the fact that  $b$ -jets nearly always contain a fast-moving hadron with non-zero bottom  $\tilde{B} \neq 0$ . Such particles are characterised by decays to many-particle final states, with lifetimes of about  $10^{-12}$  s. Hence

**Table 9.1** The number of lepton +  $N$ -jet events of the type (9.56) observed with and without  $b$ -jets. (Data from Abe *et al.*, 1995.)

N	Observed events	Observed b-jet tags	Background tags expected
1	6578	40	$50 \pm 12$
2	1026	34	$5021 \pm 6.5$
3	164	17	$5.2 \pm 1.7$
4	39	10	$1.5 \pm 0.4$

$b$ -jets typically contain multi-prong decay vertices close to the production vertex. Other jets do not usually contain such vertices.<sup>19</sup> It is not always possible to resolve such vertices from the production vertices, even with silicon vertex detectors of the type discussed in Section 4.4.3. In the CDF experiment, however, it was possible to identify at least one of the  $b$ -jets in a  $t\bar{t}$  event with about 40% efficiency. Demanding such a trigger therefore reduces the signal by a factor of order 2, but reduces the background by a factor of order 20.

#### \*9.2.3(d) Final results

The number of events of the type (9.56) obtained by CDF, with and without the  $b$ -jet tag, is shown in Table 9.1 for various numbers of observed hadron jets. As can be seen, there is a large signal compared with the expected background in the  $N = 3$  and  $N = 4$  jet cases with a  $b$ -jet tag. Since the observed dependence of the background on the various triggers is in good agreement with theoretical expectations, this is compelling evidence for the existence of the top quark. Furthermore, since four-jet events correspond to all the decay products of the top quarks being observed, it is possible to reconstruct the top quark mass from these events. The resulting mass distribution for four-jet events with a  $b$ -jet tag is shown<sup>20</sup> in Figure 9.26. As can be seen, there is a clear peak, corresponding to a top quark mass

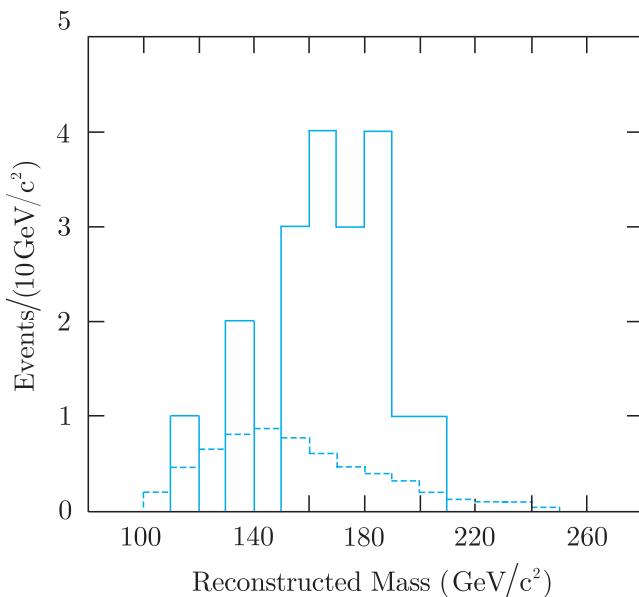
$$m_t = 176 \pm 8 \pm 10 \text{ GeV}/c^2$$

where the first error is statistical and the second is systematic. A similar result was obtained by the second experiment performed at Fermilab, called DØ.

Subsequently, the production and decays of top quark pairs have been studied in great detail in several experiments at both the

<sup>19</sup> Charmed particles decay with a similar lifetime, but to fewer particles.

<sup>20</sup> In this data sample, the cuts on the fourth jet were less stringent than those used in Table 9.1. This resulted in 19 events, of which  $7 \pm 2$  are expected to be background.



**Figure 9.26** The reconstructed  $t$ -quark mass distribution for the  $b$ -tagged ‘four-jet plus lepton’ events of type (9.56). The shape expected for background events (dotted line) is also shown. (F. Abe *et al.* 1995. Reproduced with permission from the American Physical Society.)

Tevatron and the LHC, yielding much more precise results. The current best average value inferred from their decay products using all available data is

$$m_t = 173.21 \pm 0.51 \pm 0.71 \text{ GeV}/c^2, \quad (9.59a)$$

where the first error is statistical and the second is an estimate of the systematic error, while the observed decay width is

$$\Gamma_t = 2.0 \pm 0.5 \text{ GeV} \quad (9.59b)$$

in agreement with the expected lifetime (9.50).

The top mass (9.59a) has by far the smallest percentage error of any of the quark masses. However, it is important to note that, again, it is not the mass of a ‘free quark’. The decay products of a free top quark, if such a thing existed, could not be identified exactly with any collection of leptons and hadron jets, because the quark is coloured, whereas the leptons and hadrons are not, and colour is conserved. Similarly, while the momenta of the observed jets may be a good approximation to the momenta of the coloured quarks with which they are associated, they are not identical, for reasons discussed briefly in Section 7.2.1. The *jet mass* (9.59a) should therefore be regarded as an effective mass that accounts for the invariant mass distributions of the jets and leptons produced in simultaneous decays like (9.55a) to a very good approximation. As in the case of the quarks  $u, d, s, c, b$ , a more fundamental ‘current quark mass’ can be defined, but is less well-determined, being currently estimated to be  $160^{+5}_{-4} \text{ GeV}/c^2$ .

## Problems 9

- 9.1** Show that the vertices of Figure 9.4 are the only possible  $abW$  vertices allowed by charge and lepton number conservation, where  $a$  and  $b$  can be any lepton or antilepton.
- 9.2** The mechanisms of muon decay (9.4) and the semileptonic decay of the charmed quark (9.21a) are essentially the same. Use this to estimate the rate for the charmed quark decay, assuming that its mass is  $1.5 \pm 0.1 \text{ GeV}/c^2$  and the masses of final state particles may be neglected in both decays. (*Hint:* Since  $m_c, m_\mu \ll M_W$ , a zero-range approximation is appropriate in both cases.)
- 9.3** The charmed particle decays,

$$D^+(1869) \rightarrow e^+ + \text{anything}, \quad D^0(1869) \rightarrow e^+ + \text{anything}, \\ \Lambda_c^+(2285) \rightarrow e^+ + \text{anything},$$

are usually assumed to result from the charmed quark decay  $c \rightarrow s + e^+ + \nu_e$  discussed in the previous problem. Assuming that the effects due to quark binding can be neglected, use the data given in the Particle Data Group tables to test this assumption by comparing the measured rates of these decays with each other, and with the estimated rate for the charmed quark decay obtained in the previous problem.

- 9.4** Classify the following semileptonic decays of the  $D^+(1869) = c\bar{d}$  meson as Cabibbo-allowed, Cabibbo-suppressed or forbidden in lowest-order weak interactions, by finding selection rules for the changes in strangeness, charm and electric charge in such decays:

$$\begin{array}{ll} (\text{a}) & D^+ \rightarrow K^- + \pi^+ + e^+ + \nu_e, \quad (\text{b}) & D^+ \rightarrow K^+ + \pi^- + e^+ + \nu_e, \\ & (\text{c}) & D^+ \rightarrow \pi^+ + \pi^+ + e^- + \bar{\nu}_e, \quad (\text{d}) & D^+ \rightarrow \pi^+ + \pi^- + e^+ + \nu_e. \end{array}$$

- 9.5** Use lepton universality and lepton–quark symmetry to estimate the branching ratios for the decays  $b \rightarrow c + e^- + \bar{\nu}_e$  (where the  $b$  and  $c$  quarks are bound in hadrons) and  $\tau^- \rightarrow e^- + \bar{\nu}_e + \nu_\tau$ . Ignore final states that are Cabibbo-suppressed relative to the lepton modes.
- 9.6** Draw quark diagrams for the decays (a)  $D^0 \rightarrow K^- + \pi^+$  and (b)  $D^0 \rightarrow K^+ + \pi^-$ . Estimate the ratio of their decay rates and compare this to the measured values.
- 9.7** Which of the following six decays are allowed in lowest-order weak interactions?

$$\begin{array}{ll} (\text{a}) & K^+ \rightarrow \pi^+ + \pi^+ + e^- + \bar{\nu}_e, \quad (\text{b}) & K^- \rightarrow \pi^+ + \pi^- + e^- + \bar{\nu}_e, \\ & (\text{c}) & \Xi^0 \rightarrow \Sigma^- + e^+ + \nu_e, \quad (\text{d}) & \Omega^- \rightarrow \Xi^0 + e^- + \bar{\nu}_e, \\ & (\text{e}) & \Xi^0 \rightarrow p + \pi^- + \pi^0, \quad (\text{f}) & \Omega^- \rightarrow \Xi^- + \pi^+ + \pi^-. \end{array}$$

- 9.8** Hadronic strangeness-changing weak decays approximately obey the  $\Delta I = 1/2$  rule, that is the total isospin changes by  $1/2$  in the decay. By adding a fictitious strangeness zero  $I = 1/2$  boson  $S$  in the initial state, find the prediction of this rule for the ratio

$$R \equiv \frac{\Gamma(\Xi^- \rightarrow \Lambda + \pi^-)}{\Gamma(\Xi^0 \rightarrow \Lambda + \pi^0)},$$

assuming that the state  $|\Xi^0, S^0\rangle$  is an equal mixture of states with  $I = 0$  and  $I = 1$ .

- 9.9** Show that an arbitrary  $n \times n$  unitary matrix has  $n^2$  real parameters, and hence that

$$\mathbf{U} = e^{-ia} \begin{pmatrix} \cos \theta_C e^{i\beta} & \sin \theta_C e^{i\gamma} \\ -\sin \theta_C e^{-i\gamma} & \cos \theta_C e^{-i\beta} \end{pmatrix} \quad (9.60)$$

is the most general form of a  $2 \times 2$  unitary matrix. The most general form of  $(d, s)$  mixing is

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \mathbf{U} \begin{pmatrix} d \\ s \end{pmatrix}, \quad (9.61)$$

where  $\mathbf{U}$  is an arbitrary  $2 \times 2$  unitary matrix,  $\mathbf{U}^\mathsf{T} = \mathbf{U}^{-1} = 1$ . Show that this can be reduced to the form (9.39) by adjusting the arbitrary phases of the quark states  $s, s'$  and  $d$ .

- 9.10** Show that if (9.17) were exact, then (9.43) would follow exactly, since the CKM matrix  $V_{\alpha\beta}$  is a unitary matrix.

- 9.11** Use the method of dimensions to show that

$$\frac{\Gamma(b \rightarrow q + e^- + \bar{\nu}_e)}{\Gamma(\tau^- \rightarrow \nu_\tau + e^- + \bar{\nu}_e)} = |V_{qb}|^2 \frac{m_b^5}{m_\tau^5} \quad (q = u, c), \quad (9.62)$$

if  $W$ -exchange is approximated by a zero-range interaction and the masses of all final-state particles are neglected. If quark mass corrections are taken into account, this result is modified to

$$\frac{\Gamma(b \rightarrow q + e^- + \bar{\nu}_e)}{\Gamma(\tau^- \rightarrow \nu_\tau + e^- + \bar{\nu}_e)} = |V_{qb}|^2 \frac{m_b^5}{m_\tau^5} f(m_q/m_b), \quad (9.63)$$

where

$$f(x) = 1 - 8x^2 - 24x^4 \ln(x) + 8x^6 - x^8.$$

Use this, together with the constituent quark masses of Table 3.1 and the experimental result

$$\Gamma(b \rightarrow c + e^- + \bar{\nu}_e) + \Gamma(b \rightarrow u + e^- + \bar{\nu}_e) = (6.6 \pm 0.3) \times 10^{10} \text{ s}^{-1}, \quad (9.64)$$

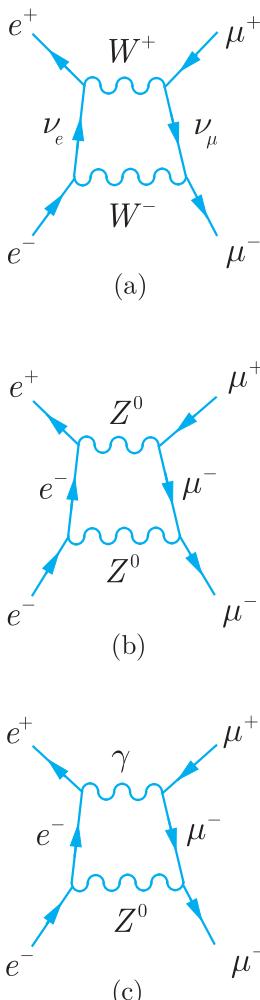
deduced from the decay  $B^- \rightarrow e^- + \bar{\nu}_e + \text{hadrons}$ , to obtain upper bounds on the magnitudes of  $V_{ub}$  and  $V_{cb}$ .

- 9.12** If the top quark were stable, the low-lying states of toponium  $t\bar{t}$  could be approximated by non-relativistic motion in a Coulomb potential (cf. Equation (7.2a))

$$V(r) = -\frac{4}{3} \frac{\alpha_s}{r}$$

with  $\alpha_s \approx 0.1$ . Use the simple Bohr model to calculate the radius of the ground state and the time taken to complete a single Bohr orbit in the ground state. Compare this with the expected lifetime of the top quark.

## Weak interactions: electroweak unification



In this chapter we continue our discussion of weak interactions by turning our attention to the unified theory of weak and electromagnetic processes. This theory was first proposed in order to solve problems associated with Feynman diagrams in which more than one  $W$  boson was exchanged, like that shown in Figure 10.1(a), which contributes to the reaction

$$e^+ + e^- \rightarrow \mu^+ + \mu^- \quad (10.1)$$

Such contributions are expected to be small because they are higher order in the weak interaction. This appears to be confirmed by experimental data, which are in good agreement with theoretical predictions that neglect them entirely. However, when these contributions are explicitly calculated, they are found to be proportional to divergent integrals; that is they are infinite.<sup>1</sup> In the unified theory, this problem is automatically solved when other higher-order diagrams, like those shown in Figures 10.1(b) and (c), are taken into account. These also give infinite contributions, but when all the diagrams of a given order are added together the divergences cancel, giving a well-defined and finite contribution overall. This cancellation is not accidental, but is a consequence of relations between the various coupling constants associated with the  $\gamma$ ,  $W^\pm$  and  $Z^0$  vertices in the unified theory, as we shall see in Section 10.1.2.

We shall say little about the net contributions of higher-order diagrams like those of Figure 10.1 because they are not only finite

<sup>1</sup> These infinites can be removed by modifying the theory in an *ad hoc* manner, but at the price of introducing extra free parameters, which is almost as bad.

**Figure 10.1** Some of the higher-order diagrams that contribute to the reaction  $e^+ + e^- \rightarrow \mu^+ + \mu^-$  in the unified theory of the electroweak interaction.

but very small compared with the dominant lowest-order diagrams. Rather, we shall concentrate initially on lowest-order diagrams, like those shown for the same reaction (10.1) in Figure 10.2, and then go on to other aspects of the unified theory, including the role of the Higgs boson, in Section 10.2.

## 10.1 Neutral currents and the unified theory

This section is devoted to weak neutral current reactions, which are those involving the emission, absorption or exchange of  $Z^0$  bosons. The existence of such reactions was predicted by the unified theory in 1968, several years before their experimental detection. They play a central role in the theory, as we shall see.

### 10.1.1 The basic vertices

In Section 9.1.1 we saw that all the charged current interactions of leptons could be understood in terms of the basic  $W^\pm$ -lepton vertices of Figure 9.4. In the same way, all known neutral current interactions can be accounted for in terms of the basic  $Z^0$ -lepton vertices of Figure 10.3. Like the  $W^\pm$ -lepton vertices, these conserve the lepton numbers  $L_e$ ,  $L_\mu$ ,  $L_\tau$  in addition to the electric charge  $Q$ , and indeed they are the only possible vertices that can be written down that satisfy these conservation laws. The corresponding quark vertices can be obtained from the lepton vertices using lepton–quark symmetry and quark mixing, in the same way that the  $W^\pm$ -quark vertices were obtained from the  $W^\pm$ -lepton vertices in Section 9.1.2. In other words, if we again confine ourselves to the first two generations, the quark vertices are obtained from the lepton vertices of Figure 10.3 by making the replacements

$$\nu_e \rightarrow u, \quad \nu_\mu \rightarrow c, \quad e^- \rightarrow d', \quad \mu^- \rightarrow s', \quad (10.2)$$

where  $d'$  and  $s'$  are the mixtures (9.15). For  $\ell = e$  or  $\mu$ , the lepton vertices of Figure 10.3 can be denoted

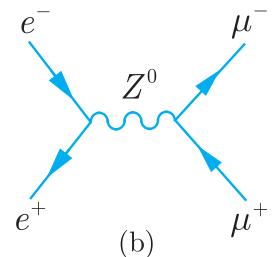
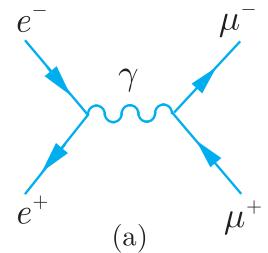
$$\nu_e \nu_e Z^0, \quad \nu_\mu \nu_\mu Z^0, \quad e^- e^- Z^0, \quad \mu^- \mu^- Z^0$$

in an obvious notation. The corresponding quark vertices are therefore

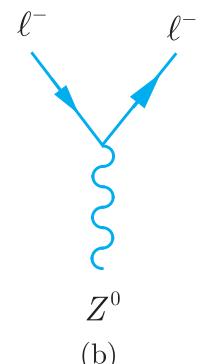
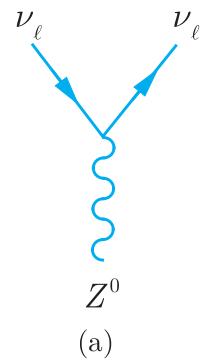
$$uuZ^0, \quad ccZ^0, \quad d'd'Z^0, \quad s's'Z, \quad (10.3)$$

in the same notation, where the latter two must be interpreted in terms of the physical states  $d$  and  $s$ . In Section 9.1.2 we had a vertex

$$ud'W = udW \cos \theta_C + usW \sin \theta_C,$$



**Figure 10.2** The two dominant contributions to the reaction  $e^+ + e^- \rightarrow \mu^+ + \mu^-$  in the unified theory.



**Figure 10.3** The basic  $Z^0$ -lepton vertices, where  $\ell = e, \mu, \tau$ .

which we interpreted as  $udW$  and  $usW$  vertices with relative strengths  $\cos\theta_C$  and  $\sin\theta_C$ , as illustrated in Figure 9.12. In the same way we have

$$\begin{aligned} d'd'Z^0 &= (d \cos\theta_C + s \sin\theta_C)(d \cos\theta_C + s \sin\theta_C)Z^0 \\ &= ddZ^0 \cos^2\theta_C + ssZ^0 \sin^2\theta_C + (dsZ^0 + sdZ^0)\sin\theta_C \cos\theta_C \end{aligned} \quad (10.4a)$$

and

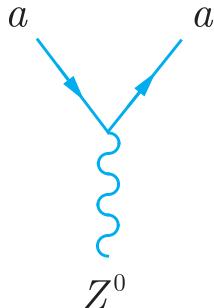
$$\begin{aligned} s's'Z^0 &= (-d \sin\theta_C + s \cos\theta_C)(-d \sin\theta_C + s \cos\theta_C)Z^0 \\ &= ddZ^0 \sin^2\theta_C + ssZ^0 \cos^2\theta_C - (dsZ^0 + sdZ^0)\sin\theta_C \cos\theta_C, \end{aligned} \quad (10.4b)$$

so that the primed vertices  $d'd'Z^0$  and  $s's'Z^0$  both contribute to the vertices  $ddZ^0$ ,  $ssZ^0$ ,  $sdZ^0$  and  $dsZ^0$  for the real particles  $d$  and  $s$ . However, the interesting point is that when both sets of contributions (10.4a) and (10.4b) are combined, one obtains

$$d'd'Z^0 + s's'Z^0 = ddZ^0 + ssZ^0,$$

so that the four vertices (10.3) can be replaced by the equivalent vertices

$$uuZ^0, \quad ccZ^0, \quad ddZ^0, \quad ssZ^0, \quad (10.5)$$



**Figure 10.4** The basic  $Z^0$ -quark vertices, where  $a = u, d, s, c$ .

which are shown in Figure 10.4.

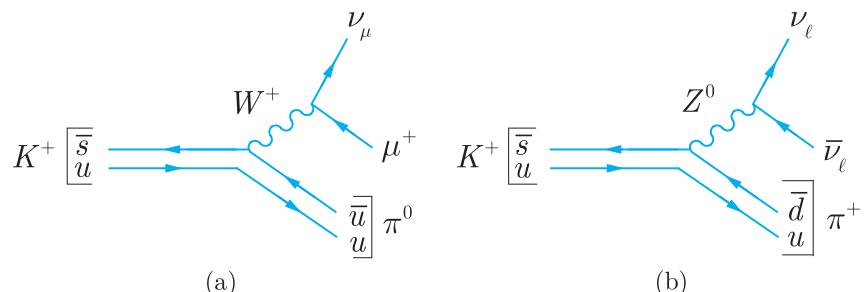
The four vertices of Figure 10.4 conserve both strangeness and charm, whereas the ‘flavour-changing’ vertices  $ucZ^0$  and  $dsZ^0$  do not occur. Thus we arrive at the conclusion that neutral current interactions conserve strangeness and charm, in contrast to the charged current interactions that do not in general conserve these quantum numbers. This result remains valid when mixing between all three generations is considered, and is confirmed by experiment. For example, the strangeness-changing decays

$$K^+ \rightarrow \pi^0 + \mu^+ + \nu_\mu \quad (10.6a)$$

and

$$K^+ \rightarrow \pi^+ + \nu_\ell + \bar{\nu}_\ell, \quad (10.6b)$$

shown in Figure 10.5, differ in that the first involves the allowed charged current vertex  $usW$  of Figure 9.13(a), whereas the second



**Figure 10.5** (a) The allowed decay  $K^+ \rightarrow \pi^0 + \mu^+ + \nu_\mu$  and (b) the forbidden decay  $K^+ \rightarrow \pi^+ + \nu_\ell + \bar{\nu}_\ell$ .

involves the forbidden neutral current vertex  $dsZ^0$ . Experimentally, no forbidden decays (10.6b) have been observed, and the upper limit on their possible decay rates is

$$\frac{\sum_{\ell} \Gamma(K^+ \rightarrow \pi^+ + \nu_{\ell} + \bar{\nu}_{\ell})}{\Gamma(K^+ \rightarrow \pi^0 + \mu^+ + \nu_{\mu})} < 10^{-7}, \quad (10.7)$$

in impressive agreement with the above scheme.

### 10.1.2 The unification condition and the $W^{\pm}$ and $Z^0$ masses

The various coupling constants that occur in electroweak interactions are not independent, but are related by the requirement that the infinities that occur in higher-order diagrams, like those of Figure 10.1, should exactly cancel in all possible processes. This is guaranteed in the unified theory provided that two equations, called the *unification condition* and the *anomaly condition*, are satisfied. The proof of this is formidable and relies on a fundamental symmetry of the theory called *gauge invariance*.<sup>2</sup> However, the form of the equations is simple. The unification condition is

$$\frac{e}{2(2\varepsilon_0)^{1/2}} = g_W \sin \theta_W = g_Z \cos \theta_W, \quad (10.8)$$

where the *weak mixing angle*  $\theta_W$  is given by

$$\cos \theta_W = M_W / M_Z \quad (0 < \theta_W < \pi/2) \quad (10.9)$$

and  $g_Z$  is a coupling constant that characterises the strength of the neutral current vertices<sup>3</sup> of Figures 10.3 and 10.4. This condition explicitly relates the weak and electromagnetic coupling constants. In contrast, the anomaly condition relates the electric charges  $Q_{\ell}$  and  $Q_a$  of the leptons  $\ell$  and quarks  $a$ , and is

$$\sum_{\ell} Q_{\ell} + 3 \sum_a Q_a = 0. \quad (10.10)$$

The sums extend over all leptons  $\ell$  and all quark flavours  $a = u, d, s, \dots$  and the factor 3 arises because there is also a sum over the three quark colour states, which has been explicitly carried out. On substituting for the quark and lepton charges, one easily finds that the anomaly condition is satisfied by the six known leptons and the six known quarks.

The unification condition (10.8) relates the strengths of the various interactions to the  $W$  and  $Z$  masses, and historically was used

<sup>2</sup> Gauge invariance is discussed in Section 10.2.1 and Appendix D.

<sup>3</sup> The strengths are not all equal, but are given by  $g_Z$  multiplied by known constants, which depend on  $\theta_W$  and the vertex in question. Their explicit form for leptons is derived in Appendix D, Section D.7.4.

to predict the latter from the former before the  $W^\pm$  and  $Z^0$  bosons were discovered. In the low-energy limit, the charged current reactions are characterised by the Fermi constant, and on substituting for  $g_W$  from (10.8) one obtains

$$M_W^2 = \frac{\sqrt{2}g_W^2}{G_F} = \frac{\pi\alpha}{\sqrt{2}G_F \sin^2 \theta_W}, \quad (10.11a)$$

which together with (10.9) implies

$$M_Z^2 = \frac{\pi\alpha}{\sqrt{2}G_F \sin^2 \theta_W \cos^2 \theta_W} \quad (10.11b)$$

for the  $Z^0$  mass. The weak mixing angle itself can be determined by comparing neutral and charged current processes at low energies  $E \ll M_W, M_Z$ . In this regime, neutral current processes are characterised by an effective zero-range coupling constant  $G_Z$  in exactly the same way that charged current reactions are characterised by the Fermi coupling constant  $G_F$ .<sup>4</sup> The neutral current coupling is given by

$$\frac{G_Z}{\sqrt{2}} = \frac{g_Z^2}{M_Z^2}, \quad (10.12)$$

by analogy with (2.17) for the Fermi constant  $G_F$ , and the ratio of these low-energy couplings can be expressed in the form

$$\frac{G_Z}{G_F} = \frac{g_Z^2}{g_W^2} \cdot \frac{M_W^2}{M_Z^2} = \sin^2 \theta_W, \quad (10.13)$$

using (10.8) and (10.9). Hence the weak mixing angle may be found by comparing the measured rates of charged and neutral current reactions at low energies, and by 1981 its value was determined in this way to be

$$\sin^2 \theta_W = 0.227 \pm 0.014.$$

This value was used to predict the  $W^\pm$  and  $Z^0$  masses by substitution into (10.11a) and (10.b) and using the values of  $\alpha$  and  $G_F$  given in (1.16) and (2.16), respectively. The resulting values were

$$M_W = 78.3 \pm 2.4 \text{ GeV}/c^2 \quad \text{and} \quad M_Z = 89.0 \pm 2.0 \text{ GeV}/c^2,$$

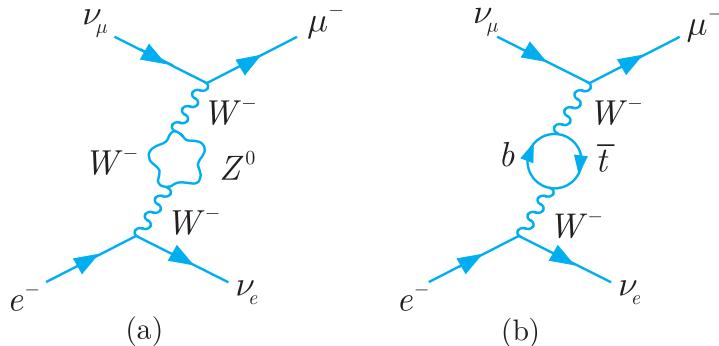
and the subsequent discovery of the  $W^\pm$  and  $Z^0$  bosons with masses compatible with these predictions (as described in Section 4.5.1) is one of the greatest triumphs of the unified theory.

Nowadays, the best value of the weak mixing angle,

$$\sin^2 \theta_W = 0.2313 \pm 0.0001, \quad (10.14)$$

---

<sup>4</sup> This was discussed in Section 2.2.1.



**Figure 10.6** Two of the higher-order contributions to inverse muon decay, which were neglected in obtaining the Fermi coupling constant  $G_F$  as a low-energy limit in Figure 2.7.

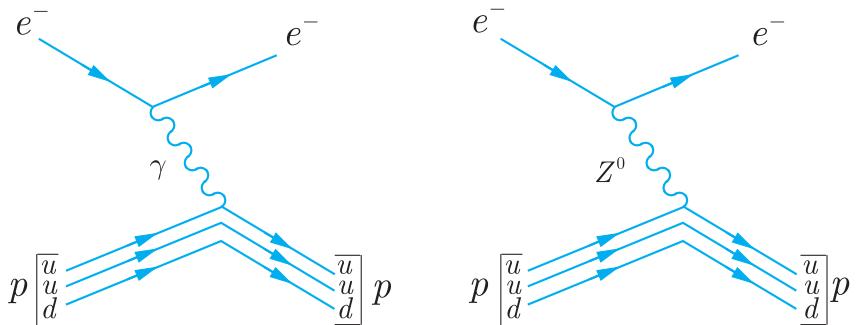
is obtained by comparing the predictions of the unified theory with a wide range of measurements on different neutral current reactions. However, substituting (10.14) into (10.11) gives

$$M_W = 77.57 \pm 0.03 \text{ GeV}/c^2 \quad \text{and} \quad M_Z = 88.47 \pm 0.04 \text{ GeV}/c^2, \quad (10.15)$$

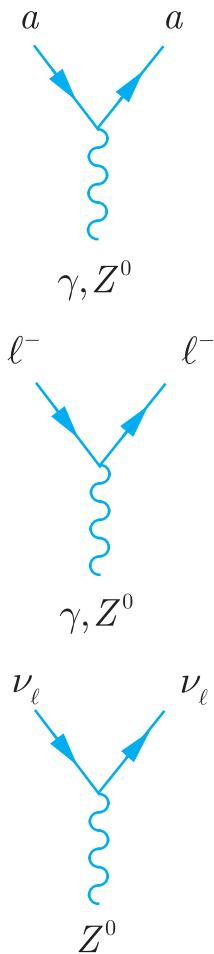
which are not in very good agreement with the best experimental values (9.1). The reason for this is that the above discussion neglects higher-order corrections. For example, in deriving (10.11) we used the relation (2.17) for the Fermi constant  $G_F$ . This latter relation was obtained by taking the low-energy limit of single  $W^\pm$  exchange only, as shown in Figure 2.7, whereas strictly speaking we should have also included the small contributions arising from higher-order diagrams. Two of the most important of these are shown in Figure 10.6. One of these involves an interaction between the  $W^\pm$  and  $Z^0$  bosons, which is predicted by the unified theory, whereas the other involves the  $t$  quark. Hence the magnitude of the higher-order corrections to (10.11), and also to other predictions obtained using just lowest-order diagrams involving single  $W^\pm$  and  $Z^0$  exchange, depends on the mass of the  $t$  quark. We shall not discuss this in detail, but merely state the important result that when higher-order corrections are taken into account, the predictions of the theory agree with experiment in all cases.

### 10.1.3 Electroweak reactions

In any process in which a photon is exchanged, a  $Z^0$  boson can be exchanged as well. This is illustrated for elastic electron–proton scattering in Figure 10.7, and follows in general from the fact that to each of the basic electromagnetic vertices there is a corresponding  $Z^0$  vertex, as illustrated in Figure 10.8. At energy and momentum transfers that are small compared with the  $Z^0$  mass, the  $Z^0$  exchange contribution can be neglected compared with the corresponding photon exchange contributions, and the reactions regarded as purely



**Figure 10.7** Photon and  $Z^0$  boson exchange contributions to electron–proton scattering.



**Figure 10.8** Summary of the  $Z^0$  and  $\gamma$  couplings to leptons and quarks in the unified electroweak theory, where  $\ell = e, \mu, \tau$  and  $a = u, d, s, \dots$

electromagnetic to a high degree of accuracy. However, at very high energy and momentum transfers,  $Z^0$  exchange contributions become comparable with photon exchange, and we are therefore dealing with genuinely electroweak processes which involve both weak and electromagnetic interactions to a comparable degree.

These points are beautifully illustrated by the cross-section for the muon pair production reaction,

$$e^+ + e^- \rightarrow \mu^+ + \mu^-, \quad (10.16)$$

which is shown in Figure 10.9. This is because the main features of this cross-section may be understood in terms of the dominant lowest-order diagrams of Figure 10.2 without detailed calculation, if we neglect the lepton masses compared with the beam energies  $E$  in the overall centre-of-mass frame. A simple dimensional estimate of the one-photon exchange contribution in Figure 10.2(a) then gives

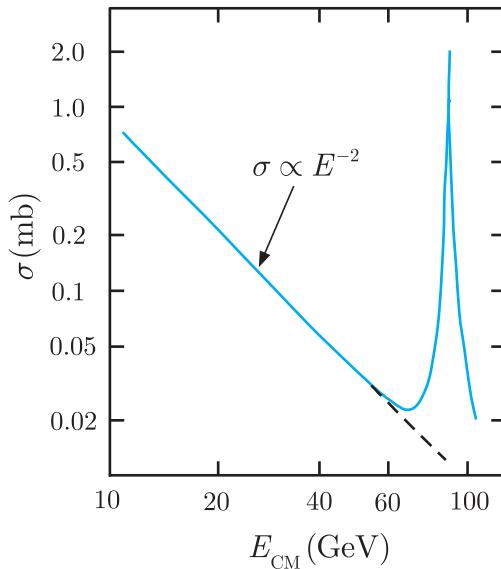
$$\sigma_\gamma \approx \alpha^2/E^2, \quad (10.17)$$

where the factor  $\alpha^2$  follows because the diagram is second order. The energy dependence  $E^{-2}$  follows because the cross-section is an area with natural dimension  $[E^{-2}]$  (see Table 1.1) and the beam energy  $E$  is the only dimensional parameter available if we neglect the lepton masses. The same argument gives

$$\sigma_Z \approx G_Z^2 E^2 \quad (10.18)$$

for the contribution of the  $Z^0$  exchange diagram in Figure 10.2(b) at energies  $E^2 \ll M_Z^2$ , where  $G_Z$  is the effective low-energy constant (10.12), which has natural dimensions  $[E^{-2}]$ . Finally, the ratio of (10.18) and (10.17) gives

$$\frac{\sigma_Z}{\sigma_\gamma} \approx \frac{G_Z^2 E^4}{\alpha^2} \approx \frac{E^4}{M_Z^4} \quad (10.19)$$



**Figure 10.9** Total cross-section for the reaction  $e^+ + e^- \rightarrow \mu^+ + \mu^-$  as a function of the total centre-of-mass energy  $E_{CM}$ . The dashed line shows the extrapolation of the low-energy behaviour (10.17) in the region of the  $Z^0$  peak.

for the relative importance of the two diagrams, where we have used (10.13) and (10.11b) and neglected factors of order unity.<sup>5</sup>

We thus see that the one-photon exchange diagram dominates at low energies, and the cross-section falls as  $E^{-2}$  by (10.17). This is in agreement with the observed behaviour shown in Figure 10.9 and justifies our neglect of the  $Z^0$  exchange contribution in our previous discussion of this process in Section 7.2. However, the relative importance of the  $Z^0$  exchange contribution in Figure 10.2(b) increases rapidly with energy, as can be seen from (10.19), and at beam energies of about 50 GeV it begins to make a significant contribution to the total cross-section. At still higher energies, the cross-section is dominated by a very large peak at a centre-of-mass energy

$$E_{CM} = 2E = M_Z, \quad (10.20)$$

corresponding to the  $Z^0$  mass. At this energy the low-energy approximation (10.18) is irrelevant and Figure 10.2(b) corresponds to the formation of physical  $Z^0$  bosons in the process



followed by their subsequent decay



to give the final-state muons. This is exactly like the formation and decay of a resonance, as shown, for example, in Figure B.4 in

<sup>5</sup> A full calculation gives the same energy dependence as (10.19), but is about a factor of four smaller.

Appendix B, except that we are now dealing with weak rather than strong interactions. As in the latter case, there is a peak in the total cross-section at the particle mass, with a width equal to the total decay rate; that is

$$\Gamma_Z = 1/\tau_Z = 2.49 \text{ GeV}, \quad (10.23)$$

where  $\tau_Z$  is the lifetime of the  $Z^0$  boson. This peak is an enormous boon to experimental physics, as we shall now discuss.

#### 10.1.4 $Z^0$ formation: how many generations are there?

The  $Z^0$  bosons produced in the reaction (10.21) will not only decay to muon pairs, as in (10.22), but also to other final states allowed by the appropriate conservation laws. Hence an electron–positron colliding-beam machine tuned to a centre-of-mass energy corresponding to the  $Z^0$  mass is an ideal ‘ $Z^0$  factory’, in which these particles can be copiously produced and studied. Two such colliders have been built: the Stanford Linear Collider (SLC), which was the first to produce results, and the ‘large electron–positron’ (LEP) collider at CERN, which produced  $Z^0$  bosons much more copiously, at a rate of order  $10^3$  per hour. This enabled the mass, width and decay modes of the  $Z^0$  to be measured with great precision, and in so doing verified that there are only three neutrino types in the sequence  $\nu_e, \nu_\mu, \nu_\tau, \dots$ , provided that their masses are small compared to the mass of the  $Z^0$  bosons. Neutrinos of this type are called *sequential neutrinos*. In the unified theory, they are the only types of neutrino allowed and by the universality of lepton interactions, they all have the same couplings to  $Z^0$  bosons.

The processes observed at the SLC and LEP colliders were

$$e^+ + e^- \rightarrow \ell^+ + \ell^- \quad (\ell = e, \mu, \tau) \quad (10.24)$$

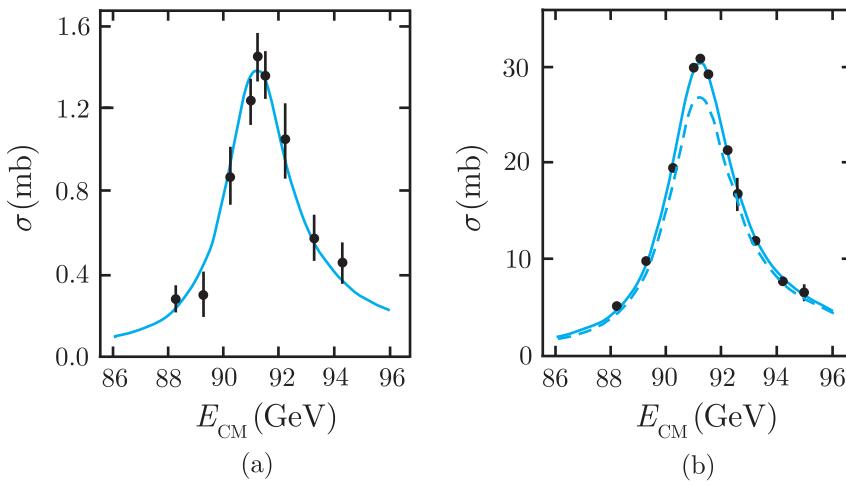
and

$$e^+ + e^- \rightarrow \text{hadrons}, \quad (10.25)$$

where in the latter case a sum over all possible hadron states is implied. In all cases there is a clear peak in the cross-section, as illustrated by the data from one experiment in Figures 10.10(a) and (b). If this peak is interpreted as due to formation and decay of  $Z^0$  bosons, as shown in Figure 10.11, then the cross-section is described by the Breit–Wigner formula (cf. Section 3.5, Equation (3.34c))

$$\sigma(e^+ + e^- \rightarrow X) = \frac{12\pi M_Z^2}{E_{CM}^2} \left[ \frac{\Gamma(Z^0 \rightarrow e^+ + e^-)\Gamma(Z^0 \rightarrow X)}{(E_{CM}^2 - M_Z^2)^2 + M_Z^2\Gamma_Z^2} \right], \quad (10.26)$$

where  $\Gamma(Z^0 \rightarrow X)$  is the decay rate of the  $Z^0$  to the observed final state  $X$ ,  $\Gamma_Z$  is the total decay rate (10.23), and electron masses are neglected compared to the centre-of-mass energy. We therefore see



**Figure 10.10** Measured cross-sections for (a)  $e^+ + e^- \rightarrow \mu^+ + \mu^-$  and (b)  $e^+ + e^- \rightarrow$  hadrons, in the region of the  $Z^0$  peak. The solid lines show the predictions of the standard model on the assumption that there are three types of light neutrinos; the dashed line in (b) shows the prediction on the assumption that there are four such neutrinos. (Adapted from Akravy et al. 1990. Reproduced with permission from Elsevier.)

that, while the positions and widths of the peaks in (10.26) are determined by the  $Z^0$  mass and total width  $\Gamma_Z$ , their heights are proportional to the products of branching ratios

$$B(Z^0 \rightarrow e^+ + e^-)B(Z^0 \rightarrow X) \equiv \frac{\Gamma(Z^0 \rightarrow e^+ + e^-)}{\Gamma_Z} \frac{\Gamma(Z^0 \rightarrow X)}{\Gamma_Z}. \quad (10.27)$$

Hence, by fitting the observed data, one obtains not only precise values

$$M_Z = 91.188 \pm 0.002 \text{ GeV}/c^2 \quad (10.28a)$$

and

$$\Gamma_Z = 2.495 \pm 0.002 \text{ GeV} \quad (10.28b)$$

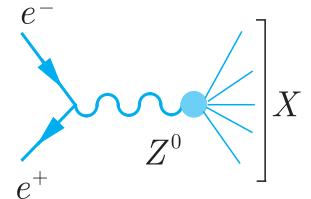
for the mass and total decay rate,<sup>6</sup> but also for the various partial decay rates  $\Gamma(Z^0 \rightarrow X)$ . In this way the hadronic decay rate is found to be

$$\Gamma(Z^0 \rightarrow \text{hadrons}) = 1.744 \pm 0.002 \text{ GeV}, \quad (10.29a)$$

while the leptonic decay rates are all consistent with the value

$$\Gamma(Z^0 \rightarrow \ell^+ + \ell^-) = 0.0840 \pm 0.0009 \text{ GeV}, \quad (10.29b)$$

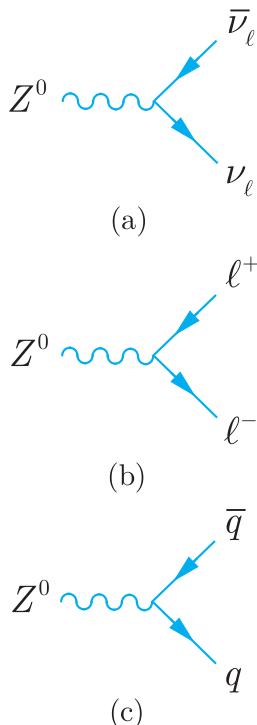
independent of the identity of the lepton type  $\ell = e, \mu, \tau$ .<sup>7</sup>



**Figure 10.11** Formation of a  $Z^0$  and its decay to an arbitrary state  $X$ .

<sup>6</sup> The values quoted in (10.28a) and (10.28b) are averages of several different determinations.

<sup>7</sup> This is exactly what is expected from the universality of leptonic interactions if the masses of the final state leptons can be neglected compared with the  $Z^0$  mass. This is in fact a very good approximation (cf. the discussion of lepton decays in Section 2.2.2 and  $W$  decays in Section 9.1.3).



**Figure 10.12** Feynman diagrams for the  $Z^0$  decays (10.30a), (10.30b) and (10.30c), respectively.

The final states observed in the reactions (10.24) and (10.25) account for just over 80% of all  $Z^0$  decays, as can be seen by comparing the observed decay rates (10.29a) and (10.29b) with the total decay rate (10.28b). The remaining decays are ascribed to final states containing only neutrinos. To see this, we note that in the unified theory, the only  $Z^0$  decays that can arise from the basic vertices of Figures 10.3 and 10.4 are

$$Z^0 \rightarrow \nu_\ell + \bar{\nu}_\ell \quad (\ell = e, \mu, \tau), \quad (10.30a)$$

$$Z^0 \rightarrow \ell^+ + \ell^- \quad (\ell = e, \mu, \tau) \quad (10.30b)$$

and

$$Z^0 \rightarrow q + \bar{q}, \quad (10.30c)$$

corresponding to the Feynman diagrams of Figures 10.12(a), (b) and (c), respectively. In the last of these cases, the quark pair is not seen directly, but ‘fragments’ into two or more jets of hadrons that are observed in the final state, as shown in Figure 10.13. On adding up the various decays we thus obtain

$$\Gamma_Z = \Gamma(Z^0 \rightarrow \text{hadrons}) + 3\Gamma(Z^0 \rightarrow \ell^+ \ell^-) + N_\nu \Gamma(Z^0 \rightarrow \nu_\ell \bar{\nu}_\ell) \quad (10.31)$$

for the total decay rate, where we have allowed for an arbitrary number  $N_\nu$  of sequential neutrinos,  $\nu_e, \nu_\mu, \nu_\tau, \dots$  but only three charged leptons  $e, \mu$  and  $\tau$ , for reasons to be discussed shortly. The total neutrino contribution is therefore

$$\begin{aligned} N_\nu \Gamma(Z^0 \rightarrow \nu_\ell \bar{\nu}_\ell) &= \Gamma_Z - \Gamma(Z^0 \rightarrow \text{hadrons}) - 3\Gamma(Z^0 \rightarrow \ell^+ \ell^-) \\ &= 0.499 \pm 0.004 \text{ GeV} \end{aligned} \quad (10.32)$$

where we have substituted the experimental values (10.28), (10.29a) and (10.29b). The decay rate to neutrino pairs cannot, of course, be measured directly, but can be calculated from the diagram in Figure 10.12(a) and is found to be<sup>8</sup>

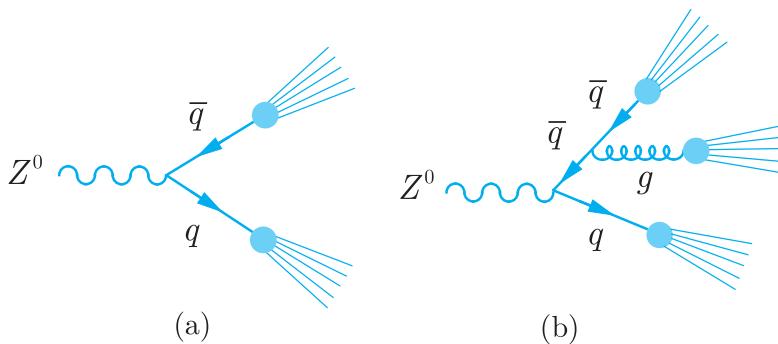
$$\Gamma(Z^0 \rightarrow \nu_\ell \bar{\nu}_\ell) = 0.166 \text{ GeV}. \quad (10.33)$$

This is compatible with the measured value (10.32) if, and only if, the number of sequential neutrinos  $N_\nu = 3$ .

This is a very significant result. Three generations of fermions ( $e^-, \nu_e, u, d$ ), ( $\mu^-, \nu_\mu, c, s$ ) and ( $\tau^-, \nu_\tau, t, b$ ) are known to exist, and each individually satisfies the anomaly condition (10.10). That is,

$$Q_e + Q_{\nu_e} + 3(Q_u + Q_d) = 0,$$

<sup>8</sup> This result is given in, for example, pp. 441 and 448 of Mandl and Shaw (2010). The decay rates  $\Gamma(Z^0 \rightarrow \ell^+ \ell^-)$  and  $\Gamma(Z^0 \rightarrow \text{hadrons})$  can also be calculated from the Feynman diagrams of Figures 10.12(b) and 10.13, respectively, giving results in agreement with the measured values (10.29).



**Figure 10.13** Decay of the  $Z^0$  into (a) two and (b) three hadron jets.

with similar results for the second and third generations. Within the standard model, there are no restrictions on adding more generations with the same properties. It was for a long time a matter of speculation whether such extra generations existed, since neutrinos are difficult to detect, and additional charged leptons and quarks would have escaped detection if they were very heavy. In particular, any such charged leptons must be too heavy to contribute to  $Z^0$  decays of the type (10.30b) or they would have been detected. However, if additional neutrinos existed in the sequence  $\nu_e, \nu_\mu, \nu_\tau, \dots$ , with masses much less than  $M_Z$ , then the  $Z^0$  would decay to them with a rate given by (10.33) according to the unified electroweak theory. Thus, in the framework of this theory, the result (10.32) restricts the number of lepton generations to three if we assume that neutrino masses are not very large. The same result can be illustrated graphically by simply comparing the predictions of the unified theory directly with the data of Figure 10.10 for the two cases  $N_\nu = 3$  and 4. From (10.31) and (10.33) we see that the total width  $\Gamma_Z$  will be increased by 0.166 GeV if  $N_\nu = 4$ , and the product of branching ratios (10.27) will be correspondingly reduced. Consequently, the peaks in the cross-sections (10.26) are predicted to be broader and lower if  $N_\nu = 4$ . This is in clear disagreement with the data, as shown in Figure 10.10.

We thus see that, within the standard model, of which the unified electroweak theory is a part, there can only be three generations of leptons and quarks if neutrinos are assumed to be light compared with the  $Z^0$  mass.

## 10.2 Gauge invariance and the Higgs boson

Gauge invariance is a fundamental symmetry associated with theories in which the force carriers are spin-1 bosons. It plays an important role in the unified electroweak theory, where it is needed to ensure the cancellation of the divergences that occur in individual Feynman diagrams, as mentioned in Section 10.1. In addition,

because the  $W^\pm$  and  $Z^0$  bosons have non-zero masses, it leads to the prediction of a new spin-0 boson – the *Higgs boson* – which, after a huge experimental programme, was finally detected in 2012. This section is devoted to a qualitative introduction to gauge invariance and its consequences for weak interactions and the unified theory. A fuller account, including mathematical details, is given in Appendix D.

There are different forms of gauge invariance, corresponding to the different interactions of particle physics, as we shall see. Their common feature is that the parameters of the corresponding symmetry transformations are allowed to vary with position  $(\mathbf{r}, t)$  in space and time.<sup>9</sup> For example, in QED one takes it as a fundamental assumption that the theory must be invariant under gauge transformations of the form

$$\psi(\mathbf{r}, t) \rightarrow \psi'(\mathbf{r}, t) = \exp[-iqf(\mathbf{r}, t)]\psi(\mathbf{r}, t), \quad (10.34)$$

where  $\psi(\mathbf{r}, t)$  is the wavefunction of a particle of charge<sup>10</sup>  $\sqrt{\varepsilon_0}q$  and  $f(\mathbf{r}, t)$  is an arbitrary continuous function. In other words, one requires that if  $\psi(\mathbf{r}, t)$  is a solution of the equation of motion,  $\psi'(\mathbf{r}, t)$  must also be a solution. This condition is not satisfied by a free or non-interacting particle. In this case, the equation of motion is the free particle Schrödinger equation

$$i\frac{\partial\psi(\mathbf{r}, t)}{\partial t} = -\frac{1}{2m}\nabla^2\psi(\mathbf{r}, t) \quad (10.35)$$

in the non-relativistic limit, and it is easy to see that if  $\psi(\mathbf{r}, t)$  is a solution,  $\psi'(\mathbf{r}, t)$  is not a solution for arbitrary  $f(\mathbf{r}, t)$ . Hence gauge invariance requires the existence of an interaction. Furthermore, if one adds the minimal interaction required to make (10.35) invariant under (10.34), it can be shown that one arrives at the correct equation of motion for a charge particle interacting with an electromagnetic field. This approach, in which the form of the interaction is inferred by adding the minimal interaction terms needed to make the equation of motion gauge invariant, is called the *principle of minimal gauge invariance*, or the *gauge principle* for short. It is discussed in some detail for the interested reader in Appendix D. Here we shall confine ourselves to a qualitative discussion of its application to weak interactions, omitting all technical details. In doing so, we shall restrict ourselves for simplicity to the first generation of

<sup>9</sup> This is in contrast, for example, to our discussion of translational invariance in Section 5.1. In that case the parameter  $\mathbf{a}$  that characterised the transformation (5.1) was assumed to be independent of position and time. Hence translational invariance is not an example of gauge invariance.

<sup>10</sup> We remind the reader that we are working in units that correspond to standard SI units when factors of  $\hbar$  and  $c$  are restored. It is more usual to discuss this topic in Heavyside–Lorentz or rationalised Gaussian units, in which case  $\varepsilon_0 = 1$  and  $q$  is just the electric charge.

leptons ( $e^-$ ,  $\nu_e$ ), since the extension to other leptons, and to quarks using lepton–quark symmetry, is straightforward and involves no new issues of principle.

### 10.2.1 Unification and the gauge principle

The gauge principle described above begs the question: what is the appropriate form of the gauge transformation? In particular, what are the appropriate forms of the gauge transformations for the weak interactions? In QED, one considers phase transformations (10.34) that change an electron state to a new electron state, symbolically  $e^- \rightarrow e^-$ . The gauge principle then leads to interactions  $e^- \rightarrow e^- \gamma$  in which a gauge boson (the photon) is emitted or absorbed. More generally, one can define gauge transformations in which not only the phase of the wavefunction but also the nature of the particle changes. Specifically, one can define a set of gauge transformations in which electrons and electron neutrinos can transform into themselves, or each other.<sup>11</sup> That is, they incorporate the transformations

$$e^- \rightarrow \nu_e, \quad \nu_e \rightarrow e^-, \quad e^- \rightarrow e^-, \quad \nu_e \rightarrow \nu_e$$

and lead, via the gauge principle, to interactions

$$e^- \rightarrow \nu_e W^-, \quad \nu_e \rightarrow e^- W^+, \quad e^- \rightarrow e^- W^0, \quad \nu_e \rightarrow \nu_e W^0$$

in which gauge bosons of the appropriate charge are emitted or absorbed. As implied by our notation, the charged bosons can be identified with the observed  $W^\pm$  bosons, leading to charged current processes of the type observed. However, we also have a  $W^0$  boson, leading to neutral current processes of the same form and strength as the charge current interactions. This is not what is observed experimentally.

The resolution of this problem lies in unification of the weak interactions with electromagnetism. If electromagnetism were introduced directly, as described above following (10.34), we would simply add the photon to the list of gauge bosons  $W^+$ ,  $W^-$ ,  $W^0$  and the problem of the  $W^0$  would remain unchanged. Instead we regard both the photon  $\gamma$  and the neutral boson  $Z^0$  as mixtures of the  $W^0$  with another neutral boson  $B^0$ . That is, we write

$$\gamma = B^0 \cos \theta_W + W^0 \sin \theta_W \tag{10.36a}$$

and

$$Z^0 = -B^0 \sin \theta_W + W^0 \cos \theta_W, \tag{10.36b}$$

where  $\theta_W$  is the weak mixing angle and the particle  $B^0$  is such that the combination (10.36a) has all the properties of the observed

---

<sup>11</sup> The precise form of these transformations is discussed in Appendix D, Section D.7.

photon. This can be achieved by requiring gauge invariance under the transformations

$$\psi_\ell(\mathbf{r}, t) \rightarrow \psi'_\ell(\mathbf{r}, t) = \exp[-ig_Z y_\ell f(\mathbf{r}, t)] \psi_\ell(\mathbf{r}, t) \quad (\ell = e, \nu_e) \quad (10.37)$$

where  $g_Z$  is the weak coupling constant introduced in Equation (10.8) and  $y_\ell \equiv y(e^-), y(\nu_e)$  are constants to be determined.

The transformations (10.37) are similar to (10.34), so the gauge principle leads to vertices

$$e^- \rightarrow e^- B^0, \quad \nu_e \rightarrow \nu_e B^0,$$

analogous to  $e^- \rightarrow e^- \gamma$ , but with the charge replaced by the couplings  $g_Z y(e^-), g_Z y(\nu_e)$ . An explicit calculation<sup>12</sup> then shows that these can be chosen so that the combination (10.36a) has exactly the couplings of a photon, provided that the condition

$$\frac{e}{2(2\varepsilon_0)^{1/2}} = g_W \sin \theta_W = g_Z \cos \theta_W$$

is satisfied. This is just the unification condition (10.8).

In this way, one arrives at a unified theory of electroweak interactions that is gauge invariant, as required. The assumed form of gauge invariance is complicated and is chosen to ensure that the resulting electromagnetic and charged current interactions have the form dictated by experiment. Once that has been done, the properties of the weak neutral current interactions arising from the exchange of the  $Z^0$  boson (10.36b) are predicted in detail. The precise agreement between these predictions and experiment is one of the greatest triumphs of the unified theory, as we have seen. However, there remained one problem, to which we now turn.

### 10.2.2 Particle masses and the Higgs field

The *Higgs boson* is a neutral spin-0 particle whose existence is predicted by the standard model. It is required because gauge invariance can be shown<sup>13</sup> to imply that the spin-1 gauge bosons have zero masses. This is acceptable for QED and QCD, where the gauge bosons are the photons and the gluons, which do indeed have zero masses. However, the  $W^\pm$  and  $Z^0$  bosons are very heavy, and not massless as they would be if gauge invariance were exact. This problem is overcome by assuming that the various particles interact with a new type of scalar field, called the *Higgs field*. The interactions of

---

<sup>12</sup> This calculation is done in Appendix D, Section D.7.3.

<sup>13</sup> This is done explicitly for the electromagnetic case in Appendix D, Section D.3.

the Higgs field with the gauge bosons are gauge invariant, but it differs from other fields in its behaviour in the so-called *vacuum state*, which contains no particles of any kind. Other fields, such as the electromagnetic field, are assumed to be zero in the vacuum state, as one would naïvely expect. However, the Higgs field has a non-zero value  $\eta_0$  in the vacuum and this value is not invariant under a gauge transformation. Because of this, the theory is no longer gauge invariant, and the gauge bosons are no longer required to have zero mass. This form of symmetry breaking, in which the gauge invariance of the interaction (as opposed to the gauge invariance of the vacuum) remains exact, is called *spontaneous symmetry breaking*.

Spontaneous symmetry breaking of this type is known in other branches of physics. It occurs whenever the vacuum state, defined as the state of lowest energy, is not unique. A familiar example is ferromagnetism. In a ferromagnetic material, the forces that couple the electron spins are rotationally invariant and at high temperatures the net magnetisation  $\mathbf{M}$  is zero. Below the Curie temperature the spins in the ground state are aligned in some definite direction, resulting in a magnetisation that breaks the rotational invariance. However,  $\mathbf{M}$  could equally well point in any other direction and the rotational invariance of the interaction manifests itself in that all the properties, other than the direction of  $\mathbf{M}$ , are independent of the actual direction.

To see how such effects can occur in particle physics, we consider the simple case of a complex scalar field  $\eta(\mathbf{r}, t)$ , which we write in terms of two real fields  $\eta_1(\mathbf{r}, t)$  and  $\eta_2(\mathbf{r}, t)$ , that is

$$\eta(\mathbf{r}, t) = \eta_1(\mathbf{r}, t) + i\eta_2(\mathbf{r}, t). \quad (10.38)$$

Let us assume that the potential energy density of  $\eta(\mathbf{r}, t)$  is given by

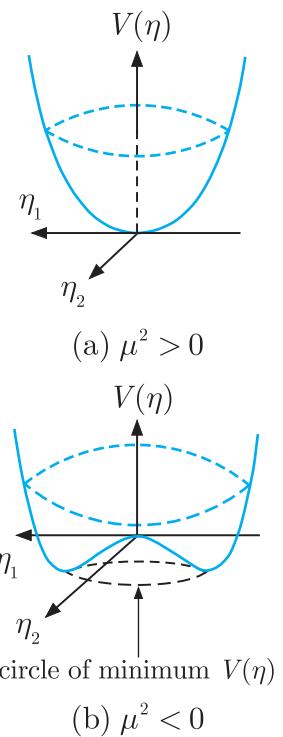
$$V(\eta) = \mu^2 |\eta(\mathbf{r}, t)|^2 + \lambda |\eta(\mathbf{r}, t)|^4. \quad (10.39)$$

Here  $\lambda$  and  $\mu^2$  are real parameters and the interactions are invariant under the transformation

$$\eta(\mathbf{r}, t) \rightarrow \eta(\mathbf{r}, t)e^{i\beta}, \quad (10.40)$$

where  $\beta$  is an arbitrary phase parameter. The vacuum state can be identified by minimising the potential energy density, since the kinetic energy density is either positive or zero. There are two cases to consider, and in both we require  $\lambda > 0$  so that the potential energy density is bounded from below:

(i)  $\mu^2 > 0$ . In this case, both terms in  $V(\eta)$  are positive definite and the corresponding potential energy density surface  $V(\eta)$  is sketched as a function of  $\eta_1(\mathbf{r}, t)$  and  $\eta_2(\mathbf{r}, t)$  in Figure 10.14(a). As can be seen,  $V(\eta)$  has a unique minimum at  $\eta(\mathbf{r}, t) = 0$  and spontaneous symmetry breaking does not occur.



**Figure 10.14** The potential energy density  $V(\eta)$ , as given by (10.39), for  $\lambda > 0$ .

(ii)  $\mu^2 < 0$ . The potential energy surface for this case is shown in Figure 10.14(b). There is a circle of absolute minima at

$$\eta(\mathbf{r}, t) = \eta_0 = \left( -\frac{\mu^2}{2\lambda} \right)^{1/2} e^{i\theta}, \quad 0 \leq \theta \leq 2\pi, \quad (10.41)$$

where the phase angle  $\theta$  specifies a direction in the complex  $\eta$ -plane. This arbitrariness in the direction of  $\theta$  is analogous to that of the direction of the magnetisation  $\mathbf{M}$  of a ferromagnet. By analogy with the latter case, spontaneous symmetry breaking corresponds to taking a particular value of  $\theta$  to represent the ground state. The actual value chosen is not significant, because of the invariance of the interactions under a phase transformation (10.40), and it is conventional to choose  $\theta = 0$ , so that the vacuum state corresponds to

$$\eta_0 = \left( -\frac{\mu^2}{2\lambda} \right)^{1/2}. \quad (10.42)$$

This is obviously not invariant under the transformation (10.40) and the symmetry is spontaneously broken.

To see how this affects gauge theories, consider the introduction of electromagnetic interactions into the above model in such a way that the equations of motion remain invariant under a gauge transformation of the standard form (10.34), that is

$$\eta(\mathbf{r}, t) \rightarrow \eta'(\mathbf{r}, t) = \exp[-iqf(\mathbf{r}, t)]\eta(\mathbf{r}, t). \quad (10.43)$$

Since the electromagnetic field vanishes in the vacuum state, this has no effect on the above discussion of the scalar field value in the vacuum. Again, there are two cases to consider:

(i)  $\mu^2 > 0$ . In this case, we again have  $\eta(\mathbf{r}, t) = 0$  in the vacuum, so there is no spontaneous symmetry breaking. Gauge invariance remains exact and the photon must be massless, as observed in nature.

(ii)  $\mu^2 < 0$ . In this case, spontaneous symmetry breaking does occur and we can again choose to represent the vacuum field by (10.42). This clearly is not invariant under the gauge transformaton (10.43), so gauge invariance is broken and the photon is no longer required to be massless. Furthermore, the photon automatically acquires mass<sup>14</sup> due to its interaction with the vacuum field  $\eta_0$ . This mechanism, whereby a gauge boson acquires mass as a result of its interaction with a non-vanishing vacuum field, was first proposed in 1964 by

---

<sup>14</sup> This is discussed briefly in Section D.3 of Appendix D, and more completely in Sections 13.1 and 13.2 of Mandl and Shaw (2010).

Higgs, and independently by Englert and Brout. It is now called the *Higgs mechanism*. It is in some ways analogous to the way a conduction electron in a semiconductor acquires an effective mass different from its mass *in vacuo* as a result of its interactions with the crystal lattice.

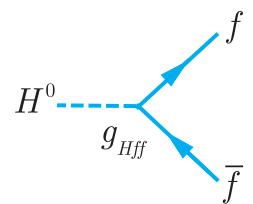
Of course the unified electroweak theory is more complicated than the above simple model. We will omit all details here and simply note that applying the Higgs mechanism to the unified theory leads to three main consequences:

- (i) The  $W^\pm$  and  $Z^0$  bosons acquire the non-zero masses discussed in Section 10.1.2. These masses arise from the interactions of the gauge fields with the non-zero vacuum expectation value of the Higgs field.
- (ii) There are electrically neutral quanta  $H^0$  associated with the Higgs field, called *Higgs bosons*, in the same way that there are quanta, that is photons, associated with the electromagnetic field, except that the Higgs bosons have non-zero mass and have spin 0 rather than spin 1. The mass of the Higgs boson is not predicted by the theory, but for any given mass, their interactions are completely specified.
- (iii) The Higgs field throws light on the origin of quark and lepton masses.

The explanation of (iii) is that in the absence of a Higgs field, gauge invariance requires that the masses of spin-1/2 fermions with parity-violating interactions are zero.<sup>15</sup> Parity is conserved in strong and electromagnetic interactions, but violated in weak interactions, so that quarks and leptons would be massless in this case. However, interactions with the Higgs field can generate fermion masses arising from the non-zero vacuum expectation value  $\eta_0$  of the Higgs field, as well as interactions with the Higgs bosons. These interactions are of a form shown in Figure 10.15, with a dimensionless coupling constant  $g_{Hff}$  related to the fermion mass  $m_f$  by

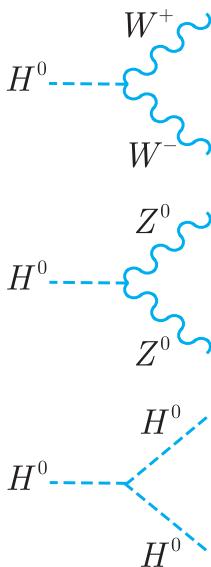
$$g_{Hff} = \sqrt{2}g_W \left( \frac{m_f}{M_W} \right). \quad (10.44)$$

This theory of fermion masses – that they are generated by interactions with the Higgs field – does not make any predictions for their values  $m_f$ . However, it can be tested by measuring the Higgs boson couplings  $g_{Hff}$ , and verifies the predictions (10.44), as we shall see.

