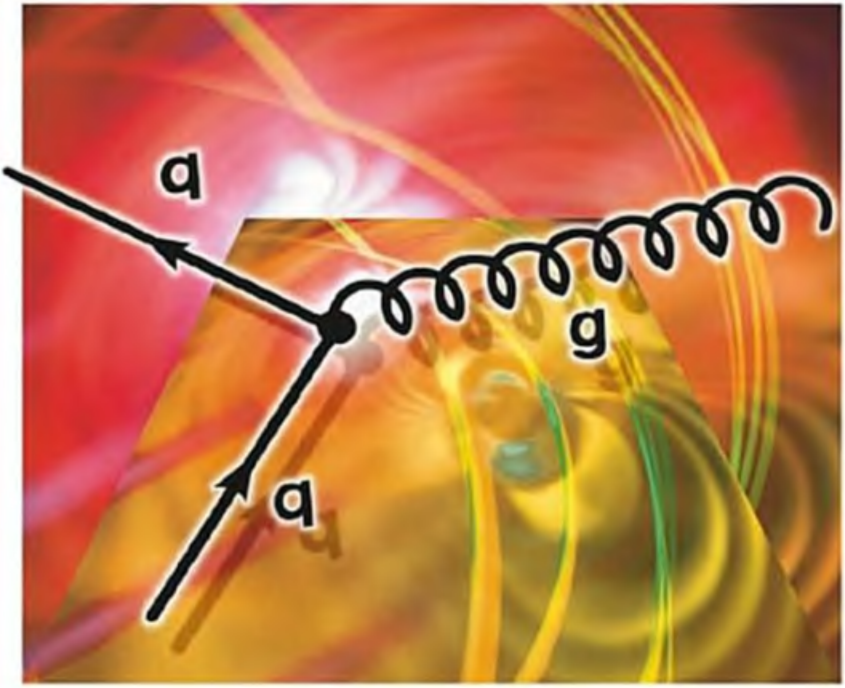
PHYSICS TEXTBOOK

David Griffiths

WILEY-VCH

Introduction to Elementary Particles

Second, Revised Edition



Copyrighted material

David Griffiths

Introduction to Elementary Particles

Second, Revised Edition

WILEY-VCH

WILEY-VCH Verlag GmbH & Co. KGaA

The Author

David Griffiths
Reed College
Portland, OR
USA

For a **Solutions Manual**, instructors should contact the editorial department at vch-physics@wiley-vch.de, stating their affiliation and the course in which they wish to use the book.

All books published by Wiley-VCH are carefully produced. Nevertheless, authors, editors, and publisher do not warrant the information contained in these books, including this book, to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Library of Congress Card No.:
applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
Die Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

© 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into a machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Typesetting Laserwords Private Limited, Chennai, India
Printing Strauss GmbH, Mörlenbach
Binding Litges & Dopf Buchbinderei GmbH, Heppenheim

Printed in the Federal Republic of Germany
Printed on acid-free paper

ISBN: 978-3-527-40601-2

Contents

Preface to the First Edition *IX*
Preface to the Second Edition *XI*
Formulas and Constants *XIII*

Introduction *1*

1 Historical Introduction to the Elementary Particles *13*
1.1 The Classical ERA (1897–1932) *13*
1.2 The Photon (1900–1924) *15*
1.3 Mesons (1934–1947) *18*
1.4 Antiparticles (1930–1956) *20*
1.5 Neutrinos (1930–1962) *23*
1.6 Strange Particles (1947–1960) *30*
1.7 The Eightfold Way (1961–1964) *35*
1.8 The Quark Model (1964) *37*
1.9 The November Revolution and Its Aftermath (1974–1983 and 1995) *44*
1.10 Intermediate Vector Bosons (1983) *47*
1.11 The Standard Model (1978–?) *49*

2 Elementary Particle Dynamics *59*
2.1 The Four Forces *59*
2.2 Quantum Electrodynamics (QED) *60*
2.3 Quantum Chromodynamics (QCD) *66*
2.4 Weak Interactions *71*
2.4.1 Neutral *72*
2.4.2 Charged *74*
2.4.2.1 Leptons *74*
2.4.3 Quarks *75*
2.4.4 Weak and Electromagnetic Couplings of W and Z *78*
2.5 Decays and Conservation Laws *79*
2.6 Unification Schemes *84*

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

3 Relativistic Kinematics *89*
3.1 Lorentz Transformations *89*
3.2 Four-vectors *92*
3.3 Energy and Momentum *96*
3.4 Collisions *100*
3.4.1 Classical Collisions *100*
3.4.2 Relativistic Collisions *101*
3.5 Examples and Applications *102*

4 Symmetries *115*
4.1 Symmetries, Groups, and Conservation Laws *115*
4.2 Angular Momentum *120*
4.2.1 Addition of Angular Momenta *122*
4.2.2 Spin $\frac{1}{2}$ *125*
4.3 Flavor Symmetries *129*
4.4 Discrete Symmetries *136*
4.4.1 Parity *136*
4.4.2 Charge Conjugation *142*
4.4.3 CP *144*
4.4.3.1 Neutral Kaons *145*
4.4.3.2 *CP* Violation *147*
4.4.4 Time Reversal and the *TCP* Theorem *149*

5 Bound States *159*
5.1 The Schrödinger Equation *159*
5.2 Hydrogen *162*
5.2.1 Fine Structure *165*
5.2.2 The Lamb Shift *166*
5.2.3 Hyperfine Splitting *167*
5.3 Positronium *169*
5.4 Quarkonium *171*
5.4.1 Charmonium *174*
5.4.2 Bottomonium *175*
5.5 Light Quark Mesons *176*
5.6 Baryons *180*
5.6.1 Baryon Wave Functions *181*
5.6.2 Magnetic Moments *189*
5.6.3 Masses *191*

6 The Feynman Calculus *197*
6.1 Decays and Scattering *197*
6.1.1 Decay Rates *197*
6.1.2 Cross Sections *199*
6.2 The Golden Rule *203*
6.2.1 Golden Rule for Decays *204*

6.2.1.1 Two-particle Decays *206*
6.2.2 Golden Rule for Scattering *208*
6.2.2.1 Two-body Scattering in the CM Frame *209*
6.3 Feynman Rules for a Toy Theory *211*
6.3.1 Lifetime of the A *214*
6.3.2 $A + A \rightarrow B + B$ Scattering *215*
6.3.3 Higher-order Diagrams *217*

7 Quantum Electrodynamics *225*
7.1 The Dirac Equation *225*
7.2 Solutions to the Dirac Equation *229*
7.3 Bilinear Covariants *235*
7.4 The Photon *238*
7.5 The Feynman Rules for QED *241*
7.6 Examples *245*
7.7 Casimir's Trick *249*
7.8 Cross Sections and Lifetimes *254*
7.9 Renormalization *262*

8 Electrodynamics and Chromodynamics of Quarks *275*
8.1 Hadron Production in e^+e^- Collisions *275*
8.2 Elastic Electron–Proton Scattering *279*
8.3 Feynman Rules For Chromodynamics *283*
8.4 Color Factors *289*
8.4.1 Quark and Antiquark *289*
8.4