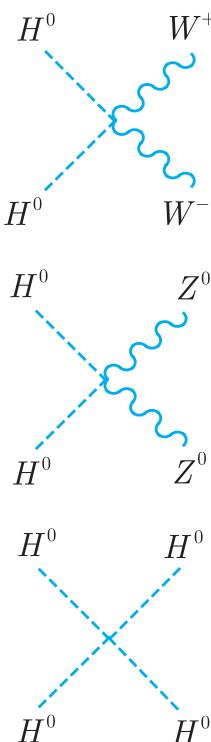


**Figure 10.15** The basic vertices for Higgs boson–fermion interactions. The fermion  $f$  can be any quark, charged lepton or neutrino.

<sup>15</sup> The reason for this is discussed in Section D.7.4 of Appendix D.



**Figure 10.16** The basic first-order vertices for Higgs boson–boson interactions.



**Figure 10.17** The basic second-order vertices for Higgs boson–boson interactions.

### 10.2.3 Properties of the Higgs boson

The basic vertices for the interaction of Higgs bosons with fermions are shown in Figure 10.15. In addition, the unified electroweak theory predicts the existence of the first and second-order boson vertices shown in Figures 10.16 and 10.17, respectively. The strength of the various Higgs boson interactions are completely specified for a given value of the Higgs mass, and, for leptons, quarks and weak bosons, they increase with the mass of the particle to which it couples. The Higgs boson therefore couples very weakly to light particles like neutrinos, electrons, muons and  $u$ ,  $d$  and  $s$  quarks, and much more strongly to heavy particles like  $W^\pm$  and  $Z^0$  bosons and  $t$  quarks. This is reflected in both the decay and production properties of the Higgs boson, as we shall now discuss.

Although the Higgs mass was not predicted in the unified theory before it was first detected in 2012, its mass was estimated by calculating the contribution to any given electroweak process from higher-order diagrams in which Higgs bosons are emitted and absorbed. These diagrams, together with other higher-order diagrams, give very small corrections to the leading order processes that we have considered throughout this chapter. Nonetheless, these corrections can be measured, and the requirement that the measured and calculated corrections agree led to the restricted range

$$M_H = 94^{+29}_{-24} \text{ GeV}/c^2, \quad (10.45)$$

corresponding the range  $46 < M_H < 152 \text{ GeV}/c^2$  at a 95% confidence limit. The observed mass

$$M_H = 125.09 \pm 0.24 \text{ GeV}/c^2 \quad (10.46a)$$

is compatible with this. In contrast, the decay width of the Higgs boson has not been measured directly, but is predicted to be

$$\Gamma_H = 4.1 \pm 0.2 \text{ MeV}, \quad (10.46b)$$

corresponding to a lifetime of about  $1.6 \times 10^{-22} \text{ s}$ . This value is obtained by calculating and adding together the rates of all the principal decay modes, which are given in Table 10.1.

The decays listed in Table 10.1 arise as follows. The Higgs boson is much more stable than the  $W^\pm$  and  $Z^0$  bosons, because its couplings to the relatively light particles to which it can decay are heavily suppressed. In lowest order, the decays allowed by energy conservation are

$$H \rightarrow f + \bar{f}, \quad (10.47)$$

where  $f$  can be any lepton or quark except the  $t$ -quark. These decays occur via the vertices shown in Figure 10.15, and if the fermion

**Table 10.1** Predicted decay widths  $\Gamma$  and branching ratios  $B$  for the main decay modes of the Higgs boson and for the rare but important decay  $H^0 \rightarrow \gamma\gamma$ . The errors arise from uncertainties in the input parameters.

Decay mode	$\Gamma(MeV)$	$B$
$b\bar{b}$	$2.35 \pm 0.11$	$(57.5 \pm 1.8) \times 10^{-2}$
$W^\pm f \bar{f}'$	$0.88 \pm 0.01$	$(21.6 \pm 0.9) \times 10^{-2}$
$gg$	$0.35 \pm 0.02$	$(8.6 \pm 0.8) \times 10^{-2}$
$\tau^+ \tau^-$	$0.26 \pm 0.01$	$(6.3 \pm 0.3) \times 10^{-2}$
$c\bar{c}$	$0.12 \pm 0.01$	$(2.9 \pm 0.3) \times 10^{-2}$
$Z^0 f \bar{f}$	$(11.0 \pm 0.1) \times 10^{-3}$	$(2.7 \pm 0.1) \times 10^{-2}$
$\gamma\gamma$	$(9.5 \pm 0.1) \times 10^{-3}$	$(2.3 \pm 0.1) \times 10^{-3}$

masses are neglected relative to the Higgs mass, the corresponding decay widths are given by<sup>16</sup>

$$\Gamma(H \rightarrow f \bar{f}) = N \alpha_W \left( \frac{m_f^2}{M_W^2} \right) M_H, \quad (10.48)$$

where  $\alpha_W^2 = g_W^2/4\pi = 0.0042$  is the characteristic strength of the weak interaction. The suppression factor  $(m_f^2/M_W^2)$  comes from the mass factor in the coupling (10.44) and is much less than unity for all  $f$ . The factor  $N$  in (10.48) is 1 for leptons but 3 for quarks, to allow for the three possible colour states in the latter case, so that

$$\Gamma(H \rightarrow b\bar{b}) : \Gamma(H \rightarrow c\bar{c}) : \Gamma(H \rightarrow \tau\bar{\tau}) = 3m_b^2 : 3m_c^2 : m_\tau^2. \quad (10.49)$$

In the case of the quarks, the masses are the current masses evaluated at  $\mu = M_H$ , which are approximately 0.6 and 3 GeV/ $c^2$  for the  $c$  and  $b$  quarks respectively. The resulting partial widths and branching ratios for these modes are given in Table 10.1. The decay widths for the other lighter quarks and leptons are very much smaller.

The first-order decays (10.47) are so heavily suppressed by the mass factors in (10.44), arising from the  $H f \bar{f}$  vertices of Figure 10.15, that higher-order decays can lead to comparable decay rates. The three higher-order decays that contribute significantly to the total decay rate are

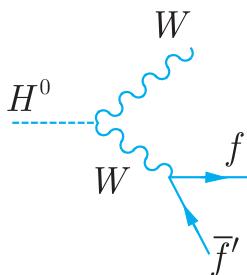
$$H^0 \rightarrow W^\pm + f + \bar{f}' \quad (10.50)$$

$$H^0 \rightarrow Z^0 + f + \bar{f} \quad (10.51)$$

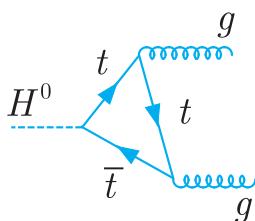
and

$$H^0 \rightarrow g + g, \quad (10.52)$$

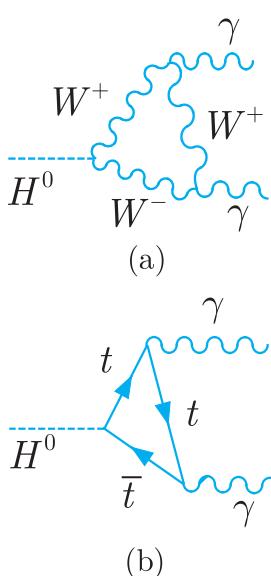
<sup>16</sup> The exact first-order result, including fermion masses, is given in Section 19.5.1 of Mandl and Shaw (2010). Note that (10.48) is often expressed in terms of the Fermi coupling, which is related to  $\alpha_W$  by (2.17).



**Figure 10.18** Mechanism of the  $H^0$  decay (10.50). The mechanism for (10.51) is obtained by replacing  $W$  by  $Z^0$  and  $\bar{f}'$  by  $\bar{f}$ .



**Figure 10.19** The dominant mechanism for the decay  $H^0 \rightarrow g + g$ .



**Figure 10.20** The dominant mechanisms for the rare decay  $H^0 \rightarrow \gamma + \gamma$ .

where  $f\bar{f}$  and  $f\bar{f}'$  are any lepton–antilepton or quark–antiquark pair allowed by the conservation laws, and where the gluons and quarks are observed as jets.<sup>17</sup> Reactions (10.50) and (10.51) are second-order electroweak processes. This is illustrated for (10.50) in Figure 10.18, while (10.52) proceeds via an intermediate  $t\bar{t}$  pair as shown in Figure 10.19. Their rates are given in Table 10.1, together with that for the rare decay

$$H^0 \rightarrow \gamma + \gamma, \quad (10.53)$$

the dominant mechanisms for which are shown in Figure 10.20. It is included because of its important role in the experimental discovery of the Higgs boson.

Before the above decays can be studied experimentally, it is necessary to produce Higgs bosons at a measurable rate, and since they couple extremely weakly to the leptons and quarks of which stable beams and targets are composed, this can only occur by processes in which heavy particles, namely  $W^\pm$ ,  $Z^0$  bosons and top quarks, act as intermediate states. For example, in electron–positron colliding beams, the dominant production processes

$$e^+ + e^- \rightarrow H^0 + Z^0, \quad (10.54a)$$

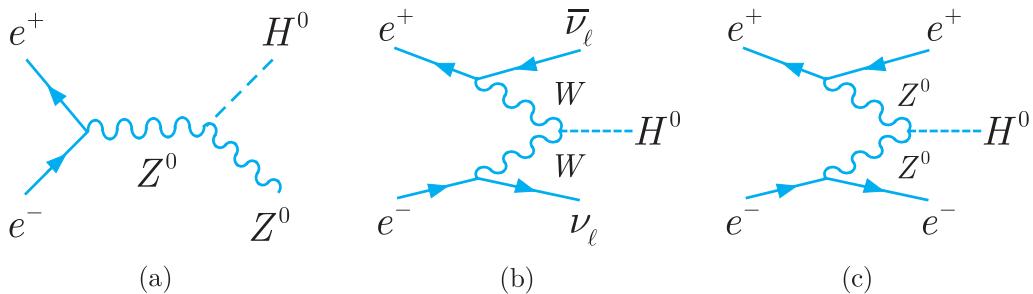
$$e^+ + e^- \rightarrow H^0 + \nu_\ell + \bar{\nu}_\ell \quad (\ell = e, \mu, \tau) \quad (10.54b)$$

$$e^+ + e^- \rightarrow H^0 + \ell + \bar{\ell} \quad (\ell = e, \mu, \tau) \quad (10.54c)$$

arise from the mechanisms shown in Figures 10.21(a), (b) and (c), respectively. The first of these reactions, often called *Higgsstrahlung*, dominates in the region close to threshold, but decreases as the square of the centre-of-mass energy  $E_{CM}$  at high energies, while (10.54b), which increases with  $E_{CM}$ , dominates at higher energies. Reaction (10.54c) is similar to (10.54b) but is much smaller, as can be seen in Figure 10.22, where the predicted cross-sections for these reactions are shown for centre-of-mass energies up to 3 TeV.

The Large Electron–Positron (LEPII) collider at CERN had a maximum energy of 208 GeV when it closed in 2002. This energy was too low for a Higgs boson to be detected, but the failure to observe reaction (10.54a) set a lower limit of 114 GeV on the then unknown Higgs mass. Nonetheless, reactions (10.54) are in principle an ideal way to study the detailed properties of Higgs bosons, which may be realised in practice if the proposed International Linear Collider mentioned in Section 4.2.3 is constructed. The bosons are produced

<sup>17</sup> Reactions (10.50) and (10.51) are sometimes written as  $H \rightarrow W + W^* \rightarrow W^\pm + f + \bar{f}'$  and  $H^0 \rightarrow Z + Z^* \rightarrow Z^0 + f + \bar{f}$ , respectively, where  $W^*$  is a virtual boson that decays to fermions, as shown in Figure 10.18, and analogously for  $Z^*$ .



**Figure 10.21** Dominant mechanisms for reactions (10.54a), (10.54b) and (10.54c).

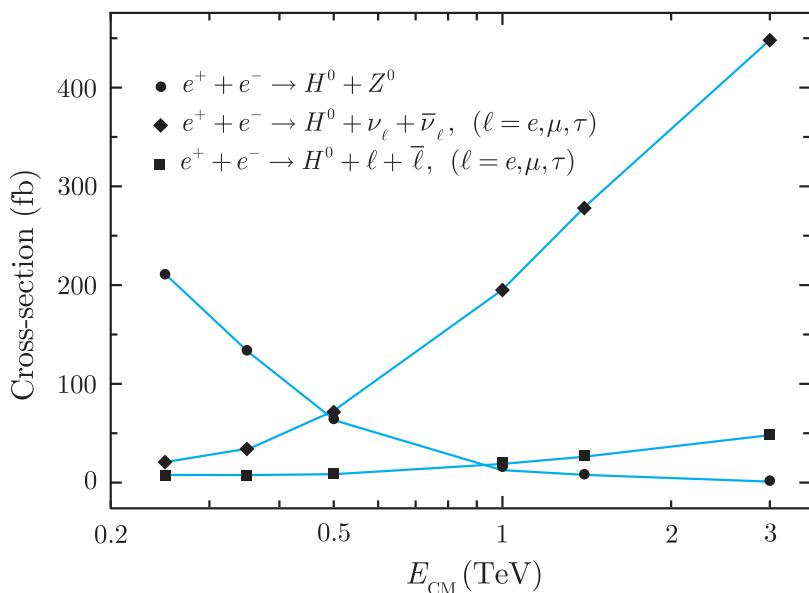
in final states containing very few other particles, so that the signal is easy to identify, and background reactions are not too large since they are second-order electroweak processes like the production processes themselves. In proton–proton collisions, the situation is less straightforward, as we shall see below.

#### 10.2.4 The discovery of the Higgs boson

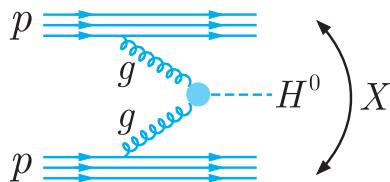
Results that first established the existence of the Higgs boson were reported in 2012 from two experiments, ATLAS and CMS, using the LHC at CERN. The ATLAS detector has been described in Section 4.5.2; CMS is a broadly similar general-purpose detector, although there are significant differences in their details.

At the LHC the reaction of interest is

$$p + p \rightarrow H^0 + X, \quad (10.55)$$



**Figure 10.22** Predicted cross-sections for reactions (10.54a), (10.54b) and (10.54c). (Based on Table 1-15 of Ajaib *et al.*, arXiv.1310.8361 (2014).)



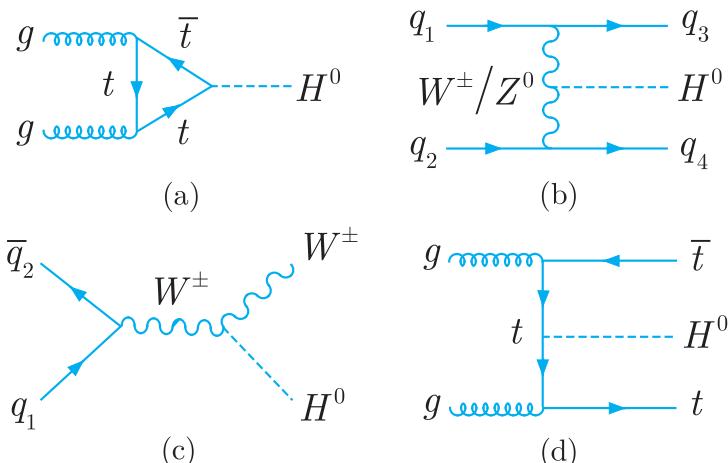
**Figure 10.23** Gluon fusion production mechanism for the reaction  $p + p \rightarrow H^0 + X$  at the LHC.

where  $X$  is any state consistent with conservation of the appropriate quantum numbers. The dominant Higgs production mechanism for this reaction at the LHC is ‘gluon fusion’, as shown in Figure 10.23, and the contribution of this process to the production cross-sections (10.55) is given by

$$\sigma(pp \rightarrow HX) = \int_0^1 \int_0^1 g(x_1)g(x_2)\sigma(gg \rightarrow H)dx_1dx_2,$$

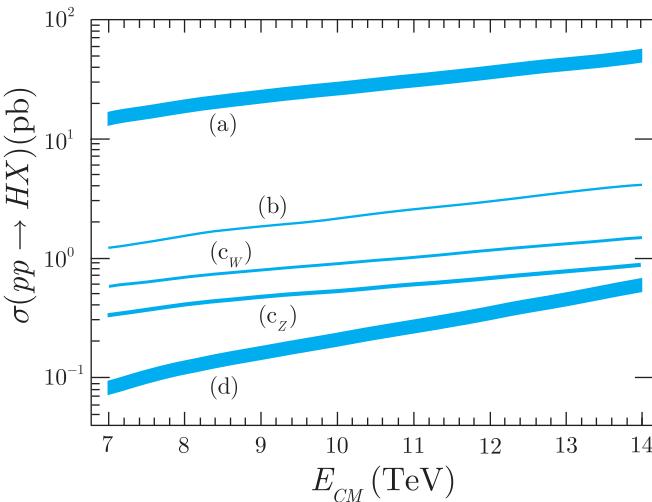
where  $g(x)$  is the gluon distribution function of the proton<sup>18</sup> and the cross-section for  $\sigma(gg \rightarrow H)$  is dominated by a Feynman diagram with a loop involving top quarks, shown in Figure 10.24(a), because of the strong coupling of  $H^0$  to the very heavy top quark. Other production processes are shown in Figure 10.24(b), (c) and (d), where for simplicity the spectator particles are not shown, and the corresponding production cross-sections for all four in the LHC energy range are shown in Figure 10.25.

In all these processes, the Higgs boson is produced together with large numbers of particles with low transverse momenta resulting from beam fragmentation, as indicated schematically in Figure 10.23. However, since the Higgs boson is heavy, its decay



**Figure 10.24** Production mechanisms for the Higgs boson at the LHC.

<sup>18</sup> The use of quark and gluon distribution functions to make analogous calculations was discussed in Section \*8.4.2.



**Figure 10.25** Predicted values for Higgs boson production cross-sections in  $pp$  collisions as a function of the centre-of-mass energy. The labelling of the curves corresponds to the diagrams in Figure 10.24, where  $c_{W,Z}$  refer to the production of  $W$  and  $Z$  bosons with the Higgs boson, respectively. The bands represent the uncertainties arising from input parameters and theoretical uncertainties. (Carena *et al.*, 2013.)

products have appreciable transverse momenta, except for the few cases when they are almost collinear with the beam. Hence in selecting Higgs boson candidates, the background from particles produced in beam fragmentation can in principle be greatly reduced by accepting only events in which the decay products have significant transverse momenta, as was done in the discovery of the  $W$  and  $Z$  mesons and the top quark, discussed in Sections 4.5.1 and 9.2.3, respectively. The main problem arises because, on the one hand, most decays of the Higgs boson result in hadron jets either directly from the quarks or gluons produced in the decay, or indirectly because most of the weak bosons emitted themselves decay to jets in the final state. On the other hand, at the same collision energies the proton–proton total cross-section is of order 100 mb, and it is very difficult to distinguish jets arising from Higgs decays from the huge number of such jets produced in proton–proton collisions by QCD interactions, like those described in Section \*8.4.2. In particular, from Table 10.1 we see that the dominant decay mode is  $H^0 \rightarrow b\bar{b}$  with the quarks fragmenting to jets, but even if these are ‘tagged’ as  $b$ -jets<sup>19</sup> it is almost impossible to distinguish those arising from Higgs decays from the much more numerous ones from other sources. This can be partly overcome by requiring that a  $W$  or  $Z$  boson be produced together with the Higgs boson in the final state,

$$p + p \rightarrow H + (W, Z) + X, \quad (10.56)$$

corresponding to the production mechanism of Figure 10.24(c). While this reduces the Higgs production rate, it reduces the jet background much more, but the latter still remains very large.

<sup>19</sup> This technique was used in the discovery of the top quark, and is discussed in Section \*9.2.3.

Because of this, and similar problems with the other prominent decay modes, the best results were obtained by choosing rare decay modes containing only photons and leptons in the final states, so that the problem of background jets is eliminated, albeit at the price of a lower counting rate.

In any experiment, the number  $N$  of Higgs bosons produced and decaying to a particular final state  $F$  is given by

$$N = L_I \sigma(pp \rightarrow H + \dots) B(H \rightarrow F), \quad (10.57)$$

where  $L_I$  is the integrated luminosity,<sup>20</sup>  $\sigma$  is the production cross-section and  $B(H \rightarrow F)$  is the branching ratio for the final state observed. The experiments that established the existence of the Higgs boson collected data in 2011 and 2012 at proton–proton centre-of-mass energies of 7 GeV and 8 GeV, where the cross-section for Higgs production is roughly 20 pb (1 pb =  $10^{-12}$  b). In addition, rare decay modes are chosen to reduce background problems, so that the branching ratio  $B(H \rightarrow F)$  is also small and high luminosity is required to produce a useful counting rate. Hence the LHC was designed as a  $pp$  rather than a  $p\bar{p}$  collider, since proton beams can be produced with much higher intensity than antiproton beams. In particular, the ATLAS and CMS experiments each had integrated luminosities of about  $5 \text{ pb}^{-1}$  at a centre-of-mass energy of 7 TeV and about  $20 \text{ pb}^{-1}$  at 8 TeV, giving of order 500 predicted Higgs boson decays for a decay leading to a final state  $F$  with a branching ratio of  $10^{-3}$ .<sup>21</sup> In practice, two rare decays are of particular importance, in that they both avoid final-state jets and lead to high mass resolution. We shall consider each in turn.

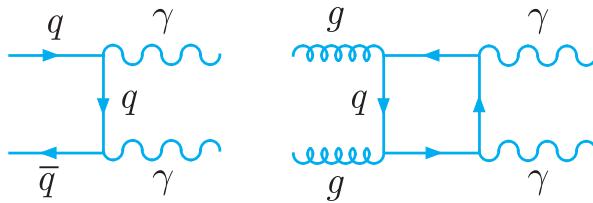
#### 10.2.4(a) $H^0 \rightarrow \gamma\gamma$

This decay has a predicted branching ratio of about  $2 \times 10^{-3}$  and the Higgs boson should appear as a narrow peak in the  $2\gamma$  mass distribution, sitting on a large predicted background from QCD production of two photons, the principal diagrams for which are shown in Figure 10.26. In the ATLAS experiment, the energies and directions of photons are detected in the electromagnetic calorimeters, and events were selected in which the transverse energy of the higher energy (lower energy) photon in the pair was required to be above 40 GeV (30 GeV). A typical event is shown in Plate 8. The invariant mass spectrum of the selected events exhibited a

---

<sup>20</sup> Luminosity and integrated luminosity are defined in Section 4.1 and discussed in more detail in Section B.1 in Appendix B.

<sup>21</sup> The discovery of the Higgs boson was initially reported on the basis of integrated luminosities of  $5 \text{ pb}^{-1}$  at 7 TeV and of  $5 \text{ pb}^{-1}$  at 8 TeV. However, the results discussed here correspond to the full luminosity stated above.



**Figure 10.26** Dominant background processes producing  $2\gamma$  final states.

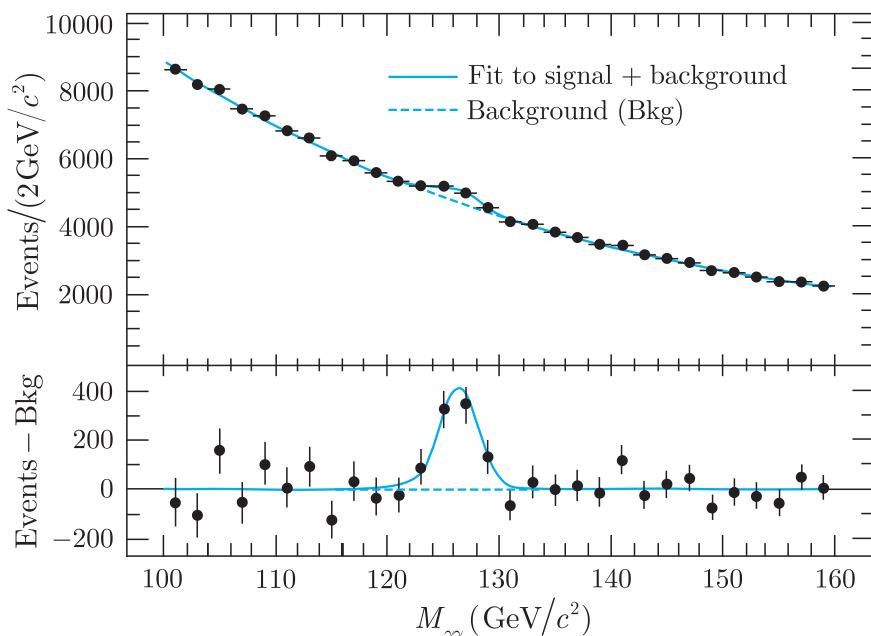
clear peak, as shown in Figure 10.27, and a detailed analysis, taking the backgrounds carefully into account, led to a Higgs mass of  $M_H = 125.98 \pm 0.42 \pm 0.28 \text{ GeV}/c^2$ , where the first error is statistical and the second systematic. Similar results were also obtained by CMS, corresponding to a mass value of  $124.70 \pm 0.31 \pm 0.15 \text{ GeV}/c^2$ .

#### 10.2.4(b) $H^0 \rightarrow 4 \text{ charged leptons}$

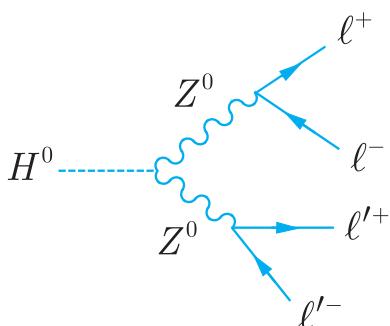
The principal mechanism for the decays

$$H^0 \rightarrow \ell^+ + \ell^- + \ell'^+ + \ell'^- \quad (\ell = e, \mu; \ell' = e, \mu) \quad (10.58)$$

is shown in Figure 10.28. The main contribution to these decays arises when one of the intermediate  $Z^0$  bosons is ‘on-shell’, that is it is a real particle. The resulting branching ratio is predicted



**Figure 10.27** Distribution of the invariant mass of  $2\gamma$  candidates obtained in the ATLAS experiment from runs at  $E_{CM} = 7 \text{ TeV}$  with an integrated luminosity of  $4.8 \text{ fb}^{-1}$ , and  $8 \text{ TeV}$  with an integrated luminosity of  $20.7 \text{ fb}^{-1}$ , taken during 2011 and 2012. (Adapted from Aad *et al.*, 2013.)

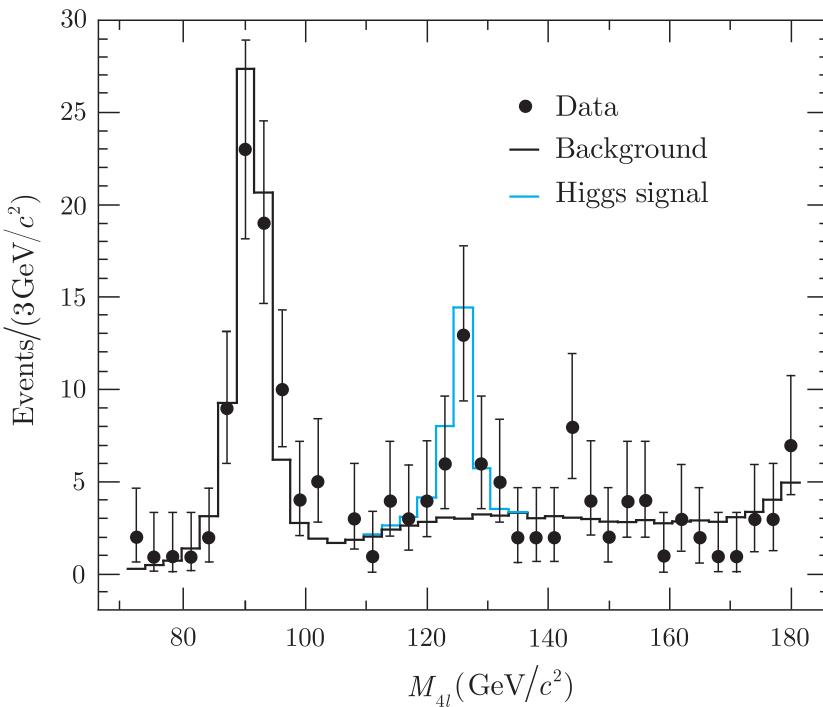


**Figure 10.28** Dominant mechanism for the decay  $H^0 \rightarrow \ell^+ + \ell^- + \ell'^+ + \ell'^-$  ( $\ell = e, \mu; \ell' = e, \mu$ ).

(cf. Problem 10.9) to be much smaller than the branching ratio for the  $2\gamma$  decay, and hence there are far fewer events produced in any given experiment. However, to compensate for this, the signal-to-background ratio is expected to be much better and the energy resolution is almost as good as in the  $2\gamma$  mode. The charged leptons were observed in the tracking detectors and in the electromagnetic calorimeters. In the CMS experiment, the electrons and muons were required to have a minimum transverse momenta of  $7\text{ GeV}/c$  and  $5\text{ GeV}/c$ , respectively; the invariant mass of one of the heaviest of the particle–antiparticle pairs was required to be in the range  $40\text{--}120\text{ GeV}/c^2$  and the other in the range  $12\text{--}120\text{ GeV}/c^2$ ; and the measured tracks of the pairs originating from  $Z$  decay were checked to ensure they were consistent with them originating from the same primary vertex. The invariant mass spectrum of the selected events is shown in Figure 10.29. The peak at around  $90\text{ GeV}/c^2$  arises from  $Z$  decays to charged lepton pairs in which one of the leptons radiates a photon that converts into a second lepton pair. More significantly, there is a peak, sitting on a relatively small background, corresponding to a mass of  $125.6 \pm 0.4 \pm 0.2\text{ GeV}/c^2$ , where again the first error is statistical and the second systematic. ATLAS obtained similar results, with a mass of  $124.51 \pm 0.52 \pm 0.06\text{ GeV}/c^2$ .

The above results clearly indicate the existence of a new state with a mass of about  $125\text{ GeV}/c^2$ , and combining the four mass determinations cited above yields the value given in (10.46a), where the statistical and systematic errors are combined in quadrature. The widths of the peaks in the invariant masses are compatible with the experimental resolution, which is too large for the relatively small predicted width (10.46b) to be verified. However, the fact that the new particle is observed to decay to two photons implies that its  $C$  parity is  $+1$  and that it has integer spin, while a subsequent study of the angular distributions in the four-charged lepton mode shows that it has spin 0, as required for the Higgs boson.

Further evidence is obtained by the comparison of the number of observed events with the number predicted by the standard model, as noted briefly above. Since the luminosity is known, one sees from



**Figure 10.29** Distribution of the invariant mass of the four charged lepton candidates obtained in the CMS experiment from experimental runs at  $E_{CM} = 7 \text{ eV}$  with an integrated luminosity of  $5.1 \text{ fb}^{-1}$  and  $8 \text{ TeV}$  with an integrated luminosity of  $19.6 \text{ fb}^{-1}$ . (Adapted from S. Chatrchyan *et. al.* 2014, with permission from the American Physical Society.)

(10.57) that the observed rate determines the product of the production cross-section and the branching ratio. In practice, it is usual to give results in terms of the ratio  $\mu$  of the measured value to that predicted by the standard model, that is

$$\mu(F) = \frac{[\sigma(pp \rightarrow H + \dots)B(H \rightarrow F)]_{\text{exp}}}{[\sigma(pp \rightarrow H + \dots)B(H \rightarrow F)]_{\text{SM}}}. \quad (10.59)$$

As well as the two-photon and four-charged lepton channels discussed above, results of comparable statistical significance for  $\mu(F)$  have been obtained for the decay channel

$$H \rightarrow W + W^* \rightarrow \ell + \nu_\ell + \ell' + \nu_{\ell'}, \quad (10.60)$$

where  $\ell$  and  $\ell'$  can be electrons or muons and the neutrinos are such as to conserve lepton numbers in the reaction. This has a much higher branching ratio than the mode with four-charged leptons, but a poorer mass resolution because the neutrinos cannot be detected. Values of  $\mu$  with less precision are also obtained for the decay  $H \rightarrow b\bar{b}$ , which has the background problems discussed above, and the rare decay  $H \rightarrow \tau\bar{\tau}$ , in which observed final states again contain neutrinos resulting from the tauon decays. In addition, the predicted and observed rates depend not only upon the decay channel  $F$  but also on the production reaction. For example, requiring that the Higgs boson be produced together with a  $W$  or  $Z$  boson selects the mechanism of Figure 10.24(c), while the

**Table 10.2** Measured values of the ratio (10.59) for the observed to predicted signal in Higgs production at the LHC, for individual final states  $F$  and for all final states combined.

Final state	ATLAS			CMS	
	$\mu$	$+\delta\mu$	$-\delta\mu$	$\mu$	$\pm\delta\mu$
$\ell\ell'\nu\nu$	1.16	0.24	0.21	0.83	0.21
$\gamma\gamma$	1.17	0.27	0.27	1.12	0.24
$\ell^+\ell^-\ell'^+\ell'^-$	1.44	0.40	0.33	1.00	0.29
$b\bar{b}$	0.52	0.40	0.40	0.84	0.44
$\tau^+\tau^-$	1.43	0.43	0.37	0.91	0.28
Combined	1.17	0.27	0.27	1.00	0.14

mechanisms of Figures 10.24(b) and 10.24(d) produce the  $H$  in association with two high  $p_T$  jets or top quarks, respectively. Choosing particular signals in this way is called *tagging* and a wide range of data is available for the various decay channels with and without various tags. A sample of these results is shown in Table 10.2, where the CMS data are tagged and the ATLAS data, apart from the channel  $H \rightarrow b\bar{b}$ , are not.<sup>22</sup> As can be seen, the values for  $\mu$  from these very varied final states, and the combined value from all the channels, are all statistically consistent with the predictions of the standard model, providing further impressive evidence that the observed state is indeed the long-sought-for Higgs boson.

At this point, we note that the above results are obtained from the analysis of data taken at 7 and 8 TeV in the period 2011–2012. Subsequently, the LHC was shut down to be upgraded to an energy of 13 TeV, where the Higgs production cross-section is about three times higher while the luminosity was increased by a factor of about two. The machine was restarted in 2015, so more precise results should be available in the near future.

Finally, in the above discussion we have considered only the simplest case of a single, electrically neutral, Higgs boson, as assumed in the standard model. However, some interesting extensions of the standard model, which are discussed in Chapter 12, require more than one Higgs particle, including electrically charged varieties, as well as other new states. The LHC was designed to search for states up to masses of a few TeV, but none have been observed so far. Whatever the truth, it is clear that experimental investigation of the Higgs sector will play a central role in developments in particle physics for many years to come.

<sup>22</sup> For further details and a more extensive set of results, see Aad *et al.* (2016) and Chatrchyan *et al.* (2015). All the experimental results cited in this section are taken from these papers and references therein.

## Problems 10

- 10.1** Show that the vertices of Figure 10.3 are the only possible vertices  $abZ^0$  allowed by charge and lepton number conservation, where  $a$  and  $b$  can be any lepton or antilepton.
- 10.2** Which of the following processes are allowed in electromagnetic interactions and which are allowed in weak interactions via the exchange of a single  $W^\pm$  or  $Z^0$ ?
- (a)  $K^+ \rightarrow \pi^0 + e^+ + \nu_e$ , (b)  $K^+ \rightarrow \pi^+ + e^+ + e^-$ ,  
 (c)  $\Sigma^0 \rightarrow \Lambda^0 + e^+ + e^-$ , (d)  $\Sigma^0 \rightarrow n + \nu_e + \bar{\nu}_e$ .
- 10.3** The reaction (10.6b) is forbidden to occur via lowest-order weak interactions like Figure 10.5(b). However, it can proceed by higher-order diagrams involving the exchange of two or more bosons. Draw examples of such diagrams. Make a simple dimensional estimate of the ratio of decay rates (10.7).
- 10.4** Draw the dominant Feynman diagrams for the reaction  $e^+ + e^- \rightarrow \nu_e + \bar{\nu}_e$ . Estimate the order of magnitude and energy dependence of the corresponding contributions to the total cross-section for centre-of-mass beam energies  $m_e \ll E \ll M_W, M_Z$ .
- 10.5** If small corrections of order  $(m_f/M_Z)^2$  are neglected, the partial width for  $Z^0 \rightarrow f\bar{f}$  arising from the lowest-order diagrams of Figure 10.12 are given by

$$\Gamma_f = \frac{G_F M_Z^3}{3\pi\sqrt{2}} \left[ \frac{1}{4} \pm q_f \sin^2 \theta_W + 2q_f^2 \sin \theta_W \right],$$

where  $q_f$  is the electric charge of the fermion in units of  $e$ ,  $\theta_W$  is the weak mixing angle and where the plus sign applies for  $f = \ell^-, d, s, b$  and the minus sign applies for  $f = \nu_\ell, u, c, t$ . Assuming that  $\sin^2 \theta_W = 0.25$ , calculate the branching ratios for  $Z^0 \rightarrow c\bar{c}$  and  $Z^0 \rightarrow b\bar{b}$ , and the decay widths to charged lepton pairs and to hadrons. Compare the latter two to the experimental values (10.29a) and (10.29b) and comment on any significant discrepancy.

- 10.6** The centre-of-mass differential cross-section for the reaction  $e^+e^- \rightarrow \mu^+\mu^-$  due to the one-photon exchange diagram in Figure 10.2(a) is given in Equation (7.15),

$$\sigma(s, \theta) = \frac{d\sigma(s, \cos \theta)}{d\cos \theta} = \frac{\pi \alpha^2}{2s} (1 + \cos^2 \theta),$$

where  $s = E_{CM}^2$ ,  $\theta$  is the scattering angle and we have, as usual, neglected the lepton masses compared to  $E_{CM}$ . When  $Z^0$  exchange is taken into account, this expression is modified to become

$$\sigma(s, \theta) = \frac{\pi \alpha^2}{2s} [F(s)(1 + \cos^2 \theta) + G(s) \cos \theta],$$

with

$$F(s) = 1 + \frac{1}{128\pi^2} \left( \frac{M_Z^2}{M_Z^2 - s} \right)^2 \left( \frac{sG_F}{\alpha} \right)^2$$

and

$$G(s) = -\frac{1}{2\sqrt{2}\pi} \left( \frac{M_Z^2}{M_M^2 - s} \right) \left( \frac{sG_F}{\alpha} \right),$$

where we have assumed  $M_Z^2 - s \gg \Gamma_Z$  and  $\sin^2 \theta_W = 1/4$ . The terms in  $\sigma(s, \theta)$  proportional to  $\alpha^2$ ,  $G_F^2$  and  $\alpha G_F$  arise from photon exchange,  $Z^0$  exchange and the interference between the two diagrams of Figure 10.2, respectively. Use these formulas to obtain expressions for the total cross-section  $\sigma_{\text{tot}}(s)$  and the forward–backward asymmetry parameter

$$A(s) \equiv \frac{A_+(s) - A_-(s)}{A_+(s) + A_-(s)}$$

in terms of  $F(s)$  and  $G(s)$ , where

$$A_+(s) = \int_0^1 \sigma(s, \theta) d\cos\theta, \quad \text{and} \quad A_-(s) = \int_{-1}^0 \sigma(s, \theta) d\cos\theta.$$

Hence evaluate the magnitude of  $A(s)$  and the percentage correction to the QED total cross-section  $\sigma_\gamma = 4\pi\alpha^2/3s$  at  $E_{CM} = M_z/3 \approx 30$  GeV generated by the  $Z^0$  exchange contribution.

### 10.7

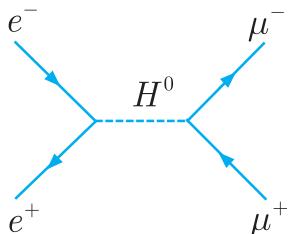
In our discussion of the reaction  $e^+ + e^- \rightarrow \mu^+ + \mu^-$ , we completely neglected the Higgs exchange diagram of Figure 10.30, compared with the dominant diagrams of Figures 10.2(a) and (b). Since the Higgs contribution is obviously largest at  $E_{CM} = M_H$ , where the contributions of Figures 10.2(a) and (b) are expected to be comparable, justify this approximation by using the appropriate  $H^0$  branching ratios to estimate the total cross-section at  $E_{CM} = M_H$  that would arise from Figure 10.30 alone, and compare with estimates of the cross-sections which would arise from Figure 10.2(a) alone. Make the same comparisons of the contributions from Figure 10.2(a) alone and Figure 10.2(b) alone at  $E_{CM} = M_Z$ .

### 10.8

Suppose that the reaction  $e^+e^- \rightarrow Z^0H^0$  was to be studied at a centre-of-mass energy of 500 GeV in a collider operating for  $10^7$  s per year. Given that the cross-section at this energy is approximately 65 fb, what instantaneous luminosity (in units of  $\text{cm}^{-2}\text{s}^{-1}$ ) would be needed to collect 2000 events in one year if the detection efficiency is 10%? To what extent would selecting events with  $b$ -quark jets present increase the statistical error and enhance the  $e^+e^- \rightarrow Z^0H^0$  signal relative to the background reaction  $e^+e^- \rightarrow Z^0Z^0$ ? The branching ratio for  $Z^0 \rightarrow b\bar{b}$  is about 15% (see Problem 10.7).

### 10.9

Estimate the branching ratios for the four-lepton decays (10.58) and (10.60) of the Higgs boson by relating them to the branching ratios for other reactions listed in Table 10.1. Use your answers to estimate the average total number of Higgs decays (10.58) for an LHC energy of 7–8 TeV, where the Higgs production cross-section is about 20 pb



**Figure 10.30** Feynman diagram for the reaction  $e^+e^- \rightarrow \mu^+\mu^-$  via the exchange of a single Higgs boson.

with an error of approximately 15% and an integrated luminosity of  $25 \text{ pb}^{-1}$ . Compare your results with the number actually observed as shown in Figure 10.29.

- 10.10** Draw the dominant lowest-order diagrams for the double Higgs boson production reaction  $e^+ e^- \rightarrow H^0 H^0 Z^0$  at a total centre-of-mass energy of 500 GeV.

---

## Discrete symmetries: $C$ , $P$ , $CP$ and $CPT$

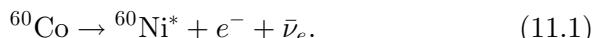
---

In Chapter 5 we introduced charge conjugation ( $C$ ) and parity ( $P$ ) as exact symmetries of the strong and electromagnetic interactions. Here we discuss them in the context of the weak interaction, where they are not conserved. In the course of this, two interconnected themes will emerge. The first is that these effects have their origin in the spin dependence of weak interactions, which is quite remarkable, as we shall see. The second is that while parity and  $C$  parity violations are large effects, there is a weaker combined symmetry, called  $CP$  invariance, which seems to be consistent with all current data on the weak interaction of leptons. However, while extending this symmetry to the weak interactions of hadrons initially leads to some striking experimental successes, it also leads to experimental evidence for deviations from  $CP$  invariance. This topic is discussed in Section 11.2, where we also briefly discuss the more fundamental symmetry of  $CPT$  invariance, which is believed to be respected in all interactions. Finally in Section 11.3, we discuss the important question of whether  $CP$  violation can be fully accounted for within the standard model, or whether some new interaction is required.

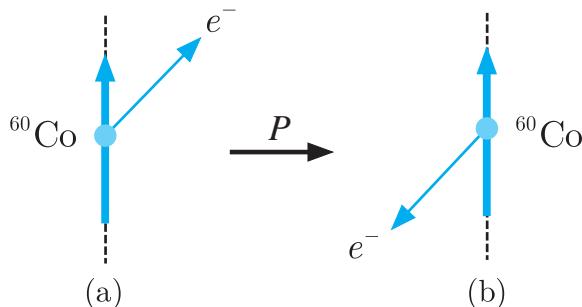
### 11.1 $P$ violation, $C$ violation and $CP$ conservation

The concept of parity was first introduced in the context of atomic physics by Wigner in 1927 and for a long time afterwards its conservation was believed to be a universal law of nature, like angular momentum conservation. The first indications that this was not the

case came from observations on the pionic decays of  $K$ -mesons,<sup>1</sup> and these led Lee and Yang in 1956 to make a thorough study of previous experiments in which parity conservation had been assumed. They came to the startling conclusion that while there was strong evidence for parity conservation in strong and electromagnetic interactions, there was no evidence for its conservation in weak interactions, and they suggested experiments in which the latter assumption could be tested. This led directly to the classic demonstration of parity violation from a study of the  $\beta$  decay of polarised cobalt-60 nuclei. This was made in 1957 by Wu and co-workers, who placed a sample of cobalt-60 inside a solenoid and cooled it to a temperature of 0.01 K. At such temperatures, the interaction of the magnetic moments of the nuclei with the magnetic field overcomes the tendency to thermal disorder, and more of the nuclear spins align parallel to the field direction. A polarised cobalt-60 nucleus produced in this way decays to an excited state of nickel-60 by the process



Parity violation is established by the observation of a ‘forward–backward decay asymmetry’, that is the fact that fewer electrons are emitted in the forward hemisphere than in the backward hemisphere with respect to the spins of the decaying nuclei. This follows because the parity transformation (5.30) reverses all particle momenta  $\mathbf{p}$  while leaving their orbital angular momenta  $\mathbf{r} \times \mathbf{p}$ , and by analogy their spin angular momenta, unchanged. Hence in the rest frame of the decaying nuclei its effect is to reverse the electron momentum while leaving the nuclear spins unchanged, as shown in Figure 11.1. Parity invariance would then require that the rates for



**Figure 11.1** Effect of a parity transformation on  ${}^{60}\text{Co}$  decay (11.1). The thick arrows indicate the direction of the spin of the  ${}^{60}\text{Co}$  nucleus, while the thinner arrows show the direction of the electron's momentum.

<sup>1</sup> Two particles, called at that time  $\tau$  and  $\theta$ , were observed to decay via the weak interaction to  $\pi\pi$  and  $\pi\pi\pi$  final states, respectively, which necessarily had different final-state parities, as we shall show in Section 11.2.1. However, the  $\tau$  and  $\theta$  had properties, including the near equality of their masses, that strongly suggested that they were in fact the same particle, which we now know to be the  $K$  meson. Analysis of this ‘ $\tau - \theta$  puzzle’ suggested that parity was not conserved in the decays. See Problem 11.1.

the two processes in Figures 11.1(a) and (b) were equal, so that equal numbers of electrons would be emitted in the forward and backward hemispheres with respect to the nuclear spins, in contradiction to what was observed.

The discovery of parity violation was a watershed in the history of weak interactions because the effect is large, and the remarkable spin structure of the weak interactions cannot be understood unless it is taken into account. This will be illustrated in Sections 11.1.2 and 11.1.3 below. Firstly, however, we turn to the evidence for  $C$  violation in muon decays.

### 11.1.1 Muon decay symmetries

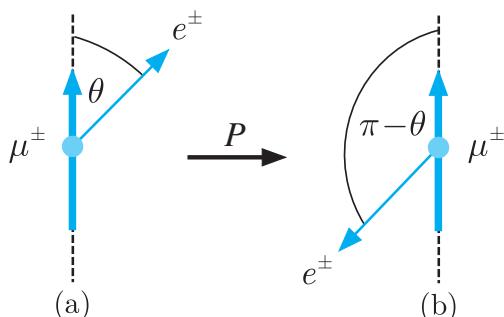
The existence of  $C$  violation and its close relationship to  $P$  violation in leptonic weak interactions are both conveniently illustrated by considering the angular distributions of the electrons and positrons emitted in the decays

$$\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu \quad \text{and} \quad \mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu \quad (11.2)$$

of polarised muons.<sup>2</sup> In the rest frame of the decaying particle these were found to be of the form<sup>3</sup>

$$\Gamma_{\mu^\pm}(\cos \theta) = \frac{1}{2} \Gamma_\pm \left( 1 - \frac{\xi_\pm}{3} \cos \theta \right), \quad (11.3)$$

where  $\theta$  is the angle between the muon spin direction and the direction of the outgoing electron or positron, as shown in Figure 11.2(a). The quantities  $\xi_\pm$  are called the *asymmetry parameters* and  $\Gamma_\pm$  are



**Figure 11.2** Effect of a parity transformation on the muon decays (11.2). The thick arrows indicate the direction of the muon spin, while the thinner arrows show the direction of the electron's momentum.

<sup>2</sup> The muons produced in the decays of charged pions are naturally polarised, as we shall see in Section 11.1.3 below.

<sup>3</sup> An account of early experiments on the decay of positive muons, including experimental details, which are omitted below, is given in Chapter 10 of Trigg (1975), where the cobalt-60 experiment is also discussed in detail.

equal to the total decay rates, or equivalently the inverse lifetimes, that is

$$\tau_{\pm}^{-1} \equiv \int_{-1}^{+1} d \cos \theta \Gamma_{\mu^{\pm}}(\cos \theta) = \Gamma_{\pm}, \quad (11.4)$$

as may easily be checked by direct substitution.

We consider now the consequences of assuming parity and charge conjugation for these decays, starting with the latter as it is the simpler. Charge conjugation transforms all particles into their antiparticles, so that  $\mu^-$  decay converts into  $\mu^+$  decay.  $C$  invariance then implies that the rates and angular distributions for these decays should be the same, that is

$$\Gamma_+ = \Gamma_- \quad (C \text{ invariance}) \quad (11.5a)$$

and

$$\xi_+ = \xi_- \quad (C \text{ invariance}) \quad (11.5b)$$

The parity transformation (5.30) preserves the identity of the particles, but reverses their momenta, while leaving their spins unchanged. Its effect on muon decay is shown in Figure 11.2, where we see that it changes the angle  $\theta$  to  $\pi - \theta$  so that  $\cos \theta$  changes sign. Hence  $P$  invariance implies

$$\Gamma_{\mu^{\pm}}(\cos \theta) = \Gamma_{\mu^{\pm}}(-\cos \theta) \quad (P \text{ invariance}) \quad (11.6a)$$

and substituting (11.3) into (11.6a) then gives the result that the asymmetry parameters vanish,

$$\xi_{\pm} = 0 \quad (P \text{ invariance}) \quad (11.6b)$$

Experimentally, the  $\mu^{\pm}$  lifetimes are equal to very high precision, so that (11.5a) is satisfied, but the measured values of the asymmetry parameters are

$$\xi_- = -\xi_+ = 1.00 \pm 0.04, \quad (11.7)$$

which shows that both  $C$  invariance (11.5b) and  $P$  invariance (11.6b) are violated.

In view of these results, a question that arises is: why do the  $\mu^+$  and  $\mu^-$  have the same lifetime if  $C$  invariance is violated? The answer lies in the principle of  $CP$  conservation, which states that the weak interaction is invariant under the combined operation  $CP$ , even though both  $C$  and  $P$  are separately violated. The  $CP$  operator transforms particles at rest to their corresponding antiparticles at rest, and  $CP$  invariance requires that these states should have identical properties. Thus, in particular, the masses of particles and antiparticles are predicted to be the same. More specifically, if we apply the  $CP$  operator to muon decays, the parity operator changes  $\theta$

to  $\pi - \theta$  as before, while the  $C$  operator changes particles to antiparticles. Hence  $CP$  invariance alone implies that the condition (11.6a) obtained from  $P$  invariance is replaced by the weaker condition

$$\Gamma_{\mu^+}(\cos \theta) = \Gamma_{\mu^-}(-\cos \theta). \quad (11.8)$$

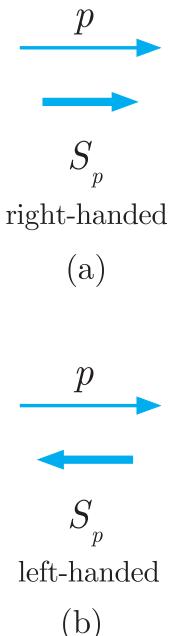
Substituting (11.3) into this equation gives

$$\Gamma_+ = \Gamma_- \quad (CP \text{ invariance}), \quad (11.9a)$$

implying equal lifetimes by (11.4), as well as

$$\xi_+ = -\xi_- \quad (CP \text{ invariance}), \quad (11.9b)$$

in agreement with the experimental results (11.7). Thus  $CP$  invariance retains the symmetry between particles and antiparticles as observed by experiment, at least for muon decays. In fact  $CP$  invariance has been verified in a wide variety of experiments involving weak interactions, and the data are consistent with exact  $CP$  conservation in the weak interactions of leptons. The weak interactions of quarks do not respect  $CP$  invariance, although it is often a very good approximation, and the observed violations are so far confined to the decays of neutral  $K$  mesons and of  $B$  mesons. This will be discussed in Sections 11.2 and 11.3. For the moment, we will concentrate on the properties of the leptons and assume that  $CP$  conservation is exact.



**Figure 11.3** Helicity states of a spin-1/2 particle. The thin arrows represent the momenta of the particles and the thick arrows their spins.

### 11.1.2 Left-handed neutrinos and right-handed antineutrinos

We turn now to the spin structure of the weak interactions, which is closely related to the symmetry properties discussed above. As this spin structure takes its simplest form for neutrinos and antineutrinos, we will discuss these first. In doing so, it is convenient to use *helicity states*, in which the spin is quantised along the direction of motion of the particle, rather than along some arbitrarily chosen ‘z direction’. For a spin-1/2 particle, the spin component along the direction of its motion can be either  $+1/2$  or  $-1/2$ , as illustrated in Figure 11.3, corresponding to positive or negative helicity respectively.<sup>4</sup> These states are called *right-handed* or *left-handed* respectively, since the spin direction corresponds to rotational motion in a right-handed or left-handed sense when viewed along the momentum direction. We will denote these states by subscripts  $R$  or  $L$ , so that  $\nu_L$  means a left-handed neutrino,  $e_R^-$  a right-handed electron and so on.

The remarkable fact about neutrinos and antineutrinos, which only interact via the weak interaction, is that only left-handed neutrinos  $\nu_L$  and right-handed antineutrinos  $\bar{\nu}_R$  have been observed in

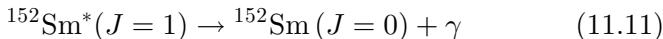
<sup>4</sup> More precisely, the helicity of a particle is defined as the spin projection along its direction of motion, divided by its maximum possible value.

nature. This obviously violates  $C$  invariance, which requires neutrinos and antineutrinos to have identical weak interactions. It also violates  $P$  invariance, which requires the states  $\nu_L$  and  $\nu_R$  to have identical weak interactions, since the parity operator reverses the momentum while leaving the spin unchanged, and so converts a left-handed neutrino into a right-handed neutrino. It is, however, compatible with  $CP$  invariance, since the  $CP$  operator converts a left-handed neutrino to a right-handed antineutrino, as illustrated in Figure 11.4.

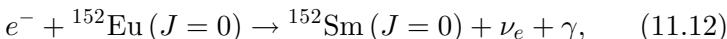
The helicity of the neutrino was first measured in an ingenious experiment by Goldhaber and co-workers in 1958. They studied electron capture in europium-152; that is



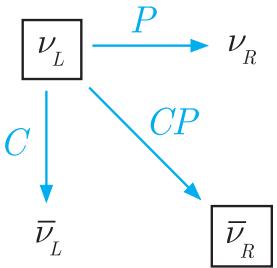
where the spins of the nuclei are shown in brackets. The excited state of samarium that is formed decays to the ground state by  $\gamma$ -emission



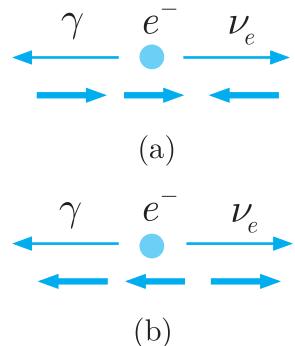
and it is these  $\gamma$  rays that were detected in the experiment. In the reaction (11.10) the electrons are captured from the  $K$ -shell and the initial state has zero momentum, so that the neutrino and the  ${}^{152}\text{Sm}^*$  nucleus recoil in opposite directions. The experiment selected events in which the photon was emitted in the direction of motion of the decaying  ${}^{152}\text{Sm}^*$  nucleus,<sup>5</sup> so that overall the observed reaction was



where the three final-state particles were collinear, and the neutrino and photon emerged in opposite directions, as shown in Figure 11.5. The helicity of the neutrino can then be deduced from the measured helicity of the photon by applying angular momentum conservation about the event axis to the overall reaction (11.12). In doing this, no orbital angular momentum is involved, because the initial electron is captured from the  $K$ -shell and the final-state particles all move along the event axis. Hence the spin components of the neutrino and photon, which can be  $\pm 1/2$  and  $\pm 1$ , respectively,<sup>6</sup> must add to give the spin component of the initial electron, which can be  $\pm 1/2$ . This gives two possible spin configurations, as shown in Figures 11.5(a)



**Figure 11.4** Effect of  $C$ ,  $P$  and  $CP$  transformations on a left-handed neutrino  $\nu_L$ . Only the states shown in boxes have been observed in nature.



**Figure 11.5** Possible helicities of the photons and neutrinos emitted in (11.12) for those events in which they are emitted in opposite directions. Measurements of the photon helicity show that only events corresponding to case (a), left-handed neutrinos, are observed.

<sup>5</sup> This was done by resonant scattering from a second samarium target. It relies on the fact that those  $\gamma$ -rays travelling in the opposite direction to the neutrino have slightly more energy than those emitted in other directions, and only the former have enough energy to excite the resonance level.

<sup>6</sup> There are only two possible spin states for physical photons because there are only two possible polarisation states for electromagnetic waves. In fact the spin states with components  $+1$  and  $-1$  in the direction of the photon momentum correspond to right and left circular polarisation respectively.

and (b), and in each case the helicities of the photon and neutrino are the same. In the actual experiment, the polarisation of the photons was determined by studying their absorption in magnetised iron,<sup>7</sup> and the results obtained were consistent with the occurrence of left-handed neutrinos only, corresponding to Figure 11.5(a).

A similar experiment for antineutrinos has been carried out involving the emission of a  $\gamma$ -ray following the positron decay of the  $^{203}\text{Hg}$  nucleus. We will not discuss this, but note that the measured photon polarisation in this case is consistent with the occurrence of right-handed antineutrinos only, in agreement with the statement at the beginning of this section.

### 11.1.3 Pion and muon decays revisited

In our discussion of the charged current interactions in Section 9.1, we completely ignored their spin dependence. This is of a special form, called a  $V$ - $A$  interaction.<sup>8</sup> Here we shall consider only the most important characteristic of this spin dependence, which is that the results discussed for neutrinos in Section 11.1.2 hold for all fermions in the ultra-relativistic limit. That is, in the limit that their velocities approach that of light, only left-handed fermions  $\nu_L, e_L^-$ , etc., and right-handed antifermions  $\bar{\nu}_R, e_R^+$  are emitted in charged current interactions. For neutrinos this is always a good approximation. For other particles it is only a good approximation for large energies  $E$ , when the contributions of the ‘forbidden’ helicity states  $e_R^-, e_L^+$ , etc., are suppressed by factors that are typically of order

$$\left(1 - \frac{v}{c}\right) \approx \frac{m^2}{2E^2}, \quad (11.13)$$

where  $m$  is the appropriate fermion mass and  $v$  is its velocity.

The above spin properties can be verified most easily for the electrons and muons emitted in weak decays, by directly measuring their spins. Here we shall assume them to hold and use them to understand some interesting features of pion and muon decays.

---

<sup>7</sup> This is discussed further in Section 7.5 of Perkins (1987), where a brief description of the apparatus is also given.

<sup>8</sup> This name is derived from the behaviour under a parity transformation of the weak interaction analogue of the electromagnetic current. The letter  $V$  denotes a proper vector, that is one whose direction is reversed by a parity transformation, such as momentum  $\mathbf{p}$ . The familiar electric current, to which photons couple, transforms as a proper vector under parity. Because parity is not conserved in weak interactions, the corresponding weak current, to which  $W^\pm$  bosons couple, has in addition to a vector component another component whose direction is unchanged by a parity transformation. Such a quantity is called an axial vector ( $A$ ); an example is orbital angular momentum  $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ .

We start by considering the pion decay modes

$$\pi^+ \rightarrow \ell^+ + \nu_\ell \quad (\ell = e, \mu). \quad (11.14)$$

In the rest frame of the decaying pion, the charged lepton and the neutrino recoil in opposite directions, and because the pion has zero spin, their spins must be opposed to satisfy angular momentum conservation about the decay axis. Since the neutrino is left-handed, it follows that the charged lepton must also be left-handed, as shown in Figure 11.6, in contradiction to the expectations for a relativistic antilepton. For the case of a positive muon this is unimportant, since it is easy to check that it recoils non-relativistically and so both helicity states are allowed. However, if a positron is emitted it does recoil relativistically, implying that this mode is suppressed by a factor that is estimated from (11.13) to be of order  $\frac{1}{2}(m_e/m_\pi)^2 = 9.3 \times 10^{-5}$ . Thus the positron decay mode is predicted to be much rarer than the muonic mode. This is indeed the case and the measured ratio

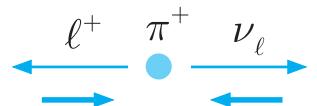
$$\frac{\Gamma(\pi^+ \rightarrow e^+ \nu_e)}{\Gamma(\pi^+ \rightarrow \mu^+ \nu_\mu)} = (1.230 \pm 0.004) \times 10^{-4} \quad (11.15)$$

is in excellent agreement with a full calculation that takes into account both the above suppression and the difference in the density of final states for the two reactions.

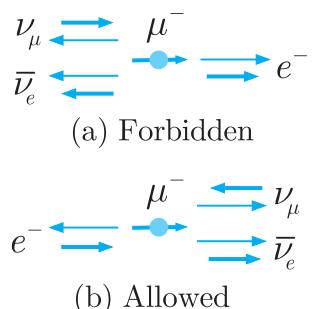
A second consequence of the helicity argument is that the muons emitted in pion decays are polarised (see Figure 11.6) and can therefore be used to measure the muon decay asymmetries discussed in Section 11.1.1. These have their origin in the spin structure of the interaction, as we shall illustrate for the highest-energy electrons emitted in the decays (11.2). These have energy

$$E = \frac{m_\mu}{2} \left( 1 + \frac{m_e^2}{m_\mu^2} \right) \gg m_e \quad (11.16)$$

and correspond to decays in which the neutrino and antineutrino are both emitted in the direction opposite to the electron. This is illustrated in Figure 11.7 for the two simplest cases in which the electron is emitted in the muon spin direction, and opposite to it. Since the neutrino and antineutrino have opposite helicities, the muon and electron must have the same spin component along the event axis in order to conserve angular momentum, implying the electron helicities shown in Figure 11.7. When combined with the fact that the relativistic electrons emitted must be left-handed, this implies that electrons cannot be emitted in the muon spin direction. We thus see that the spin structure of the interaction automatically gives rise to a forward-backward asymmetry in polarised muon decays. Of course, not all the electrons have the maximum energy (11.16), and the actual asymmetry, averaged over all electron energies, can only be calculated by using the full form of the  $V-A$  interaction. We will not



**Figure 11.6** Helicities of the charged leptons and neutrinos emitted in the pion decays (11.14).



**Figure 11.7** Muon decays in which electrons of the highest possible energy are emitted  
 (a) in the muon spin direction and (b) opposite to the muon spin direction. Case (a) is heavily suppressed, because in the limit that its velocity approaches that of light, an electron emitted in a charge current reaction must necessarily be left-handed.

pursue this further, but merely note that the resulting prediction is in excellent agreement with the measured values of (11.7).

## 11.2 CP violation and particle–antiparticle mixing

We now turn to experiments on the weak interaction decays of hadrons. In this case, small deviations from  $CP$  invariance were discovered in 1964 in neutral kaon decays and, from 2001 onwards much larger violations were also observed in the decays of both charged and neutral mesons with non-zero bottom quantum number,  $\bar{B} \neq 0$ . No violations have been observed in any other reactions, including all baryon decays, and in many cases precise experiments show that  $CP$  violation must be very small, if present at all.

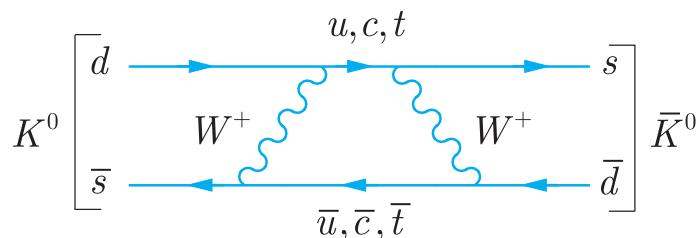
The theory of  $CP$  violation in the standard model will be discussed in Section 11.3. In this section we will discuss the experimental evidence for the distinctive pattern of results that it needs to explain. In order to understand the most important of these experiments, we will also need to introduce the phenomenon of particle–antiparticle mixing, which occurs for all neutral mesons with non-zero values of strangeness, charm or bottom, whether or not  $CP$  is conserved. We will first discuss the evidence for these various phenomena in neutral kaon decays, where they were first discovered, and then consider other cases, especially the neutral  $B$  meson system, where recent discoveries have been of great importance in testing predictions of the standard model.

### 11.2.1 CP eigenstates of neutral kaons

In Section 3.4 we saw that there are two neutral kaon states

$$K^0(498) = d\bar{s} \quad \text{and} \quad \bar{K}^0(498) = s\bar{d}, \quad (11.17)$$

which have strangeness  $S = +1$  and  $S = -1$ , respectively. However, because strangeness is not conserved in weak interactions, these states can be converted into each other by higher-order weak processes like that shown in Figure 11.8. This is in marked contrast



**Figure 11.8** Example of a process that can convert a  $K^0 = d\bar{s}$  state into a  $\bar{K}^0 = s\bar{d}$  state.

to most other particle–antiparticle systems, for which such transitions are forbidden because the particle and its antiparticle differ by quantum numbers that are conserved in all known interactions. For example, the  $\pi^+$  and  $\pi^-$  have opposite electric charges, and the neutron and antineutron have opposite baryon number. For neutral kaons, however, there is no conserved quantum number to distinguish the  $K^0$  and  $\bar{K}^0$  states when weak interactions are taken into account and the observed physical particles correspond not to the  $K^0$  and  $\bar{K}^0$  states themselves, but to linear combinations of them. This is the phenomenon of  $K^0 - \bar{K}^0$  mixing, which we shall now describe. We shall see later that a similar phenomenon occurs for some other states, including neutral  $B$  mesons and neutral  $D$  mesons.

We start by assuming that  $CP$  invariance is exact and that neutral kaons are eigenstates of the combined  $CP$  operator. To find the effect of this operator, we note that the operator  $\hat{C}$  changes a quark into its antiquark, and so interchanges the states (11.17), giving

$$\hat{C}|K^0, \mathbf{p}\rangle = -|\bar{K}^0, \mathbf{p}\rangle \quad \text{and} \quad \hat{C}|\bar{K}^0, \mathbf{p}\rangle = -|K^0, \mathbf{p}\rangle. \quad (11.18)$$

Here we have used the Dirac notation to denote a state with momentum  $\mathbf{p}$ , and the signs in (11.18) are chosen to accord with the standard convention used to discuss neutral kaon decays.<sup>9</sup> Since kaons have negative intrinsic parity, we also have

$$\hat{P}|K^0, \mathbf{p} = \mathbf{0}\rangle = -|K^0, \mathbf{p} = \mathbf{0}\rangle \quad \text{and} \quad \hat{P}|\bar{K}^0, \mathbf{p} = \mathbf{0}\rangle = -|\bar{K}^0, \mathbf{p} = \mathbf{0}\rangle, \quad (11.19)$$

which together with (11.18) gives

$$\hat{C}\hat{P}|K^0, \mathbf{p} = \mathbf{0}\rangle = |\bar{K}^0, \mathbf{p} = \mathbf{0}\rangle \quad \text{and} \quad \hat{C}\hat{P}|\bar{K}^0, \mathbf{p} = \mathbf{0}\rangle = |K^0, \mathbf{p} = \mathbf{0}\rangle \quad (11.20)$$

for the action of the combined operator  $\hat{C}\hat{P}$ . From these equations it is straightforward to verify that the  $CP$  eigenstates are

$$|K_1^0, \mathbf{p} = \mathbf{0}\rangle \equiv \frac{1}{\sqrt{2}}[|K^0, \mathbf{p} = \mathbf{0}\rangle + |\bar{K}^0, \mathbf{p} = \mathbf{0}\rangle] \quad (11.21a)$$

and

$$|K_2^0, \mathbf{p} = \mathbf{0}\rangle \equiv \frac{1}{\sqrt{2}}[|K^0, \mathbf{p} = \mathbf{0}\rangle - |\bar{K}^0, \mathbf{p} = \mathbf{0}\rangle], \quad (11.21b)$$

with

$$\hat{C}\hat{P}|K_1^0, \mathbf{p} = \mathbf{0}\rangle = |K_1^0, \mathbf{p} = \mathbf{0}\rangle \quad (11.22a)$$

and

$$\hat{C}\hat{P}|K_2^0, \mathbf{p} = \mathbf{0}\rangle = -|K_2^0, \mathbf{p} = \mathbf{0}\rangle, \quad (11.22b)$$

---

<sup>9</sup> See Equations (5.50a) and (5.50b). If the negative signs were omitted in (11.18) then some sign changes would be necessary in the equations that follow, but the physical predictions would remain the same.

respectively. According to our initial assumptions, these are the states corresponding to the observed particles, and while the  $K_1^0$  should decay entirely to states with  $CP = 1$ , the  $K_2^0$  should decay entirely into states with  $CP = -1$ . As we shall see, this leads to the predictions that the pionic decays

$$K_1^0 \rightarrow \pi^+ \pi^-, \quad \pi^0 \pi^0 \quad (11.23a)$$

and

$$K_2^0 \rightarrow \pi^+ \pi^- \pi^0, \quad \pi^0 \pi^0 \pi^0 \quad (11.23b)$$

are allowed by  $CP$  conservation, while the decays

$$K_1^0 \rightarrow \pi^+ \pi^- \pi^0, \quad \pi^0 \pi^0 \pi^0 \quad (11.24a)$$

and

$$K_2^0 \rightarrow \pi^+ \pi^-, \quad \pi^0 \pi^0 \quad (11.24b)$$

are forbidden.

The results (11.23) and (11.24) are central to the interpretation of the observed neutral kaon decays that follows. They are established by showing that the two-pion final states have  $CP = +1$ , while the three-pion final states have  $CP = -1$ . We will show this by considering each possible final state in turn.

### 11.2.1(a) Two-pion final states

Since the kaon has spin-0, the pion pair must have zero orbital angular momentum in the rest frame of the decaying particle, by angular momentum conservation. Its parity is therefore given by

$$P = P_\pi^2 (-1)^L = 1, \quad (11.25)$$

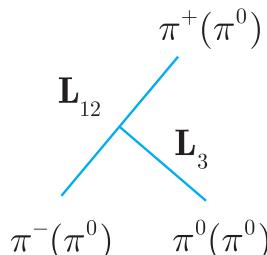
where  $P_\pi = -1$  is the intrinsic parity of the pion. For the  $\pi^0$  the  $C$  parity is given by

$$C = (C_{\pi^0})^2 = 1$$

where  $C_{\pi^0}$  is the  $C$ -parity of the neutral pion. Combining these results gives  $CP = +1$  for the  $\pi^0 \pi^0$  final state, as required. For the  $\pi^+ \pi^-$  state, the parity is again given by (11.25), while charge conjugation interchanges the two charged pions, giving

$$C = (-1)^L = 1, \quad (11.26)$$

by (5.52). Combining (11.25) and (11.26) again gives  $CP = +1$  for the  $\pi^+ \pi^-$  state, as required.



**Figure 11.9** Angular momenta in the three-pion systems  $\pi^+ \pi^- \pi^0$  and (in brackets)  $\pi^0 \pi^0 \pi^0$ .

The argument for three-pion final states is more complicated because there are two orbital angular momenta to consider, as shown in Figure 11.9. Here  $\mathbf{L}_{12}$  is the orbital angular momentum of the chosen pair in their mutual centre-of-mass frame and  $\mathbf{L}_3$  is the orbital

### 11.2.1(b) Three-pion final states

angular momentum of the third pion about the centre-of-mass of the pair in the overall centre-of-mass frame. The total orbital angular momentum

$$\mathbf{L} = \mathbf{L}_{12} + \mathbf{L}_3 = \mathbf{0},$$

since the decaying particle has spin-0, and this can only be satisfied if  $L_{12} = L_3$ . This implies that the parity of the final state is

$$P = P_\pi^3 (-1)^{L_{12}} (-1)^{L_3} = -1. \quad (11.27)$$

For the state  $\pi^0\pi^0\pi^0$ , the  $C$  parity is

$$C = (C_{\pi^0})^3 = 1$$

and so combining this with (11.27) gives  $CP = -1$  overall, as required. For the state  $\pi^+\pi^-\pi^0$ , the parity is again given by (11.27) but now charge conjugation interchanges the  $\pi^+$  and  $\pi^-$ , giving a factor  $(-1)^{L_{12}}$  in analogy with (11.26), together with a factor for the  $C$  parity of a neutral pion. That is, we have

$$C = C_{\pi^0} (-1)^{L_{12}} = (-1)^{L_{12}}$$

and combining this with (11.27) we obtain

$$CP = (-1)^{L_{12}+1}.$$

The value of  $L_{12}$  can be determined experimentally by studying the angular distribution of the pions, leading to the results  $L_{12} = 0$  and  $CP = -1$ , as required.

### 11.2.2 The discovery of $CP$ violation

Two neutral kaons are observed experimentally, called  $K^0$ -short and  $K^0$ -long, denoted  $K_S^0$  and  $K_L^0$ , respectively. They have almost equal masses<sup>10</sup> of about  $498 \text{ MeV}/c^2$ , but very different lifetimes and decay modes. The  $K_S^0$  has a lifetime of about  $9 \times 10^{-11} \text{ s}$  and principal decay modes

$$K_S^0 \rightarrow \pi^0 + \pi^0 \quad (B = 0.31) \quad K_S^0 \rightarrow \pi^+ + \pi^- \quad (B = 0.69), \quad (11.28)$$

with the branching ratios shown, while the longer-lived  $K_L^0$  has a lifetime of about  $5 \times 10^{-8} \text{ s}$  and its principal decay modes are

$$K_L^0 \rightarrow \pi^0 + \pi^0 + \pi^0 \quad (B = 0.20), \quad K_L^0 \rightarrow \pi^+ + \pi^- + \pi^0 \quad (B = 0.13) \quad (11.29)$$

and

$$K_L^0 \rightarrow \pi^\pm + l^\mp + \nu_l (\bar{\nu}_l) \quad (B = 0.67). \quad (11.30)$$

---

<sup>10</sup> The tiny mass difference between them will be determined in Section 11.2.4.

If  $CP$  is conserved, the selection rules (11.23) and (11.24) immediately suggest the identification

$$K_S^0 = K_1^0 \quad \text{and} \quad K_L^0 = K_2^0. \quad (11.31)$$

However, in 1964 Christenson, Cronin, Fitch and Turlay discovered that  $K_L^0$  also decayed to two pions

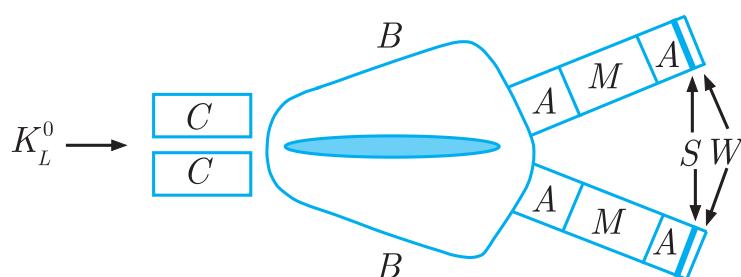
$$K_L^0 \rightarrow \pi^+ + \pi^-, \quad (11.32)$$

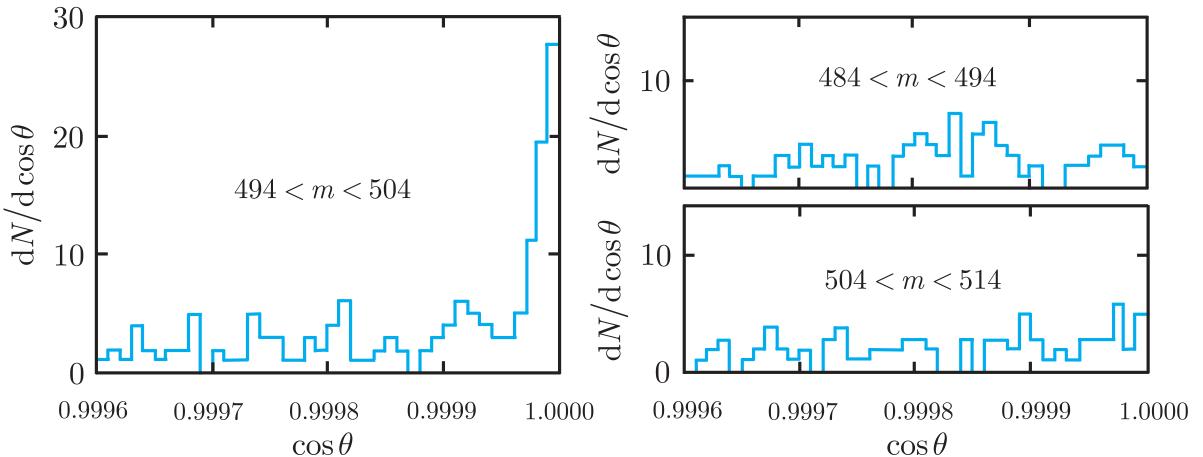
with a branching ratio of order  $10^{-3}$ . This result is clear evidence of  $CP$  violation, since two-pion final states have  $CP = 1$ , whereas the dominant three-pion final states (11.29) have  $CP = -1$ , as we showed above.

The first stage in demonstrating the existence of the decay mode (11.32) was to prepare a  $K_L^0$  beam. This was done by colliding a 30 GeV proton beam with a metal target and forming a secondary beam from the produced particles using a lead collimator. Charged particles were swept out of the latter beam by a bending magnet, while most of the photons were removed by passing it through a 4 cm thick block of lead. At this point the beam contained both  $K_S^0$  and  $K_L^0$  mesons, but by the time it reached the detection apparatus 18 m away the short-lived  $K_S^0$  component had completely decayed, leaving a beam that consisted primarily of  $K_L^0$  particles and neutrons, together with a few photons that had penetrated the lead block.

The detection apparatus used is shown schematically in Figure 11.10. It consisted of a second lead collimator to precisely define the direction of the beam, a helium-filled bag within which the observed decays occurred and a pair of symmetrically placed spectrometers to observe the pions produced in the decay. The main part of each spectrometer consisted of a pair of spark chambers separated by a bending magnet to measure the direction and momentum of any incident charged particles. However, these were only triggered when scintillation and water Čerenkov counters, also contained in the spectrometers, recorded the passage of two charged particles with velocities greater than or equal to  $0.75c$ . This automatically eliminated many background events involving slow-moving particles produced, for example, in collisions involving the residual neutrons

**Figure 11.10** Schematic diagram of the apparatus used in the discovery of  $CP$  violation. The  $K_L^0$  beam entered a helium-filled bag  $B$  through a lead collimator  $C$ , and those  $CP$ -violating decays that occurred in the shaded region were detected by the symmetrically spaced spectrometers. These each contained a pair of spark chambers  $A$  separated by a magnet  $M$ , followed by a scintillation counter  $S$  and a water Čerenkov counter  $W$ . (J.H. Christenson *et al.* 1964. Reproduced with permission from the American Physical Society.)





**Figure 11.11** Angular distribution of the  $\pi^+\pi^-$  pairs detected using the apparatus of Figure 11.10, where  $\theta$  is the angle between the line-of-flight of the centre-of-mass of the pair and the initial beam direction. Results are shown for an invariant mass range including the  $K_L^0$  mass ( $498 \text{ MeV}/c^2$ ), and for two neighbouring mass ranges. (J.H. Christenson *et al.* 1964. Reproduced with permission from the American Physical Society.)

in the beam. Even so, only some of the detected events corresponded to the sought-after decays (11.32). These were identified using the facts that if the pair of charged particles detected arose from  $K_L^0$  decays they should have opposite charges, their invariant mass should correspond to the  $K_L^0$  mass of  $498 \text{ MeV}/c^2$  and their net momentum should lie in the direction of the  $K_L^0$  beam.

The results of the experiment are illustrated in Figure 11.11. This shows the observed distributions of oppositely charged pairs as a function of the angle  $\theta$  between the net momentum of the pair and the beam direction. The point here is that decay events should occur only at zero angles to the beam and at an invariant mass of  $498 \text{ MeV}/c^2$ , whereas background events from other sources will presumably have a smooth dependence on this angle and on the invariant mass. As can be seen, there is a clear peak in the beam direction  $\theta = 0^\circ$  for the appropriate invariant mass range  $494 < m < 504 \text{ GeV}/c^2$ , which is interpreted as arising from the  $CP$ -violating decay (11.32).

### \*11.2.3 $CP$ -violating $K_L^0$ decays

Because  $CP$  is not conserved, the physical states  $K_S^0$  and  $K_L^0$  need not correspond to the  $CP$  eigenstates  $K_1^0$  and  $K_2^0$  as assumed in (11.31), but can contain small components of states with the opposite  $CP$ , that is

$$\left| K_S^0, \mathbf{p} = \mathbf{0} \right\rangle = \frac{1}{(1 + |\varepsilon|^2)^{1/2}} \left[ \left| K_1^0, \mathbf{p} = \mathbf{0} \right\rangle + \varepsilon \left| K_2^0, \mathbf{p} = \mathbf{0} \right\rangle \right] \quad (11.33a)$$

and

$$\left| K_L^0, \mathbf{p} = \mathbf{0} \right\rangle = \frac{1}{(1 + |\varepsilon|^2)^{1/2}} \left[ \varepsilon \left| K_1^0, \mathbf{p} = \mathbf{0} \right\rangle + \left| K_2^0, \mathbf{p} = \mathbf{0} \right\rangle \right], \quad (11.33b)$$

where  $\varepsilon$  is a small complex parameter. This mixing can be verified rather directly by measurements on the semileptonic decays

$$K_L^0 \rightarrow \pi^- + e^+ + \nu_e \quad \text{and} \quad K_L^0 \rightarrow \pi^+ + e^- + \bar{\nu}_e. \quad (11.34)$$

To understand this, we note that the  $K^0$  and  $\bar{K}^0$  can decay by the semileptonic reactions

$$K^0 \rightarrow \pi^- + e^+ + \nu_e \quad \text{and} \quad \bar{K}^0 \rightarrow \pi^+ + e^- + \bar{\nu}_e, \quad (11.35)$$

whereas the corresponding reactions

$$K^0 \rightarrow \pi^+ + e^- + \bar{\nu}_e \quad \text{and} \quad \bar{K}^0 \rightarrow \pi^- + e^+ + \nu_e \quad (11.36)$$

are forbidden by the  $\Delta S = \Delta Q$  selection rule (9.30). Thus the relative yields  $N^+$  and  $N^-$  of positrons and electrons in the decays (11.34) yield a direct measure of the relative probabilities of finding a  $K^0$  or  $\bar{K}^0$ , respectively, in the decaying  $K_L^0$ . From (11.33b) and (11.21) these are given by

$$N^\pm \propto |1 \pm \varepsilon|^2 (1 + |\varepsilon|^2)^{-1}$$

for  $K^0$  and  $\bar{K}^0$ , respectively, so if we neglect terms of order  $|\varepsilon|^2$ , the asymmetry

$$A \equiv (N^+ - N^-)/(N^+ + N^-) = 2\text{Re}\varepsilon \quad (11.37)$$

for a pure  $K_L^0$  beam.

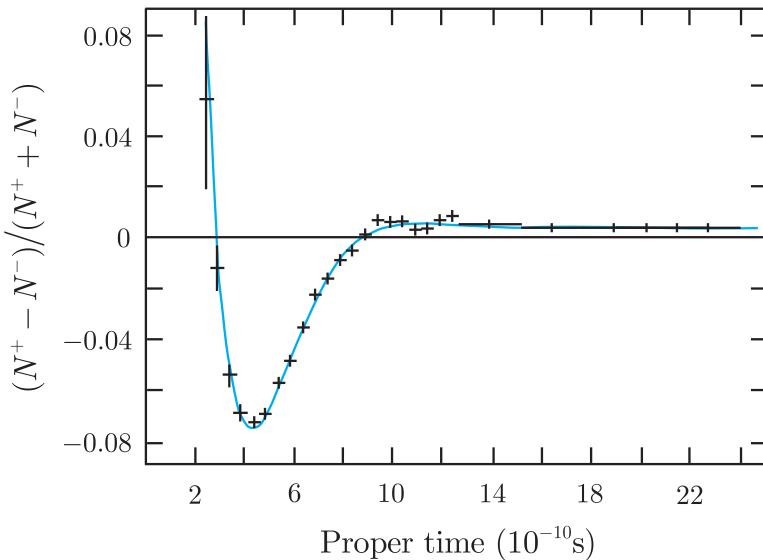
Figure 11.12 shows data on the asymmetry (11.37) as a function of proper time, that is the time in the rest frame of the decaying particles, for the case of a beam that is initially predominantly a  $K^0$  state. After the initial oscillations, which occur because the  $K_S^0$  decays more rapidly than the  $K_L^0$ , only the  $K_L^0$  state survives and there is seen to be an asymmetry whose value is  $2\text{Re}\varepsilon \approx 3.2 \times 10^{-3}$ .

The magnitude of  $\varepsilon$  can be deduced from the measured rates of the two  $CP$ -violating decays

$$\begin{aligned} K_L^0 &\rightarrow \pi^+ + \pi^- \quad (B = 2.0 \times 10^{-3}), \\ K_L^0 &\rightarrow \pi^0 + \pi^0 \quad (B = 8.6 \times 10^{-4}), \end{aligned} \quad (11.38)$$

which have both been observed with the branching ratios shown. These can occur in two different ways: either

- (i)  $CP$  violation by mixing, in which the  $CP$ -forbidden  $K_1^0$  component in the  $K_L^0$  decays via the  $CP$ -allowed process (11.23), giving a



**Figure 11.12** The charge asymmetry observed for  $K^0 \rightarrow \pi^- e^+ \nu_e$  and  $\bar{K}^0 \rightarrow \pi^+ e^- \bar{\nu}_e$  as a function of proper time, for a beam that is initially predominantly  $K^0$ . (Adapted from Gjesdal *et al.* 1974. Reproduced with permission from Elsevier.)

contribution proportional to the probability

$$|\varepsilon|^2 / (1 + |\varepsilon|^2) \approx |\varepsilon|^2$$

of finding a  $K_1^0$  component in the  $K_L^0$ ; or

(ii) *Direct CP violation*, in which the  $K_2^0$  component in the  $K_L^0$  decays directly to pion pairs via the *CP*-violating reactions (11.24b).

In practice, *CP* violation by mixing dominates, as can be seen by comparing the ratios of probability amplitudes  $\mathcal{M}$  for  $K_L^0$  and  $K_s^0$  to  $\pi^0\pi^0$  and  $\pi^+\pi^-$ . Defining

$$\eta_{00} \equiv \frac{\mathcal{M}(K_L^0 \rightarrow \pi^0\pi^0)}{\mathcal{M}(K_S^0 \rightarrow \pi^0\pi^0)}, \quad \eta_{+-} \equiv \frac{\mathcal{M}(K_L^0 \rightarrow \pi^+\pi^-)}{\mathcal{M}(K_S^0 \rightarrow \pi^+\pi^-)}, \quad (11.39)$$

it is easily seen from (11.33), that if direct *CP* violation in decay is completely neglected, these ratios are given by

$$\eta_{+-} = \eta_{00} = \varepsilon. \quad (11.40)$$

The measured values of the ratios of decay rates yields

$$|\eta_{00}| = (2.220 \pm 0.007) \times 10^{-3}, \quad |\eta_{+-}| = (2.232 \pm 0.007) \times 10^{-3}, \quad (11.41)$$

which are consistent with the prediction of *CP* violation by mixing alone. However, complete agreement with all the data on  $K_L^0$  decays also requires a very small, but non-zero, additional contribution from the direct *CP*-violating term. When this is included (11.40) becomes

$$\eta_{+-} = \varepsilon + \varepsilon', \quad \eta_{00} = \varepsilon - \varepsilon', \quad (11.42)$$

where  $\varepsilon'$  is the additional contribution from direct  $CP$  violation. A detailed analysis of all the data with this term included yields

$$|\varepsilon| = (2.228 \pm 0.011) \times 10^{-3}, \quad |\varepsilon'| = (3.67 \pm 0.24) \times 10^{-6}, \quad (11.43)$$

confirming that  $CP$  violation by mixing is the dominant mechanism for  $CP$  violation in  $K_L^0$  decays.<sup>11</sup>

#### 11.2.4 Flavour oscillations and the *CPT* theorem

Although  $CP$  is not conserved, there is good reason to believe that all interactions are invariant under the combined operation of charge conjugation  $C$ , parity  $P$  and time reversal  $T$ , taken in any order. This result is called the *CPT theorem* and can be shown to hold in any relativistic quantum theory in which signals cannot propagate faster than the speed of light. Like  $CP$ , the combined operation of  $CPT$  converts particles at rest to antiparticles at rest, and invariance under this operation requires particles and antiparticles to have the same masses and lifetimes. This is in accord with experiment and is tested by observations on so-called ‘flavour oscillations’ in neutral meson systems. Here we will concentrate on the neutral kaons, as these yield the most precise data.

When neutral kaons are produced in strong interaction processes, associated production means they are almost invariably produced with definite strangeness. For example, the neutral kaon produced in the reaction

$$S = \begin{matrix} \pi^- + p & \rightarrow K^0 + \Lambda^0 \\ 0 & 0 & 1 & -1 \end{matrix} \quad (11.44)$$

must necessarily be in a  $K^0$  state with  $S = 1$ , in order to conserve strangeness. However, if the produced particle is allowed to travel through free space and its strangeness is measured, one finds that it no longer has a definite strangeness  $S = 1$ , but has components with both  $S = 1$  and  $S = -1$  whose intensities oscillate with time. These are called *strangeness oscillations*, and they enable the mass difference between  $K_S^0$  and  $K_L^0$  particles to be measured with extraordinary precision, as we will now show.

In what follows, we shall measure time in the rest frame of the produced particle and define  $t = 0$  as the moment when it is produced. The initial state produced in the reaction (11.44) is

$$|K^0, \mathbf{p}\rangle = \frac{1}{\sqrt{2}} \left[ |K_S^0, \mathbf{p}\rangle + |K_L^0, \mathbf{p}\rangle \right],$$

---

<sup>11</sup>See Gershon and Nir (2014) for a discussion of this analysis, which also shows that  $\varepsilon'$  and  $\varepsilon$  have the same phase (as assumed in inferring (11.43) from the ratio of the real parts) of about  $45^\circ$  to a very good approximation.

where for simplicity we ignore small corrections from  $CP$  violation throughout. At later times, this will become

$$\frac{1}{\sqrt{2}} \left[ a_S(t) |K_S^0, \mathbf{p}\rangle + a_L(t) |\bar{K}_L^0, \mathbf{p}\rangle \right], \quad (11.45)$$

where

$$a_\alpha(t) = e^{-im_\alpha t} e^{-\Gamma_\alpha t/2}, \quad (\alpha = S, L) \quad (11.46)$$

and  $m_\alpha$  and  $\Gamma_\alpha$  are the mass and decay rate of the particle concerned. Here the first exponential factor is the usual oscillating time factor  $\exp(-iEt)$  associated with any stationary state, evaluated in the rest frame of the particle. The second exponential factor reflects the fact that the particles decay and ensures that the probability

$$\left| \frac{1}{\sqrt{2}} a_\alpha(t) \right|^2 = \frac{1}{2} e^{-\Gamma_\alpha t}, \quad (\alpha = S, L)$$

of finding a  $K_S^0$  or  $\bar{K}_L^0$ , decreases exponentially with a mean lifetime

$$\tau_\alpha = \Gamma_\alpha^{-1}, \quad (\alpha = S, L)$$

whose value is given in Section 11.2.2. Because  $\tau_S \ll \tau_L$ , for times  $t$  such that  $\tau_S \ll t \ll \tau_L$ , only the  $K_L^0$  component survives, implying equal intensities for the  $K^0$  and  $\bar{K}^0$  components. Here we are interested in the intensities of the  $K^0$  and  $\bar{K}^0$  components at shorter times, and to deduce these we use

$$|K_S^0, \mathbf{p}\rangle = |K_1^0, \mathbf{p}\rangle \equiv \frac{1}{\sqrt{2}} [|K^0, \mathbf{p}\rangle + |\bar{K}^0, \mathbf{p}\rangle] \quad (11.47a)$$

and

$$|K_L^0, \mathbf{p}\rangle = |K_2^0, \mathbf{p}\rangle \equiv \frac{1}{\sqrt{2}} [|K^0, \mathbf{p}\rangle - |\bar{K}^0, \mathbf{p}\rangle], \quad (11.47b)$$

where again we ignore small  $CP$ -violating effects. The expression (11.45) can then be written in the form

$$A_0(t) |K^0, \mathbf{p}\rangle + \bar{A}_0(t) |\bar{K}^0, \mathbf{p}\rangle, \quad (11.48)$$

where

$$A_0(t) = \frac{1}{2} [a_S(t) + a_L(t)] \quad (11.49a)$$

and

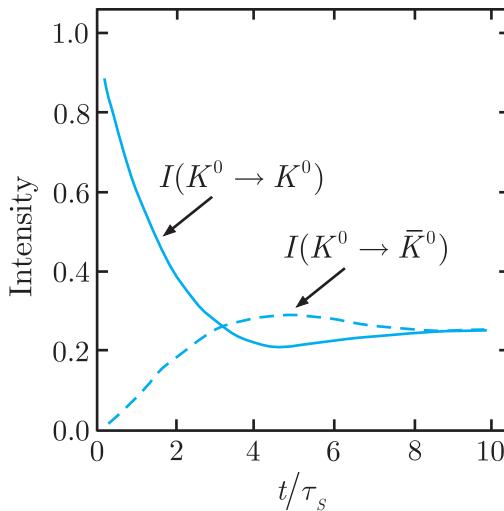
$$\bar{A}_0(t) = \frac{1}{2} [a_S(t) - a_L(t)]. \quad (11.49b)$$

The intensities of the two components are then given by

$$I(K^0 \rightarrow K^0) \equiv |A_0(t)|^2 = \frac{1}{4} \left[ e^{-\Gamma_S t} + e^{-\Gamma_L t} + 2e^{-(\Gamma_S + \Gamma_L)t/2} \cos(\Delta m t) \right] \quad (11.50a)$$

and

$$I(K^0 \rightarrow \bar{K}^0) \equiv |\bar{A}_0(t)|^2 = \frac{1}{4} \left[ e^{-\Gamma_S t} + e^{-\Gamma_L t} - 2e^{-(\Gamma_S + \Gamma_L)t/2} \cos(\Delta m t) \right], \quad (11.50b)$$



**Figure 11.13** Predicted variation with time of the intensities  $I(K^0 \rightarrow K^0)$  (solid line) and  $I(K^0 \rightarrow \bar{K}^0)$  (dashed line) for an initial  $K^0$  beam. The curves are calculated using (11.50) for  $\Delta m \cdot \tau_S = 0.5$ , where  $\Delta m$  is the mass difference (11.51) and  $\tau_S$  is the  $K_S^0$  lifetime.

where

$$\Delta m = |m_S - m_L| \quad (11.51)$$

and we have used (11.47) to explicitly evaluate the amplitudes (11.48). The resulting predictions for  $\Delta m \cdot \tau_S = 0.5$ , where  $\tau_S$  is the lifetime of the  $K_S^0$ , are shown in Figure 11.13, and the *strangeness oscillations* associated with the cosine terms in (11.50a) and (11.50b) are clearly seen. The corresponding argument starting from  $\bar{K}^0$  leads to the results

$$I(\bar{K}^0 \rightarrow \bar{K}^0) = I(K^0 \rightarrow K^0), \quad I(\bar{K}^0 \rightarrow K^0) = I(K^0 \rightarrow \bar{K}^0), \quad (11.50c)$$

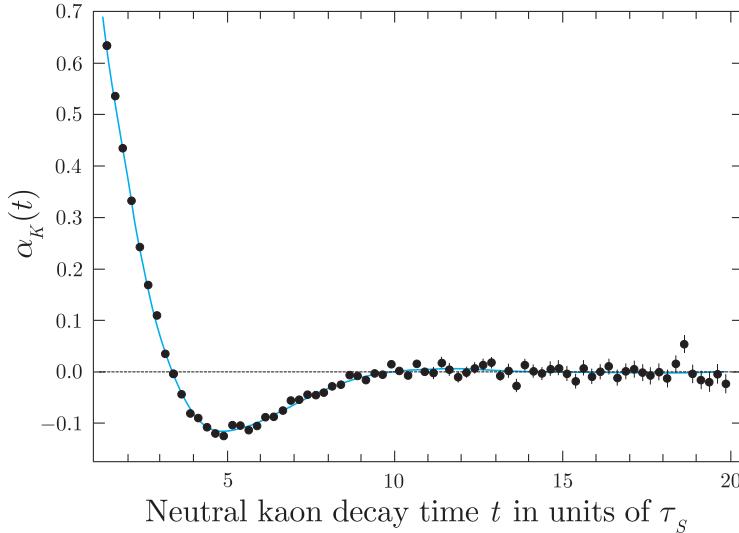
which will be required below.

The strangeness oscillations (11.50) have been observed and the mass difference (11.51) measured in several experiments. At the CPLEAR experiment at CERN, protons and antiprotons were observed to annihilate at rest at the centre of a large  $4\pi$  detector and events corresponding to the reactions



selected, where the identity of the neutral kaon produced is deduced using strangeness conservation. In addition, the particles are produced with sufficiently low momenta that the charged particles can be observed in the same detector and the time between production and decay determined.

Because of oscillations, the identity of the decaying particle as either a  $K^0$  or a  $\bar{K}^0$  will not necessarily be the same as that of the kaon produced initially, but can be identified by detecting it via the allowed leptonic decays (11.35). In this way, the numbers of events in which the nature of the neutral kaon has changed  $N_+(t)$ , or not



**Figure 11.14** Plot of the ratio  $\alpha_K(t)$  as a function of the neutral kaon decay time in units of  $\tau_S$ . The theoretical curve for  $\alpha_K(t)$  is obtained from (11.53) using a mass difference  $\Delta m = (3.483 \pm 0.006)10^{-12} \text{ MeV}/c^2$ . (Adapted from Angelopoulos *et al.*, 2001.)

changed  $N_-(t)$ , can be measured as a function of the time  $t$  between formation and decay. The resulting values for the ratio

$$\alpha_K(t) \equiv \frac{N_+(t) - N_-(t)}{N_+(t) + N_-(t)} \quad (11.52)$$

are plotted in Figure 11.14. The corresponding theoretical predictions are obtained by noting that  $K^0$  and  $\bar{K}^0$  mesons are produced in equal numbers at  $t = 0$ , so that

$$\alpha_K(t) = \frac{I(K^0 \rightarrow K^0) + I(\bar{K}^0 \rightarrow \bar{K}^0) - I(K^0 \rightarrow \bar{K}^0) - I(\bar{K}^0 \rightarrow K^0)}{I(K^0 \rightarrow K^0) + I(\bar{K}^0 \rightarrow \bar{K}^0) + I(K^0 \rightarrow \bar{K}^0) + I(\bar{K}^0 \rightarrow K^0)},$$

and substituting from (11.50) gives

$$\alpha_K(t) = \frac{2 \exp[-(\Gamma_S + \Gamma_L)t/2] \cos(\Delta mt)}{\exp(-\Gamma_S t) + \exp(-\Gamma_L t)}. \quad (11.53)$$

This is in good agreement with the measurements, as can be seen in Figure 11.14, and together with other experiments yields a value

$$\Delta m = (3.483 \pm 0.006)10^{-12} \text{ MeV}/c^2, \quad (11.54)$$

which is close to the value assumed, for illustrative purposes, in Figure 11.13.

The result (11.54) leads to a very precise confirmation that the masses of particles and antiparticles are equal. The  $K_S^0$  and  $K_L^0$  are not antiparticles, but the  $K^0$  and  $\bar{K}^0$  are, as can be seen from (11.18), and the *CPT* theorem requires

$$m_{K^0} = m_{\bar{K}^0}, \quad (11.55)$$

where the masses are defined as the expectation values of the energy for states of zero momentum. If this is so, the mass difference

(11.54) can be shown to arise solely from the possibility of transitions  $K^0 \leftrightarrow \bar{K}^0$ , whose magnitude can be calculated from diagrams like those shown in Figure 11.8. We shall not discuss this further, but merely note that the resulting agreement between the predicted and measured values confirms the identity (11.55) to better than one part in  $10^{18}$ . In contrast, the particle–antiparticle mass relation that has been most precisely tested by direct measurement is  $m_{e^+} = m_{e^-}$ , which is only verified to within an experimental error of order of one part in  $10^8$ .

We conclude by noting that similar flavour oscillations have also been observed in the  $D^0 - \bar{D}^0$ ,  $D_s^0 - \bar{D}_s^0$ ,  $B^0 - \bar{B}^0$  and  $B_s^0 - \bar{B}_s^0$  meson systems, and will be discussed for the  $B^0 - \bar{B}^0$  case in Section 11.2.6.

### 11.2.5 Direct $CP$ violation in decay rates

In Section 11.2.3, we saw that  $CP$  violation in  $K_L^0$  decays is dominated by  $CP$  violation in mixing, with only a very small contribution from direct  $CP$  violation. However, this is not so in other reactions where  $CP$  violation is observed, and for charged meson decays, where there can be no particle–antiparticle mixing without violating charge conservation, direct  $CP$  violation is the only form that can occur. One method of testing for it is by comparing the rates of  $CP$ -conjugate reactions. In particular, consider the decay  $A \rightarrow f$  of a particle  $A$  at rest to a final state  $f$ . Then under a  $CP$  transformation this will become the antiparticle decay  $\bar{A} \rightarrow \bar{f}$ , with the momenta of the final-state particles reversed. Summing over all possible final state momenta then gives

$$\Gamma(A \rightarrow f) = \Gamma(\bar{A} \rightarrow \bar{f}) \quad (11.56)$$

for the total decay rates if  $CP$  is conserved, and a non-zero value of the asymmetry parameter

$$A_{CP}(A \rightarrow f) \equiv \frac{\Gamma(\bar{A} \rightarrow \bar{f}) - \Gamma(A \rightarrow f)}{\Gamma(\bar{A} \rightarrow \bar{f}) + \Gamma(A \rightarrow f)} \quad (11.57)$$

is unambiguous evidence for direct  $CP$  violation.

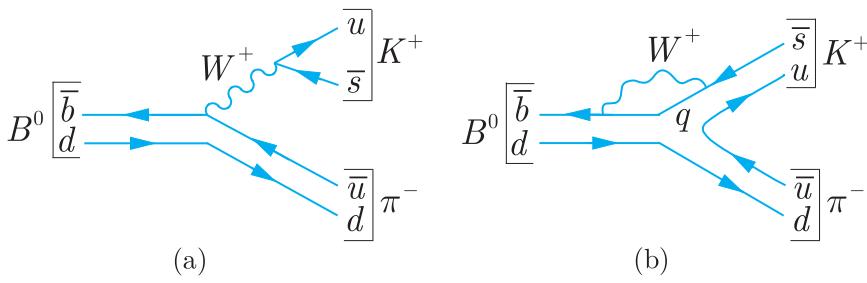
At present, non-zero values of the asymmetry parameter (11.57) have been observed in several decay modes of the lightest meson states with a non-zero bottom quantum number:

$$B^+(5279) = u\bar{b}, \quad B^0(5280) = d\bar{b} \quad (\tilde{B} = +1), \quad (11.58a)$$

$$B^-(5279) = b\bar{u}, \quad \bar{B}^0(5280) = b\bar{d} \quad (\tilde{B} = -1), \quad (11.58b)$$

and the lightest bottom-strange mesons

$$B_s^0(5367) = s\bar{b}, \quad \bar{B}_s^0(5367) = b\bar{s}, \quad (11.58c)$$



**Figure 11.15** Quark diagrams for the decay  $B^0 \rightarrow K^+ + \pi^-$  in lowest-order weak interactions. The quark  $q = \bar{u}, \bar{c}$  or  $\bar{t}$ . The corresponding mechanisms for  $\bar{B}^0 \rightarrow K^- + \pi^+$  are obtained by replacing all particles by their antiparticles.

with quantum numbers  $\tilde{B} = +1$ ,  $S = -1$  and  $\tilde{B} = -1$ ,  $S = +1$ , respectively. To illustrate this, consider the decay modes

$$B^0 \rightarrow K^+ + \pi^- \quad \text{and} \quad \bar{B}^0 \rightarrow K^- + \pi^+ \quad (\text{allowed}), \quad (11.59\text{a})$$

which proceed in lowest order by the mechanisms shown in Figure 11.15. In contrast, there are no possible diagrams leading to the decays

$$\bar{B}^0 \rightarrow K^+ + \pi^- \quad \text{and} \quad B^0 \rightarrow K^- + \pi^+ \quad (\text{forbidden}) \quad (11.59\text{b})$$

in lowest-order weak interactions. Hence neutral  $B$  mesons decaying into  $K^+\pi^-$  and  $K^-\pi^+$  automatically identify themselves as  $B^0$  and  $\bar{B}^0$  mesons, respectively. Comparison of their measured decay rates leads to the asymmetry parameter

$$A_{CP}(B^0 \rightarrow K^+ + \pi^-) = -0.082 \pm 0.006, \quad (11.60\text{a})$$

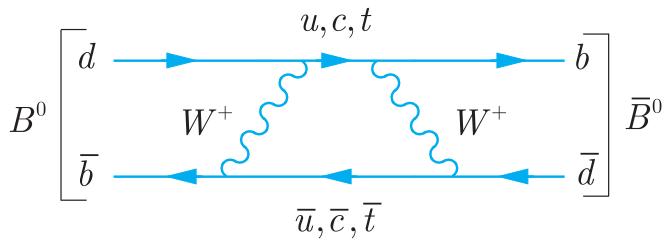
while other examples of non-zero measured asymmetries include

$$A_{CP}(B_s^0 \rightarrow K^+ + \pi^-) = 0.28 \pm 0.04 \quad \text{and} \quad A_{CP}(B^+ \rightarrow K^+ + \eta) = -0.37 \pm 0.08, \quad (11.60\text{b})$$

where  $\eta(549)$  is the neutral meson listed in Table 6.4. These values are very large compared to the contribution of direct  $CP$  violation in  $K_L^0$  decays, but were expected on the basis of the standard model. This is discussed in Section 11.3, where we will also see that  $CP$  violation in charmed meson decays is predicted to be much smaller than for  $B$  mesons and has not yet been observed at the time of writing.

### 11.2.6 $B^0 - \bar{B}^0$ mixing

We now specialise to the  $B^0 - \bar{B}^0$  system, where  $B^0$  is the meson  $B^0(5279) = d\bar{b}$ . The bottom quantum number is not conserved in weak interactions and  $B^0 - \bar{B}^0$  mixing can occur by diagrams like Figure 11.16 in analogy to  $K^0 - \bar{K}^0$  mixing (Figure 11.8), and can be described by the same formalism. In general, for any system  $M^0 - \bar{M}^0$  for which mixing of this type can occur the physical states



**Figure 11.16** Example of a process that can convert a  $B^0 = d\bar{b}$  state into a  $\bar{B}^0 = \bar{d}b$  state.

analogous to the  $K_S^0$  and  $K_L^0$  mesons, which we denote  $M_a$  and  $M_b$ , can be expanded in the form

$$|M_a\rangle = p|M^0\rangle + q|\bar{M}^0\rangle \quad (11.61a)$$

and

$$|M_b\rangle = p|M^0\rangle - q|\bar{M}^0\rangle, \quad (11.61b)$$

where  $p$  and  $q$  are complex parameters satisfying

$$|p|^2 + |q|^2 = 1. \quad (11.61c)$$

*CP conservation in mixing* is defined by

$$|p| = |q|, \quad (11.62)$$

implying equal probabilities for the particle and antiparticle components in the physical states  $M_a$ ,  $M_b$ . On substituting (11.21) into (11.33), one sees that  $K_S^0$  and  $K_L^0$  can be expanded in the form (11.61) with  $|p| \neq |q|$ , so that *CP* violation in mixing occurs, and in fact dominates over direct *CP* violation, as we have seen.

The  $B^0 - \bar{B}^0$  system has much in common with the  $K^0 - \bar{K}^0$  system, but there are also important differences. There are a large number of decay modes possible with large  $Q$  values for  $B^0$  and  $\bar{B}^0$  mesons, so that the lifetimes are much shorter than for kaons. Furthermore, the total decay rates for  $B^0$  mesons are dominated by reactions  $B^0 \rightarrow f$  in which the final state  $f$  is such that the corresponding decay  $\bar{B}^0 \rightarrow f$  is forbidden. Decays to final states  $f$  for which  $B^0 \rightarrow f$  and  $\bar{B}^0 \rightarrow f$  are both possible are relatively rare and in discussions of mixing they can, to a good approximation, be neglected. If this is done, it can be shown (see Section \*11.2.8) that mixing takes the form

$$\left|B_L^0\right\rangle = \frac{1}{\sqrt{2}}[|B^0\rangle + \xi|\bar{B}^0\rangle] \quad (11.63a)$$

and

$$\left|B_H^0\right\rangle = \frac{1}{\sqrt{2}}[|B^0\rangle - \xi|\bar{B}^0\rangle], \quad (11.63b)$$

where

$$\xi = q/p = e^{-2i\beta}, \quad (11.63c)$$

and  $\beta$  is a real phase. In (11.63a) and (11.63b),  $B_L^0$  and  $B_H^0$  are the physical particles analogous to the  $K_S^0$  and  $K_L^0$  mesons. The notations are *L* for ‘light’ and *H* for ‘heavy’, because  $B_L^0$  and  $B_H^0$

have almost identical lifetimes of about  $1.5 \times 10^{-12}$  s. On comparing (11.63) with (11.61) and (11.62), we see that in this approximation *CP* violation by mixing is completely absent. This is not an exact result, but it shows that, in contrast to the  $K^0 - \bar{K}^0$  system, *CP* violation in mixing is a small effect compared to direct *CP* violation, which is a large effect for  $B$  mesons, as shown in Section 11.2.5. This does not mean that mixing is unimportant. In those rare decays in which  $B^0$  and  $\bar{B}^0$  decay to the same final state, it leads to a third form of *CP* violation – called *CP violation in interference* – which will be discussed in the next section. As for neutral kaons, it leads to flavour oscillations that can be used to determine the magnitude of the tiny mass difference between the  $B_L^0$  and  $B_H^0$  mesons, as we will immediately see.

To obtain the form of flavour oscillations in the  $B^0 - \bar{B}^0$  system, we start from the initial states

$$|B^0, t=0\rangle = \frac{1}{\sqrt{2}} [ |B_L^0\rangle + |B_H^0\rangle ]$$

and

$$|\bar{B}^0, t=0\rangle = \frac{1}{\xi \sqrt{2}} [ |B_L^0\rangle - |B_H^0\rangle ],$$

obtained by inverting (11.63). At later times, these will evolve into states

$$|B^0, t\rangle = \frac{1}{\sqrt{2}} [ a_L(t) |B_L^0\rangle + a_H(t) |B_H^0\rangle ] \quad (11.64a)$$

and

$$|\bar{B}^0, t\rangle = \frac{1}{\xi} \left[ \frac{1}{\sqrt{2}} a_L(t) |B_L^0\rangle - \frac{1}{\sqrt{2}} a_H(t) |B_H^0\rangle \right], \quad (11.64b)$$

where

$$\begin{aligned} a_L(t) &= \exp(-im_L t) \exp(-\Gamma t/2), \\ a_H(t) &= \exp(-im_H t) \exp(-\Gamma t/2). \end{aligned} \quad (11.65)$$

Here  $m_L$  and  $m_H$  are the masses of  $B_L^0$  and  $B_H^0$ ,  $\Gamma = \hbar/\tau = (4.3 \times 10^{-12})$  MeV is their common decay width and  $t$ , as usual, is the proper time. Substituting (11.63a) and (11.63b) into (11.64a) and (11.64b) gives

$$|B^0, t\rangle = A(t) |B^0\rangle + \xi \bar{A}(t) |\bar{B}^0\rangle \quad (11.66a)$$

and

$$|\bar{B}^0, t\rangle = A(t) |\bar{B}^0\rangle + (1/\xi) \bar{A}(t) |B^0\rangle, \quad (11.66b)$$

where

$$A(t) = \frac{1}{2} [a_L(t) + a_H(t)] = e^{-\Gamma t/2} e^{-iM t} \cos(\Delta m t/2) \quad (11.67a)$$

and

$$\bar{A}(t) = \frac{1}{2} [a_L(t) - a_H(t)] = i e^{-\Gamma t/2} e^{-iM t} \sin(\Delta m t/2). \quad (11.67b)$$

Here,  $M = (m_L + m_H)/2$  and  $\Delta m = m_H - m_L$ . Since  $|\xi| = 1$ , these relations imply

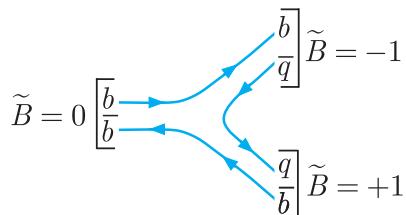
$$I(B^0 \rightarrow B^0) = I(\bar{B}^0 \rightarrow \bar{B}^0) = e^{-\Gamma t} \cos^2(\Delta m t/2) \quad (11.68a)$$

and

$$I(B^0 \rightarrow \bar{B}^0) = I(\bar{B}^0 \rightarrow B^0) = e^{-\Gamma t} \sin^2(\Delta m t / 2), \quad (11.68b)$$

where  $I(B^0 \rightarrow B^0)$  is the probability of finding a  $B^0$  at a time  $t$  in a state evolving from a  $B^0$  at  $t = 0$ , and so on.

The striking thing about (11.68) is that they are independent of the phase  $\beta$  in (11.63c) and would be unchanged if  $\beta = 0$ , when  $B_H^0$  and  $B_L^0$  would be  $CP$  eigenstates. However, in order to test these predictions, we must first confront an experimental problem.

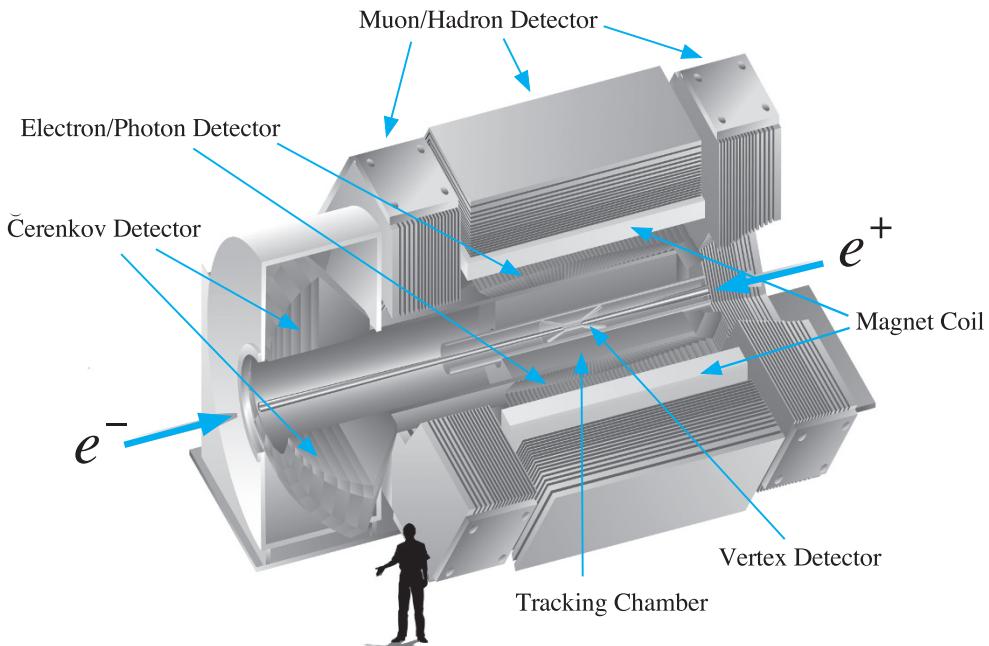


**Figure 11.17** Mechanism for the decay of the 4S  $\tilde{B} = 0$  state  $\Upsilon(10.58)$ ;  $q = u$  or  $d$ .

Because their lifetimes are so short compared to those of the kaons, it is not possible to form well-defined beams of neutral  $B$  mesons, so some other means of studying their decay modes must be found. One way of doing this is by the construction of so-called  $B$  factories. These facilities exploit the properties of the  $\Upsilon(4S) = b\bar{b}$  resonance, which has a mass of  $11.58 \text{ GeV}/c^2$  and a width of only  $20 \text{ MeV}$ . This is a state of bottomonium that is just heavy enough to decay to the lightest meson states with a non-zero bottom quantum number listed in (10.58) by the mechanism of Figure 11.17, but not heavy enough to decay to any other final states with non-zero bottom by the same mechanism.<sup>12</sup> It therefore decays almost entirely to  $B^+ B^-$  and  $B^0 \bar{B}^0$  pairs, in approximately equal numbers. In addition, it has the same quantum numbers  $J^{PC} = 1^{--}$  as the photon and so can be produced in  $e^+e^-$ -collision annihilation by the mechanism of Figure 6.7. There is correspondingly a peak in the  $e^+e^-$  annihilation cross-section, and tuning the beam energies to coincide with this peak results in a copious source of  $B^\pm$ ,  $B^0$  and  $\bar{B}^0$  mesons.

Two such  $B$  factories were constructed to study  $CP$  violation in  $B$  decays: the PEP-II facility at SLAC, California, which operated from 1999 to 2008, and the KEK-B facility in Japan, which operated from 1999 to 2010. At PEP-II, a  $3.1 \text{ GeV}$  positron beam was collided head-on with a  $9.0 \text{ GeV}$  electron beam, whereas at the KEK-B facility a  $3.5 \text{ GeV}$  positron beam was collided with an  $8 \text{ GeV}$  electron beam at a small crossing angle of  $83 \text{ mrad}$ . In each case, the centre-of-mass energy corresponded to the  $\Upsilon(4S)$  mass, while the asymmetric beam energies ensured that  $B$ -mesons were produced with enough

<sup>12</sup>Compare the discussion of charmonium and bottomonium in Sections 6.5.1 and 6.5.2.



**Figure 11.18** Schematic view of the BaBar detector used at the PEP-II facility at SLAC, California. The components of the detector are described in the text. (Photo courtesy of Stanford Linear Accelerator Center.)

momentum to travel a measurable distance before decaying. This is important, because studies of *CP* violation often require the measurement of the time between the production and decay of the  $B$  mesons. This was achieved using the dedicated detectors BaBar at PEP-II and Belle at KEKB.<sup>13</sup> The BaBar detector is shown in Figure 11.18 and, like the Belle detector, had the typical structure of a ‘layered’ detector. It was made up of five subdetectors, which from the inside outwards were:

1. A silicon vertex tracker that provided precise position information on charged tracks, and was also the sole tracking device for very low-energy particles.
2. A tracking (drift) chamber that provided the main momentum measurements for charged particles and helped in particle identification through measurements of the rate of energy loss.
3. A Čerenkov detector that provided charged particle identification.

<sup>13</sup> KEK-B has been shut down until mid-2017 to upgrade it to SuperKEKB, with beams of energies 4.0 and 7.0 GeV, and a projected luminosity about 40 higher than KEK-B.

4. An electromagnetic calorimeter that provided particle identification for electrons and photons.
5. An instrumented flux return, used together with a 1.5 T solenoidal magnet, provided particle charge and momentum measurements.

Because  $B_L^0$  and  $B_H^0$  have very similar lifetimes, it is not possible to observe a pure sample of one species analogous to the  $K_L^0$  in neutral kaon decays, so that one is always dealing with particle mixtures. This problem was dealt with by using  $B$  meson ‘tagging’. In  $B$  factories, as we have seen, pairs of  $B$  mesons result from  $\Upsilon(4S)$  decays, and recoil against each other in their mutual centre-of-mass frame. Like kaons, the states produced will evolve in time, but which is which is unclear. Let us suppose, however, that one of the mesons is observed to decay by a  $\bar{B}^0$  mode; then the  $B$  meson that recoils against it must at that instant be thrown into a  $B^0$  state, to conserve the bottom quantum number. (This is an example of the quantum mechanical phenomenon known as ‘collapse of the wavefunction’.) Away from this instant, the  $B^0$  state will evolve into an oscillating superposition of  $B^0$  and  $\bar{B}^0$  states, as specified by (11.68). A similar result applies if the identified meson decays by a  $B^0$  mode rather than a  $\bar{B}^0$  mode. This technique is called tagging and in both cases one can classify the observed pairs of decays into *mixed events*, in which the pairs both decay by  $B^0$  (or  $\bar{B}^0$ ) mode, implying that mixing has occurred in the proper time  $t$  elapsed between the two decays, and *unmixed events*, in which a  $B^0$  decay is associated with a  $\bar{B}^0$  decay, which would always be the case if there were no mixing. If the numbers of each event are measured as a function of time, then the asymmetry

$$\alpha_B(t) = \frac{N(\text{mixed}) - N(\text{unmixed})}{N(\text{mixed}) + N(\text{unmixed})} \quad (11.69a)$$

is given by

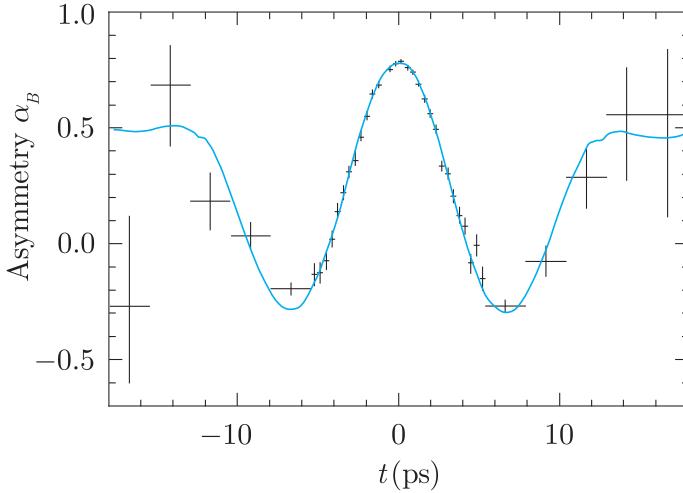
$$\alpha_B(t) = \frac{I(B^0 \rightarrow B^0) + I(\bar{B}^0 \rightarrow \bar{B}^0) - I(B^0 \rightarrow \bar{B}^0) - I(\bar{B}^0 \rightarrow B^0)}{I(B^0 \rightarrow B^0) + I(\bar{B}^0 \rightarrow \bar{B}^0) + I(B^0 \rightarrow \bar{B}^0) + I(\bar{B}^0 \rightarrow B^0)} \quad (11.69b)$$

and substituting from (11.68a) and (11.68b) yields the predicted behaviour

$$\alpha_B(t) = \cos(\Delta m t). \quad (11.70)$$

Figure 11.19 shows the asymmetry obtained in the BaBar experiment. The observed oscillations are fully compatible with (11.70), when a slowly varying background is taken into account, and when combined with data from other experiments, yield a best value of

$$m_H - m_L = (3.337 \pm 0.033) \times 10^{-10} \text{ MeV}/c^2.$$



**Figure 11.19** The measured asymmetry  $\alpha_B$  as a function of the observed time difference  $t$  between the two tagged neutral  $B$  decays as measured in the BaBar experiment. The error increases at large  $t$  as the mesons decay, and fewer events are observed. (B. Aubert *et al.* 2006. Reproduced with permission from the American Physical Society.)

### \*11.2.7 CP violation in interference

We now turn to rare decays of the type  $B^0 \rightarrow f$  and  $\bar{B}^0 \rightarrow f$ , where  $f$  is an eigenstate of  $CP$ , that is

$$\hat{C}\hat{P}|f\rangle = \eta_f|f\rangle, \quad \eta_f = \pm 1, \quad (11.71)$$

and calculate the probability that a particle tagged as, say, a  $B^0$  at  $t = 0$  decays to a final state  $f$  at time  $t$ . Because of oscillations, this can occur in two ways:  $B^0 \rightarrow B^0 \rightarrow f$  or  $B^0 \rightarrow \bar{B}^0 \rightarrow f$ , where the first step in each case denotes the evolution from  $t = 0$  to time  $t$  and the second step is the actual decay. It is the interference between these two terms that gives rise to *CP violation in interference*, as we shall see.

We will denote the decay amplitudes for  $B^0 \rightarrow f$  and  $\bar{B}^0 \rightarrow f$  by

$$\mathcal{M}_f = \langle f | H | B^0 \rangle \quad \text{and} \quad \bar{\mathcal{M}}_f = \langle f | H | \bar{B}^0 \rangle, \quad (11.72)$$

respectively. Then using (11.66), we see that the amplitude for an initial  $B^0$  state to decay to a final state  $f$  at time  $t$  is

$$\mathcal{M}[B^0(t) \rightarrow f] = \mathcal{M}_f A(t) + \xi \bar{\mathcal{M}}_f \bar{A}(t), \quad (11.73a)$$

while the corresponding result for an initial  $\bar{B}^0$  state is

$$\mathcal{M}[\bar{B}^0(t) \rightarrow f] = \bar{\mathcal{M}}_f A(t) + (1/\xi) \mathcal{M}_f \bar{A}(t). \quad (11.73b)$$

The corresponding decay probabilities are then given by

$$\begin{aligned} I[B^0(t) \rightarrow f] &= |\mathcal{M}[B^0(t) \rightarrow f]|^2, \\ I[\bar{B}^0(t) \rightarrow f] &= |\mathcal{M}[\bar{B}^0(t) \rightarrow f]|^2. \end{aligned} \quad (11.74)$$

However, before evaluating these expressions, it is useful to rewrite (11.73) in the forms

$$\mathcal{M}[B^0(t) \rightarrow f] = \mathcal{M}_f [A(t) + \lambda \bar{A}(t)] \quad (11.75a)$$

and

$$\mathcal{M}[\bar{B}^0(t) \rightarrow f] = (\mathcal{M}_f/\xi)[\bar{A}(t) + \lambda A(t)], \quad (11.75b)$$

where the complex parameter

$$\lambda = \xi \bar{\mathcal{M}}_f / \mathcal{M}_f. \quad (11.76)$$

Substituting (11.75) into (11.74) and using (11.67) to evaluate the resulting contributions gives, after some algebra (see Problem 11.8),

$$\begin{aligned} I[B^0(t) \rightarrow f] &= |\mathcal{M}_f|^2 e^{-\Gamma t} \left[ \frac{1}{2}(1 + |\lambda|^2) + \frac{1}{2}(1 - |\lambda|^2) \cos(\Delta mt) \right. \\ &\quad \left. - \text{Im } \lambda \sin(\Delta mt) \right] \end{aligned} \quad (11.77a)$$

and

$$\begin{aligned} I[\bar{B}^0(t) \rightarrow f] &= |\mathcal{M}_f|^2 e^{-\Gamma t} \left[ \frac{1}{2}(1 + |\lambda|^2) - \frac{1}{2}(1 - |\lambda|^2) \cos(\Delta mt) \right. \\ &\quad \left. + \text{Im } \lambda \sin(\Delta mt) \right], \end{aligned} \quad (11.77b)$$

where the terms proportional to  $\text{Im } \lambda$  arise from the interference between the first and second terms in (11.73a) and (11.73b), respectively. Finally, in comparing to experiment it is usual to define the asymmetry

$$\alpha_{fCP} \equiv \frac{I[\bar{B}^0(t) \rightarrow f] - I[B^0(t) \rightarrow f]}{I[\bar{B}^0(t) \rightarrow f] + I[B^0(t) \rightarrow f]} \quad (11.78)$$

and substituting (11.77a) and (11.77b) into this expression gives

$$\alpha_{fCP} = S \sin(\Delta mt) - C \cos(\Delta mt), \quad (11.79a)$$

where

$$S = \frac{2\text{Im } \lambda}{1 + |\lambda|^2} \quad \text{and} \quad C = \frac{1 - |\lambda|^2}{1 + |\lambda|^2}. \quad (11.79b)$$

At this point, we note that we have made no assumption about the amplitudes (11.72) for direct  $B^0$  and  $\bar{B}^0$  decays. If we now assume there is no direct  $CP$  violation in these decays, so that  $\hat{X}H\hat{X}^{-1} = H$ , where  $\hat{X} = \hat{C}\hat{P}$  and  $H$  is that part of the Hamiltonian that contributes to these decays, then

$$\mathcal{M}_f = \langle f | H | B^0 \rangle = \langle f | \hat{X}^{-1} \hat{X} H \hat{X}^{-1} \hat{X} | B^0 \rangle = \eta_f \langle f | H | \bar{B}^0 \rangle = \eta_f \bar{\mathcal{M}}_f.$$

Remembering that  $\eta_f^2 = 1$ ,

$$\lambda = \eta_f \xi = \eta_f e^{-2i\beta}. \quad (11.80)$$

Hence the coefficient  $C = 0$  and (11.79) reduces to

$$\alpha_{fCP}(t) = -\eta_f \sin 2\beta \sin \Delta mt. \quad (11.81)$$

Finally, if  $CP$  were exactly conserved, then the physical states (11.63) would be eigenstates of  $CP$ , requiring  $\xi = \pm 1$  and  $\alpha_{fCP} = 0$ .

Such oscillating asymmetries have been observed in several decay modes, including  $f = J/\psi K_S^0$  and  $J/\psi K_L^0$ , for which the standard

model predicts that (11.80) and hence (11.81) hold to a very good approximation, as we shall see in Section 11.3. These modes have a combined branching ratio of about  $10^{-3}$  and the final states are readily identified if the kaons are detected by the  $2\pi$  and  $3\pi$  decay modes, respectively. Since  $CP$  violation is very small for neutral kaons,  $K_S^0$  and  $K_L^0$  are to a good approximation identical to the  $K_1^0$  and  $K_2^0$  states with  $CP = +1, -1$ , respectively. The  $J/\psi$  state has  $C = P = -1$ , so that  $CP(J/\psi) = +1$ , and it has spin 1, so that the final state must have orbital angular momentum  $L = 1$ , leading to a factor  $(-1)^L = -1$  under  $CP$  from the parity transformation. Hence

$$\eta_f(J/\psi K_S^0) = CP(K_S^0) \times CP(J/\psi) \times (-1)^L = -1$$

and similarly

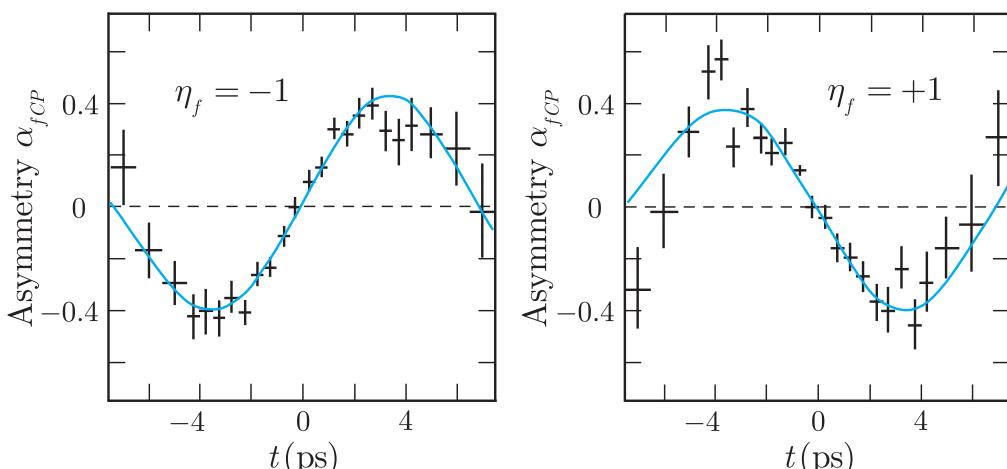
$$\eta_f(J/\psi K_L^0) = CP(K_L^0) \times CP(J/\psi) \times (-1)^L = +1.$$

The asymmetries measured in the Belle experiment are shown in Figure 11.20. As can be seen, they are compatible with (11.81) and, together with other experiments, give

$$\sin 2\beta = 0.682 \pm 0.019.$$

This does not, of course, mean that there is no contribution from direct  $CP$  violation for all final states  $f$  and, for example, for the very rare decay mode  $f = \pi^+ \pi^-$ , the more general result (11.79) is required, with

$$C = -0.31 \pm 0.05 \quad \text{and} \quad S = -0.67 \pm 0.06.$$



**Figure 11.20** The asymmetry parameter  $\alpha_{fCP}$  as measured by the Belle collaboration. The right panel is for the channel  $f = J/\psi K_L^0$  ( $CP$  eigenstate  $\eta_f = +1$ ) and the left panel is for  $f = J/\psi K_S^0$  ( $CP$  eigenstate  $\eta_f = -1$ ). (Adapted from Adachi *et al.* 2012. © American Physical Society. Reproduced with permission.)

### ★11.2.8 Derivation of the mixing formulas

In order to derive the mixing formulas (11.63a) and (11.63b), it will be useful to discuss an approach to particle decay introduced by Wigner and Weisskopf.<sup>14</sup> For simplicity, we start by considering an unstable particle  $A$  created at time  $t = 0$  and described in the Dirac notation by a quantum state

$$|\psi\rangle = a_0(t)|A\rangle + \sum_{n>0} a_n(t)|n\rangle, \quad (11.82)$$

where  $|n\rangle$  are the final states to which  $A$  can decay and

$$a_0(0) = 1, \quad a_n(0) = 0, \quad (n > 0). \quad (11.83a)$$

The time dependence of  $|\psi\rangle$  is then determined by the Schrödinger equation, in which the hermiticity of the Hamiltonian guarantees the conservation of probability, that is

$$|a_0(t)|^2 + \sum_{n>0} |a_n(t)|^2 = 1$$

A full treatment is complicated, but if we are only interested in  $a_0(t)$ , this is reproduced by the much simpler equation

$$\left(M - \frac{\Gamma}{2}\right) a_0(t) = i \frac{da_0(t)}{dt}, \quad (11.83b)$$

corresponding to exponential decay

$$a_0(t) = e^{-iMt} e^{-\Gamma t/2}, \quad (11.84)$$

where  $M$  and  $\Gamma$  are the mass and decay width and  $t$  is measured, as usual, in the rest frame of the decaying particle.

If we now consider a coupled system such as the  $B^0 - \bar{B}^0$  system, then a neutral  $B$  meson created at  $t = 0$  is described by a state

$$|\psi\rangle = a_0(t)|B^0\rangle + \bar{a}_0(t)|\bar{B}^0(t)\rangle + \sum_{n>0} a_n(t)|n\rangle, \quad (11.85a)$$

where now  $|n\rangle$  are the final states to which either  $B^0$  or  $\bar{B}^0$  can decay and again  $a_n(0) = 0$  ( $n > 0$ ), with

$$|a_0(0)|^2 + |\bar{a}_0(0)|^2 = 1, \quad (11.85b)$$

with the particular values of  $a_0(0)$  and  $\bar{a}_0(0)$  depending on whether a  $B^0$  or a  $\bar{B}^0$  is created. The equation that describes the time evolution of  $a_0(t)$  and  $\bar{a}_0(t)$ , in analogy to (11.83b), is now

$$\begin{pmatrix} M - i\Gamma/2 & M_{12} - i\Gamma_{12}/2 \\ M_{12}^* - i\Gamma_{12}^*/2 & M - i\Gamma/2 \end{pmatrix} \begin{pmatrix} a_0 \\ \bar{a}_0 \end{pmatrix} = i \frac{d}{dt} \begin{pmatrix} a_0 \\ \bar{a}_0 \end{pmatrix}, \quad (11.86)$$

---

<sup>14</sup>The application of the Wigner–Weisskopf method to neutral meson decays is discussed in Chapter 6 of Bigi and Sanda (2009).

where the matrix operator is called the *mass matrix*. In the absence of the off-diagonal terms, (11.86) reduces to two equations of the form (11.84). The diagonal elements must be equal to ensure that the expectation values of the mass and width of the  $B^0$  are equal to those of the  $\bar{B}^0$  to satisfy the *CPT* theorem. The off-diagonal terms correspond to the transition amplitudes needed to account for  $B^0 - \bar{B}^0$  mixing, where box diagrams like Figure 11.16 give rise to non-zero terms  $M_{12}$ , while  $\Gamma_{12}$  is generated by common decay states  $n = f$ , that is by the decays  $B^0 \rightarrow f$  and  $\bar{B}^0 \rightarrow f$ . In the approximation that these are neglected used in the previous sections, (11.86) reduces to

$$\begin{pmatrix} M - i\Gamma/2 & M_{12} \\ M_{12}^* & M - i\Gamma/2 \end{pmatrix} \begin{pmatrix} a_0 \\ \bar{a}_0 \end{pmatrix} = i \frac{d}{dt} \begin{pmatrix} a_0 \\ \bar{a}_0 \end{pmatrix}. \quad (11.87)$$

The coupled equations (11.87) are similar to those for coupled harmonic oscillators, for example small vibrations,<sup>15</sup> and the states  $B_L$  and  $B_H$  are the analogues of the normal modes. To find them, we must solve the eigenvalue equation

$$\begin{pmatrix} M - i\Gamma/2 & M_{12} \\ M_{12}^* & M - i\Gamma/2 \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = \lambda \begin{pmatrix} p \\ q \end{pmatrix}, \quad (11.88)$$

where the eigenvalues  $\lambda$  must satisfy the characteristic equation<sup>16</sup>

$$\det \begin{pmatrix} M - i\Gamma/2 - \lambda & M_{12} \\ M_{12}^* & M - i\Gamma/2 - \lambda \end{pmatrix} = 0,$$

which has solutions

$$\lambda_H = m_H + i\Gamma/2, \quad \lambda_L = m_L + i\Gamma/2,$$

where

$$m_H = M + |M_{12}|, \quad m_L = M - |M_{12}|. \quad (11.89)$$

The corresponding eigenvalues are found by substituting these eigenvalues into (11.88) and gives

$$\frac{q}{p} = \pm \frac{M_{12}^*}{M_{12}} = \pm e^{-2i\beta}, \quad (11.90)$$

where we have written  $M_{12} = |M_{12}|e^{i\beta}$  and where the plus and minus signs correspond to  $\lambda_H$  and  $\lambda_L$ , respectively. Ignoring possible arbitrary and irrelevant overall phase factors, the normalised eigenvectors can be chosen to be

$$\begin{pmatrix} p \\ q \end{pmatrix}_H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ e^{-2i\beta} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} p \\ q \end{pmatrix}_L = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -e^{-2i\beta} \end{pmatrix}, \quad (11.91)$$

<sup>15</sup>See, for example, Section 10.2.1 of Martin and Shaw (2015).

<sup>16</sup>See, for example, Section 10.1 of Martin and Shaw (2015).

and one can easily check that

$$\begin{pmatrix} a_0(t) \\ \bar{a}_0(t) \end{pmatrix}_{H,L} = \begin{pmatrix} p \\ q \end{pmatrix}_{H,L} \exp(-im_{H,L}t) \exp(-\Gamma t/2) \quad (11.92)$$

are the required independent solutions of (11.87). At  $t = 0$ , these are just the states  $|B_H^0\rangle$  and  $|B_L^0\rangle$  given by (11.63), as can be seen by substituting (11.89) into (11.85).

### 11.3 CP violation in the standard model

We conclude this chapter by reviewing whether the experimental observations on  $CP$  violation can be accommodated in the standard model or whether there is some other mechanism necessary. In the standard model,  $CP$  violation has its origins in the quark mixing described by (9.40), where the elements of the CKM matrix

$$V \equiv \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}$$

determine the relative strengths of the various  $\alpha\beta W$  couplings, where  $\alpha = u, c, t$  and  $\beta = d, s, b$ . The CKM matrix in general contains nine complex elements. However, the unitary nature of the matrix implies that there are nine relations between the elements, such as

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0.$$

Using these, and exploiting the freedom to define the phases of the basic quark states, the CKM matrix may be parameterised by just four quantities, three mixing angles and one phase angle, and is conveniently written in the standard form:

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{13}s_{23}e^{i\delta} & c_{12}c_{23} - s_{12}s_{13}s_{23}e^{i\delta} & c_{13}s_{23} \\ s_{12}s_{23} - c_{12}s_{13}c_{23}e^{i\delta} & -c_{12}s_{23} - s_{12}s_{13}c_{23}e^{i\delta} & c_{13}c_{23} \end{pmatrix}. \quad (11.93)$$

Here  $c_{ij} = \cos \theta_{ij}$ ,  $s_{ij} = \sin \theta_{ij}$  and the angles  $\theta_{ij} = \theta_{12}$ ,  $\theta_{13}$  and  $\theta_{23}$  can be chosen to lie in the first quadrant, so that  $s_{ij}$ ,  $c_{ij} \geq 0$ . Since the time-reversal operator acting on a state involves complex conjugation (cf. Equations (5.71) and (5.72)), a non-zero phase  $\delta$  implies  $T$  violation, and hence, if  $CPT$  invariance is assumed,  $CP$  violation. This is in contrast to the general form of the mixing matrix (9.39) for two generations only, which is purely real, so that  $CP$  violation could not arise from quark mixing if there were only two generations.

The mixing angles and the  $CP$ -violating phase must be determined from experiment and the magnitude of a particular  $CP$ -violating effect depends not only on the magnitude of  $\delta$  but also the mixing angles. Experimentally,  $s_{12} \gg s_{23} \gg s_{13}$ , and it is instructive to approximate the resulting matrix by the so-called *Wolfenstein parameterisation*:

$$V = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda + \frac{1}{2}A^2\lambda^5[1 - 2(\rho + i\eta)] & 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4(1 + 4A^2) & A\lambda^2 \\ A\lambda^3[1 - (1 - \frac{1}{2}\lambda^2)(\rho + i\eta)] & -A\lambda^2 + A\lambda^4[1 - 2(\rho + i\eta)] & 1 - \frac{1}{2}A^2\lambda^4 \end{pmatrix} + O(\lambda^6), \quad (11.94)$$

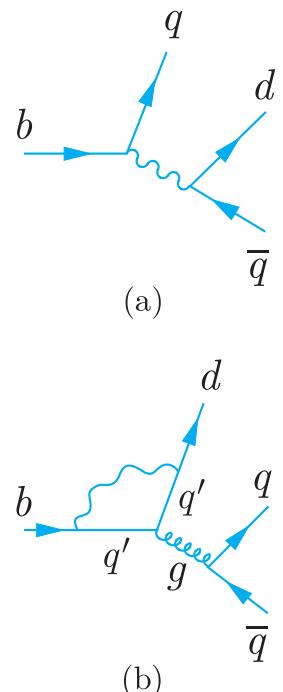
with parameters  $A$ ,  $\lambda$ ,  $\rho$  and  $\eta$ , where a non-zero value of  $\eta$  is indicative of  $CP$  violation. The quantity  $\lambda = |V_{us}| \approx 0.225$  plays the role of an expansion parameter and from (11.94) one sees that the only elements that can give rise to  $CP$  violation, up to and including terms of order  $\lambda^5 \approx 6 \times 10^{-4}$ , are

$$\begin{aligned} V_{ub} &= A\lambda^3(\rho - i\eta), & V_{cd} &= -\lambda + \frac{1}{2}A^2\lambda^5[1 - 2(\rho + i\eta)], \\ V_{td} &= A\lambda^3[1 - (1 - \frac{1}{2}\lambda^2)(\rho + i\eta)], & V_{ts} &= -A\lambda^2 + A\lambda^4[1 - 2(\rho + i\eta)]. \end{aligned} \quad (11.95)$$

The other parameters are obtained from a detailed analysis of quark mixing data<sup>17</sup> and the current values given by the Particle Data Group are

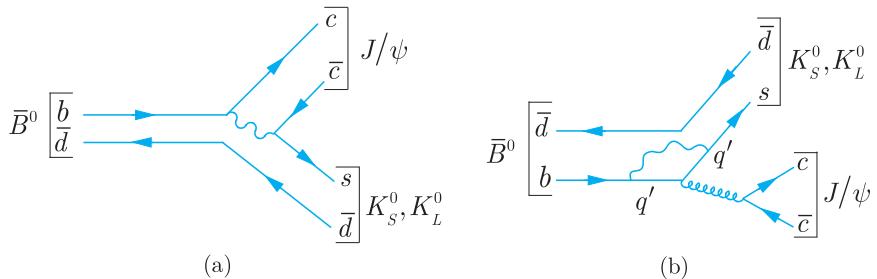
$$A \approx 0.81, \quad \rho \approx 0.12, \quad \eta \approx 0.37.$$

From this one sees immediately that direct  $CP$  violation is predicted to be largest in those  $B$  decays in which  $V_{ub}$  plays an important role. This is illustrated in Figure 11.21 for the decays for the decays  $b \rightarrow dq\bar{q}$ , where  $q = u, c$ . Diagrams of the type shown in (a) are examples of *tree diagrams*, so called because they contain no closed loops. Diagrams of the type shown in (b) are called *penguin diagrams* because, suitably orientated, they remind some people of penguins. The loop in the penguin diagrams produces a suppression factor, which depends on the particular process, but is typically in the range 0.2 – 0.3. In addition, to estimate the relative importance of tree and penguin diagrams one must also take into account the powers of the expansion parameter  $\lambda$  in the vertices involved. For example, in the case  $b \rightarrow du\bar{u}$ , the tree diagram has vertices  $ubW$  and  $udW$ , which are of order unity and  $\lambda^3$ , respectively, as can be seen from the entries in (11.94), giving a factor of  $\lambda^3$  overall, whereas the leading penguin diagram has  $q' = t$ , so that the weak vertices are  $tbW$  and  $tdW$ , which also give a factor of order  $\lambda^3$  overall. Hence the tree diagram is the most important in this case, but the penguin



**Figure 11.21** The (a) tree and (b) penguin diagrams for the decays  $b \rightarrow dq\bar{q}$ , where in (a)  $q = u, c$  and in (b)  $q = u, c, t$ .

<sup>17</sup>The details are given in Ceccucci, Ligeti and Sakai (2014).



**Figure 11.22** The leading tree and penguin diagrams for the decays  $\bar{B}^0 \rightarrow f = J/\psi K_S^0$  and  $J/\psi K_L^0$ .

diagram is only suppressed by a loop factor of order 0.2–0.3, and so also makes a significant contribution. Examples of decays involving the transition  $b \rightarrow du\bar{u}$  are  $\bar{B}_s^0 \rightarrow K^+ + \pi^-$  and  $\bar{B}^0 \rightarrow \pi^+ + \pi^-$ , for both of which large direct  $CP$ -violating effects have been observed, as noted in Sections 11.2.5 and 11.2.7, respectively.

On the other hand, the leading tree and penguin diagrams for the decays  $\bar{B}^0 \rightarrow f = J/\psi K_S^0$  and  $J/\psi K_L^0$  involve the process  $b \rightarrow sc\bar{c}$ , as shown in Figure 11.22, where the superposition of  $K_S^0$  and  $K_L^0$  produced is initially a  $\bar{K}^0$ . In this case, the tree diagram is of order  $\lambda^2$  and if it dominates there is no direct  $CP$  violation in these decays. The penguin contributions can give rise to  $CP$ -violating terms, that is terms proportional to  $(\rho - i\eta)$ , but one easily checks that they are of order  $\lambda^4$ . Together with the loop factor, this gives a suppression factor of order 0.01 for the penguin diagrams relative to the tree diagram. Hence direct  $CP$  violation is expected to be extremely small in these decay modes, as confirmed by the data on  $CP$  violation in interference for these decays discussed in Section 11.2.7.

For  $K$  and  $D$  decays, however, the main contributions to  $CP$  violation arise from  $V_{cd}$  and are of order  $\lambda^5$ , whereas  $V_{cd}$  itself is of order  $\lambda$ , so that direct  $CP$ -violating effects are predicted to be much smaller in these cases. This is consistent with the data: as we saw in Section 11.2.3, direct  $CP$  violation in neutral kaon decays is very small compared to  $CP$  violation by mixing, while  $CP$  violation in  $D$  decays has yet to be detected.

Precise quantitative predictions are more complicated, since one must not only take into account a variety of diagrams like those shown in Figure 11.21 for the decay  $b \rightarrow du\bar{u}$ , but, more problematically, the effects of the quarks being bound in mesons. We will not pursue this, but merely note that the data and predictions are consistent and the success of the mixing model in accounting for all  $CP$ -violating data in terms of a single  $CP$ -violating phase is a major triumph of the standard model. However, this is probably not the complete story on  $CP$  violation, because we will see in Chapter 12 that suggested explanations of the matter–antimatter asymmetry observed in the Universe require  $CP$  violation at a level several

orders-of-magnitude larger than that needed to account for meson decays.

Finally, in Section 2.3.3, we noted that a complete discussion of neutrino oscillations must incorporate mixing between all three neutrino states, rather than just two. The most general form of this mixing can again be written in the standard form of (11.93), that is

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{13}s_{23}e^{i\delta} & c_{12}c_{23} - s_{12}s_{13}s_{23}e^{i\delta} & c_{13}s_{23} \\ s_{12}s_{23} - c_{12}s_{13}c_{23}e^{i\delta} & -c_{12}s_{23} - s_{12}s_{13}c_{23}e^{i\delta} & c_{13}c_{23} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix},$$

where a value of the phase  $\delta$  other than 0 or  $\pi$  will give rise to  $CP$  violation and differences in the behaviour of neutrinos and antineutrinos. The experimental values of the mixing angles  $\theta_{12}$ ,  $\theta_{13}$  and  $\theta_{23}$  are now reasonably well-determined by the neutrino oscillation data discussed in Section 2.3, and are given by (2.32), (2.38) and (2.39), respectively, but the value of the phase  $\delta$  is still undetermined. It will hopefully be measured in future long-baseline oscillation experiments using beams of neutrinos and antineutrinos produced in accelerators, and a value other than 0 or  $\pi$  will imply  $CP$  violation in the neutrino sector. However, such effects, if they exist, are unlikely to be detected in leptonic decays, since their contributions will be proportional to the very small neutrino masses.

## Problems 11

- 11.1** Show that the existence of the decays  $K^+ \rightarrow \pi^+\pi^0$  and  $K^+ \rightarrow \pi^+\pi^+\pi^-$  implies that parity is violated if the kaon is assumed to have spin 0.
- 11.2** The intensity of the electrons emitted in the decay (11.1) of polarised cobalt-60 nuclei is found to be consistent with the form

$$I(v, \theta) = 1 + \alpha \frac{v}{c} \cos \theta,$$

where  $v$  is the magnitude of the electron velocity  $\mathbf{v}$  and  $\theta$  is the angle between its direction and the direction of the  ${}^{60}\text{Co}$  spin. Deduce the value of the coefficient  $\alpha$  by considering events in which the electron is emitted in the direction of the spin of the decaying nuclei. The spins of the  ${}^{60}\text{Co}$  and  ${}^{60}\text{Ni}$  nuclei are  $J = 5$  and  $J = 4$ , respectively, and the energies of the emitted particles are sufficiently small that orbital angular momenta may be neglected.

- 11.3** Neglecting the electron mass, the energy spectrum for the electrons emitted in muon decay is given by

$$\frac{d\omega}{dE_e} = \frac{2G_F^2 E_e^2}{(2\pi)^3} \left( 1 - \frac{4E_e}{3m_\mu} \right).$$

What is the most probable energy for the electron? Draw a diagram showing the orientation of the momenta of the three outgoing particles and their helicities for the case when  $E_e \approx m_\mu/2$ . Show also the spin of the muon. Integrate the energy spectrum to obtain an expression for the total decay width of the muon. Hence calculate the muon lifetime in seconds.

- 11.4** Show that the total decay rates for the reactions  $K^0 \rightarrow \pi^- e^+ \nu_e$  and  $\bar{K}^0 \rightarrow \pi^- e^- \bar{\nu}_e$  are equal if  $CP$  is conserved.  
**11.5** By decomposing the  $\pi^0 \pi^0$  state into components of definite isospin  $I$ , using the methods given in Appendix C, it can be shown that

$$\mathcal{M}(K_{S,L}^0 \rightarrow \pi^0 \pi^0) = \sqrt{\frac{2}{3}} e^{i\delta_2} \mathcal{M}_{S,L}^2 - \sqrt{\frac{1}{3}} e^{i\delta_0} \mathcal{M}_{S,L}^0, \quad (11.96)$$

where the phase factors are due to the strong interaction between the final-state pions and the superscripts (0, 2) refer to the isospin of the  $\pi\pi$  state. ( $I = 1$  is forbidden by Bose statistics.) Use this result, together with expansion (11.33), to show that the ratio  $\eta_{00}$  defined in (11.39) may be written in the form (11.42), with

$$\varepsilon' = -i\sqrt{2} \frac{\text{Im } A_2}{A_0} \exp[i(\delta_2 - \delta_0)],$$

where

$$\mathcal{M}[K^0 \rightarrow (\pi\pi)_{0,2}] \equiv A_{0,2}, \quad (11.97)$$

and by CPT invariance,

$$\mathcal{M}[\bar{K}^0 \rightarrow (\pi\pi)_{0,2}] \equiv A_{0,2}^*, \quad (11.98)$$

with  $A_2 \ll A_0$ , and where  $A_0$  may be taken as real. In deriving this result, neglect second-order terms in the small quantities  $\varepsilon$  and  $A_2$ .

- 11.6** At SuperKEKB it is planned to collide 4 GeV positrons with 7 GeV electrons to form  $\Upsilon(4S)$  particles. If the latter decay to  $B^\pm$  particles with equal energies, how far will they travel on average before decaying? (The  $B^\pm$  particles have masses of  $5.28 \text{ GeV}/c^2$  and lifetimes of  $1.64 \times 10^{-12} \text{ s}$ .)  
**11.7** Four mesons each of mass  $5280 \text{ MeV}/c^2$  are produced in a  $B$  factory and observed to decay to

$$(a) \bar{D}^0 \pi^- \mu^+ \nu_\mu, \quad (b) \rho^+ K^-, \quad (c) \rho^+ \pi^-, \quad (d) D^- D_s^+,$$

where  $\rho^+$  is the resonance discussed in Section 6.2.1. Is it possible to distinguish between a  $B^0$  and a  $\bar{B}^0$  in each case, and if so why?

- 11.8** Derive the formulas (11.77a) and (11.77b) for  $CP$  violation in mixing.  
**11.9** Compare the expansions (11.61a) and (11.61b) for mesons  $M_a$  and  $M_b$  with expansions of the form (11.33a) and (11.33b) for  $K_S^0$  and

$K_L^0$ , respectively, and show that in the case of no violation in mixing, (11.63c) implies  $\varepsilon = i \tan \beta$ .

- 11.10** Estimate the relative importance of the contributions of Figures 11.15(a) and (b) to the amplitude for the decay  $B^0 \rightarrow K^+ \pi^-$ .
- 11.11** Explain why you might expect significant direct  $CP$  violation in  $\bar{B}^0 \rightarrow \pi^+ \pi^-$  decays and discuss its possible importance in  $\bar{B}^0 \rightarrow D^+ D^-$  decays.

---

## Beyond the standard model

---

What is the path? There is no path. On into the unknown. *Goethe*<sup>1</sup>

The standard model, described in this book, evolved in response to a series of experimental discoveries, many of which were totally unexpected, over a period of several decades. At the present time, provided non-zero neutrino masses are incorporated,<sup>2</sup> all experimental observations in particle physics are consistent with the standard model, but there is no reason to suppose that there will not be more surprises in the future, as higher energy regions are explored. Furthermore, if one looks beyond the laboratory to the cosmos, there is a clear need for new physics ‘beyond the standard model’. As we shall see in Section 12.4, there is strong evidence that the particles of the standard model can only account for a small fraction of the matter in the Universe, and the observed predominance of matter over antimatter cannot be understood in the framework of the standard model. Also, although evidence has recently been reported for the existence of gravitational waves, which are a necessary consequence of any quantum theory of gravity (with an associated particle, called the *graviton*), we have no consistent theory of quantum gravity. Finally, the standard model itself embodies many assumptions and more than twenty free parameters, giving rise to many

---

<sup>1</sup> This was a favourite quotation of Niels Bohr, from *Faust*. See, for example, p. 126 of Moore (1985).

<sup>2</sup> In the original formulation of the standard model, the neutrino mass was assumed to be zero, so that neutrino oscillations were ‘beyond the standard model’. However, non-zero masses are easily incorporated, as we have assumed in earlier chapters. In fact this can be done in more than one way, as we shall see in Section 12.5 below.

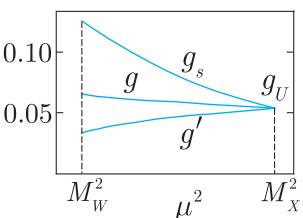
questions requiring an answer. Some examples are: can the number of parameters be reduced; why are there three families of quarks and leptons, rather than just the one that is required to describe ‘ordinary’ matter; are the quarks really point-like particles, or will they, like the proton and neutron, turn out to be composite when we can explore sufficiently short distance scales; why does the weak interaction violate  $CP$  invariance, but not the strong interaction? Many theories have been advanced to try to answer these and other questions. They typically predict the existence of hitherto unknown phenomena and, in some cases, large and ongoing experimental programmes have been mounted to test them.

In the rest of chapter we will briefly discuss a few of the ideas currently being investigated. The material is necessarily highly conjectural and our aim is to give the flavour of current research in ‘physics beyond the standard model’ without detailed argument. In doing so, we shall focus on three themes: the search for further unifications of the fundamental forces, the close relation between research in particle physics and problems in cosmology and whether neutrinos are their own antiparticles.<sup>3</sup>

## 12.1 Grand unification

With the success of the unified theory of electroweak interactions, it became natural to ask whether the strong interaction could also be united with the weak and electromagnetic interactions into a single so-called *grand unified theory (GUT)*. The much greater strength of the strong interaction compared with the electroweak interaction at presently accessible energies would seem to make this a hopeless task. However, the strength of an interaction depends on the distance over which it acts, or more precisely on the magnitude of the associated energy–momentum scale  $\mu^2$ , as discussed in Section 7.1.1. In particular, the strong interaction coupling decreases with  $\mu^2$ , as shown in Figure 7.3, while the electroweak couplings vary much more slowly, and a naïve extrapolation from their low-energy values suggests that the various couplings might become equal at an enormous value  $\mu^2 = M_X^2 c^4$ , where the so-called *unification mass*  $M_X$  is of order  $10^{15}$  GeV/ $c^2$ . This is illustrated in Figure 12.1, where  $g_s$  is related to the QCD coupling  $\alpha_s$  of Section 7.1 by

$$\alpha_s = g_s^2 / 4\pi \quad (12.1a)$$



**Figure 12.1** Idealised behaviour of the strong and electroweak couplings (12.1) as functions of the squared energy–momentum transfer  $\mu^2$  in a typical grand unified theory.

<sup>3</sup> For general introductions to particle astrophysics see, for example, Perkins (2009), Spurio (2015) and De Angelis (2015).

and the electroweak couplings are related to those defined in Sections 2.2 and 10.1.2, respectively, by<sup>4</sup>

$$g = 2\sqrt{2}g_W; \quad g' = 2\sqrt{2}g_Z. \quad (12.1b)$$

In grand unified theories all three interactions are united into a single interaction, characterised by a single coupling constant  $g_U$ , at the unification mass; differences between them emerge as one interpolates downwards to currently available energies. Of course, this interpolation assumes that nothing totally unexpected will emerge between energies of order  $10^2$  GeV and  $10^{15}$  GeV that could spoil the predictions. This is a very optimistic assumption given the history of particle physics. Nevertheless, this assumption is central to grand unified theories, which have aroused considerable interest for reasons that we will describe shortly.

There are many ways in which grand unified theories may be constructed so that they contain the predictions of both QCD and the unified electroweak theory at currently attainable energies. The earliest, and simplest, is due to Georgi and Glashow in 1974, and incorporates the known quarks and leptons into common families. For example, in the standard model the three colour states of the down quark, which are conveniently denoted

$$(d_r, d_g, d_b), \quad (12.2a)$$

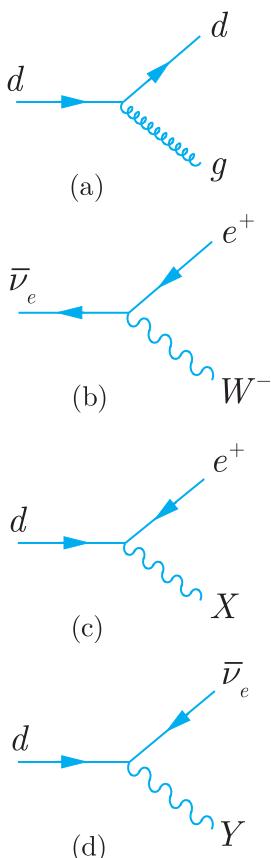
can be converted into each other by gluon emissions, as illustrated in Figure 12.2(a), while the positron and antineutrino

$$(e^+, \bar{\nu}_e) \quad (12.2b)$$

can be converted into each other by  $W^\pm$  emissions, as illustrated in Figure 12.2(b). In the Georgi–Glashow model both sets of particles are assigned to a single family<sup>5</sup>

$$(d_r, d_g, d_b, e^+, \bar{\nu}_e) \quad (12.3)$$

and, in addition to the familiar process of Figures 12.2(a) and (b), the quarks and leptons can convert into each other by processes like those shown in Figures 12.2(c) and (d). These involve the emission of two new gauge bosons  $X$  and  $Y$ , with electric charges  $-4e/3$  and  $-e/3$ , respectively, and masses of order  $10^{15}$  GeV/c<sup>2</sup>. At the unification mass, all the processes of Figure 12.2 are characterised by the single



**Figure 12.2** Two familiar processes (a, b) that can occur within the family of particles (12.3), together with two new processes (c, d), whose existence is predicted by grand unified theories.

<sup>4</sup> The precise couplings, which are assumed to become equal at the unification mass, depend on the particular grand unified theory in question. In the simplest case, discussed below, they are  $g_s$ ,  $g$  and  $\sqrt{(5/3)g_t}$ .

<sup>5</sup> Strictly speaking, it is the ‘right-handed’ states defined in Section 11.1.2 that are associated in a single family, but we will ignore such spin complications in this simple account.

‘grand unified coupling constant’  $g_U$ , whose value is such that the analogue of the fine structure constant is

$$\alpha_U \equiv \frac{g_U^2}{4\pi} \approx \frac{1}{42}. \quad (12.4)$$

However, at energies  $E \ll M_X, M_Y$ , processes involving the exchange of the  $X$  and  $Y$  bosons are heavily suppressed because of their large masses, in the same way that  $W^\pm$  exchange processes are suppressed relative to electromagnetic ones at energies  $E \ll M_W$  in the unified electroweak theory. Because of this, processes involving the exchange of  $X$  and  $Y$  particles are difficult, but perhaps not impossible, to observe at presently attainable energies, as we shall see later. Firstly, however, we briefly comment on two other striking predictions of this simplest of all grand unification schemes.

### 12.1.1 Quark and lepton charges

An attractive feature of the Georgi–Glashow model is that it offers an explanation of one of the longest-standing mysteries in physics – the equal magnitudes of the electric charges of the electron and proton. In the standard model this equality is guaranteed by assigning electric charges  $2e/3$  and  $-e/3$  to the  $u$  and  $d$  quarks, respectively, where  $-e$  is the charge of the electron, but no explanation is given for these empirical assignments. However, in the Georgi–Glashow model it can be shown that the sum of the electric charges of all the particles in any given family must be zero. Applying this to (12.3) gives

$$3Q_d + e = 0,$$

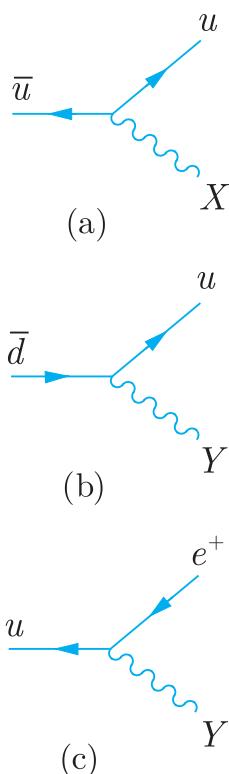
where  $Q_d$  is the charge of the down quark. Hence  $Q_d$  is determined to be  $-e/3$  in terms of the electron charge, and the factor of 3 is seen to be a consequence of the fact that the quarks have three distinct colour states. The charge of the up quark  $Q_u$  is shown to be  $2e/3$  by a similar argument, and the equality of the proton charge

$$Q_p = 2Q_u + Q_d = e$$

and the positron charge follows from the usual quark assignment  $p = uud$ .

### 12.1.2 The weak mixing angle

When the strong and electroweak interactions are extrapolated to the unification mass  $M_X$ , they are characterised by a single coupling constant  $g_U$ , as illustrated in Figure 12.1. Conversely, the three effective low-energy couplings (12.1) of the standard model can be expressed in terms of the two parameters  $g_U$  and  $M_X$  by reversing



**Figure 12.3** The three fundamental vertices involving  $X$  and  $Y$  bosons that are predicted by the Georgi–Glashow model, in addition to those shown in Figures 12.2(c) and (d).

this extrapolation. Consequently, in grand unified theories one of the three low-energy coupling constants (12.1) can be predicted, given the values of the other two. It is conventional to convert this result into a prediction of the weak mixing angle  $\theta_W$ , which is related to the coupling constants (12.1b) by (10.8), and in the Georgi–Glashow model this can be shown to give the value

$$\sin^2 \theta_W = 0.21. \quad (12.5)$$

This value is encouragingly close to the measured value (10.14), but is not in precise agreement with it.<sup>6</sup> We shall return to this point in Section 12.2.

### 12.1.3 Proton decay

The most striking prediction of grand unified theories is that the proton is unstable, and in the Georgi–Glashow model it can decay by a variety of processes involving the exchange of  $X$  and  $Y$  bosons and their antiparticles  $\bar{X}$  and  $\bar{Y}$ . These are built from a number of basic vertices, two of which are shown in Figures 12.2(c) and (d). If we restrict ourselves to the family (12.3) and their antiparticles, the model predicts three more such vertices, which are shown in Figure 12.3, and there are also another five that can be obtained from Figures 12.2(c) and (d) and Figure 12.3 by replacing all particles by their antiparticles. These give rise to proton decays such as

$$p \rightarrow \pi^0 + e^+ \quad \text{and} \quad p \rightarrow \pi^+ + \bar{\nu}_e, \quad (12.6)$$

by mechanisms like those shown in Figure 12.4. In all such processes, both the baryon and lepton numbers are not conserved, but the combination

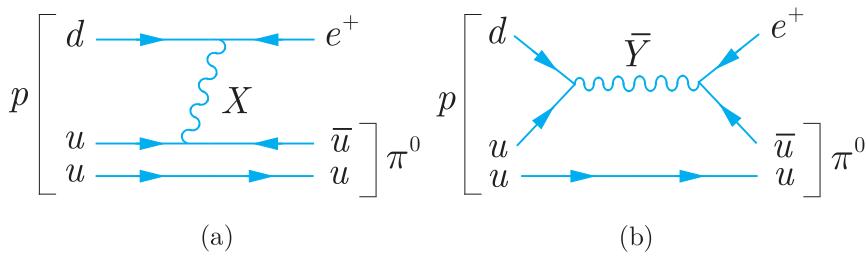
$$B - L = B - \sum_{\alpha} L_{\alpha} \quad (\alpha = e, \mu, \tau) \quad (12.7)$$

is conserved, as may be verified directly from the vertices of Figures 12.2 and 12.3.<sup>7</sup>

Although the proton is predicted to decay via processes like (12.6), its predicted lifetime is extremely long, and before quoting detailed values it is useful to make a qualitative estimate to understand why this is so. The first step is to estimate the magnitude of the effective coupling constant  $G$  for processes involving the exchange of  $X$  and  $Y$  bosons at energies  $E \ll M_X$ . At these energies,  $X$  and  $Y$  exchanges may be approximated by point interactions as shown in Figure 12.5, where on the left-hand side we have approximated the

<sup>6</sup> Another way of putting this is to say that if the measured value (10.14) is imposed, the three curves in Figure 12.1 fail to meet at the same point.

<sup>7</sup> This is done explicitly in Problem 12.2.



**Figure 12.4** Examples of processes that contribute to the proton decay reaction  $p \rightarrow \pi^0 + e^+$ .

couplings of the  $X$  and  $Y$  bosons to the lepton and quarks by their value  $g_U$  at the unification mass  $M_X$  and ignored their dependence on the range of the interaction. The effective low-energy coupling is then given by

$$\frac{G}{\sqrt{2}} \approx \frac{g_U^2}{M_X^2} = \frac{4\pi\alpha_U}{M_X^2}, \quad (12.8)$$

by analogy with (2.17) for the Fermi coupling  $G_F$ , and like  $G_F$  it has dimension  $[E^{-2}]$  in natural units. Since the rate  $\Gamma$  of any  $X$  or  $Y$  exchange process has a natural dimension  $[E]$  and is proportional to  $G^2$ , a simple dimensional estimate gives

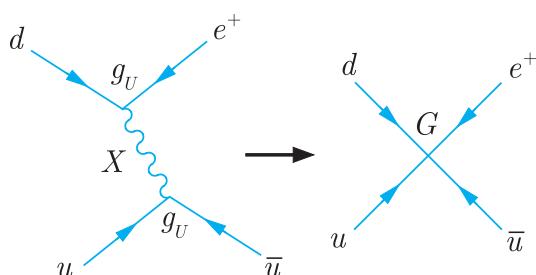
$$\Gamma \approx G^2 E^5 \approx \frac{g_U^2 E^5}{M_X^4},$$

where  $E \ll M_X$  is some energy characteristic of the process. A reasonable estimate for proton decay is  $E = M_p$ , and so we finally arrive at

$$\tau = \Gamma^{-1} \approx \frac{M_X^4}{g_U^2 M_p^5} \quad (12.9)$$

for the lifetime of the proton in natural units. Taking  $M_X$  to be  $10^{15}$  GeV/c<sup>2</sup> and  $g_U$  from (12.4) gives an estimate of  $10^{29}$  years, in agreement with the range of values

$$\tau = 10^{29} - 10^{30} \text{ years} \quad (12.10a)$$



**Figure 12.5** The zero-range approximation to an  $X$  boson exchange process.

obtained from more detailed and reliable calculations. This value is, however, very sensitive to the value chosen for the unification mass  $M_X$ , as can be seen from (12.9); in other grand unified theories in which  $M_X$  is somewhat larger, the lifetime can be as long as

$$\tau = 10^{32} - 10^{33} \text{ years.} \quad (12.10\text{b})$$

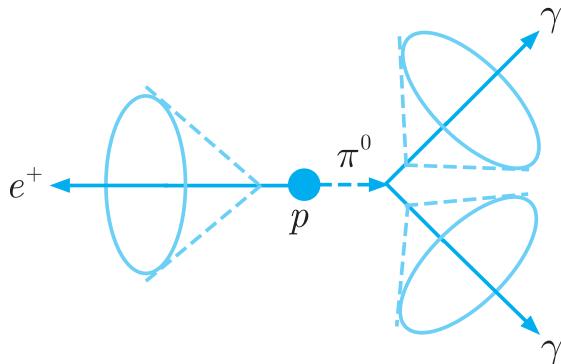
In addition, neutrons bound in stable nuclei that cannot decay by beta decay because of energy conservation would be expected to decay by baryon number violating decays, with similar lifetimes. For comparison, the age of the universe is believed to be of order  $10^{10}$  years.

If proton decays occur at all, they are clearly very rare events. For example, 300 tons of iron would only yield about one proton decay per year if the lifetime were of order  $10^{32}$  years. The problems involved in searching for such rare events, and their resolution, have been discussed in general in Section 4.6 and have been exploited to search for many possible proton decay modes. The results of these searches are expressed as lower limits on the *partial lifetimes*, which are the inverse of the decay rates for any particular decay mode  $p \rightarrow f$  and are given by

$$\tau(p \rightarrow f) = \tau_p / B(p \rightarrow f), \quad (12.11)$$

where  $\tau_p$  is the proton mean lifetime and  $B(p \rightarrow f)$  the corresponding branching ratio. The most stringent limit has been obtained for the decay  $p \rightarrow \pi^0 + e^+$  (Nishino *et al.*, 2009) using the SuperKamiokande detector discussed in Section 2.3.2(a) and shown in Figure 2.9. This detector is, as we saw, essentially a huge Čerenkov detector filled with extremely pure water, situated deep underground to shield it from cosmic rays, other than neutrinos. A search was carried out for decays originating in the central part of the detector containing 22.5 kilotonnes of water – the *fiducial volume* – over a period of years, giving a total exposure of 141 kiloton-years. If a  $p \rightarrow \pi^0 + e^+$  decay occurs, the positron would give rise to a cone of Čerenkov radiation whose radius depends on the velocity of the particle, while the particle energy can be inferred from the number of photons detected. The neutral pion would decay almost immediately into two photons, since the neutral pion lifetime is about  $10^{-16}$ s, and the photons would in turn interact to produce electrons and two further cones of Čerenkov radiation, as indicated schematically in Figure 12.6.

Because the positron and photons emitted in the decay are of relatively high energy, background events arising from radioactivity in the surroundings or the detector itself are negligible. Background events arising from atmospheric neutrinos are, however, not negligible and were expected to give rise to several thousand ‘three-ring’ events in the course of the experiment. However, these can be almost totally eliminated by imposing suitable cuts on the acceptance. If an



**Figure 12.6** The decay  $p \rightarrow \pi^0 + e^+$  as seen in the SuperKamiokande detector. The light-coloured lines show the three cones of Čerenkov radiation, as discussed in the text.

event corresponds to the decay of a proton at rest, then the energy of the positron is uniquely determined, the invariant mass of the two photons should correspond to the pion mass and momenta of the three particles produced must sum to zero. In imposing these constraints, account must be taken not only of the experimental resolution but also of the fact that most of the protons in water are bound in oxygen nuclei, giving rise to Fermi motion, that is they are not at rest. In addition, the neutral pion produced can interact with another nucleon in the oxygen nucleus and can convert to a charged pion, reducing the efficiency to detect the final state  $\pi^0 + e^+$ , which has a predicted branching ratio of about 30%. When these factors are taken into account, the expected background is reduced to 0.3 events and the fact that no events are observed puts a limit of

$$\tau(p \rightarrow \pi^0 + e^+) = \frac{\tau_p}{B(p \rightarrow \pi^0 + e^+)} > 8.2 \times 10^{33} \text{ years} \quad (12.12a)$$

on the partial lifetime at 90% confidence level. Limits on many other partial lifetimes have also been obtained in this and other experiments, including

$$\tau(n \rightarrow \pi^0 + \nu) > 1.1 \times 10^{33} \text{ y} \quad \text{and} \quad \tau(p \rightarrow K^0 + e^+) > 1.0 \times 10^{33} \text{ y}, \quad (12.12b)$$

which together suggest lower limits in the range  $(10^{32} - 10^{33}) \text{ y}$  on the proton lifetime.

In the Georgi–Glashow model, the branching ratio  $B(p \rightarrow \pi^0 e^+)$  is predicted to be about 0.3, so that (12.12a) is clearly incompatible with the predicted lifetime (12.10a). However, other models, which retain the successes of the Georgi–Glashow model, predict longer lifetimes (12.10b), and are compatible with the proton decay data, as we shall see in the following section.

## 12.2 Supersymmetry

The problems with the predictions for the proton lifetime (12.10a) and the weak mixing angle (12.5) in grand unified theories can be resolved by incorporating a new proposed symmetry between bosons and fermions called *supersymmetry*. According to this, every known elementary particle has a supersymmetric partner (called a *superpartner*), which is like it in all respects, including having the same quantum numbers, except for its spin. Spin-1/2 fermions, leptons and quarks have spin-0 superpartners, while spin-1 bosons, like photons, have spin-1/2 superpartners. The supersymmetric partners of fermions are named by adding a prefix ‘s’ to the name of the fermion, while the supersymmetric partners of the bosons are named by adding the ending ‘ino’ to the root of the normal name. This is illustrated in Table 12.1, where we list the various fermions and gauge bosons of the standard model, together with their superpartners. The latter are often collectively called *superparticles*, or *sparticles* for short. The situation for the Higgs boson is more complicated, since if one simply associates a superpartner with the standard model Higgs boson, the problems that were eliminated by the unification and anomaly conditions, discussed in Section 10.1.2, reappear and the theory becomes inconsistent. Supersymmetry requires additional spin-0 Higgs bosons, as well as their spin-1/2 superpartners, called *Higgsinos*. The simplest version of supersymmetry is called the *minimal supersymmetric standard model (MSSM)* because it contains the minimum number of new particles that are required for a consistent theory. It has three neutral and two charged spin-0 Higgs bosons, together with two neutral and two charged Higgsinos. The photino, zino and the neutral Higgsinos and are all spin-1/2 particles that interact by electroweak forces only. They are expected to mix together in the same way as neutrinos, to form four new particles,

**Table 12.1** The fermions and gauge bosons of the standard model and their superpartners, before mixing. Superparticles are distinguished from particles by a tilde over their symbol.

Particle	Symbol	Spin	Superparticle	Symbol	Spin
Quark	$q$	1/2	Squark	$\tilde{q}$	0
Electron	$e$	1/2	Selectron	$\tilde{e}$	0
Muon	$\mu$	1/2	Smuon	$\tilde{\mu}$	0
Tauon	$\tau$	1/2	Stauon	$\tilde{\tau}$	0
Neutrino	$\nu$	1/2	Sneutrino	$\tilde{\nu}$	0
$W$ boson	$W$	1	Wino	$\tilde{W}$	1/2
$Z$ boson	$Z$	1	Zino	$\tilde{Z}$	1/2
Photon	$\gamma$	1	Photino	$\tilde{\gamma}$	1/2
Gluon	$g$	1	Gluino	$\tilde{g}$	1/2

called *neutralinos*. In the same way, the charged Higgsinos can mix with the winos to form four particles, called *charginos*. Finally, a useful way of characterising particle states as particles or superparticles is provided by a new quantum number, called *R* parity, defined by

$$R \equiv (-1)^{3(B-L)+2S}, \quad (12.13)$$

where *B* is the baryon number, *L* is the lepton number and *S* is the spin. From this, one easily verifies that all the particles of the standard model have  $R = +1$ , while the sparticles have  $R = -1$ .

If supersymmetry were exact, a particle and its supersymmetric partner would have exactly the same mass. This is obviously not realised in nature or superparticles would have been detected long ago. Supersymmetry is thus at best only an approximate symmetry. Because of this, the masses of the predicted superparticles are unknown and even the simplest models incorporating supersymmetry breaking contain large numbers of free parameters whose values are unknown. However, in grand unified theories, a rough limit on the degree of symmetry breaking is suggested by the so-called *hierarchy problem*. In quantum field theory, the square of the physical mass *M* of a Higgs boson – that is the mass that can be measured – can be expressed in the form

$$M^2 = M_0^2 - \Delta M^2, \quad (12.14)$$

where  $M_0$  is the ‘bare mass’ of a hypothetical non-interacting Higgs boson and  $\Delta M^2$  is an additional contribution generated by radiative corrections. In the standard model, the lowest-order contributions to the latter are of order  $g_W^2 \lambda^2 / 16\pi^2$ , where  $\lambda$  is a cut-off parameter above which the theory is assumed to break down. In grand unified theories,  $\lambda$  must be at least as large as the unification scale. Hence  $\Delta M^2$  must be extremely large and a correspondingly huge Higgs mass can only be avoided by rather artificial ‘fine-tuning’ of the unknown bare mass to almost entirely cancel it. While this does not violate any known physical principles, it is not very natural. However, an attractive feature of supersymmetric theories is that fine-tuning is avoided by automatic cancellations between the contributions of particles and their superpartners to the radiative correction  $\Delta M^2$ , provided that the differences between their masses are not too large.<sup>8</sup> This is assumed to be the case in many ‘weak-scale’ supersymmetric theories in which the supersymmetric particles are assumed to have masses of order  $1 \text{ TeV}/c^2$  or less, since above this mass fine-tuning begins to reappear. Furthermore, when supersymmetric particles assumed to be in this mass range are taken into account, it can

---

<sup>8</sup> Fine-tuning can be avoided in other ways, for example by postulating that the Higgs boson is a composite particle with a spatial dimension of order  $\lambda^{-1}$ , but supersymmetry is usually considered the simplest.

be shown that the extrapolation of Figure 12.1 is modified in such a way that the grand unification mass  $M_X$  is increased to a value of order  $10^{16}$  GeV/ $c^2$ , while the value of the grand unified coupling constant  $g_U$  remains relatively constant. As a consequence of this, the proton lifetime increases, as expected from (12.9), to a value of order  $10^{32} - 10^{33}$  years, which is just consistent with present experimental limits. At the same time, the prediction (12.5) of the weak mixing angle is slightly modified to yield a value in agreement with the experimental value (10.14).

### 12.2.1 The search for supersymmetry

The introduction of supersymmetry enables the hoped-for unification of strong and electroweak forces to be successfully realised and offers at least the possibility of solutions to two serious problems in cosmology, as we will see in Section 12.4. These are attractive features, but to verify supersymmetry, it will of course be necessary to detect the predicted superparticles. Unfortunately, as noted earlier, it is difficult to say very much about their masses because the nature of supersymmetry breaking is unknown and even the simplest supersymmetric models contain many unknown parameters. However, according to most versions of supersymmetry,<sup>9</sup> including the MSSM,  $R$  parity is conserved, so that sparticles can only be created or destroyed in pairs. Hence the decay of a superparticle must yield at least one superparticle in the final state, and the lightest such particle must be stable. Because of this, the lightest sparticle is a possible contributor to the so-called dark matter in the universe, to be discussed in the Section 12.4.1.

There are several candidates for the identity of the lightest superparticle, but most models assume it is the lightest *neutralino*, denoted  $\bar{\chi}_0$ , which is a mixture of a photino, neutral Higgsinos and a zino, as we have seen above. If we accept this, then a simple reaction that can be studied at an  $e^+e^-$  collider is

$$e^+ + e^- \rightarrow \tilde{e}^+ + \tilde{e}^-, \quad (12.15a)$$

followed by the decays

$$\tilde{e}^\pm \rightarrow e^\pm + \bar{\chi}_0, \quad (12.15b)$$

giving an overall reaction

$$e^+ + e^- \rightarrow e^+ + e^- + \bar{\chi}_0 + \bar{\chi}_0. \quad (12.16)$$

---

<sup>9</sup> For a fuller introduction to supersymmetry, see, for example, Dawson (1997) and references therein.

The theory makes three clear predictions for this reaction sequence:

- (i) The cross-section for producing selectron pairs via (12.15a) should be comparable to that for producing ordinary charged particle pairs of the same mass.
- (ii) The selectrons decay very rapidly by (12.15b) before they can reach the detector.
- (iii) The neutralinos should escape detection because they interact extremely weakly with ordinary matter.

Hence only the final-state electron and positron should be detected in (12.16) and the events would be characterised by the fact that on average these will account for only half the energy of the initial state. Moreover, the electron and positron will not in general be emitted in opposite directions in the centre-of-mass frame, as they would be in the case of a two-body final state. Thus there are clearly defined criteria for the identification of selectron pair-production events, provided of course that there is enough energy to produce them. A similar discussion applies to the production of other pairs of charged sparticles in electron–positron collisions. The highest energy electron–positron collider to date was the large electron–positron collider LEPII at CERN, which had a maximum energy of 208 GeV when it closed in 2002. No charged sparticles were detected, setting lower limits that are typically of order  $100 \text{ GeV}/c^2$  on their masses.

Both charged and neutral superparticles can be produced at hadron colliders. At the LHC, the ATLAS and CMS experiments have studied a wide range of reactions, but to date have found no evidence for the existence of sparticles in the mass ranges studied. This failure to detect any such events sets lower limits on the masses of the different sparticles, which vary depending on the model. Here we shall just consider the simple MSSM model with  $R$  parity conservation, in which the lightest supersymmetric particle is a neutralino. For the masses of the squarks, which play a crucial role in solving the hierarchy problem, the associated lower limits are of order  $1 \text{ TeV}/c^2$  for the first two generations and of order  $100 \text{ GeV}/c^2$  for the third generation, while the lower limit on the mass of the gluino is about  $800 \text{ GeV}/c^2$ . In contrast, the lower limit on the mass of the lightest neutralino, which has only weak interactions, is about  $50 \text{ GeV}/c^2$ .

As mentioned above, supersymmetry also predicts the existence of several Higgs bosons, rather than the single one predicted by the standard model. In the MSSM<sup>10</sup> these are: two  $CP = +1$  neutral bosons, denoted  $h$  and  $H$ ; one  $CP = -1$  neutral boson denoted  $A$ ;

---

<sup>10</sup> For a fuller account of the Higgs sector in the MSSM, see Haber (2011) or Gunion *et al.* (1990).

and two charged bosons, denoted  $H^\pm$ . Ignoring radiative corrections, the mass of the lightest Higgs boson is given by

$$M_h = M_Z |\cos 2\beta| \leq M_Z, \quad (12.17a)$$

where, as the parameter  $\beta \rightarrow 0$ ,  $M_h \rightarrow M_Z$  and the lightest Higgs boson becomes identical to the Higgs boson of the standard model, while the other Higgs bosons become increasingly heavy and essentially decouple from the theory at presently attainable energies. When radiative corrections  $\varepsilon_h$  are included, (2.17a) is modified to

$$M_h = M_Z |\cos 2\beta| + \varepsilon_h < (120 - 130) \text{ GeV}/c^2, \quad (12.17b)$$

where the uncertainty in the limit reflects the uncertainty in the calculation of the radiative corrections. The mass of the Higgs boson is close to the limit (12.17b) and the agreement, within errors, between its observed properties and those predicted in the standard model, is a severe constraint upon supersymmetric theories. As more data are taken and the errors shrink, either the agreement will be maintained and the scale of supersymmetry breaking pushed to still higher values or deviations from the predictions of the standard model will be firmly established for the first time. Of course a breakdown of the standard model does not in itself prove that it is due to supersymmetry, but it would be a prime candidate.

### 12.3 Strings and things

Supersymmetry is an important component in even more ambitious schemes to unify gravity with the other forces of nature at an enormous *superunification* scale where gravitational interactions are comparable in strength with those of the grand unified strong and electroweak interactions. The problems here are mathematically formidable, not the least of them being that the divergences encountered in trying to quantise gravity are far more severe than those in either QCD or the electroweak theory, and there is at present no successful ‘stand alone’ quantum theory of gravity analogous to the former two. In order to resolve this problem, the theories that have been proposed invariably replace the idea of point-like elementary particles with tiny quantised one-dimensional *strings*, and for reasons of mathematical consistency are formulated in many more dimensions (usually 10, including one time dimension) than we observe in nature. Such theories have a single free parameter – the string tension. However, we live in a four-dimensional world and one possibility to explain this is that the extra dimensions are ‘compactified’, that is reduced to an unobservably small size. It was originally hoped that in doing this the standard model, with its many free parameters

(masses of quarks and leptons, coupling constants, mixing angles, etc.), would emerge from string theory as a unique low-energy four-dimensional theory, and thus the precise values of the parameters of the model would be explainable in terms of just a single parameter, the string tension.

Early optimism has not been sustained. In the particle picture, the structure of the corresponding quantum field theory (such as QED or QCD) is known and physical predictions may be obtained using the appropriate Feynman rules. However, in string theory the structure is not known, and there are five sets of possible Feynman rules, each operating in a 10-dimensional space–time continuum. Furthermore, string theorists have discovered that, far from being unique, there is a vast ‘landscape’ of at least  $10^{500}$  possible low-energy theories that could result after compactification, each corresponding to a universe with a different set of fundamental particles, interactions and parameters. Unless there is a method of choosing between the vast possibilities offered by this ‘landscape’, string theories have been criticised for having little or no real predictive power. For this reason they have generated a lively philosophical debate as to whether they should even be considered as scientific theories, although their proponents claim that string theories are being judged by standards that historically have not been applied to other emergent theories.

One controversial approach to the question of choice has been to invoke the so-called ‘anthropic principle’. There are various forms of this, but essentially it states that what we can expect to observe must be restricted by the conditions necessary for our presence as observers. In other words, the world is observed to be the way it is because that is the only way that humans could ever be here to consider such questions in the first place. This circular-sounding ‘principle’ has been invoked by cosmologists to explain the apparently improbable values of some cosmological constants, but it is by no means generally accepted as a way forward for string theories and other theorists believe that some form of dynamical selection will eventually be possible. All one can say at present is that there is no consensus on how the problem of choice is to be solved.

The self-consistency of string theories in 10 dimensions has been shown to imply the existence of higher-dimensional objects called *branes* (short for membranes), and it has been conjectured that using branes it will be possible to construct an even more fundamental theory in 11 dimensions in which all five supersymmetric string theories are unified. This theory has a name – *M theory* – although no-one knows if the conjecture is true, or how to construct such a theory. Nevertheless, string theory has provided some powerful theoretical tools that have contributed to a better understanding of gauge theories and their relation to gravity.

Leaving aside the mathematical difficulties of string theories, a major practical problem is that they apply at an energy scale where gravitational effects are comparable to those of the gauge interactions, that is at energies defined by the so-called *Planck mass*  $M_P$ , given by

$$M_P = \left( \frac{\hbar c}{G} \right)^{1/2} = 1.2 \times 10^{19} \text{ GeV}/c^2, \quad (12.18)$$

where  $G$  is the gravitational constant.<sup>11</sup> This energy is so large that it is difficult to think of a way that the theories could be tested at currently accessible energies, although some theorists believe that information produced even at the ‘low’ energies of the LHC (for example, the discovery of superparticles) may help to test string theories. Despite this, the appeal of string theories at present is mainly the mathematical beauty and ‘naturalness’ that their sponsors claim for them, although the history of physics is littered with failed theories whose authors doubtless considered them both beautiful and natural. Needless to say, experimentalists will remain sceptical until definite experimental tests can be suggested and carried out.

## 12.4 Particle physics and cosmology

Modern theories of particle physics and cosmology are closely related. This is because the conditions in the early universe that are implied by the standard cosmological theory can only be approached, even remotely, in high-energy particle collisions. At the same time, laboratory experiments cannot reach anything like the energies at which we might expect to see the full effects of unification, so that the conditions that existed shortly after the formation of the universe constitute a unique laboratory for the study of unification models. Here we will concentrate on a few aspects of this connection.

### 12.4.1 Dark matter

The suggestion that there might exist substantial quantities of non-luminous matter in the universe has its origins in the work of the astronomer Fritz Zwicky. In the 1930s he noted that there was a discrepancy between the mass estimates of clusters of galaxies based on luminosity calculations and those resulting from dynamical estimates based on the assumption that they were gravitationally bound systems. The latter resulted in much larger masses than those based on luminosity calculations, which Zwicky concluded implied that there

---

<sup>11</sup>This implies that strings have dimensions of order  $x_P \sim \hbar/M_p c = 1.6 \times 10^{-35} \text{ m}$ .

must be some form of non-luminous *dark matter* present in the clusters that dominated their overall mass content. This observation was reinforced by observations on the rotational motion of stars within galaxies, which were only consistent with the laws of gravity if the visible stars were surrounded by a so-called ‘halo’ of dark matter that contained the bulk of the galaxies’ total mass.

These early conclusions have been confirmed on cosmological scales by investigations using satellites and have resulted in rather precise estimates of the various contributions to the total energy of the universe. The modern description of the universe is based on the observation that it is expanding and assumes that the origin of this is a sudden explosion, the so-called *big bang*, at some time in the past. Because the universe appears isotropic at large distance scales, there can be no preferred points in space and so the big bang must have occurred everywhere at once, thus ensuring that the expansion appears the same to all observers irrespective of their locations. Key pieces of evidence for this model include: the existence of a *cosmic microwave background (CMB) radiation*, described to a very good approximation by a black-body spectrum at an effective temperature of 2.7 K; and the understanding it provided for the observed cosmic abundance of light elements.<sup>12</sup> Whether the expansion will continue indefinitely depends on the average density of the universe  $\rho$ . The critical density  $\rho_c$  at present times, below which the expansion will continue indefinitely, and above which it will eventually halt and the universe start to contract, can be written

$$\rho_c = \frac{3H_0^2}{8\pi G} \approx 0.85 \times 10^{-26} \text{ kg/m}^3 \approx 4.8 (\text{GeV}/c^2)/\text{m}^3, \quad (12.19)$$

where  $G$  is the gravitational constant and the current value for Hubble’s constant  $H_0$  has been used to evaluate (12.19). In the most popular and very successful version of the big bang idea, called the *inflationary big bang model* or the *standard cosmological model*, the density is equal to the critical density, that is the relative density  $\Omega$  is

$$\Omega \equiv \rho/\rho_c = 1. \quad (12.20)$$

The relative density is conveniently written as the sum of three components,

$$\Omega \equiv \Omega_{\text{total}} = \Omega_r + \Omega_m + \Omega_\Lambda, \quad (12.21)$$

where  $\Omega_r$  is the contribution due to radiation,  $\Omega_m$  is that due to matter and the third term  $\Omega_\Lambda$  is called the *cosmological term* or the *dark energy*. When Einstein first applied his theory of general relativity to cosmology, he introduced a new term into the equation of motion,

---

<sup>12</sup>For an accessible introduction to the big bang model, and cosmology in general, see, for example, Roos (2015).

with a magnitude characterised by a new *cosmological constant*  $\Lambda$ . This produces a repulsion, which opposed the gravitational attraction of matter, and it was later realised that the existence of such a term would follow from a uniform, non-zero energy density in the vacuum state. Such a term is necessary to account for the increasing expansion rate of the universe revealed by studies of distant supernovae and is the largest of the three terms contributing to the relative density (12.21). The radiation density can be directly measured and has a value  $\Omega_r \approx 5 \times 10^{-5}$ , that is, it is a negligible contribution to the total relative density. The first estimates of the relative matter density were obtained from the gravitational energy needed for consistency with the observations of the rotations of galaxies and the kinematics of large-scale structures in the universe mentioned earlier, and gave estimates in the range 0.24–0.30. However, the best values of both the matter and dark energy densities are obtained from precise data on the tiny temperature variations of the black-body radiation obtained by the WMAP space probe and the Planck satellite, together with data on the so-called ‘baryon acoustic oscillations’.<sup>13</sup> The current best values of relative mass density  $\Omega_m$  lie in the range 0.29–0.32, depending on the precise data set used, while the corresponding range for the relative dark energy density  $\Omega_\Lambda$  is 0.71–0.68. These analyses usually assume (12.20), but this should be a very good approximation, since when this constraint is relaxed, the value

$$\Omega_{\text{total}} = 1.0005 \pm 0.0033 \quad (12.22)$$

is obtained.<sup>14</sup>

There are two unsatisfactory features of the decomposition (12.21). One is that the origin of the largest term,  $\Omega_\Lambda$ , is totally unknown. The second concerns the composition of the matter density  $\Omega_m$ . The contribution of luminous baryonic matter<sup>15</sup> is obtained from the observed matter in the form of stars and intergalactic gas and dust and is about 0.01, so that the matter density is dominated by dark matter, as noted at the beginning of this section. There could also be sources of non-luminous baryonic matter, for example

---

<sup>13</sup>In the early universe, collisions between the hot particles would have produced acoustic waves that bounced between them. Residual patterns of these waves, left after the particles were too far apart for the waves to bounce between them, can be observed in small variations in the galaxy distribution of the present day universe and are called baryon acoustic oscillations (see, for example, Davis, 2014).

<sup>14</sup>These results are taken from Lahav and Liddle (2013), where a more detailed account of the assumptions made and data used may be found.

<sup>15</sup>Cosmologists use the term ‘baryonic matter’ to mean ‘ordinary matter including electrons as well as nucleons’.

in the form of brown dwarfs and small black holes the size of planets, and there is experimental evidence that such ‘massive, compact halo objects’ (*MACHOs*) do indeed exist in the halo of our galaxy. However, it is thought that they alone cannot account for much of the ‘missing’ matter and the total baryonic contribution to  $\Omega_m$  may be inferred from knowledge of how nuclei are formed in the universe (*nucleosynthesis*). Its value is  $\Omega_b \approx 0.05$  and since  $\Omega_m$  lies between 0.29 and 0.32, we are forced to conclude that the bulk of matter in the universe is non-baryonic. Its nature is unknown, although many candidates have been suggested. In the rest of this section, we shall discuss two of the possible candidates, while a third is considered in Section 12.4.4 below.

#### 12.4.1(a) Neutrinos

Neutrinos are one possible dark matter candidate. Such particles would have to be heavy enough to have been non-relativistic in the early stages of the universe (so-called *cold dark matter*), because if they were relativistic (*hot dark matter*) they would have rapidly dispersed, giving rise to a uniform energy distribution in space. Calculations suggest that in this case there would have been insufficient time for the observed galaxies to have formed. The contribution of massive neutrinos to the matter term can be calculated once the number of species and their masses are known. For masses in the range  $5 \times 10^{-4}$ eV to 1 MeV, the contribution of neutrinos to  $\Omega_m$  is

$$\Omega_\nu = \sum m_\nu / 42 \text{eV}/c^2, \quad (12.23)$$

which immediately gives a limit  $\Omega_\nu < 0.14$  for the contribution of the three known neutrinos to the matter density, since from (2.41) their masses are each less than about  $2 \text{eV}/c^2$ . However, this bound can be greatly improved. This is because neutrinos with masses even as small as  $0.1 \text{eV}/c^2$  can have an observable effect on the formation of large-scale structure in the universe by dampening the growth of perturbations. This enables an upper limit to be placed on  $\Omega_\nu/\Omega_m$ . Using the standard cosmological model together with major surveys of the large-scale structure mentioned above gives the value

$$\Omega_\nu < 0.0055 \quad (12.24)$$

at 95% confidence level, so that neutrinos can at most make a very small contribution to the relative density of non-baryonic dark matter required by the cosmological data. In addition, one sees from (12.23) that this corresponds to a limit

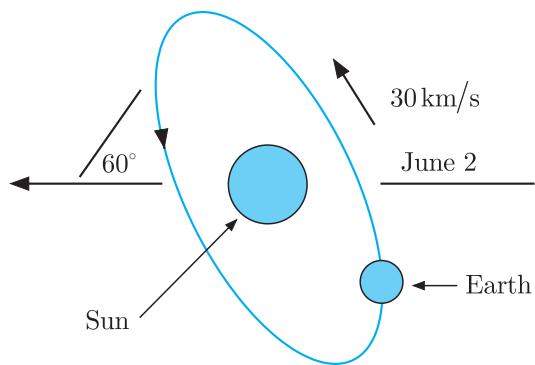
$$\sum m_\nu \leq 0.23 \text{ eV}/c^2 \quad (12.25)$$

for the sum of neutrino masses. This value is much lower than the upper bound of about  $2 \text{eV}/c^2$  obtained from tritium decay.

### 12.4.1(b) WIMPs

The most popular hypothesis is that the bulk of dark matter consists of non-baryonic cold dark matter in the form of hitherto unknown particles with masses in the range from about  $10 \text{ GeV}/c^2$  to several  $\text{TeV}/c^2$ , whose interaction with ordinary matter are roughly the same strength as the familiar weak interaction. Such particles are called '*weakly interacting massive particles*', or *WIMPs*. If they exist, they would have been produced in large numbers following the big bang, and calculations indicate that they might account for the bulk of the dark matter density observed today, provided they are stable or, at least, extremely long-lived. The problem is that, since WIMPs do not correspond to any known particles, new types of particle have to be postulated. One possibility is the lightest particle that appears in supersymmetric theories, which is stable if  $R$  parity is conserved. In most such theories this is the lightest neutralino, which interacts only via electroweak interactions. Its mass is not predicted, but we have seen in Section 12.2.1 that in the most popular supersymmetric theories it must be greater than about  $50 \text{ GeV}/c^2$  to accord with current experimental limits. However, other theories beside supersymmetry suggest different possibilities, and we shall not assume any particular identity for WIMPs in what follows, except when explicitly stated.

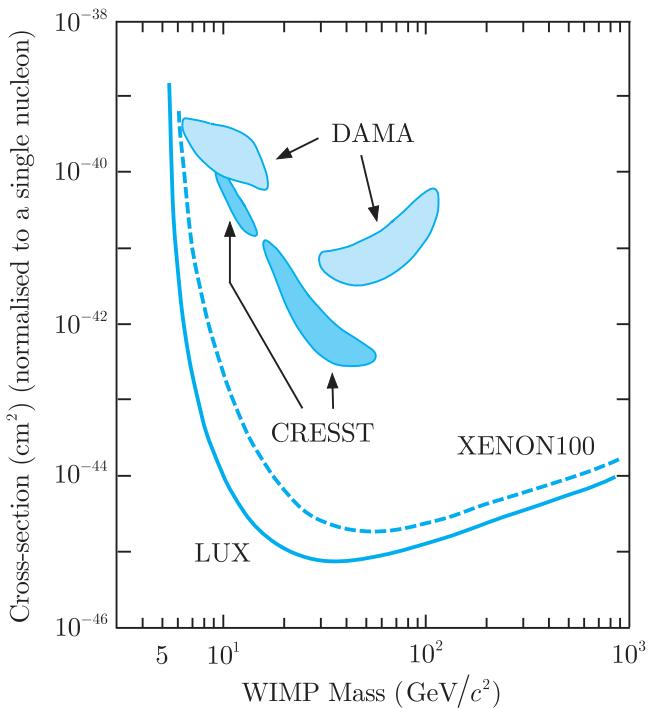
If WIMPs are the dominant form of dark matter, their number can be inferred from the estimated density of dark matter. From estimates of the dark matter density in our own galaxy, it follows that of order  $10^5$  WIMPs per second will pass through every  $\text{cm}^2$  of the Earth's surface if their mass is assumed to be of order  $60 \text{ GeV}/c^2$ . The velocities of WIMPs relative to the galaxy will be similar to those of baryonic matter, which are of order  $10^{-3}c$ , and when they interact with nuclei, the nuclei will recoil with energies in the range 5 keV to about 50 keV. These recoils can be detected in several ways, as we shall see below. The WIMP–nucleon cross-sections are unknown, but could be of order  $10^{-44} \text{ cm}^2$ , leading to just a few events per year per kilogram of target material. This is very small compared to the event rate from cosmic rays and naturally occurring radioactivity, including that in the materials of the detectors themselves. The former is minimised by working deep underground to shield the detector from cosmic rays and in areas with geological structures where radioactive rocks are absent; the latter is minimised by building detectors of extreme purity. Even with these precautions, the interpretation of possible events is difficult and can be controversial, as we will see. However, any WIMP signal obtained is expected to have a distinctive time dependence as a consequence of the Earth's motion around the Sun. The Sun moves in the galactic plane with a velocity of order  $2 \times 10^5 \text{ m/s}$ . The Earth moves around the Sun with a speed order of  $3 \times 10^4 \text{ m/s}$ , in an orbit that is inclined at about  $60^\circ$ .



**Figure 12.7** Motion of the Sun and Earth through the dark matter background. (Adapted from Spooner 1993.)

to the galactic plane, as seen in Figure 12.7. Hence the velocity of the Earth through dark matter varies by about  $\pm 15\%$  over the course of the year. This leads to a corresponding modulation in the flux and energy of WIMP collisions, resulting in a maximum event rate in May/June and a minimum in November/December. The modulation is unfortunately only a few per cent effect, and rather precise measurements will be required to detect it.

Several experiments have been built to detect WIMPs in the galactic halo by detecting the nuclear recoils when they collide with the material of a detector, as described above. This can be done in a variety of ways. The earliest experiments usually relied on the fact that in semiconductors such as GaAs, free charge will be produced that can be detected electronically. Alternatively, in a scintillator such as NaI or liquid xenon, the emission of photons can be detected using photomultipliers, while cryostatic detectors exploit the fact that in crystals at low temperatures the energy can be converted to phonons and can be detected by a very small rise in temperature. Examples are the DAMA experiment, which is a scintillation detector containing 250 kg of sodium iodide crystals, and CRESST, which is a cryostatic detector that uses a similar amount of calcium tungstate crystals at a temperature of a few thousandths of a degree kelvin. Both of these experiments are located in the Gran Sasso Laboratory deep under the Italian Alps; both reported a WIMP signal that exhibited the characteristic time dependence described above, albeit in different kinematic regions. However, these results, and those of other experiments that suggest a possible signal, are controversial and seem to be in conflict with other experiments. In general, if WIMPs are assumed to saturate the missing dark matter, the number density of WIMPs for any given mass is determined so that the number of events is determined by the WIMP mass and the WIMP–nucleon cross-section only. The failure to observe any events sets an upper limit on the WIMP–nucleon cross-section for any given WIMP mass. The most sensitive results obtained so far are from the XENON100 detector at Gran Sasso, which contains 161 kg of



**Figure 12.8** WIMP cross-section (normalised to a single nucleon) as a function of the WIMP mass.  
 (Adapted from Drees and Gerbier, 2013.)

liquid xenon, and the LUX detector at the SURF laboratory in the Homestake mine in South Dakota, which contains 170 kg of xenon in both gas and liquid form. The upper bounds on the cross-section obtained in these experiments are shown in Figure 12.8 and, as can be seen, the DAMA and CRESST results mentioned above lie well above them. Overall, there is no convincing evidence for WIMPs from direct observations<sup>16</sup> and the bounds already obtained impose serious constraints on any theory, such as supersymmetry, which purports to explain the origin of WIMPs. In addition, several further experiments are under development that will either detect WIMPs or considerably improve the existing bounds. These include the LUX-ZEPLIN experiment, which will use nine tonnes of liquid xenon as its detector, and DEEP at the SNOLAB facility in Canada, which will use 3 tonnes of liquid argon to measure the recoil energy of struck nuclei and will be sensitive to higher mass values than the majority of earlier experiments.

Not all experiments are direct searches. The Alpha magnetic spectrometer (AMS) is an experiment in orbit on the International Space Station. It uses a 7 tonne 0.15 T magnet to sort incoming particles by charge and momentum and searches for an excess of charged particles such as positrons in the energy range 500 MeV to 1 TeV, whose

<sup>16</sup>For fuller reviews of dark matter searches, see Figueroa-Feliciano (2011), Drees and Gerbier (2013) and Roos (2015).

origin could be the interaction between WIMPs. It has reported an excess of positrons over electrons that decreases from 1 to 10 GeV, in accordance with theoretical expectations, but then increases from 10 to 250 GeV. This confirms results from an earlier experiment, but in both cases other conventional interpretations have been suggested. Other experiments include searches for high-energy neutrinos emitted from the Sun. WIMPs would be expected to be slowed down and captured in the Sun over a long period of time, and the resulting higher WIMP density would then lead to a WIMP-enhanced pair annihilation rate to particles that will decay, among other things, to high-energy neutrinos. Hence, if WIMPs exist, one would expect a larger rate of high-energy neutrinos incident on the Earth from the Sun, compared to neighbouring directions, which could be observed in detectors like SuperKamiokande, described in Chapter 2 or an innovative experiment called IceCube. The latter was designed principally to search for the source of ultra-high-energy galactic neutrinos by detecting the Čerenkov radiation from the charged particles produced in neutrino collisions with nucleons. The experiment consists of strings of ‘optical modules’ containing photomultiplier tubes that convert the Čerenkov radiation to electrical signals. The modules are sited up to 2.5 km deep in the ice at the South Pole to utilise the vast quantities of frozen pure water there. No WIMP signal has yet been detected, but the interpretation of the results is complicated, since it requires an understanding of WIMP capture in the Sun and of WIMP annihilations. However, within theories such as supersymmetry, the results are a powerful constraint on WIMP parameters and are an important complement to direct searches, especially at high WIMP masses.

#### 12.4.2 Matter–antimatter asymmetry

One of the most striking facts about the universe is the paucity of antimatter compared to matter. There is ample evidence for this. For example, cosmic rays are overwhelmingly composed of matter and what little antimatter is present is compatible with its production in intergalactic collisions of matter with photons. Neither do we see intense outbursts of electromagnetic radiation that would accompany the annihilation of clouds of matter with similar clouds of antimatter. The absence of antimatter is very puzzling, because in the original big bang it would be natural to assume a total baryon number  $B = 0$ .<sup>17</sup> Assuming that baryon number is conserved, then during

---

<sup>17</sup> One could, of course, simply bypass the problem by arbitrarily assigning an initial non-zero baryon number to the universe, but it would have to be exceedingly large to accommodate the observed asymmetry, as well as being an unaesthetic solution.

the period when  $kT$  was large compared to hadron energies, baryons and antibaryons would be in equilibrium with photons via reversible reactions such as



and the number densities of protons and antiprotons would be comparable with the photon number density. This situation would continue until the temperature fell to a point where the photons no longer had sufficient energy to produce  $p\bar{p}$  pairs. Protons and antiprotons would then annihilate each other until the expansion had proceeded to a point where the density of protons and antiprotons was such that mutual annihilation became increasingly unlikely. The critical temperature is  $kT \approx 20$  MeV and at this point the ratios of baryons and antibaryons to photons ‘freezes’ to values that can be calculated to be

$$N_B/N_\gamma = N_{\bar{B}}/N_\gamma \sim 10^{-18}, \quad (12.27)$$

with of course  $N_{\bar{B}}/N_B = 1$ . These ratios would then be maintained in time, whereas the ratios currently observed are

$$N_B/N_\gamma \approx 10^{-9}, \quad N_{\bar{B}}/N_\gamma \sim 10^{-13}, \quad (12.28)$$

with  $N_{\bar{B}}/N_B \sim 10^{-4}$ , so that the prediction (12.27) fails spectacularly. However, in the equilibrium state the numbers of protons and antiprotons was so large that it would require only a very small excess of protons over antiprotons – about one in  $10^9$  – to have developed since the initial big bang for (12.27) and (12.28) to be reconciled.

The conditions whereby a baryon–antibaryon asymmetry could arise were first stated by Sakharov. It is necessary to have: (a) an interaction that violates charge conjugation  $C$  and the combined symmetry  $CP$ ; (b) an interaction that violates baryon number; and (c) a non-equilibrium situation must exist at some point to ‘seed’ the process. The reason for the first condition is that a baryon excess cannot be generated if the production of any particular particle (e.g. the proton) is balanced by an equal production of the corresponding antiparticle (the antiproton), as required by  $C$  or  $CP$  conservation. Baryon number violation is also obviously necessary and seems at first to be incompatible with the standard model in which, under present conditions, baryon number is conserved and the symmetry breaking between the electromagnetic and weak interactions is characterised by an energy scale of order 100 GeV. However, immediately after the big bang, thermal energies were very large compared to the energy scale of electroweak symmetry breaking and, if current theory is correct, the universe underwent a phase transition about  $10^{-7}$  s after the big bang from a state in which the electroweak symmetry was fully realised to the present state in which it is badly broken. Furthermore, it can be shown that during this transition, in

which the universe was in a non-equilibrium state, satisfying requirement (c), sufficient baryon number violations occurred to enable the observed matter–antimatter asymmetry to be understood, provided there is sufficient  $C$  and  $CP$  violation. Unfortunately, while both  $C$  and  $CP$  violation occur in the standard model, the  $CP$  violation is much too small to explain the observed asymmetry. In contrast, supersymmetric theories, with their host of new particles, contain several unknown mixing angles and  $CP$ -violating phases, and by adjusting these parameters it is possible to induce much larger  $CP$ -violating effects than those predicted in the standard model. However, until these, or  $CP$ -violating effects from some other source ‘beyond the standard model’ are experimentally detected, the origin of the matter–antimatter asymmetry in the universe remains an unsolved problem.

### 12.4.3 $CP$ violation and electric dipole moments

One way to detect hitherto unknown sources of  $CP$  violation is to search for non-zero electric dipole moments (EDMs) of leptons and hadrons. Consider the non-relativistic Hamiltonian  $H_{\text{EDM}}$  for the interaction of an EDM  $\mathbf{d}$  with an electric field  $\mathbf{E}$ . For an elementary particle, or a hadron, its spin  $\mathbf{J}$  is the only vector available to define a direction and so  $\mathbf{d}$  must be collinear with  $\mathbf{J}$ . Thus we can write

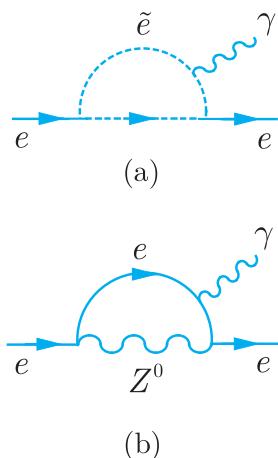
$$H_{\text{EDM}} = -\mathbf{d} \cdot \mathbf{E} = -\frac{d}{J} \mathbf{J} \cdot \mathbf{E},$$

where  $d = |\mathbf{d}|$  and  $J = |\mathbf{J}|$ . However,  $\mathbf{E}$  is even under time-reversal, while  $\mathbf{J}$  is odd, and thus  $H_{\text{EDM}}$  is odd under time-reversal. It is also odd under parity. Thus a non-zero EDM can only exist if both parity ( $P$ ) and time-reversal ( $T$ ) invariance are violated, and by implication any non-zero EDM detected would imply  $CP$  violation, assuming that  $CPT$  invariance holds.

The magnitudes of predicted electric dipole moments depend upon the assumed mechanism that causes  $CP$  violation. For example, in the standard model, a non-zero EDM could be generated by  $P$ ,  $T$  violating weak radiative corrections to the  $P$ ,  $C$  and  $T$  conserving electromagnetic interaction. For the neutron and the electron, which are the two most extensively studied, the estimates from the standard model are

$$d_n \sim 10^{-32} \text{ e cm} \quad \text{and} \quad d_e \leq 10^{-38} \text{ e cm}, \quad (12.29)$$

where  $e$  is the magnitude of the charge on the electron. For the electron EDM, for example, this is 16 orders of magnitude smaller than



**Figure 12.9** (a) An example of a diagram involving superparticles that can lead to a non-zero electron electric dipole moment in supersymmetric theories; (b) a corresponding diagram without superparticles, which is *CP* conserving in the standard model.

the electron magnetic moment and is far too small to be measured in the foreseeable future. However, in supersymmetric theories much larger *CP*-violating effects can be generated by diagrams involving virtual superparticles, like Figure 12.9(a) for the electron. In comparison, the corresponding particle diagram in Figure 12.9(b) does not violate *CP* invariance in the standard model, and a non-zero electron EDM can only arise in higher orders in perturbation theory and is consequently much smaller. Unfortunately, the magnitude of the EDM in supersymmetric theories is difficult to pin down, because it depends on the nature of supersymmetry breaking, which inevitably involves many unknown parameters. However, ‘weak scale’ supersymmetric theories, which also attempt to resolve cosmological problems like dark matter and matter–antimatter asymmetry, usually predict values in the ranges  $(10^{-25} - 10^{-28}) \text{ e cm}$  and  $(10^{-27} - 10^{-30}) \text{ e cm}$  for the neutron and electron EDMs, respectively. Other theoretical speculations also lead to much larger values than those predicted in the standard model, and the detection of electric dipole moments significantly larger than those given in (12.29) would be clear evidence for new *CP*-violating effects, whatever their origin.

For neutral spin-1/2 particles, non-zero EDMs can be detected by making use of the fact that in collinear  $\mathbf{E}$  and  $\mathbf{B}$  fields, the particle’s spin will undergo a Larmor precession<sup>18</sup> about the field direction. The frequency  $f$  of this precession is given by

$$hf = (2\mu_n B \pm 2d_n E), \quad (12.30)$$

where  $\mu_n$  is the magnetic moment,  $d_n$  is the electric dipole moment and the  $\pm$  sign refers to  $E$  being parallel (antiparallel) to  $B$ . Thus, a non-zero  $d_n$  will be revealed by an electric field-dependent shift in  $f$ . The experiments are technically very demanding, because of the relative sizes of  $\mu_n$  and  $d_n$  and innovative techniques are required to detect the frequency shift. The same technique cannot be used directly for charged particles, such as the electron, because the interaction with the electric field will accelerate the particle out of the observed region, and in practice the best measurements of the EDM of the electron have been found from studies of paramagnetic atoms and molecules with unpaired electrons. For the neutron and the electron, the present experimental limits

<sup>18</sup>A particle with a magnetic moment  $\mu$  placed in an external magnetic field  $\mathbf{B}$  will experience a torque  $\mathbf{T}$  given by  $\mathbf{T} = \mu \times \mathbf{B} = \gamma \mathbf{J} \times \mathbf{B}$ , where  $\mathbf{J}$  is the angular momentum vector and  $\gamma$  is the gyromagnetic ratio. As a result, the angular momentum vector precesses about the external field axis with frequency  $f = \mu B / J \hbar$ , known as the *Larmor frequency*.

(at 90% confidence level) are<sup>19</sup>

$$|d_n| < 2.9 \times 10^{-26} \text{ e cm} \quad \text{and} \quad |d_e| < 8.7 \times 10^{-29} \text{ e cm}. \quad (12.31)$$

These limits are much larger than those predicted by the standard model, but are already small enough to be a serious constraint on supersymmetry models. Further improved experiments will either detect non-zero EDMs or reduce these limits further.

#### 12.4.4 Axions and the strong $CP$ problem

There is another reason for measuring EDMs. In the standard model, the theory of strong interactions (QCD) is assumed to satisfy both gauge invariance and  $CP$  invariance. However, it is possible to add a term to the interaction that satisfies gauge invariance but violates  $CP$ . Whether such a term is present or not is often called ‘the strong  $CP$  problem’. If it is present, its magnitude is determined by a coupling constant  $\bar{\theta}\alpha_s$ , the size of which is not specified by the theory, and its contribution to the EDM of the neutron is  $d_n \approx 3 \times 10^{-16} \bar{\theta} \text{ e cm}$ . Hence the detection of a neutron EDM, which is too large to be explained by the standard model, could indicate either a non-zero value for  $\bar{\theta}$  or some other completely new physics, while a large EDM for the electron would definitely be proof of the latter. On the other hand, the experimental limit on the neutron EDM given in (12.31) implies that  $\bar{\theta} \leq 1 \times 10^{-10}$ , an extraordinarily small value for which there is no obvious explanation. However, it can be shown that by introducing an appropriate new symmetry into the theory, a new particle with  $J^P = 0^-$ , named the *axion*, is predicted and in turn  $\bar{\theta}$  becomes zero. There is an analogy here with the introduction of the Higgs boson, although the motivation is nowhere near as strong. The mass of the axion is unknown, but is given approximately by

$$m_a \approx \left( \frac{6 \times 10^6 \text{ GeV}}{f_a} \right) \text{ eV}/c^2, \quad (12.32)$$

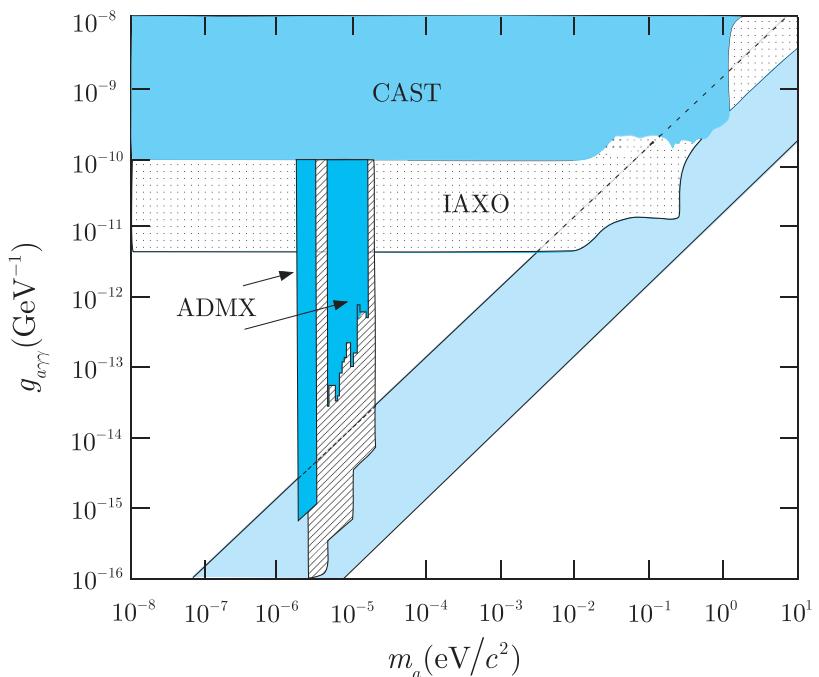
where  $f_a$  is an unknown energy scale at which the new symmetry is broken. Axions would mainly couple via the electromagnetic interaction to photons and would convert to two photons in the presence of a strong magnetic field. The strength of this coupling is given by

$$g_{a\gamma\gamma} = \alpha g_\gamma / \pi f_a, \quad (12.33)$$

where  $\alpha$  is the fine structure constant and  $g_\gamma$  is a model-dependent dimensionless constant of order unity. Equations (12.32) and (12.33)

---

<sup>19</sup>These limits have improved dramatically, and hopefully will continue to improve. For comparison, the best values available at the time of the 3rd edition of this book in 2008 were  $|d_n| < 1.2 \times 10^{-22} \text{ e cm}$  and  $|d_e| < 5.4 \times 10^{-27} \text{ e cm}$ .



**Figure 12.10** Plot of the various possible regions for dark matter axions in the  $g_{a\gamma\gamma} - m_a$  plane. The pale blue diagonal band shows the expected region of viable dark matter from conventional axion models (see text). The blue shaded areas labelled CAST and ADMX show the regions in the  $g_{a\gamma\gamma} - m_a$  plane excluded by these two experiments, while the dotted and cross-hatched areas show the regions that will be probed by the planned IAXO and ADMX experiments. (Adapted from Dafni and Iguax, 2015.)

imply a relation between the axion mass  $m_a$  and  $g_{a\gamma\gamma}$ , which is represented by the diagonal band in Figure 12.10. A light axion would be an excellent dark matter candidate as its density relative to the critical density  $\rho_c$  of the universe is given by

$$\Omega_a \approx \left( \frac{6(\mu\text{eV}/c^2)}{m_a} \right)^{7/6} \quad (12.34)$$

An axion of mass  $m_a \approx 20(\mu\text{eV}/c^2)$  would therefore account for the entire dark matter contribution  $\Omega_m \approx 0.3$ . Much smaller values of the mass would produce too large a value for  $\Omega_m$ , so a strong lower bound on the mass is  $m_a > 1(\mu\text{eV}/c^2)$ . On the other hand, much larger masses would make a negligible contribution to the mass density, but would still be relevant to the strong  $CP$  problem. In addition, various astrophysical data seem to exclude masses between about  $2 \times 10^{-2}\text{eV}/c^2$  and  $10^{-5}\text{eV}/c^2$  so that the favoured window for axion masses is now  $10^{-6}(\text{eV}/c^2) < m_a < 10^{-2}(\text{eV}/c^2)$ . However, these various bounds apply to ‘conventional’ axions, in which the mass and photon couplings are related by (12.32) and (12.33), and do not exhaust all possibilities. For example, all string-like theories have a myriad of axion and axion-like particles, although they are predicted to have masses of order  $10^{-9}\text{eV}/c^2$  and we will not consider them further here.

Axions light enough to contribute significantly to dark matter were initially thought to be so weakly coupled as to be undetectable

in conventional experiments. However, it was later shown that axions present in the Milky Way could be resonantly converted into an accurately monochromatic microwave signal in a microwave cavity permeated by a strong magnetic field. The conversion power depends on a number of quantities outside the experimenter's control, such as the axion mass, its coupling to photons and the local axion density in the galactic halo. It also depends on quantities that can be controlled, such as the strength of the magnetic field and the characteristics of the cavity. The experiments are conducted at very low temperatures of a few millikelvin to minimise spurious noise, and the searches are conducted by tuning the cavity in small overlapping steps. Several experiments have used variants of the cavity resonance technique to explore regions of  $g_{a\gamma\gamma} - m_a$  space, but to date no definite signals have been reported. The region of most interest in the  $g_{a\gamma\gamma} - m_a$  plane is that corresponding to conventional axions light enough to contribute to dark matter. One experiment – the axion dark matter experiment (ADMX) located at the University of Washington – has already begun to penetrate this region and will do so more extensively in the future, as can be seen in Figure 12.10.

Axions could also be produced in the solar interior by conversion of plasma photons to axions in the Coulomb field of charged particles, and these can be efficiently converted back to photons in the presence of a strong electromagnetic field. Dedicated 'helioscope' experiments have searched for axions using powerful magnets to effect the conversion. The best results to date come from the CERN solar axion telescope (CAST) that uses an LHC prototype magnet to produce a magnetic field of 9 T over a length of 9.3 m. CAST has been operating since 2003 and its latest bounds are shown in Figure 12.10. A significant improvement over these results is expected when a new facility, called the International Axion Observatory (IAXO), becomes operational. This will have a signal-to-noise ratio five orders-of-magnitude better than that in CAST, and will considerably extend the region of the  $g_{a\gamma\gamma} - m_a$  plane that can be explored, as shown in Figure 12.10.

Finally, we stress that the above are just a sample of the experiments performed or planned, not all of which rely on cosmological or astrophysical sources.<sup>20</sup> The search for axions is very much a work in progress.

## 12.5 Dirac or Majorana neutrinos?

In Section 1.2.2 we noted that neutral particles might, or might not, have distinct antiparticles. Thus there are two possible descriptions

---

<sup>20</sup> For a recent review of axion searches, see, for example, Graham *et al.* (2015).

of neutral leptons, that is neutrinos. Neutrinos that have distinct antiparticles are called *Dirac neutrinos*. There are then two neutrino states

$$\nu_L, \quad \nu_R \quad (L = 1)$$

and two antineutrino states

$$\bar{\nu}_L, \quad \bar{\nu}_R \quad (L = -1),$$

with opposite values of the lepton number  $L$ , which is conserved in the standard model. Neutrinos that do not have distinct antiparticles are called *Majorana neutrinos*, and there are only two states, denoted by

$$\nu_L, \quad \nu_R \quad (\text{Majorana neutrinos}).$$

In both cases, the subscripts  $L$  and  $R$  denote right- and left-handed helicity states as usual.

In the original formulation of the standard model, neutrino masses were assumed to be zero and it is not possible to experimentally distinguish between Dirac and Majorana neutrinos for two reasons. Firstly, for Dirac neutrinos with zero mass, the weak interaction only couples the left-handed neutrinos  $\nu_L$  and right-handed antineutrinos  $\bar{\nu}_R$ , as we saw in Chapter 11. Secondly, in the zero-mass limit, the Dirac equation decouples into two two-component equations describing the states  $(\nu_L, \bar{\nu}_R)$  and  $(\nu_R, \bar{\nu}_L)$ , respectively. The second pair can therefore be completely eliminated from the theory, which becomes indistinguishable from an analogous Majorana theory with the replacements  $(\nu_L, \bar{\nu}_R) \rightarrow (\nu_L, \nu_R)$ .

In contrast, as we saw in Section 11.1.3, the weak interaction couples to both helicity states of the electron, which is not massless. Similarly, for non-zero masses, the weak interaction would also couple to the other two Dirac neutrino states, albeit with relative couplings of order  $(m_\nu/E)^2$ , where  $E$  is the neutrino energy. Because of this, differences of the same order will emerge between the two descriptions of the neutrino, which can, at least in principle, be detected experimentally. In particular, in the Majorana formalism, lepton number conservation for reactions involving neutrinos is an ‘accident’ arising from the spin structure of the weak interaction for zero mass neutrinos, and tiny deviations from it are expected for non-zero neutrino masses. In most cases these will be too small to observe, and the most promising prospect for detecting them, if they exist, is by detecting neutrinoless double beta decay. This will be discussed in Section 12.5.1. First, however, we digress to mention a way in which very heavy Majorana neutrinos may play a role.

At present we have no theoretical understanding of why quarks and leptons have the particular masses they do. However, the three known neutrinos have much lower masses than the other fundamental fermions, and a possible explanation for this has been suggested in

the context of grand unified theories. In these, it is possible for both types of neutrino to coexist, and for a very small neutrino mass to emerge from the mixing of a zero-mass neutrino with a very heavy Majorana neutrino. The corresponding mass matrix for the right-handed neutrino is essentially of the form

$$\mathbf{M} = \begin{pmatrix} 0 & m_D \\ m_D & m_M \end{pmatrix}, \quad (12.35)$$

where the off-diagonal term  $m_D$ , called the Dirac mass, is of the order of the electroweak scale and  $m_M$ , the Majorana mass, is of the order of the GUT scale. Since  $m_M \gg m_D$ , the eigenvalues of  $\mathbf{M}$  are

$$|\lambda_+| \approx m_M \quad \text{and} \quad |\lambda_-| \approx m_D^2/m_M. \quad (12.36)$$

If  $\lambda_-$  is associated with the observed neutrinos, we have a natural explanation for a very small mass of order 1 eV or less, while the other eigenvalue implies a very heavy neutrino, yet to be discovered. This is called the ‘see-saw mechanism’, because from (12.36) we see that as one mass goes up, the other goes down.

### 12.5.1 Double beta decay

The double beta decay process

$$(Z, A) \rightarrow (Z + 2, A) + 2e^- + 2\bar{\nu}_e \quad (12.37)$$

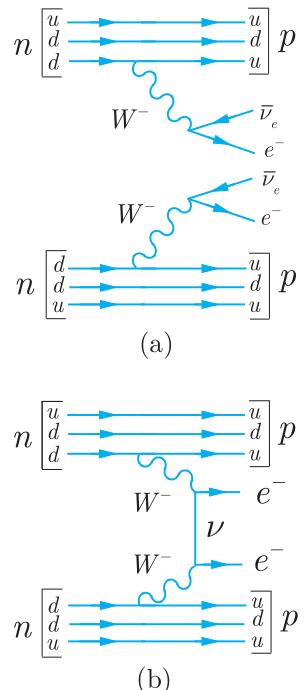
(denoted by the label  $\beta\beta 2\nu$ ), in which two neutrons decay simultaneously to give two electrons and two antineutrinos, is a second-order weak interaction allowed in the standard model, as illustrated in Figure 12.11(a). In contrast, the neutrinoless double beta decay process (labelled  $\beta\beta 0\nu$ )

$$(Z, A) \rightarrow (Z + 2, A) + 2e^- \quad (12.38)$$

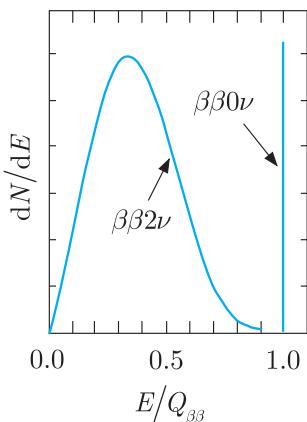
is not allowed because it violates the lepton number. However, it can occur for Majorana neutrinos of non-zero mass, which are their own antiparticles, by the mechanism illustrated in Figure 12.11(b). Observation of the reaction (12.38) would therefore be strong evidence for the existence of Majorana neutrinos, although in principle the reaction could also proceed by other mechanisms beyond the standard model, involving new hypothetical particles such as predicted by varieties of supersymmetry. On the other hand, the  $\beta\beta 2\nu$  process of (12.37) is allowed for both Dirac and Majorana neutrinos.

Because the double  $\beta$ -decay process (12.37) is second-order in the weak interaction, with a rate of order  $G_F^4$ , it can in practice only be observed if the single  $\beta$ -decay process

$$(Z, A) \rightarrow (Z + 1, A) + e^- + \bar{\nu}_e \quad (12.39)$$



**Figure 12.11** (a) Double  $\beta$ -decay  $\beta\beta 2\nu$  (allowed in the standard model); (b) neutrinoless double  $\beta$ -decay  $\beta\beta 0\nu$  (forbidden in the standard model).



**Figure 12.12** Energy spectra for the two electrons in  $\beta\beta 2\nu$  and  $\beta\beta 0\nu$  decays as a function of  $E/Q_{\beta\beta}$ , where  $E$  is their combined energy and  $Q_{\beta\beta}$  is the energy released.

is energetically forbidden, or highly suppressed. For an even- $Z$ , even- $N = A - Z$  nucleus, double  $\beta$ -decay leads to another even-even nucleus ( $Z + 2, A$ ), while single  $\beta$ -decay leads to an odd-odd nucleus ( $Z + 1, A$ ). Since odd-odd nuclei are less tightly bound than even-even nuclei, due to the pairing term in the semi-empirical nuclear mass formula,<sup>21</sup> a sequence of states can result in which single  $\beta$ -decay is forbidden by energy conservation, while double  $\beta$ -decay is allowed. If the parent nucleus ( $Z, A$ ) is also stable against  $\alpha$  and  $\gamma$ -decays, it will decay by double  $\beta$ -decay. Such decays were first directly observed in 1987 and have now been established for 11 isotopes, with mean lifetimes in the range  $10^{18} - 10^{24}$  y.

The observation of  $\beta\beta 0\nu$  decays would be a major breakthrough in particle physics because it would demonstrate the Majorana nature of neutrinos and the violation of total lepton number. It can in principle be distinguished from  $\beta\beta 2\nu$  decays by measuring the energies of the emitted electrons. In  $\beta\beta 2\nu$  decays, energy is carried away by the undetected neutrinos, resulting in a continuous spectrum for the combined energy of the electrons, whereas in  $\beta\beta 0\nu$  decays the electrons carry off all the available energy, resulting in a sharp line in their combined energy, as shown in Figure 12.12. However, a major problem is that the rate for  $\beta\beta 0\nu$  decay is expected to be much smaller than that for  $\beta\beta 2\nu$  decays, even for Majorana neutrinos, and for zero-mass neutrinos it would actually vanish.

For non-zero mass Majorana neutrinos, the decay rate  $\Gamma$  is given by

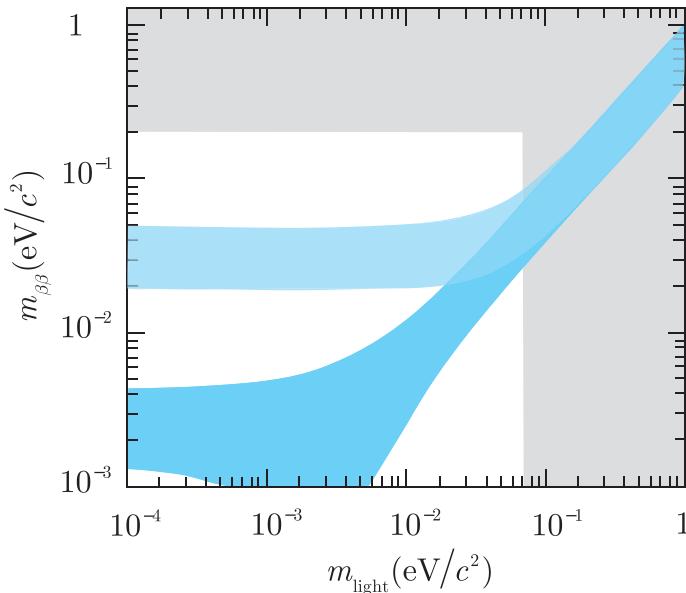
$$\Gamma = am_{\beta\beta}^2, \quad (12.40)$$

where the constant  $a$  is proportional to the coupling constant  $G_F^4$  and depends upon the particular decaying nucleus. The quantity  $m_{\beta\beta}$  is the ‘effective Majorana mass’, which reduces to

$$m_{\beta\beta} = |U_{e1}|^2 m_1 + |U_{e2}|^2 m_2 + |U_{e3}|^2 m_3. \quad (12.41)$$

Here  $U_{e1}$ ,  $U_{e2}$  and  $U_{e3}$  are the elements of the neutrino mixing matrix (11.96) that defines the electron-neutrino state,  $m_1$ ,  $m_2$  and  $m_3$  refer to the mass eigenstates  $\nu_1$ ,  $\nu_2$  and  $\nu_3$ , and, for simplicity, we have neglected  $CP$ -violating phases in the neutrino sector. From this equation it is clear that  $m_{\beta\beta}$  depends on the neutrino oscillation parameters and the unknown ordering of neutrino masses. Thus, if the three known neutrinos are indeed Majorana particles, and there are no other neutrinos, then values of  $m_{\beta\beta}$  can be found for different assumed values for the lightest neutrino mass, given the measured values of the neutrino squared mass splittings  $\Delta m_{ij}^2$  and mixing angles  $\theta_{ij}$  discussed in Chapter 2. Figure 12.13 shows  $m_{\beta\beta}$  as a function of the lightest neutrino mass  $m_{\text{light}}$  obtained in this way for

<sup>21</sup>See, for example, Section 2.6 of Martin (2008).



**Figure 12.13** The predicted values of the effective neutrino Majorana mass  $m_{\beta\beta}$ , as a function of the lightest neutrino mass,  $m_{\text{light}}$ . The dark and light blue bands correspond to the ‘normal’ and ‘inverted’ hierarchies of Figure 2.13, respectively. Two rectangular excluded regions are shown in grey. The vertical region comes from the cosmological bound  $m_{\text{light}} < 0.07$  eV (Ade *et al.*, 2014) and the horizontal one from experimental constraints on neutrinoless double beta decay described below. (Adapted from the review of Gómez-Cadenas and Martín-Albo, 2015.)

both the ‘normal’ and ‘inverted’ hierarchies shown in Figure 2.13. The widths of the bands are due to uncertainties in the neutrino oscillation parameters, and also the unknown  $CP$ -violating phases, which have been neglected in (12.41). Also shown is the upper bound on  $m_{\text{light}}$  obtained from cosmological studies, as discussed in Section 12.3, which is much more stringent than the bound of about 0.7 eV/ $c^2$  implied by direct measurements of tritium decay together with the measured values of the squared mass splittings from oscillation experiments. If this astrophysical bound is correct, then it is clear that the effective neutrino mass  $m_{\beta\beta}$  is predicted to be at most of order 50 meV/ $c^2$ .

Experimental results in nuclear physics are usually presented in terms of the half-life  $T_{1/2}$ , which is related to the mean lifetime  $\tau = 1/\Gamma$  used in particle physics by  $T_{1/2} = \tau \ln 2$ . Assuming that the exchange of light Majorana neutrinos is the dominant mechanism for neutrinoless double beta decay, the half-life may be written as

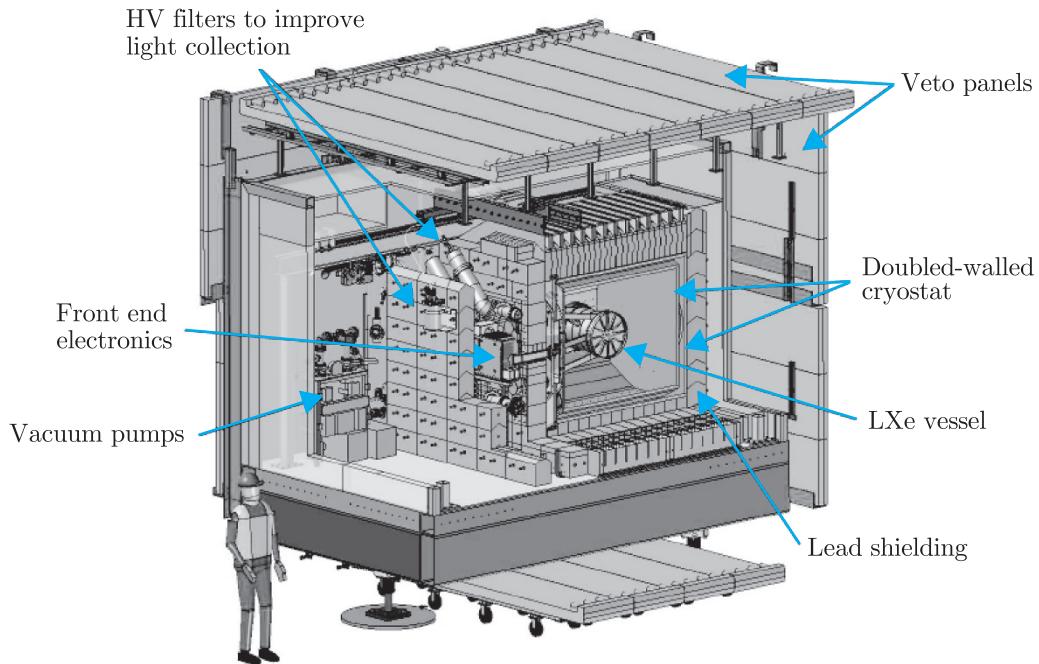
$$\left(T_{1/2}^{0\nu}\right)^{-1/2} = G_F^4 \left(\frac{m_{\beta\beta}}{m_e}\right)^2 G^{0\nu} |M^{0\nu}|^2, \quad (12.42)$$

where  $G^{0\nu}$  is a density of states (or phase space) factor,  $M^{0\nu}$  is the nuclear matrix element for the transition and we have factored out the explicit dependence on  $m_{\beta\beta}$  and the Fermi coupling constant arising from the latter term. Unfortunately  $M^{0\nu}$  can only be calculated using specific nuclear models, giving rise to significant uncertainties, which must be taken into account when translating a given value or limit on the lifetime into a value or limit for the effective mass. In addition, because of the factor  $G_F^4$  and because

the neutrino masses of interest are so small, the predicted lifetimes are extremely long, and like all low-energy weak interactions they increase rapidly with  $Q_{\beta\beta}$ , the energy released in the decay. The actual values vary from nucleus to nucleus, but for an effective neutrino mass of  $50 \text{ meV}/c^2$ , typical half-lives of order  $10^{26} - 10^{27}$  years are obtained, that is 17 orders of magnitude greater than the age of the universe. This presents huge difficulties for experiment, since it implies that to detect one  $\beta\beta0\nu$  decay per year, assuming perfect conditions of 100% detection efficiency and no interfering background, would require a mass of the decaying isotope of approximately 100 kg (see Problem 12.11). The situation confronting a real experiment is even more severe than this, because the detection efficiency is not perfect, and there are also many processes than can mimic a  $\beta\beta0\nu$  signal, including ‘ordinary’  $\beta\beta2\nu$  decays. The latter can occur if the energy resolution of the experiment is insufficient to distinguish the tail of the  $\beta\beta2\nu$  decay from the peak of the  $\beta\beta0\nu$  decay. There could also be signals arising from naturally occurring radioactivity in the material of the detectors that could easily overwhelm the signal being sought. For this reason it is crucial to work with materials with minimal radioactive content, in addition to siting the experiment deep underground to shield it from cosmic rays.

Despite these considerable difficulties, a number of experiments have emerged using different techniques. These can be roughly divided into ‘internal-source systems’, in which the decaying material is contained in the detector, and ‘external-source systems’, in which it is not. The former has been implemented using many types of detector, such as scintillators, bolometers, solid-state devices and gaseous chambers. If the decaying isotope is a semiconductor, the isotope can be both the sample and a solid-state detector that will measure the energy released in a given decay. In the bolometer type of experiment, the sample and detector are still combined, but the energy released in the decay is turned into heat and detected. Alternatively, in the external source category, the decaying sample may be surrounded by a separate detector that can observe and identify the electron tracks from the decays, as well as measure their energies. This has the advantage that a single detector can be easily used to study a range of different decaying isotopes. The three leading experiments of the current generation are EXO-200, KamLAND-Zen and GERDA, which have been operating for a few years. We will therefore concentrate on these three.

EXO-200 (Enriched Xenon Observatory) uses xenon, enriched to contain 81% of the double-beta emitter isotope  $^{136}\text{Xe}$ , which has a natural abundance of about 9%. This has a high  $Q_{\beta\beta}$  value in a region that can have low contaminations for radioactive backgrounds. It can also be dissolved in liquid scintillators, or used as a gas, to realise a homogeneous detector giving both scintillation and



**Figure 12.14** A sketch of the EXO-200 detector, cryostat and shielding. (Adapted from Auger *et al.*, 2012.)

ionization signals. The detector is shown in Figure 12.14. It consists of a cylindrical time projection chamber (TPC), about 40 cm in diameter and 44 cm in length, containing 200 kg of enriched liquid xenon (LXe). The TPC measures the coordinates and energy of ionisation products by simultaneously collecting scintillation light and charge. The xenon is held inside a thin copper vessel immersed in a cryofluid that also shields the detector from external radioactive backgrounds. The cryofluid is maintained at approximately 167 K inside a vacuum-insulated copper cryostat. Further shielding is provided by at least 25 cm of lead in all directions. The current analysis is based on a fiducial volume containing 76.5 kg of LXe and an exposure of 100 kg y. The EXO Collaboration has reported a lower limit on the half-life for the  $\beta\beta0\nu$  decay of  $^{136}\text{Xe}$  of  $T_{1/2}^{0\nu} > 1.1 \times 10^{25}$  y at a confidence level of 90% (Albert *et al.*, 2014). Using (12.42) with a variety of calculated nuclear matrix elements, this result corresponds to an upper limit on the effective Majorana neutrino mass of  $m_{\beta\beta} < (0.19 - 0.45)$  eV.

The KamLAND-Zen experiment also uses xenon, dissolved in a liquid scintillator. The basic detector is shown in Figure 4.28, except for this application a balloon containing 383 kg of  $^{136}\text{Xe}$ -enriched

xenon diluted in liquid scintillator is suspended inside the original plastic balloon, which itself contains 1 kilotonne of pure liquid scintillator. Based on an exposure of 110 kg yr, the group has reported a lower limit on the half-life of the  $\beta\beta0\nu$  decay of  ${}^{136}\text{Xe}$  of  $T_{1/2}^{0\nu} > 2.6 \times 10^{25}$  y at 90% confidence limit (Asakura *et al.*, 2014), and again using (12.42) this corresponds to an upper limit of the effective mass  $m_{\beta\beta} < (0.14 - 0.28)$  eV.

The third experiment, GERDA (GERmanium Detector Array) uses the nucleus  ${}^{76}\text{Ge}$ . This isotope has long been used in  $\beta\beta0\nu$  experiments because it can be embedded in solid-state detectors using a calorimetric approach with high-purity germanium diodes. However, there are disadvantages as well. As in the EXO-200 experiment, the isotope used here has a quite low natural abundance of about 7% and so isotope enrichment is necessary before the diodes can be constructed; also the  $Q_{\beta\beta}$  value is not very high and lies in a region where background contamination is possible. Therefore experiments using this isotope must pay special attention to screening all the materials close to the detectors, among other precautions. The GERDA experiment operates germanium diodes made of isotopically modified material, enriched to about 86% in  ${}^{76}\text{Ge}$  suspended in a liquid argon cryogenic bath. The experiment achieves very low backgrounds by using ultra-pure shielding against environmental radiation. The germanium detectors are suspended in strings into the cryostat where 64 m<sup>3</sup> of liquid argon is used both as a coolant and shield. The vessel is surrounded by a large tank filled with high-purity water, which further shields the inner volumes from local produced radiation. Based on an exposure of 18 kg y, the group has reported a lower limit on the half-life of the  $\beta\beta0\nu$  decay of  ${}^{76}\text{Ge}$  of  $T_{1/2}^{0\nu} > 2.1 \times 10^{25}$  y at 90% confidence limit, which corresponds to  $m_{\beta\beta} < (0.2 - 0.4)$  eV.

Figure 12.13 shows that there is a long way to go before neutrinoless double beta decay experiments can confront theoretical predictions obtained assuming neutrinos are indeed Majorana particles. All three experiments discussed above are currently being upgraded with this aim, and there are at least five new experiments being designed or under construction, using a range of detection methods and isotopes, with the same goal, to reach the region predicted by the inverted hierarchy. This will require increasing the effective mass of the decaying isotope by at least an order-of-magnitude to the 100 kg region, plus a similar-sized reduction in the background. If a signal is still not found, isotope mass in the 1000 kg range will be required, with another reduction in background. The logistic and financial challenges to obtain such large qualities of isotopes, which are usually rare on Earth, will be truly formidable. On the other hand, neutrinoless beta decay may occur by other mechanisms ‘beyond the standard model’ or, in the case of Dirac neutrinos, it may not occur at all.

## Problems 12

- 12.1** List the independent, continuously variable parameters of the standard model, whose values must be determined from experiment. How many are there in total?
- 12.2** Show that it is possible to assign values of  $(B - L)$  ( $B$  is the baryon number and  $L$  is the lepton number) to the  $X$  and  $Y$  bosons in such a way that  $B - L$  is conserved for all the fundamental vertices of Figures 12.2(c) and (d) and Figures 12.3. What are the values of  $B - L$  required?
- 12.3** Find two possible mechanisms for the decay  $p \rightarrow \pi^+ + \bar{\nu}_e$  by suitably relabelling the lines in Figure 12.4(a) and (b). Check that the electric charge  $Q$  and the difference  $(B - L)$  are both conserved at the vertices in the resulting diagrams, using the values for the latter quantity deduced in Problem 12.2.
- 12.4** What would be the energy of a positron if it were emitted from the decay  $p \rightarrow \pi^0 + e^+$  of a proton at rest? If such an event were observed in the SuperKamiokande detector, what would be the emission angle of the resulting Čerenkov radiation, given that the refractive index of water is 1.33?
- 12.5** Consider a star moving with velocity  $v$  in a circular orbit of radius  $R$  about the centre of a spiral galaxy. Calculate the dependence of  $v$  on  $R$  for the following extreme cases.
- The total mass of the galaxy, like the luminous mass, is concentrated almost entirely at the centre of the galaxy.
  - The mass of the galaxy is distributed almost entirely in a spherical halo of dark matter, assumed to be of constant density and radius  $R_h > R$ .
- Compare your results with the observation that  $v$  is almost independent of  $R$  for stars observed in the outer arms of spiral galaxies.
- 12.6** Consider a collision between a cosmic ray proton and a photon in the cosmic microwave background whose energy corresponds to the ambient temperature of 2.7 K. Estimate the minimum proton energy required to produce (a) an  $e^+ e^-$  pair, and (b) a proton–antiproton pair.
- 12.7** In the text, it is stated that if dark matter consisted entirely of WIMPs of mass  $60 \text{ GeV}/c^2$ , the number of WIMPs passing through each square centimetre of the Earth's surface would be of order  $10^5$ , with velocities of order  $2 \times 10^3 \text{ m/s}$ . Use these assumptions to deduce the approximate value of the local density of dark matter and compare it to the average dark matter density of the universe.
- 12.8** A WIMP of mass  $m_w$  and velocity  $v \ll c$  scatters elastically from a stationary nucleus of mass  $m_A$ . Show that the maximum kinetic energy imparted to the nucleus is

$$T_{\max} = \frac{2m_A m_w^2}{(m_A + m_w)^2} v^2.$$

For a given  $m_w$ , what value of  $m_A$  leads to the maximum value of  $T_{\max}$  and what is its value if  $m_w = 60 \text{ GeV}/c^2$  and  $v = 220 \text{ km/s}$ ?

- 12.9** Show that the electric field  $\mathbf{E}$  is odd under parity but even under time reversal, as stated in the discussion of electric dipole moments in Section 12.4.3. How does the magnetic field  $\mathbf{B}$  behave under  $P$  and  $T$ ?
- 12.10** Following the supernova explosion 1987A at a distance  $d = 1.5 \times 10^{21}$  m, a burst of  $\bar{\nu}_e$  interactions was observed in a terrestrial detector. The interactions occurred over a time interval of 7 seconds and the antineutrinos had a range of energies between 7 MeV and 11 MeV. Use this to estimate an upper limit on the  $\bar{\nu}_e$  mass.
- 12.11** Show that the half-life  $T \equiv T_{1/2}$  is related to the mean life  $\tau$  by  $T = \tau \ln 2$  and that the exponential decay law can be written in the form

$$N(t) = N_0 \left(\frac{1}{2}\right)^{t/T},$$

where  $N(t)$  is the number of particles at time  $t$  and  $N_0 = N(t = 0)$ . How many kilograms of  ${}^{76}\text{Ge}$  would be required to give rise to one  $\beta\beta 0\nu$  decay a year, assuming a half-life of  $5 \times 10^{26}$  y?

# A

---

## Relativistic kinematics

---

Particle physics utilises relativistic kinematics in many ways: to relate energies, momenta and scattering angles in different frames of reference; to deduce the masses of unstable particles from measurements on their decay products; to work out threshold energies for the production of new particles; and in a variety of other applications. In this appendix we briefly illustrate such calculations and introduce the concept of *invariant mass*, which can often be used to simplify them. The reader is assumed to be familiar with the elementary theory of special relativity, as discussed, for example, in Chapters 11 to 13 of Kleppner and Kolenkow (1973).

### A.1 The Lorentz transformation for energy and momentum

If a particle has coordinates  $(\mathbf{r}, t) \equiv (x, y, z, t)$  in an inertial frame  $S$ , then its coordinates  $(\mathbf{r}', t')$  in a second inertial frame  $S'$  moving with uniform velocity relative to  $S$  are given by a Lorentz transformation. If the two frames coincide at  $t = 0$ , and frame  $S'$  is moving with uniform speed  $v = |\mathbf{v}|$  in the positive  $x$ -direction relative to  $S$ , this takes the form<sup>1</sup>

$$x' = \gamma(v)(x - vt), \quad y' = y, \quad z' = z, \quad t' = \gamma(v)(t - vx/c^2), \quad (\text{A.1})$$

where  $\gamma(v) \equiv (1 - \beta^2)^{-1/2}$  and  $\beta \equiv v/c$  as usual. The transformation laws for energy and momentum are then obtained as follows.

---

<sup>1</sup> In this appendix we retain factors of  $c$  throughout. The corresponding equations in natural units are trivially obtained by setting  $c = 1$ .

Consider a particle  $i$  of rest mass  $m_i$  moving with uniform velocity  $\mathbf{u}_i$  in  $S$  so that its energy  $E_i$  and momentum  $\mathbf{p}_i$  are given by<sup>2</sup>

$$E_i = m_i c^2 \gamma(u_i), \quad \mathbf{p}_i = m_i \mathbf{u}_i \gamma(u_i), \quad (\text{A.2})$$

where  $u_i = |\mathbf{u}_i|$ . For simplicity, we will take  $\mathbf{u}_i$  to be in the  $x$ -direction, that is in the same direction as  $\mathbf{v}$ . Then it is straightforward to show from the definition of velocity and the transformation (A.1) that the velocity of the particle in  $S'$  is<sup>3</sup>

$$u'_i = \frac{u_i - v}{1 - u_i v/c^2}. \quad (\text{A.3})$$

Hence

$$\gamma(u'_i) \equiv \left(1 - u'^2_i/c^2\right)^{-1/2} = \gamma(u_i) \gamma(v) (1 - u_i v/c^2)$$

and the energy and momentum in the  $S'$  frame are given by

$$E'_i = m_i c^2 \gamma(u'_i) = \gamma(v) (E_i - vp_i) \quad (\text{A.4a})$$

and

$$p'_i = m_i u'_i \gamma(u'_i) = \gamma(v) (p_i - vE_i/c^2), \quad (\text{A.4b})$$

using (A.2) and (A.3). These results are easily generalised to an arbitrary direction for the particle's velocity and momentum, when they become

$$(p'_i)_x = \gamma(v)[(p_i)_x - vE_i/c^2], \quad (\text{A.5a})$$

$$(p'_i)_y = (p_i)_y, \quad (p'_i)_z = (p_i)_z \quad (\text{A.5b})$$

and

$$E'_i = \gamma(v)[E_i - v(p_i)_x]. \quad (\text{A.5c})$$

Furthermore, since (A.5a) to (A.5c) are linear in both energy and momentum, they imply the corresponding results

$$p'_x = \gamma(v)(p_x - vE/c^2), \quad (\text{A.6a})$$

$$p'_y = p_y, \quad p'_z = p_z \quad (\text{A.6b})$$

and

$$E' = \gamma(v)(E - vp_x), \quad (\text{A.6c})$$

for the total energy and momentum

$$E = E_1 + E_2 + \cdots + E_N, \quad \mathbf{p} = \mathbf{p}_1 + \mathbf{p}_2 + \cdots + \mathbf{p}_N \quad (\text{A.7})$$

of  $N$  particles, which are sufficiently far apart for their mutual interaction energies to be neglected.

---

<sup>2</sup> Energy always means the total energy, including the rest energy, as distinct from the kinetic energy  $T_i \equiv E_i - m_i c^2$ .

<sup>3</sup> This is just the usual transformation law for velocities; see, for example, Section 12.4 of Kleppner and Kolenkow (1973).

## A.2 The invariant mass

The *invariant mass*  $W$  of a set of  $N$  particles is defined by

$$W^2 c^4 \equiv E^2 - \mathbf{p}^2 c^2, \quad (\text{A.8})$$

where  $E$  and  $\mathbf{p}$  are the total energy and momentum given by (A.7). As its name implies,  $W$  has the same value in any reference frame, since

$$(E')^2 - (\mathbf{p}')^2 c^2 = (E)^2 - (\mathbf{p})^2 c^2,$$

as may be verified from (A.5). It is most easily evaluated in the centre-of-mass (CM) frame where

$$\mathbf{p} = \sum_i \mathbf{p}_i = \mathbf{0}, \quad (\text{A.9})$$

so that

$$Wc^2 = E_{CM} \quad (\text{A.10a})$$

is the total energy in the CM frame. For a single particle  $i$ , the invariant mass  $W_i$  is identical with the rest mass  $m_i$ , that is

$$W_i = m_i, \quad (\text{A.10b})$$

since the CM frame is identical with the rest frame and  $E_{CM}$  is just  $m_i c^2$ .

Many applications of relativistic kinematics can be simplified by evaluating the invariant mass in the frame of interest using (A.7) and (A.8), and equating the results to the values obtained in the CM frame (A.10a) and (A.10b). We will illustrate this by some examples.

### A.2.1 Beam energies and thresholds

The reference frames most commonly met in particle physics are those of the laboratory and the centre-of-mass, and as the first example we will relate the energies in these two frames. Specifically, we shall consider a beam of particles  $b$  with momentum  $\mathbf{p}_L$  incident upon target particles  $t$ , which are at rest in the laboratory, so that  $\mathbf{p}_t = \mathbf{0}$ . The energies of the particles in the laboratory frame are

$$E_L = (m_b^2 c^4 + \mathbf{p}_L^2 c^2)^{1/2}, \quad E_t = m_t c^2 \quad (\text{A.11})$$

and the invariant mass  $W$  is given by

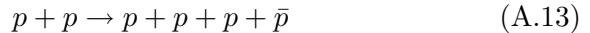
$$W^2 c^4 = (E_L + m_t c^2)^2 - \mathbf{p}_L^2 c^2 = m_b^2 c^4 + m_t^2 c^4 + 2m_t c^2 E_L,$$

using (A.7) and (A.8). On comparing with (A.10a) we immediately obtain

$$E_{CM} = \left( m_b^2 c^4 + m_t^2 c^4 + 2m_t c^2 E_L \right)^{1/2}, \quad (\text{A.12})$$

which is identical with the result Equation (4.1a) quoted for the total centre-of-mass energy  $E_{CM}$  in Chapter 4.

The total centre-of-mass energy is important because it determines the energy available for particle production. For example, the antiproton production reaction



can only occur if

$$E_{CM} \geq 4m_P c^2. \quad (\text{A.14})$$

Here the minimum value corresponds to all four final-state particles produced at rest, which is only possible in the centre-of-mass frame (A.9). In the laboratory frame, the beam energy  $E_{\min}$  corresponding to the minimum centre-of-mass energy  $4m_P c^2$  is given by

$$4m_P c^2 = (2m_P^2 c^4 + 2m_P c^2 E_{\min})^{1/2}$$

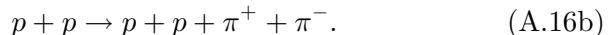
from (A.12) with  $m_b = m_t = m_P$ , leading to

$$E_L \geq E_{\min} = 7m_P c^2 = 6.6 \text{ GeV} \quad (\text{A.15})$$

as the range of beam energies corresponding to (A.14). Consequently, in an experiment in which a beam of protons is incident upon a static proton target, the beam energy must be at least 6.6 GeV if antiprotons are to be produced. Here  $E_{\min} = 6.6 \text{ GeV}$  is called the threshold energy for antiproton production,<sup>4</sup> and the corresponding thresholds for the production of other particles may be calculated in a similar way. In doing this, account must be taken of the appropriate conservation laws in identifying the lightest final state containing the desired particle. For example, for positive and negative pion production in proton–proton collisions, the lightest final states that conserve baryon number and electric charge are



and



These lead to threshold energies of 1.23 GeV and 1.54 GeV for  $\pi^+$  and  $\pi^-$  mesons, respectively, corresponding to beam momenta of  $0.79 \text{ GeV}/c$  and  $1.22 \text{ GeV}/c$ . In practice, to get observable counting rates, one must work at energies somewhat above threshold.

---

<sup>4</sup> This energy is somewhat reduced if a nuclear target is used, owing to the motion of the target nucleons within the nucleus (see Problem A.5). Antiprotons were first detected in 1955 using a proton beam incident upon a copper target. A simple account of this classic experiment may be found, for example, in Section 3.9 of Hughes (1985).

### A.2.2 Masses of unstable particles

A simple but important use of the invariant mass is in determining the masses of very short-lived particles by measurements on their decay products. If a particle  $A$  decays to  $N$  particles  $i = 1, 2, \dots, N$  in the final state, that is

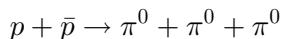
$$A \rightarrow 1 + 2 + \dots + N,$$

then the invariant mass (A.8) of the final-state particles is given by

$$W^2 c^4 = \left( \sum_i E_i \right)^2 - \left( \sum_i \mathbf{p}_i c \right)^2 = E_A^2 - (\mathbf{p}_A c)^2 = m_A^2 c^4 \quad (\text{A.17})$$

using energy-momentum conservation. Hence the mass of the decaying particle is equal to the invariant mass of its decay products, and the latter can be used to determine the former if the particle is too short-lived for its mass to be measured directly.

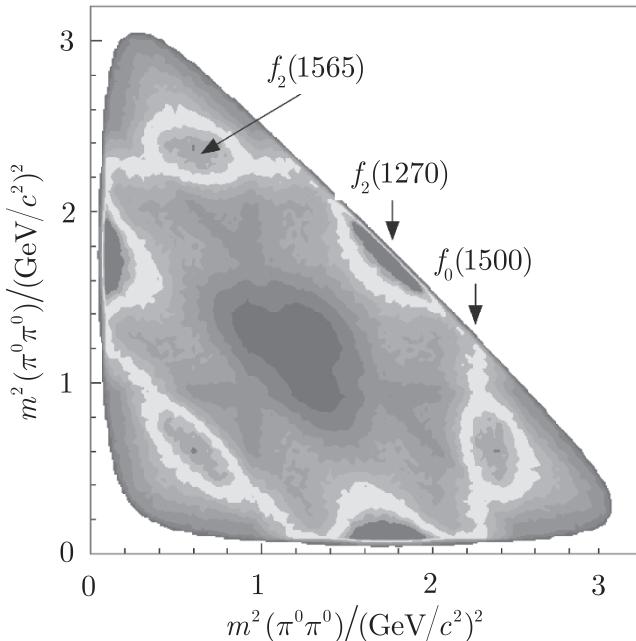
A related example is the identification of resonances from three-particle final states in reactions of the type  $1 + 2 \rightarrow 3 + 4 + 5$ , using so-called ‘Dalitz plots’.<sup>5</sup> From measurements of the kinematic variables for particles 3, 4 and 5, we can form the squared invariant masses  $m_{34}^2$ ,  $m_{45}^2$  and  $m_{35}^2$ . If we plot the first two along the  $x$  and  $y$  axes, then the third is a constant along lines at  $45^\circ$  because the sum of the squared invariant masses is a constant for a given initial energy, as shown in Problem A.7. The physically allowed kinematic variables lie inside a well-defined region on the plot and in the absence of resonances it can be shown that this region is uniformly populated with events.<sup>6</sup> Resonance behaviour in two of the final-state particles gives rise to a band of higher density parallel to one of the coordinate axes, or along a  $45^\circ$  line. This is illustrated in Figure A.1, which shows a high-statistics Dalitz plot for 712 000 final-state events in the reaction



obtained in an experiment at CERN studying meson spectroscopy. The plot has a high degree of symmetry because the three final-state particles are identical and therefore each event is entered six times. Clear enhancements due to the presence of mesons resonances can be seen, three of which are labelled with the appropriate state.

<sup>5</sup> There are several varieties of these. The original form was published by Dalitz (1953) and was used to study the so-called ‘ $\tau - \theta$  puzzle’ – see Footnote 1 of Chapter 11.

<sup>6</sup> For a proof, see, for example, p. 310 of Burcham and Jobes (1995).



**Figure A.1** Dalitz plot for the reaction  $p + \bar{p} \rightarrow \pi^0 + \pi^0 + \pi^0$  using data obtained from the crystal barrel experiment at the LEAR accelerator at CERN. (Adapted from C. Amsler 1998, with permission from the American Physical Society.)

### \*A.3 Transformation of the scattering angle

As a final example of relativistic kinematics, we consider the transformation of scattering angles between the laboratory frame, where they are measured in fixed-target experiments, and the centre-of-mass frame, which is almost invariably used in theoretical discussions. In doing this we will find that for very high beam energies the scattering angles in the laboratory tend to be very small, and most of the particles produced in the interaction emerge in a narrow cone about the incoming beam direction.

We consider a reaction of the general form

$$b(E_L, \mathbf{p}_L) + t(m_t c^2, \mathbf{0}) \rightarrow P(E, \mathbf{q}) + \dots, \quad (\text{A.18})$$

where

$$E = (m_P^2 c^4 + \mathbf{q}^2 c^2)^{1/2} \quad (\text{A.19})$$

and the dots indicate that one or more particles are produced along with the particle  $P$ . Without loss of generality, we can choose axes so that

$$\mathbf{p}_L = (p_L, 0, 0), \quad \mathbf{q} = (q \cos \theta_L, q \sin \theta_L, 0), \quad (\text{A.20})$$

that is so that the beam is in the  $x$ -direction and the momentum  $\mathbf{q}$  lies in the  $xy$  plane. The angle  $\theta_L$  between the beam direction and the direction of the particle  $P$  is called the *production angle*, or *scattering angle*, and we want to find its value in the centre-of-mass frame.

The first step is to find the velocity  $v$  of the Lorentz transformation, which takes us from the laboratory frame to the centre-of-mass

frame. For any Lorentz transformation along the beam direction (i.e. the  $x$ -axis) the beam and target momenta become

$$p'_b = \gamma(v)(p_L - vE_L/c^2), \quad p'_t = -m_t v \gamma(v)$$

by (A.4), while in the centre-of-mass frame

$$p'_b + p'_t = 0$$

by definition. Combining these results gives

$$v = \frac{p_L c^2}{E_L + m_t c^2} \quad (\text{A.21})$$

as the velocity of the transformation.

The next step is to work out the momentum  $\mathbf{q}'$  of the produced particle  $P$  in the centre-of-mass frame and hence the production or scattering angle between it and the beam direction. From (A.5) the components of the momentum are

$$q'_x = \gamma(v)(q \cos \theta_L - vE/c^2), \quad q'_y = q \sin \theta_L, \quad q'_z = 0, \quad (\text{A.22})$$

where  $v$  is given by (A.21). Writing

$$\mathbf{q}' = (q' \cos \theta_C, q' \sin \theta_C, 0), \quad (\text{A.23})$$

where  $q' = |\mathbf{q}'|$  and  $\theta_C$  is the angle between  $\mathbf{q}'$  and the  $x$ -axis, we immediately obtain

$$\tan \theta_C = \frac{q'_y}{q'_x} = \frac{1}{\gamma(v)} \frac{q \sin \theta_L}{q \cos \theta_L - vE/c^2}. \quad (\text{A.24})$$

Since the beam direction is along the  $x$ -axis, this is the desired result, which relates the centre-of-mass scattering angle  $\theta_C$  to measured quantities in the laboratory frame.

The dominance of small laboratory scattering angles  $\theta_L$  at high energies follows from the inverse of (A.24)

$$\tan \theta_L = \frac{1}{\gamma(v)} \frac{q' \sin \theta_C}{q' \cos \theta_C + vE'/c^2}, \quad (\text{A.25})$$

where

$$E' = \left( m_P^2 c^4 + q'^2 c^2 \right)^{1/2}$$

is the energy of the produced particle in the centre-of-mass frame. This result follows from (A.24), since to go from the centre-of-mass to the laboratory frame requires the inverse transformation to (A.22) with the velocity reversed, that is  $v \rightarrow -v$ . At high energies

$$E_L \approx p_L c \gg m_b c^2, m_t c^2,$$

so that

$$v \approx c(1 - m_t c/p_L) \approx c$$

by (A.21) and

$$\gamma(v) = (1 - v^2/c^2)^{-1/2} \approx (p_L/2m_t c)^{1/2}.$$

On substituting these values into (A.25), together with

$$E' = m_P c^2 \gamma(u), \quad q' = m_P u \gamma(u),$$

where  $u$  is the magnitude of the particle's velocity in the centre-of-mass frame, we obtain

$$\tan \theta_L \approx \left( \frac{2m_t c}{p_L} \right)^{1/2} \frac{u \sin \theta_C}{u \cos \theta_C + c} \quad (\text{A.26})$$

as the appropriate form at high energies. The presence of the factor  $p_L^{-1/2}$  then implies that  $\theta_L$  will be very small, unless  $u \approx c$  and  $\cos \theta_C \approx -1$ . Because these conditions are not usually satisfied, most particles are produced in a narrow cone about the beam direction in the laboratory frame, as we have asserted. A similar result applies to decay angles: when a high-energy particle decays, the decay products emerge predominantly at very small angles to the initial beam. This is demonstrated by a very similar argument to the above, and is left as an exercise for the reader (see Problem A.8).

## Problems A

- A.1** Derive an expression for the muon energy in the decay  $\pi^+ \rightarrow \mu^+ + \nu_\mu$  in terms of the pion, muon and neutrino masses for the case of pions at rest.

- A.2** A particle of mass  $m$  decays in flight to give two photons of energies  $E_1$  and  $E_2$ . Show that the opening angle  $\theta$  between the photon directions is given by

$$\cos \theta = 1 - m^2 c^4 / 2E_1 E_2.$$

(Several short-lived particles – for example  $\pi^0(140)$ ,  $\eta^0(547)$  – can decay in this way, and since they have different masses the above relation gives a simple means of identifying them.)

- A.3** Calculate the lowest energy at which  $\Lambda(1116)$  baryons can be produced in the strong interaction of negative pions with protons at rest.

- A.4** A neutral particle  $X^0$  decays to two charged particles  $A^+$  and  $B^-$ . The momentum components of the decay products in  $\text{GeV}/c$  are measured to be:

	$p_x$	$p_y$	$p_z$
$A^+$	-0.488	-0.018	2.109
$B^-$	-0.255	-0.050	0.486

Is the decay  $K_S^0(498) \rightarrow \pi^+ \pi^-$  or  $\Lambda(1116) \rightarrow p \pi^-$ ?

- A.5** Show that when the motion of nucleons within the nucleus is taken into account, the threshold energy  $E'$  for producing antiprotons in

the collision of protons with a nuclear target is given approximately by

$$E'_{\min} \approx (1 - p/m_P c) E_{\min} \approx 5.2 \text{ GeV},$$

where  $E_{\min} \approx 6.6 \text{ GeV}$  is the threshold energy for producing antiprotons on a proton target, and  $p = |\mathbf{p}|$ , the magnitude of the internal momenta of the nucleons, is taken to be typically of order  $0.2 \text{ GeV}/c$ .

- A.6** (a) Find the maximum values of velocity  $v$  and  $\gamma$  if the non-relativistic approximation for energy,  $E \approx mc^2 + p^2/2m$ , is to be used with an error less than  $\varepsilon p^2/2m$ . Hence show that for an error of 1%, the momentum satisfies  $p \approx 0.2mc$ .  
 (b) Find the minimum values of velocity  $v$  and  $\gamma$  if the relativistic approximation for energy,  $E \approx pc$ , is to be used with an error of less than  $\varepsilon pc$ . Hence show that for an error of 1%, the momentum satisfies  $p \approx 7mc$ .
- A.7** Three particles, labelled 1, 2, 3, are produced in a given experiment at a fixed centre-of-mass energy. If

$$m_{ij}^2 c^4 \equiv (E_i + E_j)^2 - (\mathbf{p}_1 + \mathbf{p}_2)^2 c^2,$$

show that the sum of the squared invariant masses is

$$m_{12}^2 c^4 + m_{23}^2 c^4 + m_{31}^2 c^4 = E_{CM}^2 + \sum_{i=1}^3 m_i^2 c^4.$$

- A.8** When high-energy particles decay, the decay products emerge predominantly at very small angles  $\theta_L$  to the initial beam direction. Verify this by showing that for decays of the form

$$b(E_L, \mathbf{p}_L) \rightarrow P(E, \mathbf{q}) + \dots$$

the decay angle  $\theta_L$  defined by

$$\mathbf{p}_L \cdot \mathbf{q} = |\mathbf{p}_L| |\mathbf{q}| \cos \theta_L$$

is given at high beam energies  $E \gg m_b c^2$  by

$$\tan \theta_L \approx \frac{m_b c^2}{\sqrt{2} E_L} \frac{u \sin \theta_C}{u \cos \theta_C + c},$$

where  $u$  is the magnitude of the velocity of particle  $P$  in the laboratory frame and  $\theta_C$  is the decay angle in the rest frame of the decaying particle.

---

## Amplitudes and cross-sections

---

In this appendix we define the various cross-sections and scattering amplitudes used in the book, and derive the Breit–Wigner resonance formulas.

### B.1 Rates and cross-sections

In a typical fixed-target scattering experiment, a beam of particles (e.g. negative pions), ideally monoenergetic, is directed on to a target (e.g. liquid hydrogen) and the rates of production of various particles are measured. These can be regarded as the sums of contributions resulting from the interactions of individual beam particles with individual target particles (i.e. individual  $\pi^- p$  interactions in our example) provided the following conditions are satisfied:

- (a) The target particles are separated by distances that are much greater than the de Broglie wavelength of the incident particles, so that interference effects between waves scattered from different target particles can be neglected.
- (b) The target is small enough, or of sufficiently low density, for multiple scattering to be neglected, so that any particles produced in the interaction will almost certainly leave the target without further interactions.
- (c) The collision energy is sufficiently high that the binding energies of the particles in the target can be neglected.
- (d) The beam density is low enough that mutual interactions of its particles can be neglected.

These conditions are usually satisfied in particle physics experiments, and we will focus our attention on interactions between individual beam and target particles.

For each initial state, a number of final states may be possible. For example, in  $\pi^- p$  collisions the reactions

$$\pi^- + p \rightarrow \pi^- + p, \quad (\text{B.1a})$$

$$\pi^- + p \rightarrow K^0 + \Lambda, \quad (\text{B.1b})$$

$$\pi^- + p \rightarrow \pi^+ + \pi^- + n \quad (\text{B.1c})$$

and

$$\pi^- + p \rightarrow \pi^+ + \pi^- + \pi^0 + n, \quad (\text{B.1d})$$

as well as many others compatible with the conservation laws, can all be observed. However, whatever reaction we choose, the rate will be proportional to:

- (i) the number  $N$  of particles in the target illuminated by the beam;
- (ii) the rate per unit area at which beam particles cross a small surface placed in the beam at rest with respect to the target and perpendicular to the beam direction.

The latter rate is called the *flux*, denoted by  $J$ , and is given by

$$J = n_b v_i, \quad (\text{B.2})$$

where  $n_b$  is the density of particles in the beam and  $v_i$  their velocity in the rest frame of the target. Hence the rate  $W_r$  at which a specific reaction  $r$  occurs in a particular experiment can be written in the form

$$W_r = J N \sigma_r, \quad (\text{B.3a})$$

where  $\sigma_r$ , the constant of proportionality, is called the *cross-section* for reaction  $r$ . This decomposition is very convenient, because the dependence of the observed rate on the densities and geometries of the beam and target is absorbed into the factor

$$L \equiv J N, \quad (\text{B.3b})$$

called the *luminosity*  $L$ , which from (B.2) and (B.3a) has dimensions  $[\text{length}]^{-2} [\text{time}]^{-1}$ . A related quantity that is also used is the *integrated luminosity*, defined as the integral of  $L$  over time. The cross-section is independent of  $L$ , but characteristic of the particular reaction  $r$ . It follows from (B.3a) that  $\sigma_r$  has the dimensions of an area, and the rate per target particle  $J \sigma_r$  at which the reaction occurs is equal to the rate at which beam particles would hit a surface of area  $\sigma_r$  placed in the beam at rest with respect to the target, and perpendicular to the beam direction. Since the area of such a surface is unchanged by a Lorentz transformation in the beam

direction, the cross-section has the same value in both the laboratory and centre-of-mass frames, which are related by just such a transformation.

In the case of a colliding beam experiment, the ‘target’ is itself another beam, but the conditions (a) to (d) above are still generally valid. Equation (B.3a) still holds, but it is more convenient to express it in a different form. For example, we consider a collider with beams of particles of types 1 and 2, which may or may not be the same, circulating in opposite directions and made to collide at a given point. The beams will be composed of bunches, and we define  $N_1$  and  $N_2$  to be the numbers of particles per bunch in the two beams and  $f$  to be the frequency of collisions between bunches. Since the beams do not have sharp cut-offs at their edges, we will assume Gaussian-shaped beams with transverse dimensions  $s_x$  and  $s_y$ , which yield a cross-sectional area  $A = 4\pi s_x s_y$ . In one collision between bunches, a particle in beam 1 crosses  $N_2/A$  particles in beam 2, regarded as the target, and because it is in a bunch of  $N_1$  particles, the luminosity is

$$L = f N_1 N_2 / A \quad (\text{B.4a})$$

and the number of events per second in which a specific reaction  $r$  occurs is

$$R = \sigma_r f N_1 N_2 / A. \quad (\text{B.4b})$$

Note that, in this case, the luminosity may be increased not just by increasing the number of particles in the bunches, but also by increasing the frequency of revolution of the beams. In the case of the LHC,  $s_x = s_y = 1.6 \times 10^{-3}$  cm,  $N_1 = N_2 = 1.15 \times 10^{11}$  and the time between bunches is  $2.5 \times 10^{-8}$  s, so that  $L \approx 10^{34} / (\text{cm}^2 \text{s})$ . Finally, there could be a reduction factor in the luminosity if the beams do not collide head-on at the interaction point; at the LHC this is about 5%.

## B.2 The total cross-section

The quantity  $\sigma_r$  is better named the *partial cross-section*, because it is the cross-section for a particular reaction  $r$ . The *total cross-section*  $\sigma$  is defined by

$$\sigma \equiv \sum_r \sigma_r, \quad (\text{B.5})$$

where the sum extends over all reactions  $r$  that satisfy the relevant conservation laws, such as the reactions listed in (B.1) for  $\pi^- p$  scattering. The total reaction rate is then

$$W \equiv \sum_r W_r = L \sigma, \quad (\text{B.6})$$

by (B.3) and (B.4). Since every reaction removes a particle from the incoming beam, it is the total cross-section that determines how the beam intensity is reduced on traversing a fixed target that is too long for condition (b) above to be satisfied.

To explore this further, we consider a beam of cross-sectional area  $A$  traversing a stationary target of length  $d$ , and focus our attention on a thin segment of the target of thickness  $dx$  at a distance  $x$  within it, as shown in Figure B.1. The reduction in the beam intensity on crossing this segment is equal to the reaction rate  $dW$  within the segment, that is

$$AdJ(x) = -dW, \quad (\text{B.7a})$$

where  $dJ$  is the change in the flux  $J(x)$  between  $x$  and  $x + dx$ . By (B.3a) the rate is also given by

$$dW = J(x)\sigma n_t A dx, \quad (\text{B.7b})$$

where  $n_t$  is the density of particles in the target. Combining (B.7a) and (B.7b) gives

$$dJ(x) = -n_t \sigma J(x) dx,$$

and if the incident flux is  $J(0) \equiv J_i$ , a simple integration gives

$$J(x) = J_i \exp(-x/l_c) \quad (0 \leq x \leq d), \quad (\text{B.8})$$

where the collision length

$$l_c \equiv 1/n_t \sigma \quad (\text{B.9})$$

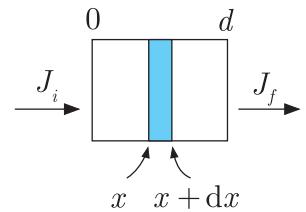
is the mean free path travelled by a beam particle in an infinitely long target ( $d = \infty$ ) before it interacts. Beyond the target, the transmitted flux

$$J_f \equiv J(x \geq d) = J_i \exp(-d/l_c). \quad (\text{B.10})$$

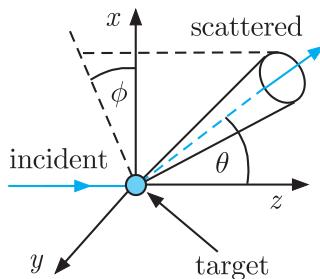
These results have two important consequences. The first is that multiple scattering within the target can only be neglected if the target thickness  $d$  is much less than the collision length  $l_c$ . The basic cross-section formulas (B.3) and (B.6) obviously depend on this, since they treat the flux  $J$  as a constant. The second consequence is that the measured depletion of a beam flux on passing through a long target ( $d \sim l_c$ ) gives a relatively simple determination of the total cross-section using (B.10).

### B.3 Differential cross-sections

We next consider the angular distributions of the particles produced in a scattering reaction. These are described by writing the cross-section  $\sigma_r$  as a sum of contributions corresponding to different directions of emission of the final-state particles. For processes with two-body final states, such as (B.1a) and (B.1b), only the direction



**Figure B.1** Beam of particles with initial flux  $J_i$  being transmitted through a target of length  $d$  with reduced flux  $J_f < J_i$ .



**Figure B.2** Geometry of the differential cross-section: a beam of particles is incident along the  $z$ -direction and collides with a stationary target at the origin. The differential cross-section is proportional to the rate for particles to be scattered into a small solid angle  $d\Omega$  in the direction  $(\theta, \phi)$ .

of one particle needs to be specified, since the direction of the other follows uniquely from energy and momentum conservation.<sup>1</sup> The direction of the chosen particle – for example the  $\pi^-$  in (B.1a) or the  $K^0$  in (B.1b) – is specified by a polar angle  $\theta$  and an azimuthal angle  $\phi$  defined in a convenient reference frame in which the  $z$ -axis coincides with the beam direction. The angle  $\theta$  between the particle direction and the beam direction is called the *scattering angle* in an elastic reaction like (B.1a), or the *production angle* in an inelastic reaction like (B.1b).

The angular distribution of the chosen particle (e.g.  $\pi^-$  or  $K^0$ ) produced in a two-body reaction  $r$  is described by the *differential cross-section*

$$\frac{d\sigma_r(\theta, \phi)}{d\Omega}, \quad (\text{B.11})$$

defined by

$$dW_r \equiv L \frac{d\sigma_r(\theta, \phi)}{d\Omega} d\Omega, \quad (\text{B.12})$$

where  $dW_r$  is the measured rate for the particle to be emitted into an element of solid angle  $d\Omega = d\cos\theta d\phi$  in the direction  $(\theta, \phi)$ , as shown in Figure B.2. The reaction cross-section  $\sigma_r$  can then be recovered by integrating the differential cross-section over all solid angles,

$$\sigma_r = \int_0^{2\pi} d\phi \int_0^1 d\cos\theta \frac{d\sigma_r(\theta, \phi)}{d\Omega} \quad (\text{B.13})$$

as can be seen by integrating (B.12) to give the total rate  $W_r$  and comparing with (B.3).

The differential cross-section (B.11) depends not only on the energy (which is not explicitly shown in (B.11)) and the angles but also on the spin states of the particles involved. In practice, most experiments are done using unpolarised beams and targets, and the spins of the particles in the final state are not measured. The cross-sections measured in such experiments are called *unpolarised cross-sections*, and for them the question of spin dependence does not arise. In addition, there is complete cylindrical symmetry about the beam direction (i.e. the  $z$ -direction), with nothing to distinguish physically between different directions in the  $xy$  plane, and hence between different values of the azimuthal angle  $\phi$ . Thus unpolarised cross-sections are independent of  $\phi$ , that is

$$\frac{d\sigma_r(\theta, \phi)}{d\Omega} \rightarrow \frac{d\sigma_r(\theta)}{d\Omega},$$

<sup>1</sup> For example, in the centre-of-mass frame the particles are emitted in opposite directions, since the sum of their momenta must be zero.

and it is customary to define an unpolarised differential cross-section

$$\frac{d\sigma_r}{d\cos\theta} \equiv \int_0^{2\pi} d\phi \frac{d\sigma_r}{d\Omega} = 2\pi \frac{d\sigma_r}{d\Omega}, \quad (\text{B.14})$$

with

$$\sigma_r = \int_{-1}^1 d\cos\theta \frac{d\sigma_r}{d\cos\theta}. \quad (\text{B.15})$$

The cross-sections discussed in this book are almost exclusively unpolarised cross-sections.

## \*B.4 The scattering amplitude

In this section we will relate the experimentally defined cross-sections to theoretical calculations of the probability of a transition from an initial quantum state  $i$  to a final quantum state  $f$ . In doing so, it is convenient to consider a single beam particle interacting with a single target particle, assumed to be at rest, and to confine the whole system in a large cube defined by

$$-L/2 \leq x, y, z \leq L/2, \quad (\text{B.16})$$

so that the volume  $V = L^3$ . The incident flux (B.2) is then given by

$$J = n_b v_i = v_i / V,$$

and the luminosity  $L = J$  so that (B.12) becomes

$$dW_r = \frac{v_i}{V} \frac{d\sigma_r(\theta, \phi)}{d\Omega} d\Omega. \quad (\text{B.17})$$

Expressions for the differential cross-sections are found by comparing (B.17) with theoretical predictions for  $dW_r$ . In doing this, we shall impose the periodic boundary conditions,

$$\begin{aligned} \psi(-L/2, y, z) &= \psi(L/2, y, z), \\ \psi(x, -L/2, z) &= \psi(x, L/2, z), \\ \psi(x, y, -L/2) &= \psi(x, y, L/2), \end{aligned} \quad (\text{B.18})$$

on the wavefunctions, and take the limit  $V = L^3 \rightarrow \infty$  at the end of the calculation. The results are independent of  $V$  in this limit, and also of the detailed form of the boundary conditions chosen.

We start by considering a particularly simple case: the scattering of a non-relativistic particle by a potential  $V(\mathbf{r})$ , using lowest-order

perturbation theory. We are then interested in transitions from an initial state described by a wavefunction

$$\psi_i = \frac{1}{\sqrt{V}} \exp(i \mathbf{q}_i \cdot \mathbf{r}) \quad (\text{B.19a})$$

to final states described by wavefunctions

$$\psi_f = \frac{1}{\sqrt{V}} \exp(i \mathbf{q}_f \cdot \mathbf{r}), \quad (\text{B.19b})$$

where the final momentum  $\mathbf{q}_f$  lies within a small solid angle  $d\Omega$  located in the direction  $(\theta, \phi)$ . In non-relativistic quantum mechanics, the transition rate for any process is given in first-order perturbation theory by the Born approximation<sup>2</sup>

$$dW_r = 2\pi \left| \int \psi_f^* V(\mathbf{r}) \psi_i d^3 \mathbf{r} \right|^2 \rho(E_f), \quad (\text{B.20})$$

where the *density of states factor*  $\rho(E_f)$  is specified below. Substituting (B.19) then gives

$$dW_r = \frac{2\pi}{V} \rho(E_f) |\mathcal{M}_{if}|^2, \quad (\text{B.21})$$

where the scattering amplitude  $\mathcal{M}_{if}$  is

$$\mathcal{M}_{if} = \int V(\mathbf{r}) \exp[i(\mathbf{q}_i - \mathbf{q}_f) \cdot \mathbf{r}] d^3 \mathbf{r}. \quad (\text{B.22})$$

The density of states factor  $\rho(E)$  is defined by setting  $\rho(E) dE$  equal to the number of possible quantum states of the final-state particles that have a total energy between  $E$  and  $E + dE$ . It is found by first evaluating  $\rho(q)$ , where  $\rho(q)dq$  is the number of possible final states with  $q = |\mathbf{q}|$  lying between  $q$  and  $q + dq$ , and then changing variables using

$$\rho(q) \frac{dq}{dE} dE = \rho(E) dE. \quad (\text{B.23})$$

The possible values of the momentum  $\mathbf{q}$  are restricted by the boundary conditions (B.18) to be

$$q_x = \left( \frac{2\pi}{L} \right) n_x, \quad q_y = \left( \frac{2\pi}{L} \right) n_y, \quad q_z = \left( \frac{2\pi}{L} \right) n_z,$$

where  $n_x$ ,  $n_y$  and  $n_z$  are integers. Hence the number of final states with momenta lying in the momentum space volume

$$d^3 \mathbf{q} = q^2 dq d\Omega,$$

---

<sup>2</sup> For a discussion of this formula see, for example, Section 10.2.2 of Mandl (1992) or pp. 397–399 of Gasiorowicz (1974). Note that we are using natural units, so that  $\hbar = 1$ .

corresponding to momenta pointing into the solid angle  $d\Omega$  with magnitude between  $q$  and  $q + dq$ , is given by

$$\rho(q) dq = \left(\frac{L}{2\pi}\right)^3 d^3 \mathbf{q} = \frac{V}{8\pi^3} q^2 dq d\Omega. \quad (\text{B.24})$$

The derivative

$$\frac{dq}{dE} = \frac{1}{v}, \quad (\text{B.25})$$

where  $q = mv$  for a non-relativistic particle, and substituting (B.24) and (B.25) in (B.23) gives

$$\rho(E_f) = \frac{V}{8\pi^3} \frac{q_f^2}{v_f} d\Omega \quad (\text{B.26})$$

for the density of states evaluated in the final energy  $E_f$ . The desired expression for the differential cross-section can now be found by substituting (B.26) into (B.21) and comparing with (B.17). On doing this, and remembering that for elastic scattering energy conservation requires  $v_i = v_f = v$  and  $q_i = q_f = mv$ , we obtain

$$\frac{d\sigma}{d\Omega} = \frac{m^2}{4\pi^2} |\mathcal{M}_{if}|^2, \quad (\text{B.27})$$

where  $\mathcal{M}_{if}$  is given by the Fourier transform of the potential, (B.22).

The result (B.27) has been obtained for non-relativistic elastic scattering in lowest-order perturbation theory. Nonetheless, it requires very little modification to extend it to any reaction of the form

$$a(\mathbf{q}_a, m_a) + b(\mathbf{q}_b, m_b) \rightarrow c(\mathbf{q}_c, m_c) + d(\mathbf{q}_d, m_d), \quad (\text{B.28})$$

where  $m_a$ , etc., specify the spin states of the particles, and we have indicated their momenta in the centre-of-mass frame. To see this, we first note that any theoretically calculated reaction rate involves a factor proportional to the density of final states  $\rho(E_f)$ ; hence (B.21) may be taken as a *definition* of the scattering amplitude  $\mathcal{M}_{if}$ , apart from an overall phase factor, which is irrelevant in what follows. The density of states must of course be evaluated using relativistic kinematics. The only place where we have used non-relativistic kinematics in the evaluation above was in finding the derivative (B.25). It is straightforward to show that (B.25) also holds if relativistic kinematics are used, and thus the same expression (B.26) results for the density of states. Substituting into (B.21) and comparing with (B.17) then gives

$$\frac{d\sigma}{d\Omega} = \frac{1}{4\pi^2} \frac{q_f^2}{v_i v_f} |\mathcal{M}_{if}|^2 \quad (\text{B.29})$$

as the general expression for the differential cross-section for a reaction of the type (B.28), evaluated in the centre-of-mass frame. It

reduces to our previous result (B.27) for non-relativistic elastic scattering on setting  $v_i = v_f = v$  and  $q_i = q_f = mv$ .

The result (B.29) applies to the scattering of polarised particles as symbolised in (B.28), and the amplitude

$$\mathcal{M}_{if}(\theta, \phi) \equiv \mathcal{M}_{if}(\theta, \phi; m_a, m_b, m_c, m_d)$$

depends on the particular spin states involved. If the beam and target are unpolarised, one must *average* over the initial spin states, whereas if all final-state particles are detected, independent of their spin projections, one must *sum* over the final spin states. Hence in an unpolarised scattering experiment, the spin-averaged reaction rates must be multiplied by a factor  $(2s_f + 1)$  for each final-state particle of spin  $s_f$ . In addition, since the cross-section only depends on the angle  $\theta$  in this case, we may integrate over the azimuthal angle  $\phi$  using (B.14) to give

$$\frac{d\sigma}{d\cos\theta} = \frac{(2s_c + 1)(2s_d + 1)}{2\pi} \frac{q_f^2}{v_i v_f} \overline{|\mathcal{M}_{if}|^2} \quad (\text{B.30})$$

for the unpolarised differential cross-section, where the final factor is the spin average of the squared matrix element  $|\mathcal{M}_{if}|^2$ .

## B.5 The Breit–Wigner formula

An unstable particle may be characterised by its lifetime at rest  $\tau$ , or equivalently by its natural decay width  $\Gamma = \hbar/\tau$ . Since we use natural units  $\hbar = 1$ , the decay width is equal to the decay rate

$$\Gamma = 1/\tau \quad (\text{B.31a})$$

and we shall not distinguish between them in what follows. In general, the particle will decay to a number of different channels  $f$  (for example,  $Z^0 \rightarrow e^+e^-, \mu^+\mu^-, \dots$ ) and the total decay width/rate  $\Gamma$  is the sum of the partial decay widths/rates  $\Gamma_f$  for the various channels. That is,

$$\Gamma = \sum_f \Gamma_f, \quad (\text{B.31b})$$

while the *branching ratio*, the fraction of decays leading to a particular channel  $f$ , is given by

$$B_f \equiv \Gamma_f/\Gamma. \quad (\text{B.31c})$$

The lifetimes and branching ratios for ‘long-lived’ particles with  $\tau \geq 10^{-16}$  s can be measured directly. However, some very heavy states, and particles that decay by strong interactions, have lifetimes that are much shorter than this. In such cases, the decay widths are inferred from the observation of peaks in cross-sections and mass

distributions, as we saw in Chapter 3, Section 3.5. In Section B.5.1 we give a derivation of the Breit–Wigner formula, which was used in Section 3.5 to describe the shape of such peaks, followed in Section B.5.2 by a derivation of the general form of the Breit–Wigner formula for resonance formation and decay, which was used in the important case of  $Z^0$  production in electron–positron colliding beam experiments in Section 10.1.4.

### \*B.5.1 Decay distributions

We begin by discussing how unstable states are described in quantum mechanics. If we denote the wavefunction of the unstable state by  $\psi_0$  and the wavefunctions of the possible final states by  $\psi_n (n \geq 1)$ , then we can always choose these to be orthonormal, that is

$$\begin{aligned} \int \psi_m^* \psi_n dx &= 1 \quad \text{if } m = n \\ &= 0 \quad \text{if } m \neq n \end{aligned}$$

where  $x$  denotes all relevant variables. The wavefunction  $\psi_0$  describes the system at  $t = 0$ . At a later time  $t$  the wavefunction  $\Psi$  will be a linear superposition of the  $\psi_n$ , which is conveniently written in the form

$$\Psi(\mathbf{r}, t) = \sum_{n=0}^{\infty} a_n(t) e^{-iE_n t} \psi_n(\mathbf{r}), \quad (\text{B.32})$$

where

$$a_0(0) = 1, \quad a_n(0) = 0 \quad (n \geq 1) \quad (\text{B.33})$$

and

$$E_n \equiv H_{nn} = \int \psi_n^* H \psi_n dx.$$

The exponential factors in (B.32) are inserted purely for later convenience, since the coefficients  $a_n(t)$  remain to be determined. In addition, the energy of the initial state  $E_0 \pm \Delta E$  has an uncertainty  $\Delta E \approx 1/\tau$  by the energy–time uncertainty principle, so that  $E_n \neq E_0$  in general. Because of this, there are always many quantum states  $\psi_n$  contributing to any decay channel  $i$  and the individual coefficients  $a_n(t)$  ( $n \neq 0$ ) remain very small even when  $t \gg \tau$ , the lifetime of the decaying particle.

The time dependence of the coefficients  $a_n(t)$  is determined by substituting (B.32) into the Schrödinger equation

$$i \frac{\partial \Psi}{\partial t} = H\Psi$$

to give

$$\sum_m \left\{ i \frac{da_m}{dt} e^{-iE_m t} \psi_m + E_m a_m e^{-iE_m t} \psi_m \right\} = \sum_m a_m e^{-iE_m t} H \psi_m.$$

Multiplying this equation by  $\psi_n^*$ , and integrating using the orthonormality relations above, projects out the term  $da_n/dt$ :

$$i \frac{da_n}{dt} = \sum_{m \neq n} H_{nm} e^{-i(E_m - E_n)t} a_m. \quad (\text{B.34})$$

So far we have made no approximations, and the rigorous way forward would be to solve the set of differential equations (B.34) with the initial conditions (B.33) and an appropriate Hamiltonian, as originally done by Breit and Wigner. Here, following Fermi,<sup>3</sup> we will obtain the same results in an approximate way by working to first order in  $H_{nm}$ , which are assumed to be small when  $n \neq m$ . In this case we have, for  $n \geq 1$ ,

$$i \frac{da_n}{dt} = H_{n0} e^{-i(E_0 - E_n)t} a_0, \quad (\text{B.35})$$

since  $a_m$  is of order  $H_{0m}$  for  $m \neq 0$ . Also, we specialise to the rest frame of the decaying particle so that  $E_0 = M$ , where  $M$  is its rest mass, and we assume  $a_0(t) = e^{-\Gamma t/2}$ , consistent with the exponential decay law

$$|a_0(t)|^2 = e^{-\Gamma t} = e^{-t/\tau},$$

where  $\tau$  is the lifetime at rest. Substituting this form for  $a_0(t)$  into (B.35) and solving with the boundary conditions (B.33) then gives

$$ia_n = -iH_{n0} \left\{ \frac{\exp \{-i[(M - E_n) - i\Gamma/2]t\} - 1}{(E_n - M) + i\Gamma/2} \right\}.$$

For large values of  $t \gg 1/\Gamma$ , the exponential term tends to zero, so that the probability  $P_n$  of finding a final quantum state  $n$  is given by

$$P_n = |a_n(\infty)|^2 = \frac{|H_{n0}|^2}{(E_n - M)^2 + \Gamma^2/4}.$$

This is conveniently rewritten in the form

$$P_n = \frac{2\pi}{\Gamma} |H_{n0}|^2 P(E_n - M), \quad (\text{B.36})$$

where the distribution function

$$P(E - M) = \frac{\Gamma/2\pi}{(E - M)^2 + \Gamma^2/4} \quad (\text{B.37a})$$

peaks sharply at  $E = M$  with width  $\Gamma$ , and is normalised to unity:

$$\int_{-\infty}^{\infty} P(E - M) dE = 1. \quad (\text{B.37b})$$

---

<sup>3</sup> See pp. 152–157 of Fermi (1950).

Equation (B.36) is the probability for decay to a specific quantum state  $n$ . In general, a particle can decay to several channels  $f$ , and to obtain the probability for any given channel we must sum over the many quantum states  $n$  contained in that channel. If we ignore spin, the probability  $P_f(E) dE$  for decay to a particular channel  $f$  with energy in the interval  $E$  to  $E + dE$  is therefore

$$P_f(E) dE = \frac{2\pi}{\Gamma} |H_{f0}|^2 P(E - M) \rho_f(E) dE,$$

where the *density of states*  $\rho_f(E)$  is the number of quantum states with energy between  $E$  and  $E + dE$  in channel  $f$ . If  $\Gamma$  is small,  $\rho_f(E)$  is sharply peaked at  $E = M$ , because of the factor  $P(E - M)$ , and it is reasonable to approximate  $H_{f0}$  and  $\rho_f(E)$  by their values at  $E = M$ . We then obtain

$$P_f(E) dE = \frac{2\pi}{\Gamma} |H_{f0}|^2 P(E - M) \rho_f(M) dE$$

for the energy distribution of the decays of a particle of mass  $M$  into a particular channel  $f$ . The total probability for decay to channel  $f$  is equal to the branching ratio (B.31c) and can be evaluated by integrating (B.38) over all energies using (B.37b). In this way one obtains

$$\Gamma_f = 2\pi |H_{f0}|^2 \rho_f(M), \quad (\text{B.39a})$$

so that (B.38) becomes

$$P_f(E) = \frac{1}{2\pi} \frac{\Gamma_f}{(E - M)^2 + \Gamma^2/4} \quad (\text{B.40})$$

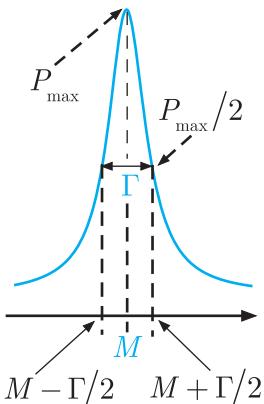
on substituting (B.37a). This is the Breit–Wigner formula for the decay distribution of a spin-0 particle. It also holds in the general case where the initial particle is unpolarised, but has spin  $s_i$ , provided the decay width (B.39a) is replaced by

$$\Gamma_f = \frac{2\pi}{2s_i + 1} \overline{|H_{f0}|^2} \rho_f(M), \quad (\text{B.39b})$$

where  $\overline{|H_{f0}|^2}$  is the spin average of the squared matrix element and the density of states  $\rho_f(M)$  includes a sum over all final particle spin states.

The Breit–Wigner formula (B.40) represents the final result of this subsection, and holds in the rest frame of the decaying particle. Since this is identical with the centre-of-mass frame of the decay products, their invariant mass  $W = E$ , and (B.40) can be written in the form

$$P_f(W) = \frac{1}{2\pi} \frac{\Gamma_f}{(W - M)^2 + \Gamma^2/4}, \quad (\text{B.41})$$



**Figure B.3** The Breit–Wigner formula (B.41).

which holds in any frame of reference. This is just the Breit–Wigner formula (3.30) used to extract resonance masses and widths in Section 3.5. Its characteristic shape is plotted in Figure B.3, where  $P_{\max} \equiv P_f(W = M)$ .

### \*B.5.2 Resonant cross-sections

We now turn to the case where an isolated unstable particle (or resonance) is formed in a scattering process and subsequently decays to a final channel  $f$ , as shown in Figure B.4. For steady incoming beams, the resonant state will be formed at a constant rate. However, this will eventually be balanced by its subsequent decay, leading to a constant average number of resonances present, independent of time. The rate of production of the final channel  $f$  is then determined by the partial decay rate  $\Gamma_f$  of the resonance to that channel.

To solve this problem exactly, we should solve (B.34) with the boundary conditions

$$a_1(0) = 1, \quad a_n(0) = 0 \quad (n \neq 1) \quad (\text{B.42})$$

at  $t = 0$ , where  $n = 1$  refers to the initial state and  $n = 0$  to the resonant state, as usual. In practice, we shall again obtain the results in an approximate way, following Fermi. To first order in  $H_{nm}$ , (B.34) gives

$$i \frac{da_0(t)}{dt} = H_{01} e^{-i(E_1 - E_0)t}, \quad (\text{B.43})$$

since  $a_n = O(H_{1n})$  for  $n \neq 1$ , and

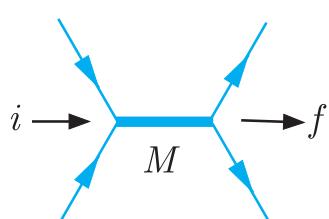
$$|a_1|^2 = 1 - \sum_{n \neq 1} |a_n|^2.$$

Equation (B.43) describes the ‘feeding’ of the resonance state from the initial state. However, it ignores the decays of the resonance, which involve an additional power of  $H_{0n}$ , but can involve a very large number of quantum states  $n$ . These can be taken into account by modifying (B.43) to

$$i \frac{da_0(t)}{dt} = H_{01} e^{-i(E_1 - E_0)t} - i \left( \frac{\Gamma}{2} \right) a_0(t), \quad (\text{B.44})$$

where the additional term ensures that  $|a_0(t)|^2$  would decay exponentially if the first term were switched off by, for example, stopping the beam. The solution of (B.44) with  $a_0 = 0$  at  $t = 0$  may be found by rewriting it as

$$i \frac{d}{dt} \left( a_0 e^{\Gamma t/2} \right) = H_{01} e^{-i(E_1 - E_0 + i\Gamma/2)t}$$



**Figure B.4** Formation and decay of an unstable particle of mass  $M$  in a two-body scattering process, where the initial channel  $i$  and the final channel  $f$  may, or may not, be the same.

and integrating over time. For times  $t \gg 1/\Gamma$ , this gives

$$a_0(t) = \frac{H_{01}e^{-i(E_1-E_0)t}}{E_1 - E_0 + i\Gamma/2}$$

and the partial decay rate into the final channel  $f$  is therefore

$$W_f = \Gamma_f |a_0(t)|^2 = \frac{|H_{01}|^2 \Gamma_f}{(E - E_0)^2 + \Gamma^2/4}, \quad (\text{B.45})$$

where we have set the initial energy  $E_1$  identically equal to  $E$  as usual.

It remains to relate (B.45) to the centre-of-mass cross-section. From (B.17) we have

$$W_f = \frac{v_i}{V} \sigma_{if},$$

where  $v_i$  is the relative velocity of the particles in the initial state.

In addition, from (B.39a) and (B.26) we have

$$\Gamma_i = 2\pi |H_{10}|^2 \rho_i(E_0) = \frac{V q_i^2}{\pi v_i} |H_{10}|^2,$$

where  $q_i$  is the initial momentum and the density of states (B.26) has been integrated over angles to accord with the density of states used in (B.39a). Substituting these equations into (B.45) gives

$$\sigma_{if} = \frac{\pi}{q_i^2} \frac{\Gamma_i \Gamma_f}{(E - E_0)^2 + \Gamma^2/4}, \quad (\text{B.46a})$$

where  $E_0 = M$ , the resonance mass, if we work in the centre-of-mass frame. This is the *Breit–Wigner resonance formula*. In the general case where the initial particles have spins  $s_1$  and  $s_2$  and the unstable particle has spin  $j$ , (B.46a) is generalised to

$$\sigma_{if} = \frac{\pi}{q_i^2} \frac{2j+1}{(2s_1+1)(2s_2+1)} \frac{\Gamma_i \Gamma_f}{(E - E_0)^2 + \Gamma^2/4}. \quad (\text{B.46b})$$

It is often written in slightly different forms, which are all compatible in the region of the peak if  $\Gamma \ll E_0$ . In particular, in order to express (B.46b) as a function of  $E^2$ , we can make the approximation

$$\frac{1}{(E - E_0)^2 + \Gamma^2/4} = \frac{4E_0^2}{(E^2 - E_0^2)^2 + E_0^2 \Gamma^2}$$

in the vicinity of the resonance, when it becomes

$$\sigma_{if} = \frac{\pi}{q_i^2} \frac{2j+1}{(2s_1+1)(2s_2+1)} \frac{4E_0^2 \Gamma_i \Gamma_f}{(E^2 - E_0^2)^2 + E_0^2 \Gamma^2}. \quad (\text{B.46c})$$

Equations (B.46b) and (B.46c) are just Equations (3.34b) and (3.34c) given in Section 3.5, except that in Section 3.5 we have used  $W$  and  $M$  rather than  $E$  and  $E_0$  to denote the centre-of-mass energy and the resonance mass, respectively.

Finally, in the common situation where  $\Gamma \ll E_0$ , the *narrow width approximation* is often used to simplify calculations. There are different forms of this, but one is to make the replacement

$$\frac{1}{(E - E_0)^2 + \Gamma^2/4} \rightarrow \frac{2\pi}{\Gamma} \delta(E - E_0) \quad (\text{B.47a})$$

in (B.46a) and (B.46b), where  $\delta$  is the Dirac delta function<sup>4</sup> and the constant  $2\pi/\Gamma$  multiplying it is chosen to ensure that both sides of (B.47a) give the same result, to a very good approximation, when integrated over all energies (cf. Problem B.3). Alternatively, if we instead express the Breit–Wigner formula as a function of  $E^2$  rather than  $E$ , (B.47a) becomes

$$\frac{4E_0^2}{(E^2 - E_0^2)^2 + E_0^2\Gamma^2} \rightarrow \frac{4\pi E_0}{\Gamma} \delta(E^2 - E_0^2), \quad (\text{B.47b})$$

where we have used the standard result

$$|a| \delta[a(x - a)] = \delta(x - a) \quad (\text{B.48})$$

for any constant  $a$  to write  $\delta(E - E_0)$  in terms of  $E^2$ . In all these formulas,  $E_0$  is the resonance mass and is frequently denoted by  $E_0 = M$ .

## Problems B

- B.1** A liquid hydrogen target of volume  $10^{-4} \text{ m}^3$  and density  $71 \text{ kg/m}^3$  is exposed to a wide uniform monoenergetic beam of negative pions of flux  $10^7 \text{ particles}/(\text{m}^2 \text{ s})$  and the reaction  $\pi^- p \rightarrow K^0 \Lambda$  is observed. If the cross-section for this reaction is  $0.4 \text{ mb}$ , what is the rate of production of  $\Lambda$  particles?
- B.2** A high-energy beam of neutrons of intensity  $10^6 \text{ /s}$  traverses a target of  $^{238}\text{U}$  in the form of a thin foil whose density per unit area is  $10^{-1} \text{ kg/m}^2$ . If the elastic and inelastic cross-sections are  $1.4\text{b}$  and  $2.0\text{b}$ , respectively, calculate: (a) the attenuation of the beam; (b) the

---

<sup>4</sup> The delta function is defined by the condition

$$\int f(x) \delta(x - a) dx \equiv f(a),$$

where  $a$  is a real constant, implying  $\delta(x - a) = 0$  if  $x \neq a$ , and

$$\int_b^c \delta(x - a) dx = 1, \quad b < a < c.$$

It is discussed in, for example, Section 13.3.2 of Martin and Shaw (2015), where derivations of its important properties, including (B.48), may be found.

rate of inelastic reactions; and (c) the flux of elastically scattered neutrons 5 m from the target, averaged over all scattering angles. (In practice the neutrons are emitted at small angles relative to the beam direction.)

- B.3** (a) Verify by explicit integration that the integral

$$I = \int_0^\infty \frac{dE}{(E - E_0)^2 + \Gamma^2/4}$$

is left unchanged by the substitution (B.47a) in the limit that  $\Gamma/E_0 \rightarrow 0$  at fixed  $E_0$ .

- (b) Show that (B.47a) and (B.47b) are compatible close to the resonance if  $\Gamma \ll E_0$ .



---

## The isospin formalism

---

The discussion of isospin in Section 6.1 was based on the assumed analogy with ordinary angular momentum and spin. In this appendix we give a more formal treatment in order to justify this analogy and to derive a number of other results that were cited in the text. Our approach will be based on the fact that many properties of angular momentum can be derived from the commutation relations

$$[\hat{J}_x, \hat{J}_y] = i\hat{J}_z, \quad [\hat{J}_y, \hat{J}_z] = i\hat{J}_x, \quad [\hat{J}_z, \hat{J}_x] = i\hat{J}_y, \quad (\text{C.1})$$

where  $\hat{J}_{x,y,z}$  are the components of any angular momentum operator (e.g. total or orbital angular momentum, or spin).<sup>1</sup> An equivalent, and in some ways more convenient, set of commutation relations that follows from (C.1) is

$$[\hat{J}_+, \hat{J}_-] = 2\hat{J}_3, \quad [\hat{J}_3, \hat{J}_\pm] = \pm\hat{J}_\pm, \quad (\text{C.2})$$

where  $\hat{J}_\pm$  and  $\hat{J}_3$  are the linear combinations

$$\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y, \quad \hat{J}_3 = \hat{J}_z. \quad (\text{C.3})$$

Hence, if we can identify the appropriate isospin operators

$$\hat{I}_+, \quad \hat{I}_-, \quad \text{and} \quad \hat{I}_3 \quad (\text{C.4})$$

that satisfy the commutation relations

$$[\hat{I}_+, \hat{I}_-] = 2\hat{I}_3, \quad [\hat{I}_3, \hat{I}_\pm] = \pm\hat{I}_\pm, \quad (\text{C.5})$$

---

<sup>1</sup>This approach to angular momentum is summarised in, for example, Sections 5.2, 5.4 and 5.8 of Mandl (1992).

then the same properties will follow for isospin. These include the multiplet structure of (6.7) and (6.8), and the rules for the addition of isospin quoted in Section 6.1.1.

## C.1 Isospin operators

The isospin operators  $\hat{I}_\pm$ , and  $\hat{I}_3$  are defined by specifying their actions on a complete set of quantum states. This is simplest for the operator  $\hat{I}_3$ , since in Section 6.1.1 we showed that the corresponding quantum number

$$I_3 \equiv Q - Y/2 \quad (\text{C.6})$$

was equivalent to

$$I_3 = \frac{1}{2} (N_u - N_d), \quad (\text{C.7})$$

where  $N_u$  and  $N_d$  are the quark numbers defined in Equation (3.3). Hence for a  $u$  quark state we have, for example,

$$\hat{I}_3 |u, \Psi\rangle = \frac{1}{2} |u, \Psi\rangle$$

for an arbitrary space and spin wavefunction  $\Psi$ . The definition of the operators  $\hat{I}_+$  and  $\hat{I}_-$  is less obvious, but can be inferred from the analogy between isospin and ordinary spin.<sup>2</sup> Here we shall just write them down and justify them afterwards by showing that they do indeed lead to the desired commutation relations (C.5). Specifically, we define

$$\hat{I}_+ u = 0, \quad \hat{I}_- u = d, \quad \hat{I}_3 u = \frac{1}{2} u, \quad (\text{C.8a})$$

$$\hat{I}_+ d = u, \quad \hat{I}_- d = 0, \quad \hat{I}_3 d = -\frac{1}{2} d, \quad (\text{C.8b})$$

$$\hat{I}_+ \bar{d} = 0, \quad \hat{I}_- \bar{d} = -\bar{u}, \quad \hat{I}_3 \bar{d} = \frac{1}{2} \bar{d}, \quad (\text{C.8c})$$

and

$$\hat{I}_+ \bar{u} = -\bar{d}, \quad \hat{I}_- \bar{u} = 0, \quad \hat{I}_3 \bar{u} = -\frac{1}{2} \bar{u}, \quad (\text{C.8d})$$

together with

$$\hat{I}_+ r = \hat{I}_- r = \hat{I}_3 r = 0 \quad (\text{C.8e})$$

for all other quarks and antiquarks, collectively denoted  $r$ . In all these relations we have suppressed the space and spin variables, which are assumed to be left unchanged. For example,  $\hat{I}_- u = d$  is really shorthand for

$$\hat{I}_- |u, \Psi\rangle = |d, \Psi\rangle,$$

---

<sup>2</sup> Specifically, the quark states ( $u, d$ ) are assumed to be analogous to the spin-up and spin-down states  $\alpha$  and  $\beta$  of an electron, and the isospin operators  $\hat{I}_3, \hat{I}_\pm$  are assumed to be analogous to the spin operators  $\hat{S}_3$  and  $\hat{S}_\pm = \hat{S}_x \pm i\hat{S}_y$ . The result  $\hat{I}_+ d = u$ , for example, then follows from the analogous relation  $\hat{S}_+ \beta = \alpha$ , which can be easily derived using (5.21) and (5.22).

which tells us that the operator  $\hat{I}_-$  replaces the  $u$  quark by a  $d$  quark with the same combined space and spin wavefunction  $\Psi$ .

To complete the definition of the isospin operators we need to generalise (C.8) to an arbitrary state  $|A\rangle$  composed of quarks and antiquarks. Such a state can be written as a linear combination

$$|A\rangle = \sum_i \beta_i |A_i\rangle \quad (\text{C.9})$$

of product states

$$|A_i\rangle = a, ab, abc, \dots, \quad (\text{C.10})$$

where  $a, b, c, \dots$  can be any quark or antiquark  $u, \bar{u}, d, \bar{d}, \dots$ , and  $\beta_i$  are arbitrary constants. The action of the operators on the product states  $|A_i\rangle$  is fixed by the distributive laws:

$$\hat{I}_\alpha(ab) = (\hat{I}_\alpha a)b + a(\hat{I}_\alpha b), \quad (\text{C.11a})$$

$$\hat{I}_\alpha(abc) = (\hat{I}_\alpha a)bc + a(\hat{I}_\alpha b)c + ab(\hat{I}_\alpha c) \quad (\text{C.11b})$$

and so on, where  $\alpha = +, -$  or  $3$ . The definition is finally completed by requiring that the  $\hat{I}_\alpha$  are linear operators, so that

$$\hat{I}_\alpha |A\rangle = \hat{I}_\alpha \sum_i \beta_i |A_i\rangle = \sum_i \beta_i \hat{I}_\alpha |A_i\rangle, \quad (\text{C.12})$$

where  $|A\rangle$  is the arbitrary state (C.9).

The above relations completely define the isospin operators, but we have not shown that they satisfy the crucial commutation relations (C.5). To do this, we note that if, for example, the relation

$$[\hat{I}_+, \hat{I}_-] a = 2\hat{I}_3 a \quad (\text{C.13})$$

holds for an arbitrary single-particle state  $a = u, \bar{u}, d, \dots$ , then it follows from (C.11) and (C.12) that

$$[\hat{I}_+, \hat{I}_-] |A\rangle = 2\hat{I}_3 |A\rangle$$

holds for an arbitrary state (C.9). Hence, to verify the first of the commutators (C.5), it is only necessary to verify (C.13). For the case  $a \equiv u$ , we have

$$[\hat{I}_+, \hat{I}_-] u = \hat{I}_+ \hat{I}_- u - \hat{I}_- \hat{I}_+ u = \hat{I}_+ d = u$$

and

$$2\hat{I}_3 u = u$$

by (C.8a) and (C.8b). Hence (C.13) is satisfied for  $a \equiv u$ , and it is equally easy to verify it for the other possible cases. This establishes the first of the commutation relations (C.5) and the others follow in a similar manner.

## C.2 Isospin states

The properties of isospin states can be derived using the commutation relations (C.5) in exactly the same way that the properties of angular momentum states can be derived from the corresponding commutation relations (C.2). In particular, it can be shown<sup>3</sup> that there exist families of  $2I + 1$  states  $|I, I_3\rangle$  with

$$I_3 = I, I - 1, \dots, -I,$$

which are simultaneous eigenstates of the operators  $I_3$  and

$$\hat{I}^2 \equiv \frac{1}{2} [\hat{I}_+ \hat{I}_- + \hat{I}_- \hat{I}_+] + \hat{I}_3^2, \quad (\text{C.14})$$

with eigenvalues given by

$$\hat{I}^2 |I, I_3\rangle = I(I + 1) |I, I_3\rangle \quad (\text{C.15a})$$

and

$$\hat{I}_3 |I, I_3\rangle = I_3 |I, I_3\rangle. \quad (\text{C.15b})$$

The action of the operators  $\hat{I}_{\pm}$  on these states is also fixed from the commutation relations and is given by

$$\hat{I}_{\pm} |I, I_3\rangle = C_{\pm}(I, I_3) |I, I_3 \pm 1\rangle, \quad (\text{C.16a})$$

where the coefficients are

$$C_{\pm}(I, I_3) = +[(I \mp I_3)(I \pm I_3 + 1)]^{1/2}. \quad (\text{C.16b})$$

As the action of  $\hat{I}_{\pm}$  shifts the values of  $I_3$  by  $\pm 1$ , the operators  $\hat{I}_{\pm}$  are known as *shift* or *ladder operators*. These equations are extremely useful in applications, and in particular they can be used to show (cf. Problem C.2) that the  $(2I + 1)$  states  $|I, I_3\rangle$  have the same energy if isospin is conserved, that is if

$$[H, \hat{I}_{\pm}] = [H, \hat{I}_3] = 0. \quad (\text{C.17})$$

## C.3 Isospin multiplets

The practical applications of isospin symmetry rest on the assignment of the observed hadrons to isospin multiplets. Here we shall discuss this a little more fully than in the main text, with a view, in particular, to establishing the quark content of the neutral pion

---

<sup>3</sup> The proofs of the following results are identical with those of the corresponding results for angular momentum given in, for example, Sections 5.2 and 5.8 of Mandl (1992).

(6.23). In doing so we will write the members of isospin multiplets inside ‘column vectors’ in the order

$$\begin{pmatrix} |a; I, I_3 = I\rangle \\ |a; I, I_3 = I - 1\rangle \\ \vdots \\ |a; I, I_3 = -I\rangle \end{pmatrix}, \quad (\text{C.18})$$

where  $a$  is some convenient label used to identify a particular multiplet. In this notation, the basic quark and antiquark isodoublets are defined by

$$\begin{pmatrix} u \\ d \end{pmatrix}, \quad \begin{pmatrix} \bar{d} \\ -\bar{u} \end{pmatrix}, \quad (\text{C.19})$$

which tells us that the quarks and antiquarks are assigned to the isospin states  $|q, I, I_3\rangle$  and  $|\bar{q}, I, I_3\rangle$  according to

$$u = \left| q; \frac{1}{2}, \frac{1}{2} \right\rangle, \quad d = \left| q; \frac{1}{2}, -\frac{1}{2} \right\rangle, \quad (\text{C.20})$$

and

$$\bar{d} = \left| \bar{q}; \frac{1}{2}, \frac{1}{2} \right\rangle, \quad \bar{u} = - \left| \bar{q}; \frac{1}{2}, -\frac{1}{2} \right\rangle, \quad (\text{C.21})$$

respectively. These assignments are, of course, chosen to be consistent with both the definitions (C.8) and the fundamental results (C.15) and (C.16) (cf. Problem C.3). The remaining quarks and antiquarks

$$r = s, \bar{s}, c, \bar{c}, \dots, \quad (\text{C.22})$$

are assigned to isosinglets, with  $I = I_3 = 0$ , as in Table 6.2.

### C.3.1 Hadron states

The isospin assignments of the hadron states  $q\bar{q}$  and  $qqq$  can be deduced from the basic quark assignments using the rules for the addition of isospin, which are identical to those for angular momentum. For example, since strange quarks and antiquarks are isosinglets with  $I = I_3 = 0$ , adding them to any state does not change its isospin properties. Hence from (C.19) we can infer that the meson states

$$\begin{pmatrix} u\bar{s} \\ d\bar{s} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} s\bar{d} \\ -s\bar{u} \end{pmatrix}$$

form isodoublets. With the usual quark assignments (3.16), this implies that the kaons in turn form isodoublets

$$\begin{pmatrix} K^+ \\ K^0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \bar{K}^0 \\ -K^- \end{pmatrix}. \quad (\text{C.23})$$

A less trivial example concerns mesons which are bound states of a  $u$  or  $d$  quark with a  $\bar{u}$  or  $\bar{d}$  antiquark. Since these quarks and antiquarks belong to isodoublet states (C.19) with  $I = 1/2$ , we expect

to find three isotriplet states with  $I = 1$  and  $I_3 = 1, 0$  and  $-1$ , and an isosinglet state with  $I = 0$  and  $I_3 = 0$ . The problem is to identify which isospin states  $|I, I_3\rangle$  correspond to which combinations  $q\bar{q}$ . In such problems, it is often convenient to start with the state with the largest possible values of  $I$  and  $I_3$ . This has  $I = I_3 = 1$  in this case, and since only the  $u$  quark and  $\bar{d}$  antiquark have  $I_3 = \frac{1}{2}$ , the corresponding  $q\bar{q}$  combination must be  $u\bar{d}$ . We therefore start from the identification

$$|I = 1, I_3 = 1\rangle = -u\bar{d}, \quad (\text{C.24a})$$

where the sign is arbitrary and is chosen to agree with standard conventions. The other  $I = 1$  states are obtained by using the operator  $\hat{I}_-$ , which decreases  $I_3$  by one unit while leaving  $I$  unchanged. Explicitly, we have

$$\hat{I}_- |I = 1, I_3 = 1\rangle = \sqrt{2} |I = 1, I_3 = 0\rangle$$

by (C.16) and

$$\hat{I}_- (-u\bar{d}) = -(\hat{I}_- u)\bar{d} - u(\hat{I}_- \bar{d}) = -d\bar{d} + u\bar{u}$$

by (C.10) and (C.8). Comparing these equations gives

$$|I = 1, I_3 = 0\rangle = \frac{1}{\sqrt{2}} (u\bar{u} - d\bar{d}) \quad (\text{C.24b})$$

and acting on both sides with  $\hat{I}$  again leads to

$$|I = 1, I_3 = -1\rangle = d\bar{u}. \quad (\text{C.24c})$$

The only remaining state is the isosinglet, which must be of the form

$$|I = 0, I_3 = 0\rangle = \alpha u\bar{u} + \beta d\bar{d} \quad (\text{C.25})$$

in order to ensure that  $I_3 = 0$ . The normalisation condition requires

$$\alpha^2 + \beta^2 = 1,$$

so that one more condition is required to determine (C.25) up to an arbitrary overall phase factor. From (C.16) we have

$$\hat{I}_- |I = 0, I_3 = 0\rangle = 0 = \hat{I}_+ |I = 0, I_3 = 0\rangle$$

and either of these leads to

$$|I = 0, I_3 = 0\rangle = \frac{1}{\sqrt{2}} (u\bar{u} + d\bar{d}), \quad (\text{C.26})$$

where the overall phase has again been chosen to coincide with the most popular convention. Alternatively, the same result can be obtained by requiring that the new state is orthogonal to (C.24b), since the latter is already known. Equations (C.24b) and (C.26) were used in Chapter 6 (where they appeared as (6.23) and (6.24a) respectively) to determine the quark content of the neutral pion.

Specifically, on identifying pions with the states (C.24), we obtain an isotriplet

$$\begin{pmatrix} -\pi^+ \\ \pi^0 \\ \pi^- \end{pmatrix}, \quad (\text{C.27})$$

with the quark assignments

$$\pi^+ = u\bar{d}, \quad \pi^0 = \frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d}), \quad \pi^- = d\bar{u}. \quad (\text{C.28})$$

Similar arguments can be applied to  $qqq$  states, leading to baryon multiplets like

$$(\Lambda), \quad \begin{pmatrix} p \\ n \end{pmatrix}, \quad \begin{pmatrix} \Sigma^+ \\ \Sigma^0 \\ \Sigma^- \end{pmatrix}, \quad \begin{pmatrix} \Delta^{++} \\ \Delta^+ \\ \Delta^0 \\ \Delta^- \end{pmatrix}, \quad (\text{C.29})$$

corresponding to  $\Lambda$  (1116),  $N$  (940),  $\Sigma$  (1193) and  $\Delta$  (1232) states with  $I = 0, 1/2, 1, 3/2$ , respectively. We shall not pursue this because the quark content of baryons follows unambiguously from their quantum numbers  $B$ ,  $Q$ ,  $S$ ,  $C$ ,  $B$  and  $T$  alone, as we saw in Section 6.2.2, and because there are no minus signs in baryon multiplets like those that occurred in the meson multiplets (C.23) and (C.27), because for baryons there are no antiquarks involved.

## C.4 Branching ratios

Up to now we have concentrated on the use of isospin as a symmetry to classify hadron states. It also leads to many useful relations between the rates of reactions involving members of the same isospin multiplets. Such applications are especially useful in strong interactions, where isospin is conserved, and in particular they lead to predictions for branching ratios for resonance decays.

To illustrate this, we start by calculating the branching ratios of the  $\Delta^+$ (1232) resonance, which has two decay modes

$$\Delta^+ \rightarrow p\pi^0, \quad n\pi^+ \quad (\text{C.30})$$

From (C.29) we see that the  $\Delta^+$  has  $I = 3/2$  and  $I_3 = 1/2$ , and so by isospin conservation must decay to a pion–nucleon state with the same quantum numbers, which we denote

$$\left| N\pi; \frac{3}{2}, \frac{1}{2} \right\rangle. \quad (\text{C.31})$$

To predict the branching ratios we must relate this state to specific charge states like  $p\pi^0$  and  $n\pi^+$ . Since nucleons have  $I = 1/2$  and pions have  $I = 1$ , possible  $N\pi$  states have  $I = 3/2$  or  $1/2$ . As usual, the easiest state to identify is the one with the maximum values

$I = I_3 = 3/2$ , since the only possible combination which gives  $I = 3/2$  is

$$\left| N\pi; \frac{3}{2}, \frac{3}{2} \right\rangle = \left| N; \frac{1}{2}, \frac{1}{2} \right\rangle |\pi; 1, 1\rangle. \quad (\text{C.32})$$

The other possible states can be deduced from this by using the shift operators and orthogonality. For example, to obtain (C.31) we act on both sides of (C.32) with  $\hat{I}_-$  using (C.16). This gives

$$\hat{I}_- \left| N\pi; \frac{3}{2}, \frac{3}{2} \right\rangle = \sqrt{3} \left| N\pi; \frac{3}{2}, \frac{1}{2} \right\rangle$$

for the left-hand side and

$$\hat{I}_- \left[ \left| N; \frac{1}{2}, \frac{1}{2} \right\rangle |\pi; 1, 1\rangle \right] = \left| N; \frac{1}{2}, -\frac{1}{2} \right\rangle |\pi; 1, 1\rangle + \sqrt{2} \left| N; \frac{1}{2}, \frac{1}{2} \right\rangle |\pi; 1, 0\rangle$$

for the right-hand side. Comparing these, we obtain the desired results

$$\begin{aligned} \left| N; \frac{3}{2}, \frac{1}{2} \right\rangle &= \frac{1}{\sqrt{3}} \left| N; \frac{1}{2}, -\frac{1}{2} \right\rangle |\pi; 1, 1\rangle + \sqrt{\frac{2}{3}} \left| N; \frac{1}{2}, \frac{1}{2} \right\rangle |\pi; 1, 0\rangle \\ &= -\frac{1}{\sqrt{3}} |n\pi^+\rangle + \sqrt{\frac{2}{3}} |p\pi^0\rangle, \end{aligned} \quad (\text{C.33})$$

where in the second equation we have identified the pion and nucleon charge states. Thus we see that isospin conservation predicts that the  $\Delta^+$  has a probability of  $1/3$  to decay into  $n\pi^+$  and a probability of  $2/3$  to decay into  $p\pi^0$ . This is in excellent agreement with experiment.

We next turn to the  $N^+$  resonances, which are excited states of the proton with the isospin quantum numbers  $I = 1/2, I_3 = 1/2$ , and can decay to pion–nucleon final states via

$$N^+ \rightarrow p\pi^0, n\pi^+. \quad (\text{C.34})$$

In addition, they can decay to more complicated final states allowed by conservation laws, like  $p\pi^+\pi^-$ ,  $n\pi^+\pi^0$ , etc. We will not consider these here, but just use isospin conservation to predict the relative decay rates for the two reactions (C.34).

To do this, we must construct a pion–nucleon final state with the same quantum numbers  $I = 1/2, I_3 = 1/2$  as the decaying  $N^+$  state. This must be of the general form

$$\left| N\pi; \frac{1}{2}, \frac{1}{2} \right\rangle = \alpha \left| N; \frac{1}{2}, -\frac{1}{2} \right\rangle |\pi; 1, 1\rangle + \beta \left| N; \frac{1}{2}, \frac{1}{2} \right\rangle |\pi; 1, 0\rangle,$$

since there are no other  $\pi N$  combinations with  $I_3 = 1/2$ . The coefficients  $\alpha$  and  $\beta$  are then determined, up to an arbitrary phase factor, by requiring that the state be normalised and orthogonal to the  $I = 3/2, I_3 = 1/2$  state (C.33). In this way we obtain

$$\left| N\pi; \frac{1}{2}, \frac{1}{2} \right\rangle = \sqrt{\frac{2}{3}} \left| N; \frac{1}{2}, -\frac{1}{2} \right\rangle |\pi; 1, 1\rangle - \sqrt{\frac{1}{3}} \left| N; \frac{1}{2}, \frac{1}{2} \right\rangle |\pi; 1, 0\rangle, \quad (\text{C.35})$$

where we have chosen the overall phase in accord with popular convention. On identifying the pion and nucleon charge states, (C.27) and (C.29), this becomes

$$\left|N\pi; \frac{1}{2}, \frac{1}{2}\right\rangle = -\sqrt{\frac{2}{3}} \left|n\pi^+\right\rangle - \sqrt{\frac{1}{3}} \left|p\pi^0\right\rangle,$$

implying

$$\Gamma(N^+ \rightarrow n\pi^+)/\Gamma(N^+ \rightarrow p\pi^0) = 2$$

for the ratio of the decay rates. This is in contrast to the ratio

$$\Gamma(\Delta^+ \rightarrow n\pi^+)/\Gamma(\Delta^+ \rightarrow p\pi^0) = 1/2,$$

which follows for  $\Delta^+$  resonances from (C.33) and is one of the means by which  $N^+$  and  $\Delta^+$  resonances can be experimentally distinguished.

## C.5 Spin states

Finally, we stress again that the isospin formalism described above has been constructed in exact analogy to the theory of angular momentum. Conversely, all the results that we have derived for isospin states  $|A : I, I_3\rangle$  apply equally well to ordinary spin states  $|A : S, S_3\rangle$ . In particular, Equations (C.31), (C.33) and (C.35) apply equally well to the combined spin states of a spin-1/2 particle  $A$  and a spin-1 particle  $B$ , when they become

$$\left|AB; \frac{3}{2}, \frac{3}{2}\right\rangle \equiv \left|AB; S = \frac{3}{2}, S_z = \frac{3}{2}\right\rangle = \left|A; \frac{1}{2}, \frac{1}{2}\right\rangle \left|B; 1, 1\right\rangle, \quad (\text{C.36a})$$

$$\left|AB; \frac{3}{2}, \frac{1}{2}\right\rangle = \sqrt{\frac{1}{3}} \left|A; \frac{1}{2}, -\frac{1}{2}\right\rangle \left|B; 1, 1\right\rangle + \sqrt{\frac{2}{3}} \left|A; \frac{1}{2}, \frac{1}{2}\right\rangle \left|B; 1, 0\right\rangle \quad (\text{C.36b})$$

and

$$\left|AB; \frac{1}{2}, \frac{1}{2}\right\rangle = \sqrt{\frac{2}{3}} \left|A; \frac{1}{2}, -\frac{1}{2}\right\rangle \left|B; 1, 1\right\rangle - \sqrt{\frac{1}{3}} \left|A; \frac{1}{2}, \frac{1}{2}\right\rangle \left|B; 1, 0\right\rangle. \quad (\text{C.36c})$$

This last equation was used in Section 6.2.4 to derive the magnetic moments of the  $J^P = \frac{1}{2}^+$  baryons  $aab$ , with the identification  $A = b$  and  $B = aa$ .

## Problems C

- C.1** Use the relations (C.1) and (C.3) to verify the commutation relations (C.2).
- C.2** Use the relations (C.16) and (C.17) to show that the  $2I + 1$  members of an isomultiplet  $|I, I_3\rangle$  all have the same energy if isospin is conserved.

- C.3** Use the definitions (C.8) to verify that (C.15) and (C.16) are satisfied for the quark and antiquark assignments (C.19).
- C.4** A resonance  $X^0(1520)$  decays via the strong interaction to the final states  $n\pi^0$  and  $p\pi^-$  with branching ratios of approximately 18% and 36%, respectively. What is its isospin?
- C.5** The  $\Lambda(1405)$  resonance decays by the strong interaction to  $\Sigma\pi$  final states with a branching ratio of 42%. What are the individual branching ratios for the charged modes  $\Sigma^+\pi^-$ ,  $\Sigma^0\pi^0$  and  $\Sigma^-\pi^+$  if the  $\Lambda(1405)$  is an isosinglet state?



---

## Gauge theories

---

Gauge theories have a particular type of symmetry, called *gauge invariance*. In this appendix we shall explain what is meant by gauge invariance and discuss its consequences. This is important because QED, QCD and the unified electroweak theory are all examples of gauge theories, with somewhat different forms of gauge invariance in each case. The simplest type of gauge invariance is that found in electromagnetism and it is on this case that we initially concentrate in Sections D.1 to D.5. In particular, in Section D.4, we shall use it to introduce the so-called *gauge principle* that makes gauge invariance the fundamental requirement from which the detailed properties of the interaction are deduced. This principle is applied to the strong interactions (QCD) and the unified electroweak theory, in Sections D.6 and D.7, respectively.

Another thing that will emerge in Section D.3 is that gauge invariance implies that the photon is strictly massless. While this is not a problem for electromagnetic interactions, it is a problem for the analogous case of the electroweak theory, where the  $W^\pm$  and  $Z^0$  are massive bosons. We have already mentioned briefly in Section 10.2.2 that non-zero masses can be generated by spontaneous symmetry breaking. This is illustrated in Section D.5 for the simple case of hypothetical massive photons.

The topics of this appendix are usually discussed in more advanced books, where advantage can be taken of students' greater knowledge to use explicitly Lorentz covariant Lagrangian field theory.<sup>1</sup> However, in keeping with the level of the rest of this book, we

---

<sup>1</sup> This approach may be found in many more advanced books, for example Chapters 12 and 13 of Mandl and Shaw (2010) or Chapter 14 of Halzen and Martin (1984).

have avoided four-vector notation and based our discussion directly on the relevant equations of motion.

## D.1 Electromagnetic interactions

The concept of gauge invariance has its origin in the study of electromagnetic interactions, and in order to introduce it we first consider the wave equation for a particle of mass  $m$  and charge  $q$  moving non-relativistically in an electromagnetic field.<sup>2</sup> In the case of an electrostatic field  $\phi(\mathbf{r})$  this is obtained by making the substitution

$$\mathbf{p} \rightarrow -i\boldsymbol{\nabla} \quad (\text{D.1})$$

in the classical Hamiltonian

$$H = \frac{\mathbf{p}^2}{2m} + q\phi(\mathbf{r})$$

to give the familiar Schrödinger equation

$$H\Psi(\mathbf{r}, t) = \left[ -\frac{1}{2m}\boldsymbol{\nabla}^2 + q\phi(\mathbf{r}) \right] \Psi(\mathbf{r}, t) = i\frac{\partial\Psi(\mathbf{r}, t)}{\partial t}. \quad (\text{D.2})$$

The wave equation for a general electromagnetic field is obtained in a similar way from the classical Hamiltonian

$$H = \frac{1}{2m}(\mathbf{p} - q\mathbf{A})^2 + q\phi, \quad (\text{D.3})$$

where  $\phi(\mathbf{r}, t)$  and  $\mathbf{A}(\mathbf{r}, t)$  are the scalar and vector potentials defined by<sup>3</sup>

$$\mathbf{E} = -\boldsymbol{\nabla}\phi - \frac{\partial\mathbf{A}}{\partial t}, \quad \mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}. \quad (\text{D.4})$$

The Hamiltonian (D.3) is justified by the fact that it can be shown (see Problem D.1) to lead to the familiar classical equation of motion

$$m\frac{d\mathbf{v}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \quad (\text{D.5})$$

<sup>2</sup> In this appendix, we work in units that coincide with the Heavyside–Lorentz or rationalised Gaussian electromagnetic units when factors of  $\hbar$  and  $c$  are restored. These units, in which  $\epsilon_0 = 1$ , are almost invariably used when this topic is discussed. To convert to the units used elsewhere in the book – which coincide with SI units when factors of  $\hbar$  and  $c$  are restored – make the replacements  $q \rightarrow q/\sqrt{\epsilon_0}$ ,  $\phi \rightarrow \sqrt{\epsilon_0}\phi$ ,  $\mathbf{A} \rightarrow \sqrt{\epsilon_0}\mathbf{A}$ ,  $\mathbf{E} \rightarrow \sqrt{\epsilon_0}\mathbf{E}$  and  $\mathbf{B} \rightarrow \sqrt{\epsilon_0}\mathbf{B}$ .

<sup>3</sup> See, for example, p. 229 of Grant and Phillips (1990).

for a charged particle moving with velocity  $\mathbf{v}$  in electric and magnetic fields given by (D.4). On making the substitution (D.1), the Hamiltonian (D.3) also leads to the desired wave equation

$$H\Psi = \left[ -\frac{1}{2m} (\nabla - i\mathbf{q}\mathbf{A})^2 + q\phi \right] \Psi = i\frac{\partial\Psi}{\partial t},$$

which is conveniently written in the form

$$i\left(\frac{\partial}{\partial t} + iq\phi\right)\Psi = -\frac{1}{2m} (\nabla - i\mathbf{q}\mathbf{A})^2 \Psi \quad (\text{D.6})$$

for what follows.

## D.2 Gauge transformations

The wave equation (D.6) determines the motion of a charged particle in an arbitrary electromagnetic field described by the potentials  $(\phi, \mathbf{A})$ . However, these potentials are not unique, since for any choice  $(\phi, \mathbf{A})$  there are many other choices  $(\phi', \mathbf{A}')$  that give rise to the same values for the electric and magnetic fields. In particular, one easily shows that substituting

$$\phi \rightarrow \phi' = \phi + \frac{\partial f}{\partial t}, \quad \mathbf{A} \rightarrow \mathbf{A}' = \mathbf{A} - \nabla f \quad (\text{D.7})$$

in (D.4) leaves the electric and magnetic fields unchanged, where  $f(\mathbf{r}, t)$  is an arbitrary scalar function.

The replacement of  $(\phi, \mathbf{A})$  by  $(\phi', \mathbf{A}')$  in accordance with (D.7) is called a *gauge transformation* and a theory whose physical predictions are unaltered by such a transformation is said to be *gauge invariant*. Because observable quantities depend only on the fields  $\mathbf{E}$  and  $\mathbf{B}$ , and not on the particular choice of potentials used to describe them, it is a fundamental requirement that any theory formulated in terms of these potentials should be gauge invariant. The wave equation (D.6) seems to contradict this requirement, since (D.6) and (D.7) together lead to a completely different wave equation,

$$i\left(\frac{\partial}{\partial t} + iq\phi' - iq\frac{\partial f}{\partial t}\right)\Psi = -\frac{1}{2m} (\nabla - i\mathbf{q}\mathbf{A}' - i\mathbf{q}\nabla f)^2 \Psi,$$

in terms of the potentials  $(\phi', \mathbf{A}')$ . This problem is resolved by noting that in obtaining this equation we assumed that the wavefunction  $\Psi$  remained unchanged when the electromagnetic potentials underwent the gauge transformation (D.7). If instead we assume that  $\Psi$  simultaneously transforms as

$$\Psi(\mathbf{r}, t) \rightarrow \Psi'(\mathbf{r}, t) = \exp[-iqf(\mathbf{r}, t)] \Psi(\mathbf{r}, t), \quad (\text{D.8})$$

then from (D.7) and (D.8) we can show that

$$\left( \frac{\partial}{\partial t} + iq\phi' \right) \Psi' = e^{-iqf} \left( \frac{\partial}{\partial t} + iq\phi \right) \Psi$$

and

$$(\nabla - iq\mathbf{A}')^2 \Psi' = (\nabla - iq\mathbf{A}')e^{-iqf}(\nabla - iq\mathbf{A})\Psi = e^{-iqf}(\nabla - iq\mathbf{A})^2 \Psi.$$

The right-hand sides of these equations are related by (D.6), and thus the left-hand sides lead to an identical equation

$$i \left( \frac{\partial}{\partial t} + iq\phi' \right) \Psi' = -\frac{1}{2m} (\nabla - iq\mathbf{A}')^2 \Psi'$$

for the transformed potentials and wavefunction defined by (D.7) and (D.8). We thus arrive at the conclusion that the fundamental equation of motion is indeed gauge invariant, provided that we extend our definition of a gauge transformation to include (D.8) in addition to (D.7).

### D.3 Gauge invariance and the photon mass

So far we have discussed the equation of motion of a charged particle in the electromagnetic potentials  $(\phi, \mathbf{A})$ , but we have not discussed the equations of motion of the potentials themselves. In this section we shall consider these equations and show that they are only gauge invariant if free photons have zero mass.

The equations of motion of the electromagnetic field in free space are Maxwell's equations

$$\text{div } \mathbf{B} = 0, \quad \text{curl } \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (\text{D.9a})$$

and

$$\text{div } \mathbf{E} = 0, \quad \text{curl } \mathbf{B} = \frac{\partial \mathbf{E}}{\partial t}. \quad (\text{D.9b})$$

These can be re-expressed in terms of the potentials by substituting (D.4) for  $\mathbf{E}$  and  $\mathbf{B}$ . The first two equations (D.9a) are automatically satisfied for any choice of potentials  $(\phi, \mathbf{A})$ . The second two equations (D.9b) lead to Maxwell's equations for the potentials themselves, which are conveniently written, for what follows, in the form:

$$\left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi - \frac{\partial}{\partial t} \left( \frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A} \right) = 0 \quad (\text{D.10a})$$

and

$$\left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{A} + \nabla \left( \frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A} \right) = 0. \quad (\text{D.10b})$$

These equations are gauge invariant and their interpretation is facilitated by ‘choosing a gauge’. That is, we take advantage of the ambiguity (D.7) in the choice of electromagnetic potentials to impose a convenient subsidiary condition. In particular, for any set of fields  $(\tilde{\phi}, \tilde{\mathbf{A}})$  satisfying (D.10), we can always find a function  $f(\mathbf{r}, t)$  such that the transformed fields

$$\phi = \tilde{\phi} + \frac{\partial f}{\partial t}, \quad \mathbf{A} = \tilde{\mathbf{A}} - \nabla f$$

satisfy the so-called *Lorentz condition*

$$\frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A} = 0. \quad (\text{D.11})$$

Imposing this condition is called ‘choosing a Lorentz gauge’ and we are free to make this choice without restricting or altering the physical content of the theory in any way. When we do so, (D.10) become

$$\left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi = 0, \quad \left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{A} = 0. \quad (\text{D.12})$$

Their interpretation follows directly from comparison with the Klein–Gordon equation (1.5) for a particle of mass  $m$ , which is

$$\left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \Psi + m^2 \Psi = 0 \quad (\text{D.13})$$

in natural units. Equations (D.12) are of exactly the same form as (D.13), except that the associated particles – photons – have zero mass.

Photons with non-zero masses (or any other massive spin-1 particles) would be described by the *Proca equations*

$$\left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi - \frac{\partial}{\partial t} \left( \frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A} \right) + m_\gamma^2 \phi = 0 \quad (\text{D.14a})$$

and

$$\left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{A} + \nabla \left( \frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A} \right) + m_\gamma^2 \mathbf{A} = 0, \quad (\text{D.14b})$$

as we shall shortly demonstrate. Firstly, we note that the Proca equations differ from the gauge-invariant Maxwell equations (D.10) by the addition of ‘mass terms’  $m_\gamma^2 \phi$  and  $m_\gamma^2 \mathbf{A}$ , which are obviously not invariant under (D.7). Hence the Proca equations are only gauge-invariant in the case  $m_\gamma = 0$ , when they reduce to Maxwell’s equations. In other words, gauge invariance requires the photon to be massless.

It remains to show that Equations (D.14) really do describe particles of mass  $m_\gamma \neq 0$  and to consider how this would affect the behaviour of the electric and magnetic fields (D.4). To do this we

differentiate (D.14a) with respect to time, take the divergence of (D.14b) and add the two together. After a short calculation, this gives

$$m_\gamma^2 \left( \frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A} \right) = 0.$$

Consequently, if  $m_\gamma \neq 0$  the Lorentz condition (D.11) necessarily holds (there is no choice!)<sup>4</sup> and the Proca equations reduce to the Klein–Gordon equations<sup>5</sup>

$$\left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi + m_\gamma^2 \phi = 0 \quad (\text{D.15a})$$

and

$$\left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{A} + m_\gamma^2 \mathbf{A} = 0 \quad (\text{D.15b})$$

for particles of mass  $m_\gamma \neq 0$ . Such a mass could modify the characteristic long-range behaviour of the electromagnetic fields derived from the potentials. For example, in the case of electrostatics,  $\phi(\mathbf{r}, t) \rightarrow \phi(\mathbf{r})$  and (D.15a) becomes

$$\nabla^2 \phi(\mathbf{r}) - m_\gamma^2 \phi(\mathbf{r}) = 0$$

which, following the discussion of Section 1.4.2, leads to a Coulomb potential of the Yukawa form

$$\phi(\mathbf{r}) = -\frac{e^2}{4\pi} \frac{e^{-r/R}}{r},$$

where  $R = 1/m_\gamma$  or  $\hbar/m_\gamma c$  in practical units. Laboratory tests of the Coulomb law provide the upper limit  $m_\gamma \leq 10^{-14} \text{ eV}/c^2$ , while similar studies on the variation of the magnetic field of Jupiter over large distances give  $m_\gamma \leq 6 \times 10^{-17} \text{ eV}/c^2$ . In short, the prediction of zero photon mass, as required by gauge invariance, has been confirmed experimentally to quite extraordinary accuracy.

## D.4 The gauge principle

In Section D.2, invariance of the wave equation (D.6) under a gauge transformation (D.7) of the electromagnetic potentials was shown

<sup>4</sup> For a gauge-invariant theory there is always a choice, since a gauge transformation (D.7) can change a set of potentials that satisfy the Lorentz condition into one that does not, or vice versa.

<sup>5</sup> In essence, the Proca equations combine the Klein–Gordon equations (D.15) and the Lorentz condition (D.11) into a single set of equations, provided the mass is non-zero. The Lorentz condition effectively reduces the four functions  $(\phi, \mathbf{A}) \equiv (\phi, A_x, A_y, A_z)$  to three independent components, corresponding to the three possible spin states of a massive spin-1 particle.

to require that the wavefunction also underwent a gauge transformation (D.8). With the development of new gauge theories – QCD and the unified electroweak theory – it has become fashionable to reverse this argument for reasons that will become clear shortly. In this approach, invariance of the wave equation under the gauge transformation (D.8) is taken as the fundamental requirement and used to infer the form of the interactions. This is called the *principle of minimal gauge invariance*, or the *gauge principle* for short, and we shall introduce it by considering the electromagnetic interactions of relativistic spin-1/2 particles.<sup>6</sup>

In Chapter 1, we saw that in the absence of interactions a relativistic spin-1/2 particle obeys the Dirac equation (1.10), that is

$$i \frac{\partial \Psi}{\partial t} = -i \boldsymbol{\alpha} \cdot \nabla \Psi + \beta m \Psi, \quad (\text{D.16})$$

where we have introduced the notation

$$\boldsymbol{\alpha} \cdot \nabla \equiv \sum_{i=1}^3 \alpha_i \frac{\partial}{\partial x_i}$$

and set  $\hbar = 1$  and  $c = 1$  as usual. However, this equation is not invariant under the fundamental gauge transformation (D.8), since the transformed wavefunction  $\Psi'$  satisfies

$$i \left( \frac{\partial}{\partial t} + iq \frac{\partial f}{\partial t} \right) \Psi' = -i \boldsymbol{\alpha} \cdot (\nabla + iq \nabla f) \Psi' + \beta m \Psi', \quad (\text{D.17})$$

as is easily verified by substituting (D.8) into (D.16). This problem is remedied by adding just those terms involving the electromagnetic potentials that are needed to make the Dirac equation (D.16) gauge invariant. To see how to do this, we note that the troublesome terms in (D.17) involve  $\partial f/\partial t$  and  $\nabla f$ , and originate from the derivatives  $\partial \Psi/\partial t$  and  $\nabla \Psi$  in (D.16), and that the same derivatives of  $f(\mathbf{r}, t)$  occur in the gauge transformations (D.7) of  $\phi$  and  $\mathbf{A}$ , respectively. Hence it should be possible to add one suitably chosen term involving a potential to each of the derivatives in (D.16) in such a way that the troublesome terms cancel when one makes a gauge transformation. This leads to the equation

$$i \left( \frac{\partial}{\partial t} + iq\phi \right) \Psi = -i \boldsymbol{\alpha} \cdot (\nabla - iq\mathbf{A}) \Psi + \beta m \Psi, \quad (\text{D.18})$$

which differs from (D.16) by the so-called minimal substitutions

$$\frac{\partial}{\partial t} \rightarrow \frac{\partial}{\partial t} + iq\phi, \quad \nabla \rightarrow \nabla - iq\mathbf{A} \quad (\text{D.19})$$

---

<sup>6</sup> We start with the relativistic case because the non-relativistic case is algebraically more complicated. (We will comment on the non-relativistic case shortly.)

and we leave it as an exercise for the reader to verify that (D.18) is indeed gauge invariant. In the familiar case of an electron,  $q = -e$  and (D.18) becomes

$$i \left( \frac{\partial}{\partial t} - ie\phi \right) \Psi = -i\boldsymbol{\alpha} \cdot (\boldsymbol{\nabla} + ie\mathbf{A}) \Psi + \beta m \Psi. \quad (\text{D.20})$$

This is precisely the equation of motion that is used for the electron in quantum electrodynamics with such spectacular success.<sup>7</sup>

## D.5 The Higgs mechanism

In Section D.3, we showed that Maxwell's equations in free space imply that photons have zero mass and that, in the absence of interactions, introducing a hypothetical non-zero photon mass necessarily violates gauge invariance. However, in Section 10.2.2, we stated that when the electromagnetic field interacts with a complex scalar field, spontaneous symmetry breaking can generate a non-zero mass without violating the gauge invariance of the interactions. In this section we will give a simple account of how this arises.<sup>8</sup> We will do this by, firstly, in Sections D.5.1 and D.5.2, exploring the equations of motion of the electromagnetic fields in the presence of interactions, and then discuss, in Section D.5.3, the result of spontaneous symmetry breaking itself.

### D.5.1 Charge and current densities

In classical physics, the sources of the electromagnetic field are the electric charge and current densities  $\rho(\mathbf{r}, t)$  and  $\mathbf{j}(\mathbf{r}, t)$ , respectively, which satisfy the continuity condition

$$\frac{d\rho}{dt} + \boldsymbol{\nabla} \cdot \mathbf{j} = 0 \quad (\text{D.21})$$

---

<sup>7</sup> If we had applied our present argument to a non-relativistic particle, we would have started from the Schrödinger equation

$$i \frac{\partial \Psi}{\partial t} = -\frac{1}{2m} \nabla^2 \Psi,$$

which is also not invariant under (D.8). The same minimal substitution (D.19) would then have led us to the wave equation (D.6), which we have already shown to be gauge invariant.

<sup>8</sup> Our discussion completes that of Section 10.2.2. For a more rigorous account, which requires the formalism of quantum field theory, see, for example, Section 13.2 of Mandl and Shaw (2010).

in order to guarantee electric charge conservation. When these sources are included, the first pair of Maxwell's equations, (D.9a), remains unchanged, while the second pair, (D.9b), become

$$\operatorname{div} \mathbf{E} = \rho, \quad \operatorname{curl} \mathbf{B} = \mathbf{j} + \frac{\partial \mathbf{E}}{\partial t}, \quad (\text{D.9c})$$

in our units. On expressing the  $\mathbf{E}$  and  $\mathbf{B}$  field in terms of the potentials using (D.4), Equations (D.9a) are again automatically satisfied, while (D.9c) lead to Maxwell's equations for the potentials in the presence of sources. These are conveniently written in the forms

$$\left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi - \frac{\partial}{\partial t} \left( \frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A} \right) = \rho \quad (\text{D.22a})$$

and

$$\left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{A} + \nabla \left( \frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A} \right) = \mathbf{j}. \quad (\text{D.22b})$$

They differ from our previous Maxwell equations (D.10a) and (D.10b) solely by the addition of the source terms, and since the values of the electric charges and currents are unaffected by gauge transformations (D.7) on the potentials, (D.22a) and (D.22b) are, like (D.10a) and (D.10b), explicitly gauge invariant.

In quantum physics, the classical charge and current densities are replaced by charge and current densities expressed in terms of fields of the associated charged particles. For point-like spin-1/2 fermions, these take the form

$$\rho = q\Psi^\dagger \Psi, \quad \mathbf{j} = q\Psi^\dagger \boldsymbol{\alpha} \Psi, \quad (\text{D.23})$$

where  $\boldsymbol{\alpha}$  are the Dirac matrices (1.9) and  $\Psi^\dagger$  is the ‘Hermitian conjugate spinor’

$$\Psi^\dagger \equiv (\Psi_1^*, \Psi_2^*, \Psi_3^*, \Psi_4^*). \quad (\text{D.24})$$

It is straightforward to show that the charge and current densities (D.23) satisfy the continuity condition (D.21) by substituting (D.23) into (D.21) and using the Dirac equation (D.16) together with the equation

$$-i \frac{\partial \Psi^\dagger}{\partial t} = i\nabla \Psi^\dagger \cdot \boldsymbol{\alpha} + \beta m \Psi^\dagger, \quad (\text{D.25})$$

obtained by taking the Hermitian conjugate of (D.16), and using the fact that the matrices  $\boldsymbol{\alpha}$  and  $\beta$  must be Hermitian to ensure that the Hamiltonian (1.8) is Hermitian. In addition, the charge and current densities are both invariant under the combined gauge transformations (D.7) and (D.8). Hence, substituting (D.23) into (D.22) gives an equation of motion for the electromagnetic field that is gauge invariant as required. This is precisely the equation of motion used for the electromagnetic field in QED, where the charge  $q = -e$  for the electron.

### D.5.2 Spin-0 bosons

In order to understand the Higgs mechanism, we first need to extend the above formalism to include charged spin-0 bosons rather than spin-1/2 fermions. Such particles are described by a complex field  $\eta(\mathbf{r}, t)$ , which satisfies the Klein–Gordon equation (1.5), that is

$$\left( \frac{\partial^2}{\partial t^2} - \nabla^2 + m^2 \right) \eta(\mathbf{r}, t) = 0. \quad (\text{D.26})$$

In this case, charge and current densities that satisfy the continuity condition (D.21) are given by

$$\rho = iq \left( \eta^* \frac{\partial \eta}{\partial t} - \frac{\partial \eta^*}{\partial t} \eta \right), \quad \mathbf{j} = -iq (\eta^* \nabla \eta - \eta \nabla \eta^*), \quad (\text{D.27})$$

as can be verified by substituting (D.27) into (D.21) and using the Klein–Gordon equation (D.26). Implicit in this is that for particles of charge  $q$ , the probability amplitude for position,  $\rho_P$ , analogous to  $\Psi^* \Psi$  for the Schrödinger equation or  $\Psi^\dagger \Psi$  for the Dirac equation, is given by

$$\rho_P = -i \left( \eta^* \frac{\partial \eta}{\partial t} - \frac{\partial \eta^*}{\partial t} \eta \right). \quad (\text{D.28})$$

This requires that  $\rho_P(\mathbf{r}, t)$  must not only satisfy a continuity equation of the form (D.21), but must also be positive definite. One can easily verify that this is true for positive energy solutions of the form (1.6a), but not for negative energy solutions of the form (1.6b), for which  $\rho_P(\mathbf{r}, t) < 0$ . Here we shall simply confine ourselves to positive energy solutions and note that a more rigorous treatment using quantum field theory, which incorporates antiparticles of charge  $-q$  associated with the negative energy solutions, yields the same expressions (D.27) for the electric charge and current densities.

The most striking thing about the densities (D.27) is that they contain derivatives, essentially because the Klein–Gordon equation contains second-order derivatives, in contrast to the Dirac equation, which is first-order. Because of these derivatives, the charge and current densities are not invariant under the gauge transformation (cf. (10.43))

$$\eta(\mathbf{r}, t) \rightarrow \exp[-iqf(\mathbf{r}, t)]\eta(\mathbf{r}, t), \quad (\text{D.29})$$

but rather transform as

$$\rho \rightarrow \rho + 2q^2 \eta^* \eta \frac{\partial f}{\partial t}, \quad \mathbf{j} \rightarrow \mathbf{j} - 2q^2 \eta^* \eta \nabla f. \quad (\text{D.30})$$

Thus, in this case, the proposed equations of motion for the electromagnetic fields (D.22a) and (D.22b) do not satisfy the fundamental requirement of gauge invariance under the combined gauge transformations (D.7) and (D.29). To resolve this problem, we invoke the gauge principle discussed in Section D.4 and add suitably chosen

interaction terms to (D.22a) and (D.22b) to ensure that the resulting equations are gauge invariant. This is achieved by modifying (D.22a) and (D.22b) to be

$$\left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi - \frac{\partial}{\partial t} \left( \frac{\partial \phi}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{A} \right) = \rho - 2q^2 \eta^* \eta \phi \quad (\text{D.31a})$$

and

$$\left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{A} + \boldsymbol{\nabla} \left( \frac{\partial \phi}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{A} \right) = \mathbf{j} - 2q^2 \eta^* \eta \mathbf{A}. \quad (\text{D.31b})$$

We leave it to the reader to verify that the right-hand sides of (D.31a) and (D.31b), and hence the equations themselves, are gauge invariant. These are precisely the equations of motion for the electromagnetic field used to describe the quantum electrodynamics of spin-0 bosons with charge  $q$ . However, if spontaneous symmetry breaking occurs, it profoundly alters the physical interpretation of these equations, as we shall immediately see.

### D.5.3 Spontaneous symmetry breaking

In Section 10.2.2, we saw that when spontaneous symmetry breaking occurs the field  $\eta(\mathbf{r}, t)$  has a non-zero value in the vacuum state that can, without loss of generality, be chosen to be (cf. (10.42))

$$\eta_0 = \frac{v}{\sqrt{2}} \equiv \left( -\frac{\mu^2}{2\lambda} \right)^{1/2} \geq 0. \quad (\text{D.32})$$

This clearly violates gauge invariance and to understand its consequences, we rewrite the theory in terms of a new field  $\sigma(\mathbf{r}, t)$  that vanishes in the vacuum state, defined by

$$\eta(\mathbf{r}, t) \equiv \frac{1}{\sqrt{2}} [v + \sigma(\mathbf{r}, t)]. \quad (\text{D.33})$$

In addition, as in Section (D.3), we again ‘choose a gauge’ that makes the physical interpretation more transparent. In this case, however, instead of imposing the Lorentz condition (D.11), we note that for any set of fields  $(\hat{\phi}, \hat{\mathbf{A}}, \hat{\eta})$ , we can always find a gauge function  $f(\mathbf{r}, t)$  such that the transformed field

$$\eta(\mathbf{r}, t) = \exp[-iqf(\mathbf{r}, t)] \hat{\eta}(\mathbf{r}, t),$$

and hence the corresponding field  $\sigma(\mathbf{r}, t)$ , is real, and hence

$$\eta^*(\mathbf{r}, t) = \eta(\mathbf{r}, t). \quad (\text{D.34})$$

Imposing (D.34) defines the *unitary gauge*, in which the charge and current densities (D.27) vanish:

$$\rho(\mathbf{r}, t) = 0, \quad \mathbf{j}(\mathbf{r}, t) = \mathbf{0}. \quad (\text{D.35})$$

Substituting (D.33) and (D.35) into (D.31a) and (D.31b) then gives

$$\left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi - \frac{\partial}{\partial t} \left( \frac{\partial \phi}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{A} \right) + q^2 v^2 \phi = -2q^2 v \sigma \phi - q^2 \sigma^2 \phi \quad (\text{D.36a})$$

and

$$\left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{A} + \boldsymbol{\nabla} \left( \frac{\partial \phi}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{A} \right) + q^2 v^2 \mathbf{A} = -2q^2 v \sigma \mathbf{A} - q^2 \sigma^2 \mathbf{A} \quad (\text{D.36b})$$

as the new equations of motion for the electromagnetic field. On comparing with (D.14a) and (D.14b), we see that they are identical with the Proca equations for ‘photons’ with non-zero mass

$$m_\gamma = q v, \quad (\text{D.37})$$

whose interactions with the real scalar field  $\sigma(\mathbf{r}, t)$  are specified by the terms on the right-hand side of (D.36). In addition, the vanishing of the electric charge and current densities implies that the quanta of the field  $\sigma(\mathbf{r}, t)$  – the Higgs bosons – are electrically neutral.

This is the essence of the Higgs mechanism, whereby a theory of charged bosons interacting with massless photons is converted by spontaneous symmetry breaking into a theory of massive photons interacting with neutral Higgs bosons. This mechanism is not realised in nature for photons, which are assumed to be exactly massless. An analogous mechanism is, however, used in the more complicated unified electroweak theory to generate masses for the  $W^\pm$  and  $Z^0$  bosons, as discussed in Section 10.2.

## D.6 Quantum chromodynamics

In QED, the gauge principle approach we have outlined in Section D.4 is an interesting way of looking at an interaction whose form was originally determined in other ways. In contrast, in QCD it was used to infer the detailed form of interactions that were previously unknown. To do this, one starts from new types of gauge transformations that are similar to, but more complicated than, the gauge transformations we have discussed previously. In particular, they not only change the phase of the wavefunction as in (D.8) but also the colour state. These colour states have been discussed in Section 6.3, but we repeat the salient points here for convenience.

**Table D.1** Some of the coefficients in the commutation relations (D.39) of the colour charges  $\hat{F}_i$  ( $i = 1, 2, \dots, 8$ ). All the other non-zero coefficients may be deduced from these by using the fact that  $f_{ijk}$  is totally antisymmetric.

$ijk$	123	147	156	246	257	345	367	458	678
$f_{ijk}$	1	$1/2$	$-1/2$	$1/2$	$1/2$	$1/2$	$-1/2$	$\sqrt{3}/2$	$\sqrt{3}/2$

There are three independent colour wavefunctions  $\chi_C = r, g, b$ , which are conveniently represented by the colour spinors (6.48):

$$r = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad g = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (\text{D.38})$$

These are acted upon by eight independent colour charge operators  $\hat{F}_1, \hat{F}_2, \dots, \hat{F}_8$ , which are represented by the matrices (6.49) and are assumed to be conserved:

$$[\hat{F}_i, H] = 0 \quad (i = 1, 2, \dots, 8).$$

The commutation relations among the  $\hat{F}_i$  were not stated in Section 6.4, but are required below. They follow directly from (6.49) and are

$$[\hat{F}_i, \hat{F}_j] = i \sum_k f_{ijk} \hat{F}_k, \quad (\text{D.39})$$

where some of the non-zero coefficients  $f_{ijk}$  are given in Table D.1 and the rest follow from the fact that they are totally antisymmetric:

$$f_{ijk} = -f_{jik} = f_{jki} = -f_{kji} = f_{kij} = -f_{ikj}. \quad (\text{D.40})$$

Because the charges  $\hat{F}_i$  do not commute with each other, they cannot all have definite values simultaneously. In fact, the colour states  $\chi_C = r, g, b$  are chosen to be eigenstates of the charges  $\hat{F}_3$  and  $\hat{F}_8$ , with eigenvalues given by (6.51) and Table 6.13. The remaining charges  $\hat{F}_1, \hat{F}_2, \hat{F}_4, \dots, \hat{F}_7$  change the colour state, as illustrated, for example, by (6.52):

$$\hat{F}_1 r = \frac{1}{2} g, \quad \hat{F}_1 g = \frac{1}{2} r, \quad \hat{F}_1 b = 0. \quad (\text{D.41})$$

Finally, the overall state of any relativistic quark is given by the product of a Dirac spinor  $\psi(\mathbf{r}, t)$  with a colour wavefunction

$$\Psi \equiv \psi(\mathbf{r}, t) \chi_C, \quad (\text{D.42})$$

or a linear combination of such wavefunctions.

The basic idea behind QCD is that the colour charges described above act as the sources of the strong interaction, in the same way that the electric charge acts as the source of the electromagnetic interaction. This idea is implemented by modifying the form of

gauge invariance imposed. In QED we demanded invariance under the gauge transformation

$$\Psi(\mathbf{r}, t) \rightarrow \Psi'(\mathbf{r}, t) = \exp[-iqf(\mathbf{r}, t)]\Psi(\mathbf{r}, t) \quad (\text{D.8})$$

on the wavefunctions of quarks or leptons of charge  $q$ . In QCD we demand invariance under the gauge transformations

$$\Psi \rightarrow \Psi' = \exp \left[ -ig_s \sum_i \hat{F}_i \omega_i(\mathbf{r}, t) \right] \Psi \quad (\text{D.43})$$

on the wavefunctions (D.42) of coloured quarks, where  $g_s$  is a coupling constant that will determine the strength of the interaction,<sup>9</sup> and there are now eight arbitrary gauge functions  $\omega_i(\mathbf{r}, t)$  ( $i = 1, 2, \dots, 8$ ), corresponding to the eight colour charges<sup>10</sup>  $\hat{F}_1, \hat{F}_2, \dots, \hat{F}_8$ .

We now use invariance under (D.43) to determine the form of the interaction, following a similar argument to that given for QED in Section D.4. Indeed, if we pretend for a moment that the charges  $\hat{F}_i$  commute (i.e. that  $f_{ijk} = 0$ ), the argument leading from (D.16) to (D.20) goes through with only trivial modifications. This gives

$$i \left( \frac{\partial}{\partial t} + ig_s \sum_i \hat{F}_i \phi_i \right) \Psi = -i\boldsymbol{\alpha} \cdot \left( \boldsymbol{\nabla} - ig_s \sum_i \hat{F}_i \mathbf{A}_i \right) \Psi + \beta m \Psi \quad (\text{D.44})$$

as the equation of motion analogous to (D.20), where  $(\phi_i, \mathbf{A}_i)$  ( $i = 1, 2, \dots, 8$ ) are eight new gluon fields, which transform as

$$\phi_i \rightarrow \phi'_i = \phi_i + \frac{\partial \omega_i}{\partial t}, \quad \mathbf{A}_i \rightarrow \mathbf{A}'_i = \mathbf{A}_i - \nabla \omega_i \quad (f_{ijk} = 0), \quad (\text{D.45})$$

in analogy with (D.7) for the electromagnetic field  $(\phi, \mathbf{A})$ . However,  $f_{ijk} \neq 0$  and the transformation law (D.45) has to be modified if (D.44) is to remain invariant when this is taken into account. In practice, it is sufficient to restrict ourselves to infinitesimal transformations, corresponding to small  $\omega_i(\mathbf{r}, t)$ <sup>11</sup> when (D.45) becomes

$$\phi_i \rightarrow \phi'_i = \phi_i + \frac{\partial \omega_i}{\partial t} + g_s \sum_{jk} f_{ijk} \omega_j \phi_k \quad (\text{D.46a})$$

<sup>9</sup> It is related to  $\alpha_s$  used in Section 7.1 by  $\alpha_s = g_s^2/4\pi$ , in analogy to  $\alpha = e^2/4\pi\varepsilon_0$  for the electromagnetic interaction, where  $\varepsilon_0 = 1$  in this appendix.

<sup>10</sup> The parallel between (D.8) and (D.43) would be even more transparent if we wrote  $q = eQ$ , where  $e$  characterises the strength of the electromagnetic interaction and  $Q = 0, -1/3$  or  $-1$ , the charge of the quark or lepton concerned.

<sup>11</sup> This is because finite transformations can be treated as the limit as  $N \rightarrow \infty$  of  $N$  successive small transformations.

and

$$\mathbf{A}_i \rightarrow \mathbf{A}'_i = \mathbf{A}_i - \nabla \omega_i + g_s \sum_{jk} f_{ijk} \omega_j \mathbf{A}_k. \quad (\text{D.46b})$$

The proof that this is correct will be given shortly. Firstly, we consider its physical interpretation.

The key to the interpretation of (D.46) lies in the form of the interaction terms in (D.44). For example, for a red quark,  $\Psi = \psi(\mathbf{r}, t) r$ , (D.44) contains interaction terms like

$$-g_s \hat{F}_1 \phi_1 \psi r = -\frac{g_s}{2} \phi_1 \psi g,$$

in which the colour charge of the quark changes. By colour charge conservation, this is only possible if the gluon fields  $(\phi_i, \mathbf{A}_i)$  can carry away colour charge; that is, the gluons are coloured. Finally, if the colour charges  $\hat{F}_i$  are the sources of the gluon fields  $(\phi_i, \mathbf{A}_i)$ , and gluons are coloured, this implies that gluons interact with other gluons. This has been discussed qualitatively already in Section 7.1.<sup>12</sup> Here we note that the existence of such interactions and their detailed form are determined by the transformation law (D.46). To see this, we note that free gluons, like free photons, would be described by equations of the form (D.10)

$$\left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi_i - \frac{\partial}{\partial t} \left( \frac{\partial \phi_i}{\partial t} + \nabla \cdot \mathbf{A}_i \right) = 0 \quad (\text{D.47a})$$

and

$$\left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{A}_i + \nabla \left( \frac{\partial \phi_i}{\partial t} + \nabla \cdot \mathbf{A}_i \right) = 0. \quad (\text{D.47b})$$

These equations are invariant under the transformations (D.45), but they are not invariant under the required gauge transformation (D.46) because of the final terms proportional to  $f_{ijk}$ . So once again one is forced to add the minimal interaction terms necessary to make a free-particle wave equation gauge invariant. These terms can be shown to lead to gluon–gluon interactions of the type shown in Figure 7.2. However, we shall not pursue this further, because the algebra is complicated and we have not used the detailed results. Rather, we reiterate that the principle of minimal gauge invariance predicts the existence and form of both quark–gluon and gluon–gluon interactions, as outlined above, and that these predictions are in agreement with a large body of data and are not contradicted by any known experiment.

We conclude this section by proving that (D.44) is invariant under the gauge transformations (D.43) and (D.46) for infinitesimal

---

<sup>12</sup>See Figure 7.1, Equation (7.1) and the accompanying discussion.

$\omega_i(\mathbf{r}, t)$ , as stated earlier.<sup>13</sup> For infinitesimal  $\omega_i(\mathbf{r}, t)$ , (D.43) reduces to

$$\Psi \rightarrow \Psi' = \left( 1 - ig_s \sum_j \hat{F}_j \omega_j \right) \Psi, \quad (\text{D.48})$$

where we have relabelled the summation variable  $i \rightarrow j$  for later convenience. By inspection, this implies that (D.44) will be invariant if

$$\left( \frac{\partial}{\partial t} + ig_s \sum_i \hat{F}_i \phi'_i \right) \Psi' = \left( 1 - ig_s \sum_j \hat{F}_j \omega_j \right) \left( \frac{\partial}{\partial t} + ig_s \sum_i \hat{F}_i \phi_i \right) \Psi \quad (\text{D.49a})$$

and

$$\begin{aligned} \boldsymbol{\alpha} \cdot \left( \nabla - ig_s \sum_i \hat{F}_i \mathbf{A}'_i \right) \Psi' = \\ \left( 1 - ig_s \sum_j \hat{F}_j \omega_j \right) \boldsymbol{\alpha} \cdot \left( \nabla - ig_s \sum_i \hat{F}_i \mathbf{A}_i \right) \Psi. \end{aligned} \quad (\text{D.49b})$$

The next step is to write the gauge field transformations in the general form

$$\phi_i \rightarrow \phi'_i = \phi_i + \frac{\partial \omega_i}{\partial t} + \Delta \phi_i \quad (\text{D.50a})$$

and

$$\mathbf{A}_i \rightarrow \mathbf{A}'_i = \mathbf{A}_i - \nabla \omega_i + \Delta \mathbf{A}_i \quad (\text{D.50b})$$

and use (D.49) to identify the unknown functions  $\Delta \phi_i$  and  $\Delta \mathbf{A}_i$ .

On substituting (D.48) and (D.50a) into (D.49a), we obtain

$$ig_s \sum_i \hat{F}_i \Delta \phi_i = (ig_s)^2 \sum_{ij} [\hat{F}_i, \hat{F}_j] \omega_j \phi_i,$$

where we have neglected terms that are second order in the infinitesimal quantities  $\omega_i$  and  $\Delta \phi_i$ . Using the commutation relations (D.39) gives

$$\sum_i \hat{F}_i \Delta \phi_i = -g_s \sum_{ijk} f_{ijk} \hat{F}_k \omega_j \phi_i = g_s \sum_{ijk} f_{ijk} \hat{F}_i \omega_j \phi_k,$$

where we have interchanged the summation labels  $i \rightarrow k$  in the last line, using the antisymmetry property (D.40). This implies

$$\Delta \phi_i = g_s \sum_{jk} f_{ijk} \omega_j \phi_k, \quad (\text{D.51})$$

---

<sup>13</sup>This proof is included for completeness and is not required in what follows. The reader who wishes to omit it should proceed directly to Section D.7.

which, together with (D.50a), is equivalent to the desired transformation law (D.46a). A similar argument applied to (D.49b) establishes the transformation law (D.46b).

## D.7 Electroweak interactions

The application of the gauge principle to weak interactions and its role in their unification with the electromagnetic interaction have been discussed qualitatively, for massless electrons and neutrinos, in Section 10.2.1. Here we will supply some of the details that were omitted in that account. In particular, we will specify in more detail the form of the gauge invariance assumed and we will explicitly derive the crucial unification condition (10.8).

The discussion proceeds as follows. We first set up a combined description of electron and neutrino states that is similar to the description of coloured quark states used in Section D.6. This is then exploited to define a set of gauge transformations that lead, via the gauge principle, to the known charged current weak interactions. However, the corresponding neutral current interactions are incorrect, and this problem is resolved by unification with the electromagnetic interaction, leading to the unification condition (10.8). In the final section, we indicate how the spin structure of the interaction and parity violation, which are ignored in the previous discussion, are incorporated into the picture.

Finally, although our discussion is confined to leptons, it is easily extended to quarks using lepton–quark symmetry and quark mixing. This is left as an exercise for the reader (cf. Problems D.6 and D.7).

### D.7.1 Weak isospin

To formulate a gauge theory of weak interactions, we require a set of gauge transformations that transform electrons and electron neutrinos into themselves or into each other, as outlined in Section 10.2.1. In order to do this, we recall that in QCD we saw how to define gauge transformations that transform different quark colour states into each other. We therefore begin by setting up a combined description of electrons and neutrinos, which is closely analogous to the description of coloured quark states used in Section D.6. Specifically, we introduce two independent flavour wavefunctions  $\chi^f = \nu, e$ , which are conveniently represented by the flavour spinors

$$\nu = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad (\text{D.52})$$

in analogy to the quark colour spinors (D.38). Just as the colour spinors are acted on by colour charge operators, the flavour spinors

are acted upon by analogous operators, called *weak isospin* operators. These are represented by two-dimensional matrices in the same way that the colour charges are represented by three-dimensional matrices (6.49). There are three such independent weak isospin operators

$$\hat{I}_i^W = \frac{1}{2}\tau_i \quad (i = 1, 2, 3), \quad (\text{D.53a})$$

with the matrices<sup>14</sup>

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (\text{D.53b})$$

The commutation relations among the  $\hat{I}_i^W$  follow directly from (D.53) and are

$$[\hat{I}_i^W, \hat{I}_j^W] = i \sum_k \varepsilon_{ijk} \hat{I}_k^W, \quad (\text{D.54})$$

where the coefficients  $\varepsilon_{ijk}$  are totally antisymmetric:

$$\begin{aligned} \varepsilon_{123} &= \varepsilon_{312} = \varepsilon_{231} = 1, \\ \varepsilon_{132} &= \varepsilon_{213} = \varepsilon_{321} = -1, \\ \varepsilon_{ijk} &= 0 \text{ if } i = j, j = k \text{ or } i = k. \end{aligned} \quad (\text{D.55})$$

The states  $\chi^f = \nu, e$  are eigenstates of  $\hat{I}_3^W$ :

$$\hat{I}_3^W \nu = \frac{1}{2}\nu, \quad \hat{I}_3^W e = -\frac{1}{2}e, \quad (\text{D.56})$$

while  $\hat{I}_1^W$  and  $\hat{I}_2^W$  interchange electron and neutrino states:

$$\hat{I}_1^W \nu = \frac{1}{2}e, \quad \hat{I}_1^W e = \frac{1}{2}\nu, \quad \hat{I}_2^W \nu = \frac{i}{2}e, \quad \hat{I}_2^W e = -\frac{i}{2}\nu, \quad (\text{D.57})$$

as one easily verifies from (D.52) and (D.53). Finally, the overall state of any relativistic lepton is given by the product of a Dirac spinor<sup>15</sup>  $\psi(\mathbf{r}, t)$  with a flavour spinor  $\chi^f$ :

$$\Psi = \psi(\mathbf{r}, t)\chi^f,$$

or by a linear combination of such states, where we restrict ourselves to the first generation of leptons ( $\ell = \nu_e, e$ ) only.

<sup>14</sup>The weak isospin operators, like the colour charge operators, must be Hermitian if they are to correspond to physical observables. The most general Hermitian matrices in two or three dimensions can be written as linear combinations of the unit matrix and the  $\tau_i$  or  $\lambda_i$ , respectively.

<sup>15</sup>We remind the reader that we are for the moment ignoring the spin structure of the interaction. When this is taken into account, the present formalism is applied to ‘left-handed’ spin states, as we shall see in Section D.7.4.

## D.7.2 Gauge invariance and charged currents

We now have a description of electron and electron neutrino states that differs from our description of coloured states (D.38) to (D.42) by the replacements:

$$\Psi = \psi(\mathbf{r}, t)\chi_C \rightarrow \Psi = \psi(\mathbf{r}, t)\chi^f,$$

$$\hat{F}_i(1, 2, \dots, 8) \rightarrow I_i^W(1, 2, 3)$$

and

$$f_{ijk} \rightarrow \varepsilon_{ijk}.$$

These can be used to adapt the results derived for QCD to the case of weak interactions. Specifically, we assume gauge invariance under a set of transformations

$$\Psi \rightarrow \Psi' \exp \left( -ig \sum_i \hat{I}_i^W f_i(\mathbf{r}, t) \right) \Psi \quad (\text{D.58})$$

analogous to (D.43), where  $g$  is an arbitrary coupling constant and  $f_i(\mathbf{r}, t)$  is a set of arbitrary functions analogous to the  $\omega_i(\mathbf{r}, t)$  of (D.43). The gauge principle then leads to the equation of motion

$$i \left( \frac{\partial}{\partial t} + ig \sum_i \hat{I}_i^W \phi_i^W \right) \Psi = -i\boldsymbol{\alpha} \cdot \left( \nabla - ig \sum_i \hat{I}_i^W \mathbf{W}_i \right) \Psi \quad (\text{D.59})$$

analogous to (D.44), where the leptons are assumed to have zero mass and the gauge fields transform as

$$\phi_i^W \rightarrow \phi'_i = \phi_i^W + \frac{\partial f_i}{\partial t} + g \sum_{jk} \varepsilon_{ijk} f_j \phi_k^W \quad (\text{D.60a})$$

and

$$\mathbf{W}_i \rightarrow \mathbf{W}'_i = \mathbf{W}_i - \nabla f_i + g \sum_{jk} \varepsilon_{ijk} f_j \mathbf{W}_k \quad (\text{D.60b})$$

for infinitesimal  $f_i(\mathbf{r}, t)$ , by analogy with (D.46).

We have now arrived at a set of gauge-invariant equations of motion incorporating interaction terms of a specific form. Their interpretation is best brought out by considering neutrino and electron states separately. For neutrino states  $\Psi \equiv \psi(\mathbf{r}, t)\nu$ , (D.59) becomes

$$\begin{aligned} i \left( \frac{\partial}{\partial t} + i \frac{g}{2} \phi_3^W \right) \psi(\mathbf{r}, t)\nu - \frac{g}{2} (\phi_1^W + i\phi_2^W) \psi(\mathbf{r}, t)e \\ = -i\boldsymbol{\alpha} \cdot \left( \nabla - i \frac{g}{2} \mathbf{W}_3 \right) \psi(\mathbf{r}, t)\nu - \frac{g}{2} \boldsymbol{\alpha} \cdot (\mathbf{W}_1 + i\mathbf{W}_2) \psi(\mathbf{r}, t)e \\ \text{or} \\ i \left( \frac{\partial}{\partial t} + \boldsymbol{\alpha} \cdot \nabla \right) \psi(\mathbf{r}, t)\nu = \frac{g}{2} (\phi_3^W - \boldsymbol{\alpha} \cdot \mathbf{W}_3) \psi(\mathbf{r}, t)\nu \\ + 2g_W (\phi^+ - \boldsymbol{\alpha} \cdot \mathbf{W}^+) \psi(\mathbf{r}, t)e, \end{aligned} \quad (\text{D.61a})$$

where we have separated out the interaction terms and introduced the linear combinations

$$\phi^\pm = \frac{1}{\sqrt{2}} (\phi_1^W \pm i\phi_2^W), \quad \mathbf{W}^\pm = \frac{1}{\sqrt{2}} (\mathbf{W}_1^W \pm i\mathbf{W}_2^W) \quad (\text{D.62})$$

and the coupling constant

$$g_W = g/(2\sqrt{2}) \quad (\text{D.63})$$

for later convenience. The corresponding equation for the electron case is

$$\begin{aligned} i \left( \frac{\partial}{\partial t} + \boldsymbol{\alpha} \cdot \boldsymbol{\nabla} \right) \psi(\mathbf{r}, t) e &= -\frac{g}{2} (\phi_3^W - \boldsymbol{\alpha} \cdot \mathbf{W}_3) \psi(\mathbf{r}, t) e \\ &\quad + 2g_W (\phi^- - \boldsymbol{\alpha} \cdot \mathbf{W}^-) \psi(\mathbf{r}, t) \nu. \end{aligned} \quad (\text{D.61b})$$

The interpretation of these equations is now relatively straightforward. If we consider (D.61a) for the neutrino, we see that the term proportional to  $g_W$  couples it to an electron, while the corresponding term in (D.61b) for the electron couples it to a neutrino. These are just the charged current interactions required by experiment, involving the transitions  $\nu \rightarrow e^- W^+$ ,  $e^- \rightarrow \nu W^-$ , etc. In addition, (D.61a) and (D.61b) contain neutral current interactions involving  $(\phi_3, \mathbf{W}_3)$ , which couple the electron to an electron and a neutrino to a neutrino, corresponding to transitions  $e^- \rightarrow e^- W^0$  and  $\nu \rightarrow \nu W^0$ . These terms are of the same form and strength (apart from a factor of  $\sqrt{2}$ ) as the charged current interactions. This is not what is required by experiment, as emphasised in Section 10.2.1.

### D.7.3 The unification condition

The problem of the erroneous neutral current terms in (D.61) is resolved by unification with the electromagnetic interaction. To do this, we incorporate the electric charge  $q = Qe$  into the theory by introducing the *weak hypercharge*  $Y^W$  defined by

$$Q \equiv I_3^W + Y^W, \quad (\text{D.64})$$

which is similar to the definition of ordinary hypercharge  $Y$  in (6.5). (Note, however, that the definitions of  $Y$  and  $Y^W$  differ by a factor of 2.) We then treat the weak hypercharge as a source. That is, we require gauge invariance under the transformations

$$\Psi \rightarrow \Psi' = \exp \left[ -ig' Y^W \omega(\mathbf{r}, t) \right] \Psi, \quad (\text{D.65})$$

where  $g'$  is a coupling constant to be determined and  $\omega(\mathbf{r}, t)$  is an arbitrary function. For the neutrino and electron ( $\nu_e, e^-$ ), we have  $I_3^W = (\frac{1}{2}, -\frac{1}{2})$  by (D.56), so that  $Y^W = (-\frac{1}{2}, -\frac{1}{2})$  and (D.65) reduces to the transformation

$$\Psi \rightarrow \Psi' = \exp [ig' \omega(\mathbf{r}, t)/2] \Psi \quad (\text{D.66})$$

in both cases. This is similar in form to the electromagnetic gauge transformation<sup>16</sup> (D.8) that led, via the gauge principle, to the equation of motion (D.18). Equation (D.66) leads to an analogous equation of motion

$$i \left( \frac{\partial}{\partial t} - i \frac{g'}{2} \phi^B \right) \Psi = -i \boldsymbol{\alpha} \cdot \left( \nabla + i \frac{g'}{2} \mathbf{B} \right) \Psi \quad (\text{D.67a})$$

or more conveniently

$$i \left( \frac{\partial}{\partial t} + \boldsymbol{\alpha} \cdot \nabla \right) \Psi = -i \frac{g'}{2} (\phi^B - i \boldsymbol{\alpha} \cdot \mathbf{B}) \Psi, \quad (\text{D.67b})$$

where the leptons are again assumed to have zero mass and the new gauge fields  $\phi^B(\mathbf{r}, t), \mathbf{B}(\mathbf{r}, t)$  transform as

$$\phi^B \rightarrow \phi'^B = \phi^B + \frac{\partial \omega}{\partial t}, \quad \mathbf{B} \rightarrow \mathbf{B}' = \mathbf{B} - \nabla \omega, \quad (\text{D.68})$$

by analogy with (D.7).

The next step is to identify the electromagnetic fields  $(\phi, \mathbf{A})$  with appropriate linear combinations of the fields  $(\phi_3^W, \mathbf{W}_3)$  and  $(\phi^B, \mathbf{B})$ , as outlined in Section 10.2. Since we require gauge invariance under both (D.58) and (D.66) – so far we have treated them separately – we must add the interaction terms in (D.67) resulting from (D.66) to the interaction terms in (D.61a) and (D.61b) resulting from (D.58). For the electron case  $\Psi = \psi(\mathbf{r}, t)e$ , this gives

$$\begin{aligned} i \left( \frac{\partial}{\partial t} + \boldsymbol{\alpha} \cdot \nabla \right) \psi(\mathbf{r}, t)e &= 2g_W (\phi^- - \boldsymbol{\alpha} \cdot \mathbf{W}^-) \psi(\mathbf{r}, t)\nu \\ &\quad - \frac{1}{2} (g\phi_3^W + g'\phi^B) \psi(\mathbf{r}, t)e + \frac{1}{2} \boldsymbol{\alpha} \cdot (g\mathbf{W}_3 + g'\mathbf{B}) \psi(\mathbf{r}, t)e. \end{aligned} \quad (\text{D.69})$$

If we write the electromagnetic field  $\mathbf{A}$  and the  $Z^0$  field  $\mathbf{Z}$  as arbitrary linear combinations

$$\begin{aligned} \mathbf{A}(\mathbf{r}, t) &= \mathbf{B}(\mathbf{r}, t) \cos \theta_W + \mathbf{W}_3(\mathbf{r}, t) \sin \theta_W \\ \text{and} \end{aligned} \quad (\text{D.70a})$$

$$\mathbf{Z}(\mathbf{r}, t) = -\mathbf{B}(\mathbf{r}, t) \sin \theta_W + \mathbf{W}_3(\mathbf{r}, t) \cos \theta_W,$$

with analogous equations for the scalar fields  $(\phi, \phi^Z)$ ,

$$\begin{aligned} \phi(\mathbf{r}, t) &= \phi^B(\mathbf{r}, t) \cos \theta_W + \phi_3^W(\mathbf{r}, t) \sin \theta_W \\ \text{and} \end{aligned} \quad (\text{D.70b})$$

$$\phi^Z(\mathbf{r}, t) = -\phi^B(\mathbf{r}, t) \sin \theta_W + \phi_3^W(\mathbf{r}, t) \cos \theta_W,$$

---

<sup>16</sup>In the unified theory, the electromagnetic gauge transformations (D.8) are not imposed directly, but emerge indirectly as an appropriate combination of (D.58) and (D.66). This is because the electric charge is not introduced directly, but indirectly as the combination (D.64) of  $Y^W$  and  $I_3^w$ .

then (D.69) becomes

$$\begin{aligned} i \left( \frac{\partial}{\partial t} + \boldsymbol{\alpha} \cdot \boldsymbol{\nabla} \right) \psi(\mathbf{r}, t) e &= 2g_W (\phi^- - \boldsymbol{\alpha} \cdot \mathbf{W}^-) \psi(\mathbf{r}, t) \nu \\ &\quad - \frac{1}{2} (g \sin \theta_W + g' \cos \theta_W) (\phi - \boldsymbol{\alpha} \cdot \mathbf{A}) \psi(\mathbf{r}, t) e \quad (\text{D.71a}) \\ &\quad - \frac{1}{2} (g \cos \theta_W - g' \sin \theta_W) (\phi^Z - \boldsymbol{\alpha} \cdot \mathbf{Z}) \psi(\mathbf{r}, t) e. \end{aligned}$$

The same argument for the neutrino case  $\Psi = \psi(\mathbf{r}, t)\nu$ , using (D.61a), (D.67) and (D.70), gives

$$\begin{aligned} i \left( \frac{\partial}{\partial t} + \boldsymbol{\alpha} \cdot \boldsymbol{\nabla} \right) \psi(\mathbf{r}, t) \nu &= 2g_W (\phi^+ - \boldsymbol{\alpha} \cdot \mathbf{W}^+) \psi(\mathbf{r}, t) e \\ &\quad + \frac{1}{2} (g \sin \theta_W - g' \cos \theta_W) (\phi - \boldsymbol{\alpha} \cdot \mathbf{A}) \psi(\mathbf{r}, t) \nu \quad (\text{D.71b}) \\ &\quad + \frac{1}{2} (g \cos \theta_W + g' \sin \theta_W) (\phi^Z - \boldsymbol{\alpha} \cdot \mathbf{Z}) \psi(\mathbf{r}, t) \nu. \end{aligned}$$

We can now identify the conditions whereby the electromagnetic interactions in (D.71a) and (D.71b), proportional to the electromagnetic field combinations  $(\phi - \boldsymbol{\alpha} \cdot \mathbf{A})$ , coincide with those of quantum electrodynamics. They are

$$g \sin \theta_W = g' \cos \theta_W,$$

which guarantees that the electromagnetic field does not couple directly to neutrinos, and

$$\frac{1}{2}(g \sin \theta_W + g' \cos \theta_W) = e,$$

which ensures that the coupling of the electromagnetic field has the strength required in QED, as in (D.20). These equations are equivalent to the unification condition

$$e = g \sin \theta_W = g' \cos \theta_W \quad (\text{D.72})$$

or alternatively

$$\frac{e}{2\sqrt{2}} = g_W \sin \theta_W = g_Z \cos \theta_W,$$

where

$$g_W = \frac{g}{2\sqrt{2}}, \quad g_Z = \frac{g'}{2\sqrt{2}}. \quad (\text{D.73})$$

This is identical with the unification condition (10.8), except that we are using Heavyside–Lorentz rather than SI electromagnetic units. When it is imposed, the electromagnetic interactions coincide with those used in QED with such spectacular success, while the neutral current interactions proportional to  $(\phi^Z - \boldsymbol{\alpha} \cdot \mathbf{Z})$  are uniquely predicted in terms of the electric charge  $e$  and the weak mixing angle  $\theta_W$ . However, before discussing this further we must take the spin structure of the interaction into account.

### D.7.4 Spin structure and parity violation

For zero-mass fermions, the Dirac equation (D.16) can be shown to separate into two independent equations<sup>17</sup>

$$i \frac{\partial \psi^L}{\partial t} = -i \boldsymbol{\alpha} \cdot \boldsymbol{\nabla} \psi^L, \quad i \frac{\partial \psi^R}{\partial t} = -i \boldsymbol{\alpha} \cdot \boldsymbol{\nabla} \psi^R, \quad (\text{D.74})$$

where  $\psi^{L,R}$  are Dirac spinors corresponding to the left and right-handed spin states of Figure 11.3, respectively. Hence in discussing the spin structure of weak interactions we can treat the left- and right-handed spin states completely separately.

In Chapter 11 we saw that when fermion masses can be neglected, the charged current weak interactions involve left-handed spin states only, in violation of parity conservation. We therefore assume that the preceding Sections D.1, D.2 and D.3 apply to left-handed spin states only. That is, (D.71) is assumed to hold for left-handed spin states only. It now remains to incorporate the right-handed spin states. To do this, we note that (D.71) has been obtained by combining the left-handed spin states of electrons and neutrinos into weak isodoublets.

$$\Psi^L = \psi^L(\mathbf{x}, t) \chi^f,$$

with  $I_3^W = 1/2, -1/2$  for neutrinos and electrons, respectively, and  $Y^W = -1/2$  in both cases by (D.64). Right-handed states are instead assigned to weak isosinglets with  $I_i^W = 0 (i = 1, 2, 3)$ . In this case, both electrons and neutrinos are described by separate Dirac spinors  $\psi_e^R(\mathbf{r}, t), \psi_\nu^R(\mathbf{r}, t)$  and the gauge transformations (D.58) reduce to

$$\psi_e^R \rightarrow \psi_e'^R = \psi_e^R, \quad \psi_\nu^R \rightarrow \psi_\nu'^R = \psi_\nu^R.$$

In other words, the wavefunctions remain unchanged, and no interactions with  $W$  bosons of the type discussed in Section D.7.2 are generated by the gauge principle. However, by (D.64),  $Y^W = (0, -1)$  for the neutrino and the electron, respectively, so that the gauge transformations (D.65) become

$$\psi_\nu^R \rightarrow \psi_\nu'^R = \psi_\nu^R, \quad \psi_e^R \rightarrow \psi_e'^R = e^{ig'\omega} \psi_e^R.$$

The argument of Section (D.7.3) then leads to the equations

$$i \left( \frac{\partial}{\partial t} + \boldsymbol{\alpha} \cdot \boldsymbol{\nabla} \right) \psi_\nu^R(\mathbf{r}, t) = 0 \quad (\text{D.75a})$$

and

$$i \left( \frac{\partial}{\partial t} + \boldsymbol{\alpha} \cdot \boldsymbol{\nabla} \right) \psi_e^R(\mathbf{r}, t) = -g'(\phi^B - \boldsymbol{\alpha} \cdot \mathbf{B}) \psi_e^R(\mathbf{r}, t)$$

---

<sup>17</sup>See, for example, Section 5.6 of Kane (1987).

in analogy to (D.67). The second of these equations becomes

$$i \left( \frac{\partial}{\partial t} + \boldsymbol{\alpha} \cdot \boldsymbol{\nabla} \right) \psi_e^R = -e (\phi - \boldsymbol{\alpha} \cdot \mathbf{A}) \psi_e^R + e \tan \theta_W (\phi^Z - \boldsymbol{\alpha} \cdot \mathbf{Z}) \psi_e^R, \quad (\text{D.75b})$$

using the mixing equations (D.70) and the unification condition (D.72). The first term on the right-hand side of (D.75b) corresponds to the usual electromagnetic interaction, while the second term is a weak neutral current interaction. In other words, for electrons (but not neutrinos) both right-handed and left-handed spin states are predicted to have neutral current interactions.

Equations (D.71) and (D.75) represent the final predictions for the electroweak interactions of electrons and neutrinos. Essentially, the gauge transformations have been chosen throughout to reproduce the known properties of the electromagnetic and weak charged current interactions. The form of the neutral current interactions, including their spin structure, is predicted uniquely in terms of the mixing angle  $\theta_W$ . The agreement between these predictions and the measured properties of neutral current interactions is one of the great triumphs of the unified theory.

Finally, we note that this analysis is based on the fact that the Dirac equation can be separated into two equations (D.74) in the limit of zero fermion mass. This enabled us to ascribe different gauge transformation properties to the left- and right-handed states, which are treated completely separately. If finite masses are assumed, then the left- and right-handed states do not decouple, and the analysis breaks down. This implies that, in the absence of a Higgs field, different interactions for the left- and right-handed states (i.e. parity violation) can only be incorporated into a gauge theory having massless fermions. This result can, however, be avoided if the fermion masses, like the  $W^\pm$  and  $Z^0$  masses, arise from interactions with the Higgs field.

## Problems D

**D.1** Show that Hamilton's equations of motion

$$\dot{x}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial x_i} \quad (i = 1, 2, 3)$$

lead to the equations of motion (D.5) for a charged particle in an electromagnetic field if the Hamiltonian is of the form (D.3), where  $\mathbf{r} = (x_1, x_2, x_3) \equiv (x, y, z)$ .

**D.2** Show that Maxwell's equations in free space (D.9) reduce to (D.10) when expressed in terms of the potentials (D.4).

- D.3** Show that for any electromagnetic potentials  $(\phi, \mathbf{A})$  there are many possible choices of gauge-transformed potentials

$$\phi = \tilde{\phi} + \frac{\partial f}{\partial t}, \quad \mathbf{A} = \tilde{\mathbf{A}} - \nabla f,$$

such that the Lorentz condition (D.11) is satisfied.

- D.4** Verify that the Dirac equation in the presence of an electromagnetic field (D.18) is gauge invariant, as stated in the text.

- D.5** Show that the gauge transformation (D.8) used in our discussion of QED can be written as two successive electroweak gauge transformations of the form (D.58) and (D.66).

- D.6** Our treatment of the electroweak interactions of massless leptons may be extended to massless quarks by invoking lepton–quark symmetry between the doublets  $(\nu_e, e^-)$  and  $(u, d)$  where we neglect quark mixing.

(a) What are the appropriate weak isospin and hypercharge quantum numbers  $(I_3^W, Y^W)$  for  $u$  and  $d$  quarks in both left-handed and right-handed states?

(b) Show that with these assignments, the unification condition (D.72) leads to the appropriate electromagnetic interactions for  $(u, d)$  quarks. (*Hint:* This does not require complicated algebra, but just appropriate modifications of the results of Sections D.7.3 and D.7.4 to take account of the different  $(I_3^W, Y^W)$  values.)

- D.7** How would you modify your answer to Problem D.6 to take account of mixing between the quark doublets  $(u, d)$  and  $(c, s)$ ? In particular, obtain the equations analogous to (D.71a) and (D.71b) for left-handed  $d$  and  $s$  quarks, instead of electrons. Hence confirm that the charged current couplings are consistent with (8.17) and that there are no terms corresponding to strangeness-changing neutral currents, in agreement with the discussion of Section 10.1.1.

---

## Answers to selected questions

---

Full solutions to all problems are available on the book's website. Here we give the final result for some selected problems where it is very brief.

### Problems 1

- 1.1** (a)  $\bar{\nu}_e + e^+ \rightarrow \bar{\nu}_e + e^+$ ; (b)  $p + n \rightarrow p + n$ ; (c)  $p + \bar{p} \rightarrow e^+ + e^- + \gamma$ .  
**1.8**  $r \approx \hbar c/mc^2 = 368 \text{ fm}$ .  
**1.10**  $\tau = 1.245 \times 10^{-10} \text{ s}$ .  
**1.11**  $\sigma = 1.40 \times 10^{-10} \text{ mb} = 0.14 \text{ pb}$ .  
**1.12** (a)  $r_c = 2.8 \text{ fm}$ . (b)  $R = \frac{3}{10} \frac{\alpha}{m} = \frac{3r_c}{m}$ .

### Problems 2

- 2.1** Reactions (a) and (d) are forbidden; (b) and (c) are allowed.  
**2.8**  $E \geq 3.8 \text{ MeV}$ .  
**2.9**  $L_{13} \approx 1.2 \text{ km}, L_{12} \approx 31.5 \text{ km}$ . At 10 km a two-component mixing model would fail badly.  
**2.10**  $0 \leq \Delta(m^2 c^4) \leq 3.8 \times 10^{-3} (\text{eV})^2$ .  
**2.11**  $\lambda \approx 7 \times 10^{12} \text{ km}$ , i.e. about  $10^7$  times the solar radius.

### Problems 3

- 3.1** Reactions (a), (b) and (c) are forbidden; (d) is allowed.  
**3.2** (a) weak, (b) electromagnetic, (c) strong, (d) weak, (e) strong, (f) electromagnetic.  
**3.3** Mean distance travelled is 17 m.  
**3.5**

$$(2, 1, 0, 1, 0) = cuu, \quad (0, 1, -2, 1, 0) = css, \quad (0, 0, 1, 0, -1) = b\bar{s}, \\ (0, -1, 1, 0, 0) = \bar{s}\bar{d}u, \quad (0, 1, -1, 1, 0) = csd, \quad (-1, 1, -3, 0, 0) = sss.$$

- 3.6** Reactions (b) and (d) are allowed; (a) and (c) are forbidden.
- 3.7**  $\tau(\Lambda\gamma) \approx (10^{-19} - 10^{-22})$  s.
- 3.8** Decay modes:  $\Sigma^+ \rightarrow n\pi^+, p\pi^0$ . Lifetime in the range  $(10^{-7} - 10^{-13})$  s.
- 3.12**

$$\begin{array}{cc} \text{Baryons} & \text{Mesons} \\ \left[ \begin{array}{cccccc} C & 3 & 2 & 1 & 0 \\ Q & 2 & 2, 1 & 2, 1, 0 & 2, 1, 0, -1 \end{array} \right] & \left[ \begin{array}{cccccc} C & 1 & 0 & -1 \\ Q & 1, 0 & 1, 0, -1 & 0, -1 \end{array} \right] \end{array}$$

- 3.13** First two quantum number combinations are compatible; second two are not.

## Problems 4

- 4.1**  $L = 2.4$  m.
- 4.4** Luminosity  $L = 6.44 \times 10^{31} /(\text{cm}^2 \text{s})$ .
- 4.5** 848 photons per second.
- 4.6** Distance = 8.3 m.
- 4.7** Thickness of iron required is 1 m.
- 4.8** Effective interaction length = 2.3 m.
- 4.9** Final energy  $E = 1.51$  GeV. For muons, energy losses are negligible at this energy.
- 4.11**  $E_p = \frac{1}{2}m_p v^2 = m_p I e^{5/3} / 4m_e = 632$  keV.
- 4.12** Minimum flight path is  $L \approx 4.8$  m.
- 4.13** Minimum detector length 7.5 m.
- 4.14** Multiplication factor is  $2^{22} = 4.2 \times 10^6 = 10^{6.6}$ .
- 4.18** Total cross-section is 2.406 nb; total number of events produced is  $2.41 \times 10^5$ .

## Problems 5

- 5.6**  ${}^1S_0$ ,  ${}^3S_1$ , with  $J^P = 0^+$  and  $1^+$ , respectively.
- 5.8**  $P_\eta = -1$ ;  $\pi\pi$  decays violate parity conservation.
- 5.9** (a)  ${}^1S_0$  and  ${}^3S_1$ , with  $J^{PC} = 0^{-+}$  and  $1^{--}$ , respectively; (b)  ${}^1P_1$ ,  ${}^3P_2$ ,  ${}^3P_1$  and  ${}^3P_0$  with  $J^{PC}$  values  $1^{+-}$ ,  $2^{++}$ ,  $1^{++}$  and  $0^{++}$ , respectively.
- 5.10**  $P = (-1)^2 P_\pi^2 = 1$ ;  $C = 1$ .
- 5.12**  ${}^3S_1 \rightarrow {}^1P_1$ ,  ${}^3P_0 \rightarrow {}^1S_0$ ,  ${}^3P_1 \rightarrow {}^1P_1$ ,  ${}^3P_2 \rightarrow {}^1D_2$ . No further restrictions from  $C$  invariance.
- 5.14** Ground state energy is 13.6 eV;  ${}^3S_1 - {}^3S_0$  energy splitting is  $1.72 \times 10^{-5}$  eV.
- 5.15** Ratio is 0.356.

## Problems 6

- 6.2** Quantum numbers are  $B = 1, Q = 1, S = 0, C = 1, \tilde{B} = 0$  and  $T = 0$ , so quark content  $\Sigma_c^+(2455) = udc$ . It has partners  $\Sigma_c^{++} = uuc$  and  $\Sigma_c^0 = ddc$ .
- 6.3**  $\delta = 1.5 \pm 0.1$  MeV/ $c^2$  and a distance  $r$  of order 1 fm.
- 6.4**  $f_2^0 \rightarrow \pi^0\gamma$  violates  $C$  parity and hence is forbidden as an electromagnetic interaction;  $\rho^0 \rightarrow \pi^0\pi^0$  cannot satisfy angular momentum conservation and the Pauli principle combined and hence is absolutely forbidden.
- 6.7** Eight  $J^P = 1/2^+$  baryons and one with  $J^P = 3/2^+$ .
- 6.10**  $\mu_{\Sigma^0} = 0.83 \mu_N$ .
- 6.14**  $\Gamma_e = 1.05$  keV, and by universality  $\Gamma_\mu = \Gamma_e$ ;  $\Gamma = 31.5$  keV.

## Problems 7

- 7.2**  $V(0.001 \text{ fm})/V(0.1 \text{ fm})$  is of order 1/3.
- 7.3**  $|V(pp)/V(q\bar{q})| \simeq 1/30$  at  $r = 3 \text{ fm}$ .
- 7.5** The  $q^2$  values are of order  $-4 \times 10^{-2} (\text{GeV}/c)^2$  and  $-4 \times 10^4 (\text{GeV}/c)^2$ , and the ratios of the invariant scattering amplitudes are of order  $2 \times 10^5$  and unity, for distances of closest approach of 1 fm and  $10^{-3} \text{ fm}$ .
- 7.10**  $R \approx 2.17, 3.58$  and  $3.89$  at  $E_{CM} = 2.8, 5$  and  $15 \text{ GeV}$ , respectively. When  $E_{CM}$  is above the  $t\bar{t}$  threshold,  $R = 5(1 + \alpha_s/\pi)$ .

## Problems 8

- 8.6** The fraction is approximately 0.5. The ‘missing momentum’ is ascribed to be carried by gluons.
- 8.10** (a)  $\sigma_{\max} = 760 \text{ nb}$ , (b)  $\sigma_{p\bar{p}} = \frac{\pi M \Gamma}{s} \sigma_{\max} \int_0^1 \frac{u(x_u) d(M^2/x_u s)}{x_u} dx_u$ .

## Problems 9

- 9.2**  $\Gamma(c \rightarrow se^+ \nu_e) = (1.7 - 3.3) \times 10^{11} \text{ s}^{-1}$ .
- 9.4** (a) Cabibbo-allowed, (b) forbidden, (c) forbidden and (d) Cabibbo-suppressed.
- 9.5** (a)  $B(b \rightarrow c + e^- + \bar{\nu}_e) = 1/9$ , (b)  $B(\tau^- \rightarrow e^- + \bar{\nu}_e + \nu_\tau) = 1/5$ .
- 9.7** Decays (b), (d) and (f) are allowed.
- 9.8** Predicted ratio is 2.
- 9.11**  $|V_{ub}| \leq (3.4 \pm 0.1) \times 10^{-2}$  and  $|V_{cb}| \leq (4.9 \pm 0.2) \times 10^{-2}$ .
- 9.12** Radius of ground state  $a = 1.7 \times 10^{-2} \text{ fm}$ ; orbit time  $t = 2.7 \times 10^{-24} \text{ s} \gg$  lifetime of top quark ( $\approx 4 \times 10^{-25} \text{ s}$ ).

## Problems 10

- 10.2** (a) Weak; (c) electromagnetic; (b) and (d) forbidden.
- 10.3**

$$\frac{\sum_\ell \Gamma(K^+ \rightarrow \pi^+ \nu_\ell \bar{\nu}_\ell)}{\Gamma(K^+ \rightarrow \pi^0 \mu^+ \nu_\mu)} \sim \mathcal{O}(10^{-11}).$$

- 10.5**  $B(Z^0 \rightarrow c\bar{c} = 0.12)$ ;  $B(Z^0 \rightarrow b\bar{b} = 0.15)$ ;  $\Gamma(Z^0 \rightarrow \ell^+ \ell^-) = 0.249 \text{ GeV}$ ;  $\Gamma(Z^0 \rightarrow \text{hadrons}) = 1.632 \text{ GeV}$ . The calculated decay rate to charged leptons is in very good agreement with the measured value (10.29b); the calculated decay rate to hadrons has an error of about 7% compared to the experimental value given in (10.29a), which probably arises mainly from gluon emission.

- 10.6**  $A = -0.07$  and  $\sigma_{\text{tot}} = \sigma_\gamma (1 + 2.2 \times 10^{-3})$ , i.e. about 0.2% correction.

- 10.7**

$$\sigma_{H^0}(e^+ e^- \rightarrow \mu^+ \mu^-) \approx \frac{1 \times 10^{-11}}{M_H^2},$$

whereas

$$\sigma_\gamma(e^+e^- \rightarrow \mu^+\mu^-) \approx \frac{2 \times 10^{-4}}{M_H^2}; \quad \sigma_{Z^0}(e^+e^- \rightarrow \mu^+\mu^-) \approx \frac{4 \times 10^{-2}}{M_Z^2}.$$

- 10.8** Instantaneous luminosity must be  $3.1 \times 10^{34}/(\text{cm}^2 \text{s})$ . Selecting events with  $b$ -quark jets present increases the statistical error by about 20% and improves the signal-to-background ratio by a factor of about 2.

## Problems 11

- 11.2**  $\alpha = -1$ .
- 11.3** The most probable energy is  $E_e = m_\mu/2$  and the figure coincides with Figure 11.5(a). The lifetime is  $\tau = \hbar/\Gamma \approx 2.2 \times 10^{-6} \text{ s}$ .
- 11.6** Distance travelled is  $d = 1.4 \times 10^{-2} \text{ cm}$ .
- 11.7** (a), (b) and (d) are  $B^0$  decays; (c) could be either  $B^0$  or  $\bar{B}^0$ .
- 11.10** The largest contribution comes from the penguin diagrams of Figure 11.15(b), and the tree diagram is suppressed relative to it by a factor of order 10.

## Problems 12

- 12.2**  $(B - L)$  is conserved provided the value  $-2/3$  is assigned to both  $X$  and  $Y$  bosons.
- 12.3** The appropriate relabellings are
- $$(a) \quad e^+ \rightarrow \bar{\nu}_e, \quad X \rightarrow Y, \quad \bar{u} \rightarrow \bar{d}; \quad (b) \quad e^+ \rightarrow \bar{\nu}_e, \quad \bar{u} \rightarrow \bar{d}$$
- for Figures 12.4(a) and (b), respectively.
- 12.4** The positron energy is  $\varepsilon = (m_p^2 - m_\pi^2)/2m_p = 459 \text{ MeV}$ . The angle of emission  $\theta = 41^\circ$ .
- 12.5** (a)  $v \propto R^{-1/2}$ ; (b)  $v \propto R$ . The observed behaviour implies the mass enclosed by a radius  $R$  is roughly proportional to  $R$ .
- 12.6** (a)  $E_P \geq 0.4 \times 10^{10} \text{ GeV}$ ; (b)  $E_P \geq E$ , where  $E \approx 1.4 \times 10^{13} \text{ GeV}$ .
- 12.7** The local density of dark matter implied is  $\rho_L \approx 0.3 (\text{GeV}/c^2)/\text{cm}^3$ , whereas the average dark matter density in the universe is  $\rho_m = \rho_c \Omega_m \approx 1.4 \times 10^{-6} (\text{GeV}/c^2)/\text{cm}^3$ .
- 12.8** Maximum value occurs when  $m_A = m_w$ . Then  $T_{\max} \approx 16 \text{ keV}$ .
- 12.9**  $\mathbf{B} \rightarrow \mathbf{B}$  and  $\mathbf{B} \rightarrow -\mathbf{B}$  under  $P$  and  $T$ , respectively.
- 12.10**  $m < 15 \text{ eV}/c^2$ .
- 12.11** 90 kg of  ${}^{76}\text{Ge}$ .

## Problems A

- A.1**  $E_\mu = \frac{(m_\pi^2 + m_\mu^2 - m_\nu^2)c^2}{2m_\pi}$ .
- A.3** Threshold energy is 0.908 GeV, corresponding to a threshold momentum of 0.897  $\text{GeV}/c$ .
- A.4** The decay is  $\Lambda \rightarrow p\pi^-$ .

## Problems B

- B.1** Production rate is 1.70 /s.
- B.2** (a) Attenuation factor:  $8.6 \times 10^{-5}$ , (b) inelastic collision rate: 51 /s, (c) flux:  $0.11 /(\text{m}^2 \text{s})$ .

## Problems C

- C.4 The isospin is  $I = 1/2$ .  
C.5 All three branching ratios are 14%.

## Problems D

- D.6 First part only: left-handed:  $(u, d)$  states have  $I_3^W(u, d) = (\frac{1}{2}, -\frac{1}{2})$  and  $Y^W = (\frac{1}{6}, \frac{1}{6})$ ; right-handed:  $(u, d)$  states have  $I_3^W(u, d) = (0, 0)$  and  $Y^W = (\frac{2}{3}, -\frac{1}{3})$ .

---

## References

---

- G. Aad *et al.*, (ATLAS Collaboration), *Physics Letters* **B726** (2013) 88
- G. Aad *et al.*, (ATLAS Collaboration), CERN-PH-EP-2015-125, arXiv:1507.04548v1. *European Physical Journal C* **76** (2016) 6
- R. Aaij *et al.*, *Physical Review Letters* **115** (2015) 072001
- V.M. Abasov *et al.*, (D0 Collaboration) (2016) arXiv 1602.07588
- F. Abe *et al.*, *Physical Review Letters* **74** (1995) 2626
- K. Abe *et al.*, *Physical Review Letters* **48** (1982) 1526
- K. Abe *et al.*, *Physical Review Letters* **111** (2013) 211803
- Y. Abe *et al.*, (Double Chooz Collaboration), *Physical Review* **D86** (2012) 052008
- M. Ablikim *et al.*, (BESIII Collaboration), *Physical Review Letters* **110**, 252001 (2013)
- I. Adachi *et al.*, (Belle Collaboration) *Physical Review Letters* **108** (2012) 171802, arXiv:1201.4643
- P.A.R. Ade *et al.*, (Planck Collaboration), *Astronomy and Astrophysics* **571** (2014) A16 arXiv:1303.5076
- A. Ajaib *et al.*, (Higgs Working Group Report), arXiv:1310.8361 (2014)
- M.Z. Akrawy *et al.*, *Physics Letters* **B240** (1990) 497
- J.B. Albert *et al.*, (EXO-200 Collaboration), *Nature* **510** (2014) 229–234, arXiv:1402.6956
- M.H. Alston *et al.*, *Physical Review Letters* **6** (1961) 520
- C. Amsler, *Reviews of Modern Physics* **70** (1998) 1293
- C. Amsler *et al.*, *Quark Model* (2013), Section 15 of K.A. Olive *et al.*, (Particle Data Group) (2014)
- F.P. An *et al.*, (Daya Bay Collaboration), *Chinese Physics* **C37** (2013) 011001
- C.D. Anderson, *Physical Review* **43** (1933) 491
- E.N. da C. Andrade, *Sir Isaac Newton*, Fontana Books (1961)
- A. Angelopoulos *et al.*, (CPLEAR Collaboration), *European Physical Journal* **C22** (2001) 55
- K. Asakura *et al.*, (KamLAND-Zen Collaboration), *Results from KamLAND-Zen*, arXiv:1409.0077 (2014)
- Y. Ashie, J. Hosaka, Y. Itow *et al.*, (SuperKamiokande Collaboration), *Physical Review Letters* **93** (2004) 101801
- B. Aubert *et al.*, *Physical Review* **D73** (2006) 012004
- M. Auger *et al.*, (EXO Collaboration), *Journal of Instrumentation* **7** (2012) P05010 (arXiv:1202.2192)
- A.B. Balantekin and W.C. Haxton, *Progress in Particle and Nuclear Physics* **71** (2013) 150
- V. Barnes *et al.*, *Physical Review Letters* **12** (1964) 204
- O. Behnke, A. Geiser and M. Lisovyi, *Progress in Particle and Nuclear Physics* **84** (2015) 1; arXiv:1506.07519 (2015)
- H.J. Behrend *et al.*, *Physics Letters* **183B** (1987) 400
- S. Bethke, G. Dissertori and G.P. Salam, *Quantum Chromodynamics* (2015), Section 9 of K.A. Olive *et al.*, (Particle Data Group) (2014, updated 2015)
- I.I. Bigi and A.I. Sanda, *CP Violation*, 2nd edition, *Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology No. 28*, Cambridge University Press (2009)
- P.E. Bosted *et al.*, *Physical Review Letters* **68** (1992) 3841
- R. Brandelik *et al.*, *Physics Letters* **B76** (1978) 361
- R. Brandelik *et al.*, *Physics Letters* **B97** (1980) 453
- W.E. Burcham and M. Jobes, *Nuclear and Particle Physics*, Longman Scientific and Technical (1995)
- M. Campbell, *Proceedings of the 1989 International Symposium on Lepton and Photon Interactions at High Energies*, Stanford University, USA

- M. Carena *et al.*, *Status of Higgs Boson Physics* (2013), Section 11 of K.A. Olive *et al.*, (Particle Data Group) (2014)
- A. Ceccucci, Z. Ligeti and Y. Sakai, *The CKM Quark-Mixing Matrix*, Section 12 of K.A. Olive *et al.*, (Particle Data Group) (2014)
- S. Chatrchyan *et al.*, (CMS Collaboration), *Journal of High Energy Physics* **10** (2011a) 007; arXiv:1108.0566
- S. Chatrchyan *et al.*, (CMS Collaboration), *Physical Review Letters* **107** (2011b) 132001
- S. Chatrchyan *et al.*, (CMS Collaboration), *Physical Review Letters* **109** (2012) 222301
- S. Chatrchyan *et al.*, (CMS Collaboration), *Physical Review* **D89** (2014) 092007
- S. Chatrchyan *et al.*, (CMS Collaboration), *European Physics Journal* **C75** (2015) 212; arXiv:1412.8662
- Y. Chen *et al.*, *Physical Review* **D73** (2006) 014516
- J.H. Christenson *et al.*, *Physical Review Letters* **13** (1964) 138
- F.E. Close, *Introduction to Quarks and Partons*, Academic Press (1979)
- T. Dafni and F.J. Iguaz, *Proceedings of Science (Technology and Instrumentation in Particle Physics)*, Amsterdam (2014), arXiv:1501.01456 (2015)
- R.H. Dalitz, *Philosophical Magazine* **44** (1953) 1068
- T. Davis, *Baryon Acoustic Oscillations*, *Physics World* **July 2014** 24
- S. Dawson, *SUSY and Such*, NATO Advanced Study Institute Series B Physics **365** (1997) 33; aeXiv:hep-ph/9612229
- A. De Angelis, *Introduction to Particle and Astroparticle Physics*, Springer (2015)
- M. Drees and G. Gerbier, *Dark Matter* (2013), Section 25 of K.A. Olive *et al.*, (Particle Data Group) (2014)
- E. Fermi, *Nuclear Physics*, University of Chicago Press (1950)
- E. Figueroa-Feliciano, *Progress in Particle and Nuclear Physics* **66** (2011) 661
- B. Foster, A.D. Martin and M.G. Vinet, *Structure Functions* (2013), Section 19 of K.A. Olive *et al.*, (Particle Data Group) (2014)
- A. Gando *et al.*, (KamLAND Collaboration), *Physical Review* **D83** (2011) 052002
- S. Gasiorowicz, *Quantum Mechanics*, John Wiley & Sons (1974)
- S. Gasiorowicz, *The Structure of Matter*, Addison-Wesley (1979)
- T. Gershon and Y. Nir, *CP Violation in the Quark Sector*, Section 13 of K.A. Olive *et al.*, (Particle Data Group) (2014)
- C. Giunti and C.W. Kim, *Fundamentals of Neutrino Physics and Astrophysics*, Oxford University Press (2007)
- S. Gjesdal *et al.*, *Physics Letters* **B52** (1974) 113
- S. Godfrey and J. Napolitano, *Reviews of Modern Physics* **71** (1999) 1411
- J.J. Gómez-Cadenas and J. Martin-Albo, *Proceedings of Science (Gran Sasso Summer Institute 2014 Hands-On Experimental Underground Physics at LNGS)*, INFN-Gran Sasso (2014), arXiv:1502.00581 (2015)
- K. Gottfried and V.F. Weisskopf, *Concepts of Particle Physics*, Vol. 2, Oxford University Press (1986)
- P.W. Graham *et al.*, *Annual Review Nuclear Particle Science* **65** (2015) 485
- I.S. Grant and W.R. Phillips, *Electromagnetism*, 2nd edition, John Wiley & Sons (1990)
- D.R. Green, *The Physics of Particle Detectors*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology No.12, Cambridge University Press (2000)
- C. Grupen and B. Shwartz, *Particle Detectors*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology No.26, Cambridge University Press (2011)
- J.F. Gunion *et al.*, *The Higgs Hunter's Guide*, Perseus Publishing, Cambridge, MA (1990)
- H.E. Haber, *Theory of the Higgs Boson: The Standard Model and Beyond*, Lectures at the Idpasc Higgs School, Foz do Arelho, Portugal (2011). www.idpasc.lip.pt/file.php/1/higgs2011/lectures/Haber.pdf
- F. Halzen and A.D. Martin, *Quarks and Leptons*, John Wiley & Sons (1984)
- I.S. Hughes, *Elementary Particles*, 2nd edition, Cambridge University Press (1985)
- J.D. Jackson, *Classical Electrodynamics*, 2nd edition, John Wiley & Sons, New York (1975)
- G.L. Kane, *Modern Elementary Particle Physics*, Addison-Wesley (1987)
- O. Klein, *Proceedings of the Symposium on Les Nouvelles Theories de la Physique*, Warsaw (1938)
- O. Klein, *Nature* **161** (1948) 897
- K. Kleinknecht, *Detectors for Particle Radiation*, 1st edition, Cambridge University Press (1986)
- K. Kleinknecht, *Detectors for Particle Radiation*, 2nd edition, Cambridge University Press (1998)
- E. Klempert and J.-M. Richard, *Reviews of Modern Physics* **82** (2010) 1095
- D. Kleppner and R.J. Kolenkow, *Introduction to Mechanics*, McGraw-Hill (1973)

- G.F. Knoll, *Radiation Detection and Measurement*, 3rd edition, John Wiley & Sons, New York (1999)
- K.S. Krane, *Introductory Nuclear Physics*, John Wiley & Sons (1988)
- O. Lahav and A.R. Liddle, *The Cosmological Parameters* (2013), Section 24 of K.A. Olive *et al.*, (Particle Data Group) (2014)
- C.M.G. Lattes *et al.*, *Nature* **159** (1947) 694
- L.M. Lederman with D. Teresi, *The God Particle: If the Universe Is the Answer, What Is the Question?* Houghton Mifflin Harcourt (1993)
- LHCb Collaboration, CERN Report CERN-LHCb-CONF-2016-004 (2016)
- Z.Q. Liu *et al.*, (Belle Collaboration), *Physical Review Letters* **110** (2013) 252002
- F. Mandl, *Quantum Mechanics*, 2nd edition, John Wiley & Sons (1992)
- F. Mandl and G. Shaw, *Quantum Field Theory*, 2nd edition, John Wiley & Sons, London (2010)
- B.R. Martin, *Nuclear and Particle Physics*, 2nd edition, John Wiley & Sons (2008)
- B.R. Martin and G. Shaw, *Mathematics for Physicists*, John Wiley & Sons (2015)
- R. Moore, *Niels Bohr*, MIT Press (1985)
- K. Nakamura and S.T. Petcov, *Neutrino Mass, Mixing and Oscillations* (2014), Section 14 of K.A. Olive *et al.*, (Particle Group) (2014)
- H. Nishino *et al.*, (Super-Kamiokande Collaboration), *Physical Review Letters* **102** (2009) 141801 (arXiv:0903.0676)
- NRC99 Report of the Board of Physics and Astronomy of the National Research Council, USA (1999) *Nuclear Physics: The Core of Matter, The Fuel of Stars*, National Academies Press, Washington, DC, USA
- K.A. Olive *et al.*, (Particle Data Group), *Chinese Physics* **C38** (2014) 090001
- S.L. Olsen, *Frontiers of Physics* **10** (2015) 101401
- S. Orito, *Proceedings of the 1979 International Symposium on Lepton and Photon Interactions at High Energies*, Fermilab (1979)
- W. Pauli, *Handbuch der Physik* **24** (1933) 246
- D.H. Perkins, *Introduction to High Energy Physics*, 3rd edition, Addison-Wesley (1987)
- D.H. Perkins, *Particle Astrophysics*, 2nd Edition, Oxford University Press (2009)
- A.C. Phillips, *The Physics of Stars*, John Wiley & Sons, Chichester (1994)
- B. Povh, K. Rith, C. Scholz and F. Zetsche, *Particles and Nuclei. An Introduction to the Physical Concepts*, Springer (1999)
- G.D. Rochester and C.C. Butler, *Nature* **160** (1947) 855
- M. Roos, *Introduction to Cosmology*, 4th edition, Wiley-Blackwell (2015)
- B. Sadoulet, *Proceedings of the 1983 International Symposium on Lepton and Photon Interactions at High Energies*, Cornell University, USA
- L.I. Schiff, *Quantum Mechanics*, 3rd edition, McGraw-Hill (1968)
- R.H. Schindler *et al.*, *Physical Review* **D21** (1980) 2716
- W.G. Seligman, FERMILAB-Thesis-1997-21 (1997)
- S.-H. Seo (RENO Collaboration), TAUP2013 International Workshop, September 9–13, 2013, Asilomar, California, USA
- A.F. Sill *et al.*, *Physical Review* **D48** (1993) 29
- M. Spiro, *Proceedings of the 1983 International Symposium on Lepton and Photon Interactions at High Energies*, Cornell University, USA
- N. Spooner, *Astronomy Now* **7** (1993) 40
- M. Spurio, *Particles and Astrophysics: A Multimessenger Approach*, Springer (2015)
- Y. Suzuki, *Physica Scripta* **T121** (2005) 33
- G.L. Trigg, *Landmark Experiments in 20th Century Physics*, Crane, Russak and Co. Ltd, New York (1975)
- J. Updike, *Telephone Poles and Other Poems*, André Deutch, London (1964)
- R.C. Walker *et al.*, *Physical Review* **D49** (1994) 5671
- D.H. Weingarten, *Scientific American* (February 1996) 104
- G.H. Wichmann, *Quantum Physics*, McGraw-Hill (1967)
- S.-L. Wu, *Physics Reports* **107** (1984) 59

---

# Index

---

- absorption length 88, 112  
accelerators 79–85  
AMS detector 366–367  
angular momentum 129–135  
    commutation relations 408–409  
anomaly condition 279  
anthropic principle 359  
antiparticles 3–9  
    relative parity of 137–138  
Askaryan effect 108  
associated production 63–65  
asymptotic freedom 195–197, 199–201  
ATLAS detector 84, 117–118, Plate 4  
axions 371–373
- BaBar detector 332–334  
baryon number 56–57  
    violation in grand unified theories 350–353  
baryons  
    in the quark model 55, 72–73, 134–135, 164–167, 175–177  
    colour wave function of 179–180, 182–184  
    exotic quantum numbers 72–75  
    exotic states 207–208  
    magnetic moments of 167–169  
    mass splittings within multiplets 161, 169–170, 172–174  
    pentaquarks 181, 207–208  
    resonances 71–72  
    tables of states 55, 166–167, 176  
beauty quantum number 56  
 $\beta$ -decay processes 26–28, 57–58, 254, 375–380  
Bethe–Bloch formula 89–92  
 $B$  factories 332–334  
big bang model 361–363  
Bjorken scaling 224–228  
    deviations from 228–231  
 $B$  mesons 65, 175  
     $B_H^0 - B_L^0$  mass difference 334  
     $B^0 - \bar{B}^0$  mixing and oscillations 329–332, 334–335  
     $CP$  violation in  $B$  decays 328–329, 335–337, 342  
Born approximation 18–19, 221, 398
- bottomonium 65, 184–185, 188–189  
bottom quantum number 56  
Bragg curve 91  
branching ratios 29, 400  
    isospin predictions of 414–416  
Breit–Wigner formula 68, 70  
    for decay distributions 401–404  
    for resonance formation 404–406  
bremsstrahlung 92–93  
bubble chamber 63–64
- Cabibbo allowed/suppressed decays 257  
Cabibbo angle 256  
Callan–Gross relation 227  
calorimeters 109–112  
cascade baryon 164, 261  
CAST detector 372–373  
CDF detector 268–269  
CEBAF accelerator 81  
centre-of-mass energy 78, 385–386  
Čerenkov counters 105–107  
Čerenkov radiation 40, 105–106  
    coherent 107–108  
charge conjugation 142–144  
    nonconservation in muon decay 310–312  
    of  $\pi^0$  and  $\eta$  mesons 144–145  
    positronium analysis 145–149  
charged weak currents 250–254, 436–437  
charmed particles 64–65, 257–258  
charmonium 64–65, 184–189  
charm quantum number 56  
charm threshold 186  
CKM matrix 264, 341  
cloud chamber 7–8, 61  
CMS detector 117, Plate 3  
cold dark matter 363  
colliding beam accelerators (colliders) 83–85  
colliding beam experiments 78  
collision length 88  
colour 177–184  
    charges 178–181, 430  
    confinement 178–181, 195

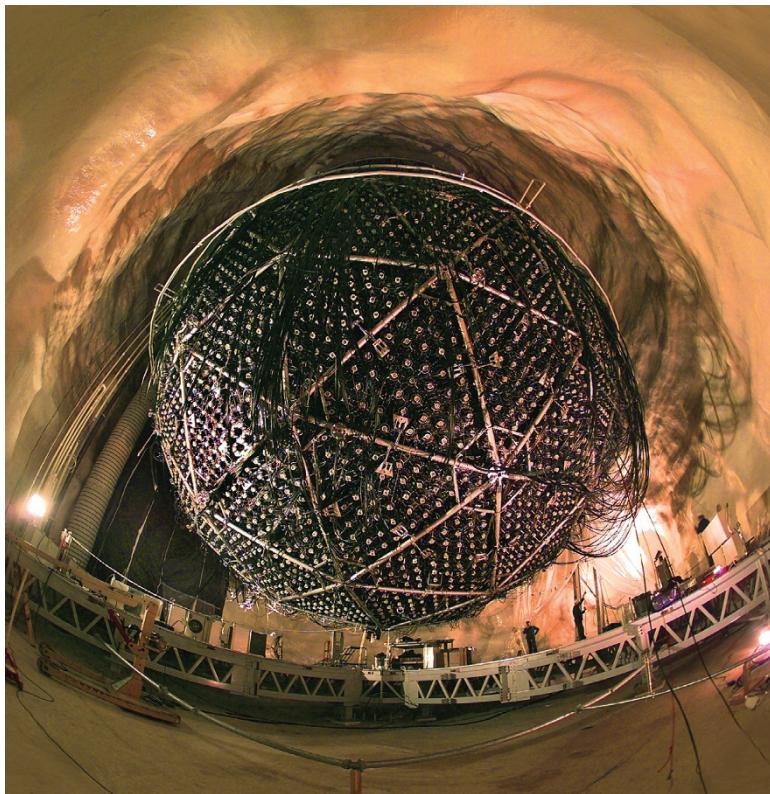
- colour (*Continued*)  
 evidence for in  $e^+e^-$  annihilation 214–215  
 hypercharge and isospin 178  
 and the Pauli principle 182–184  
 role in QCD 193–195  
 wavefunction 179–180, 430
- commutation relations  
 angular momentum 408–409  
 colour charges 430  
 isospin 408–409, 411  
 Compton scattering 21, 94  
 confinement 178–181, 195  
 conservation laws  
 angular momentum 129–132  
 baryon number 57  
 $B-L$  350  
 charge conjugation 142–144, 310–312  
 colour 178  
 $CP$  311–314  
 isospin 155–158  
 lepton numbers 25, 30  
 linear momentum 128  
 parity 135–137  
 cosmic microwave background radiation 362  
 cosmic rays 28, 39–40  
 cosmological constant 362  
 $C$  parity 143–144  
 fermion–antifermion pairs 144, 185  
 $\pi^0$  and  $\eta$  decays 144–145  
 positronium 145–149  
 violation in weak interactions 310–314
- $CP$  conservation 311–314  
 $CP$  nonconservation  
 in  $B$  decays 328–329, 335–337  
 and electric dipole moment 369–371  
 in  $K$  decays 319–324  
 in neutrino sector 343  
 in standard model 340–343  
 and strong  $CP$  problem 371  
 $CPT$  theorem 324, 327–328  
 CRESST detector 365–366  
 cross-sections 392–397  
   Born approximation for 18–19, 221, 398  
   Breit–Wigner formula 68, 70, 404–406  
     elastic electron–proton 217–222  
     Rutherford formula 218, 221–222  
   cyclic accelerators 80–83
- Dalitz plot 387–388  
 DAMA detector 365–366  
 dark energy 361  
 dark matter 360–362  
   search for cold dark matter 362–367,  
     372–373  
   decay width 67, 400
- deep inelastic scattering  
 charged lepton–proton 222–231  
 neutrino–proton 231–234  
 $\Delta S = \Delta Q$  rule 260  
 detailed balance 151–152  
 dibaryon 180–181  
 differential cross-section 395–397  
 dipole fit to proton form factors 220–221  
 Dirac equation 4–6, 424–425, 440  
 Dirac magnetic moment 6  
 $D$  mesons 64  
 double  $\beta$  decay 375–380  
 Drell–Yan mechanism 239–241  
 drift chamber 99–100
- elastic scattering 9  
 electric dipole moment 369–371  
 electromagnetic interactions  
   basic vertices 10–12  
   gauge invariance of 420–421  
   typical lifetimes 57–58  
 electromagnetic showers 110–112  
 electron–electron scattering 12–13  
 electron neutrino 25–28, 148–149  
   oscillations 39–47  
 electron number 25  
 electron–positron annihilation  
   evidence for colour in 214–215  
   to hadrons 185–187, 210–215, 284–287  
   in hole theory 10–11  
   to muon pairs 186, 212, 276–277, 284–285  
   to photons 13–14, 138, 148–149  
 electron–positron pair production 14–15, 94  
 electron–proton scattering  
   elastic 217–222  
   form factors 218–221  
   inelastic 222–231  
 electroweak interactions 281–284  
   gauge theory of 287–293, 434–441  
 electroweak unification 276–277, 279–281, 434–441  
 emulsions 59  
 energy losses by particles in matter  
   ionisation losses 89–92  
   radiation losses 92–93  
     short-range interactions with nuclei 86–89  
 $\eta$ -meson 144–145, 162–163  
 EXO-200 detector 378–379  
 exotic hadrons 74–75, 181, 201–208  
 exotic quantum numbers 72–75
- Fermi coupling constant 19, 32  
 Feynman diagrams 9–15  
   electromagnetic vertices 10–12, 250–252  
   QCD vertices 194–195  
   weak vertices 251–253, 255–257

- Feynman rules 13  
 fixed-target accelerators 83  
 fixed-target experiments 78  
 flavour independence of strong interaction 194  
 flavour oscillation 324–328, 331–335  
 form factors of proton 218–221  
 fragmentation 211–212, 237
- gaseous ionisation detectors 97–103  
 gauge bosons 1–2, 194, 248  
 gauge theories  
   gauge principle 288–290, 418, 423–425  
   Higgs boson 287–289  
   electroweak interaction 281–284, 289–290, 434–441  
   electromagnetic interaction 193–194, 420–421  
   photon mass 421–423  
   QCD 193–201, 429–434  
   unification condition 279, 437–440
- gauge transformations  
   electromagnetic 194, 420–423  
   electroweak 289–290, 434–441  
   quantum chromodynamics 193, 429–434
- Geiger–Müller counter 103  
 Geiger–Müller region 102–103  
 Georgi–Glashow model 348–351  
 GERDA detector 380  
 glueballs 195, 202–204  
 gluino 354, 357  
 gluons 194–195  
   spin determination in  $e^+e^-$  annihilation 213  
 grand unified theories (GUTs) 347–354  
 graviton 346
- hadronic showers 112  
 hadrons 55–58, 162–167, 174–177, 184–189  
   allowed and exotic quantum numbers 72–75,  
     158–161, 203  
   baryon magnetic moments 167–169  
   colour wavefunction 179–180, 182–185, 430  
   exotic states 74–75, 181, 201–208  
   mass splittings 169–174  
   resonances 66–73  
 helicity 312–316  
 hierarchy problem 355–357  
 Higgs boson  
   discovery 297–304  
   production 296–300  
   properties 294–297  
   in supersymmetry 354–355, 357–358  
 Higgs field and particle masses 290–293  
 Higgsinos 354–356  
 Higgs mechanism 292–293, 425–429  
 hole theory 6–11  
 hot dark matter 363  
 hybrid states 195, 202–205
- hypercharge 157  
 hypernuclei 180
- IceCube detector 367  
 inelastic electron and muon scattering 222–231  
 inelastic neutrino scattering 231–235  
 inflationary big bang model 361  
 integrated luminosity 79, 393  
 interaction lengths 88  
 interactions of photons in matter 93–95  
 internal quantum numbers 57  
 intrinsic parity 136  
 invariant mass 385–388  
 inverse  $\beta$  decay 27–28  
 inverse muon decay 29, 31–33  
 ionisation chamber 97–99  
 ionisation energy losses 89–92  
 isospin 156–161  
   branching ratio predictions 414–416  
   formal theory of 408–414  
   hadron multiplets 161–164  
   supermultiplets 164–167  
   weight diagrams 162
- jet chamber 100  
 jets  
   in  $e^+e^-$  annihilation 211–214, 237–238  
   in  $pp$  collisions 242–243  
 J-PARC accelerator 80  
 $J/\Psi$  meson 64, 185–189
- KamLAND detector 44–46, 121–123, 378–379  
 Klein–Gordon equation 4–5, 17, 422–423, 427–428  
 $K$  mesons (kaons) 61–63, 156, 158, 162–163, 256–257.  
   See also neutral  $K$  mesons
- lattice gauge theory 196–197, 201–202, 208, 244  
 LEP accelerator 82  
 lepton 1, 24–30  
   decays 29, 33–35  
   pair production in  $e^+e^-$  collisions 276–277, 282–284  
   pair production in  $pp$  collisions 239–241  
   parity 137–138  
 leptonic interactions 30–33  
   universality 33–35, 258–259  
 lepton number 25  
   conservation 30, 48–50  
 lepton–quark symmetry 254–259, 262–264, 277–279  
 LHC accelerator 82, Plate 2  
 LHCb detector 119–121  
 linear accelerators (linacs) 80–81  
 Lorentz condition 422–423  
 Lorentz gauge 422  
 Lorentz transformations 383–384, 388–390  
 luminosity 79, 393

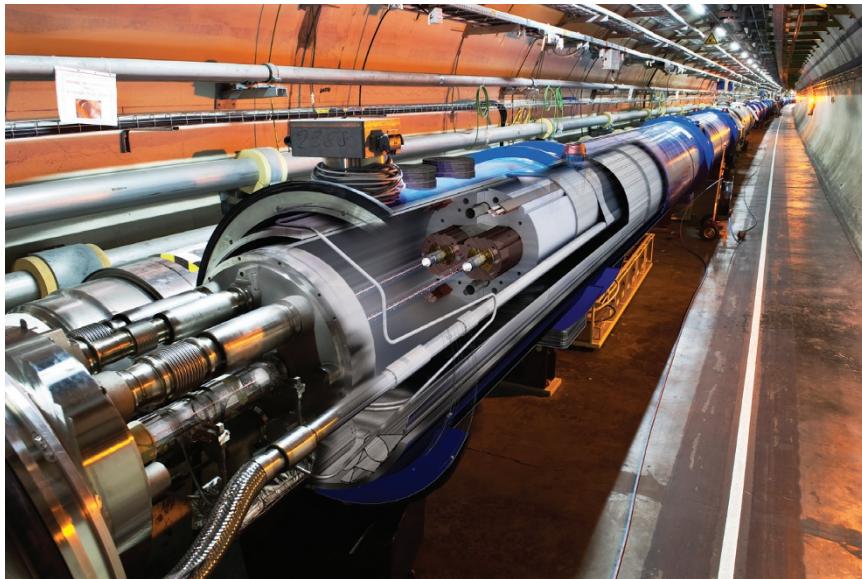
- MACHOs 363–367  
 magnet designs 80–82  
 magnetic moments  
     baryons 6, 9, 167–169  
     deuteron 133  
     Dirac form 6  
     electron and positron 6, 147–148  
     muon 28  
     quark 167  
 Majorana neutrinos 373–375  
     masses 375–376  
 matter–antimatter asymmetry in universe 367–369  
 mesons  
     charmonium and bottomonium 64–65, 185–186,  
     188–189  
     exotic states 74–75, 181, 201–206  
     pseudoscalar mesons 162–164  
     in quark model 56, 72–73, 134, 139, 162–164,  
     174–175  
     tables of states 65, 163, 174  
     vector mesons 162–164  
         weight diagrams 162, 164, 175–176  
 minimum ionisation region 90–91  
 molecular states 181, 205–206  
 MSW effect 37–38, 45  
*M* theory 359  
 multiquark states 74, 181, 207–208  
 multiwire proportional chamber 99–100  
 muon 24–25, 28–30  
     decay 29, 32–34, 310–312, 314–316  
 muon neutrino oscillation 38–42  
 natural units 19–21  
 neutralino 355–357, 364  
 neutral  $K$  mesons 62–63  
      $CP$  violation 319–324  
      $CP$  eigenstates 316–317  
      $K_L^0$  and  $K_S^0$  states 321–322  
      $K_L^0$ – $K_S^0$  mass difference 327  
      $K^0$ – $\bar{K}^0$  mixing 317, 321–322  
         strangeness oscillation 324–328  
     neutral particle beams 85–86  
     neutrino 25–30  
         Dirac or Majorana 373–375  
         helicity of 312–314  
         limit on number of 284–286  
         mixing 35–38  
         oscillations 38–46  
         sequential 284  
         solar 42–44, Plate 1  
     neutrino–electron scattering 29–30, 44  
     neutrino masses 26, 46–48, 363  
         and double-beta decay 375–380  
         see-saw mechanism 375  
     neutrino–nucleon scattering 27–28, 231–234, 249  
     neutron  
         decay 26, 57–58, 254  
         magnetic moment 6, 9, 168  
         nuclear cross-sections 88  
 omega-minus particle 164, 261–262  
 order of Feynman diagram 13  
 OZI rule 187, 198–199  
 parity 135–137  
     charged pion 140–141  
     intrinsic 136  
     lepton–antilepton pairs 137–138  
     nonconservation in weak interactions 308–312  
     photon 141–142  
     quarks and hadrons 139–140  
      $\tau - \theta$  puzzle 309  
 particle beams 77, 85–86  
 particle detectors 7, 59, 63–64, 95–112  
 particle exchange  
     range of force 15–16  
     Yukawa potential 17  
     zero-range approximation 18–19  
 particle interactions with matter 86–95  
 parton distributions 229–231, 236–237  
     scaling violations 228–231  
 parton model 226–228  
     gluon spin determination 213  
     neutrino scattering 231–233  
     quark charge determinations 234–236  
 Pauli spin matrices 131  
 penguin diagrams 341–342  
 pentaquarks 181, 207–208  
 photino 354  
 photoelectric effect 94  
 photomultiplier 104–105  
 photon  
      $C$  parity 144  
     interactions with matter 93–95  
     mass and gauge invariance 421–423  
     parity 141–142  
 pion–proton scattering 71–72, 87  
 pions 58–61  
      $C$  parity of neutral pions 144  
     decays 59–61, 144–145, 314–316  
     parity of charged pion 140–141  
     role in nuclear forces 60–61  
     spin of charged pion 152  
 Planck mass 360  
 positron 6–8  
 positronium 145–149  
 principle of detailed balance 151–152  
 Proca equations 422–423

- proportional chambers 99  
 proton  
   decay in grand unified theories 350–353  
   form factors 218–222  
   magnetic moment 6, 167–169  
   structure functions 224–231  
 pseudorapidity 238  
 pseudoscalar mesons 162–163
- quantum chromodynamics 193–201  
    $e^+e^-$  total cross-section prediction 214–215  
   gauge theory of 429–434  
   scaling violations 228–231  
 quantum electrodynamics 13–15  
 quantum fluctuations 200–201  
 quark–antiquark potential 189–191  
 quark confinement 155, 178–182  
 quark–gluon plasma 208–210  
 quark mixing 255–258, 263–265, 340–343  
 quark model 52–53  
   allowed and exotic quantum numbers 72–75,  
 158–163, 203  
   angular momentum in 134–135  
   charmonium and bottomonium 64–65, 184–189  
   isospin in 156–160  
   light baryons 164–167  
   light mesons 162–164  
   magnetic moments in 167–169  
   parity in 139–140  
 quarks 53–57, 262–267  
   CKM matrix 264, 341  
   colour charges 178–181, 430  
   constituent and current 174, 243–245  
   distribution functions 228–231, 236–237  
   electric charge determination 234–236  
   parity 139  
   quark masses 53–54, 245  
   quark mixing 255–258, 263–265  
    $Q$  value 58  
   running mass 245
- radiation energy losses 92–93  
 radiation length 92–93  
 range  
   of particle exchange force 15–16  
   of particles in matter 94–95  
   zero-range approximation 16, 18–19  
 rapidity 238  
 relativistic kinematics 383–390  
 relativistic wave equations 3–6  
 resistive plate chamber 102  
 resonances 66–73. *See also* Breit–Wigner formula  
 RHIC accelerator 84–85  
 rotational invariance 129–132
- R* parity 355–356  
 running coupling constant of QCD 197–198  
 running quark mass 245  
 Rutherford cross-section formula 218, 221–222
- scale invariance 224–226  
 scaling violations in QCD 228–231  
 scattering amplitude 18–19, 397–400  
 scintillation counter 104–105  
 sea quarks 228, 230–231, 236  
 secondary particle beams 85–86  
 see-saw mechanism 375  
 selection rules in weak interactions 259–262  
 selectron 354, 357  
 semiconductor detector 103–104  
 shower counters 110–112  
 sigma baryons 159–161, 164–168, 259–260  
 silicon microstrip detector 103–104  
 smuon 354  
 SNO detector 43–44, Plate 1  
 solar neutrinos 42–44, 47–48  
 spark chamber 103  
 spectrometer 96  
 spectroscopic notation 133–135  
 spin matrices 131  
 spinors 5, 182, 430, 434  
 spontaneous symmetry breaking 291–293, 428–429  
 squark 354–355, 357  
 standard cosmological model 361  
 STAR detector 113, 118–119, Plate 5  
 stauon 354  
 storage ring accelerators 83  
 strangeness oscillations 324–328  
 strangeness quantum number 55–56  
 strange particles 61–64  
 streamer chamber 103  
 string theories 358–360  
 strong coupling constant 195, 197–199  
 strong  $CP$  problem 371  
 structure functions 224–231  
 SuperKamiokande detector 39–41  
 superparticles 354–355  
   experimental searches for 355–358  
 supersymmetry 354–358  
    $CP$  violation in 370  
   electric dipole moments in 370–371  
 symmetries  
   charge conjugation 142–145, 310–312  
    $CP$  invariance 310–314, 328–329, 335–337,  
 340–343  
   gauge invariance 279, 287–290, 418–423, 429–434,  
 436–440  
   parity 135–137, 308–312, 440–441  
   rotational invariance 129–132

- symmetries (*Continued*)  
time reversal 149–152  
translational invariance 127–129
- synchrotron accelerator 81–84  
synchrotron radiation 81–82
- tauon 24–25, 29–30, 33–35  
tetraquarks 74–75, 181–182, 205–207  
Tevatron accelerator 82  
time projection chamber 100–101  
time reversal 149–152  
top quantum number 56  
top quark 263–267  
discovery 267–273  
total cross-section 87–88, 394–395  
transition radiation detector 108–109  
translational invariance 127–129  
tree diagrams 341
- UA1 detector 114–115  
unification condition 279, 437–439  
unification mass in GUTS 347, 352, 356  
unified electroweak interaction 276–277, 279–284,  
437–439  
unitary gauge 428  
units and dimensions 19–21  
universality of lepton interactions 33–35
- vacuum polarization effects 200  
 $V-A$  interaction 314  
valence quarks 228–230  
vector mesons 162–164  
virtual processes 12
- W* boson  
decays 113, 253–254, 258–259  
discovery 113–117  
exchange and Fermi coupling 32–35  
mass and width 117, 279–281  
weak charged current reactions 250–262  
lepton vertices 250–253  
quark vertices 254–256  
weak hypercharge 437  
weak isospin 434–435  
weak mixing angle 279–280, 349–350, 356  
weak neutral current reactions 277–278, 281–287  
lepton vertices 277  
quark vertices 277–278  
weight diagrams 162, 164, 175–176  
Weinberg angle, *see* weak mixing angle  
WIMPs 364–365  
wino 354  
wire chambers 99–102  
Wolfenstein parameterisation 341
- X* boson 348, 350–351
- Y* boson 348, 350–351  
Yukawa potential 17–19  
Yukawa theory of nuclear forces 60
- Z* bosons  
decays 113, 285–287  
discovery 113–116  
mass and width 116, 279–281, 284–286  
zero-range approximation 18–19  
zino 354



**Plate 1** The Sudbury neutrino detector during construction, showing the large number of photomultiplier tubes surrounding the sphere. (Photograph by Roy Kaltschmidt, courtesy of Lawrence Berkeley National Laboratory.)

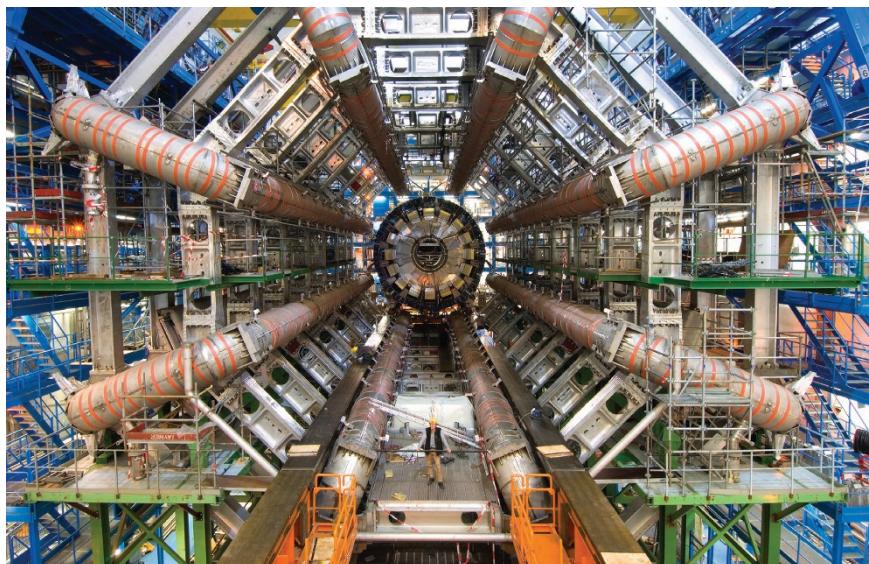


**Plate 2** LHC tunnel and beam lines. The blue cylinders contain dipole magnets and the liquid helium system required to cool them to superconducting temperatures. The cutaway section is an artist's impression, and shows the internal structure with the magnets and the two beam lines. (Reproduced by permission of CERN.)

**Plate 3** Fish-eye view of the CMS detector before being closed. (Reproduced by permission of CERN.)

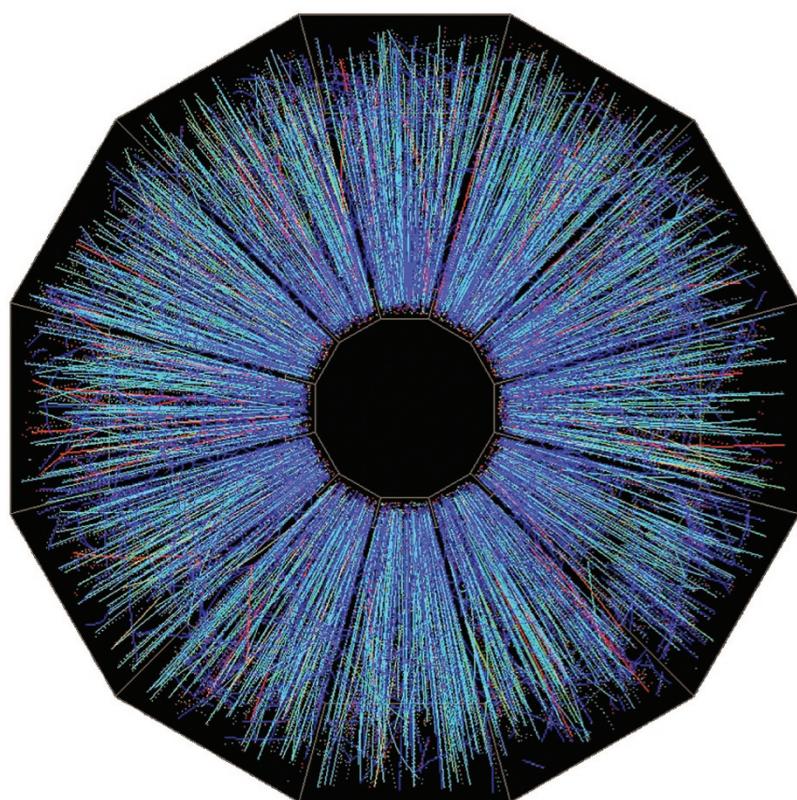


**Plate 4** ATLAS detector: view along the beam direction during the construction phase showing the eight barrel toroids installed, with a calorimeter at the end before it was moved into the middle of the detector. In the completed detector (see Figure 4.25), the central cavern is filled with subdetectors.  
(Reproduced by permission of CERN.)



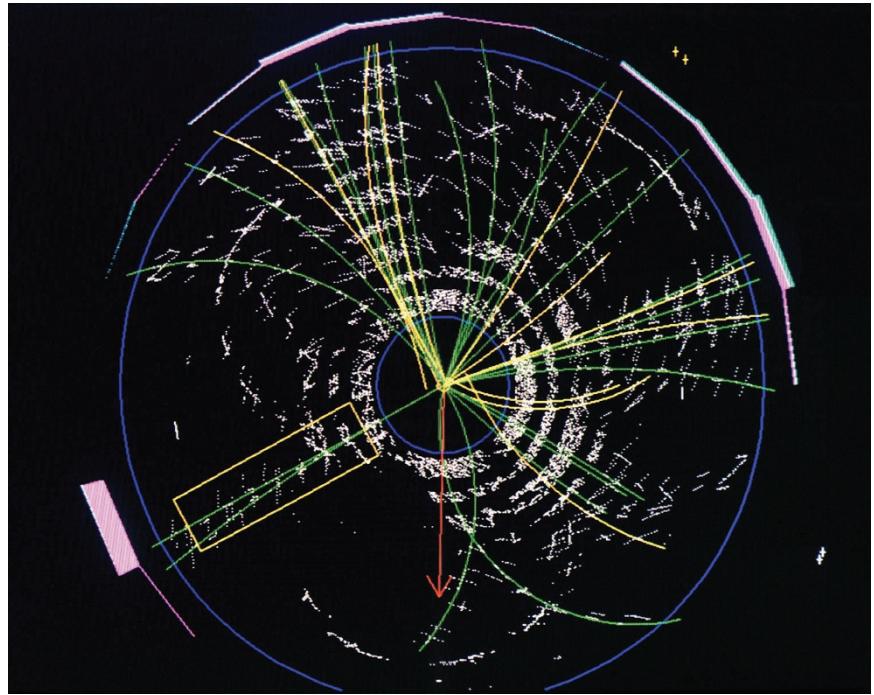


**Plate 5** The STAR detector, showing an end cap of the time projection chamber during construction (cf. Figure 4.26). (Photograph courtesy of Brookhaven National Laboratory.)

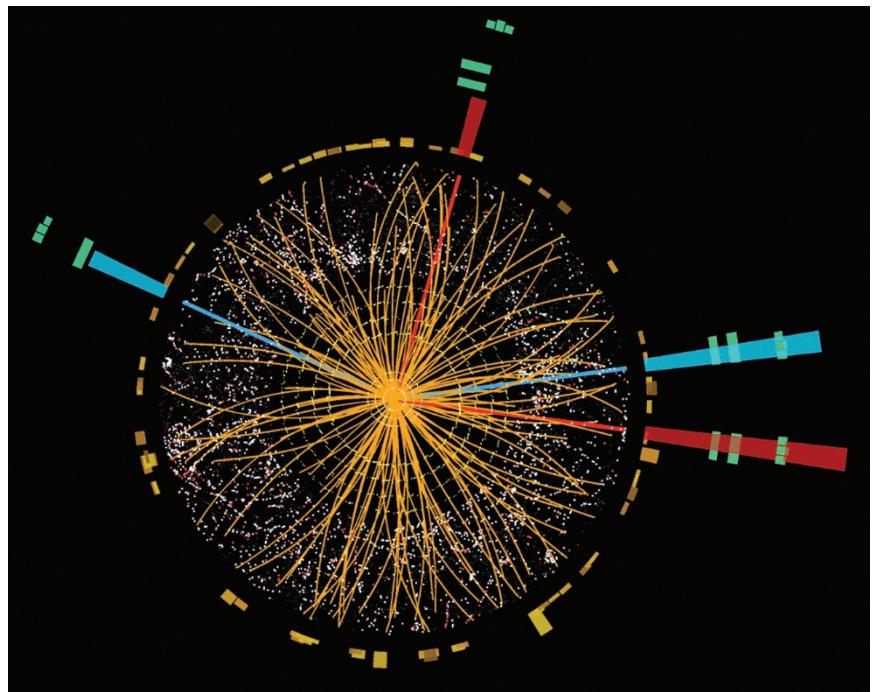


**Plate 6** View of a 200 GeV gold-gold interaction in the STAR detector at the RHIC accelerator. (Photograph courtesy of Brookhaven National Laboratory.)

**Plate 7** A  $t\bar{t}$  event at CDF. The event consists of four jets, two of them from  $b$  quarks, a positron (outlined by yellow rectangle) and a neutrino (direction denoted by red arrow). (Photograph courtesy of Fermilab and Brookhaven National Laboratory.)



**Plate 8** ATLAS event display of an  $H^0 \rightarrow$  four lepton candidate event with  $m_{4\ell} = 124.5$  GeV. The masses of the lepton pairs are 70.6 GeV and 44.7 GeV. The tracks and clusters of the two electron pairs are coloured red and blue, respectively. (Reproduced by permission of CERN.)



## **WILEY END USER LICENSE AGREEMENT**

Go to [www.wiley.com/go/eula](http://www.wiley.com/go/eula) to access Wiley's ebook EULA.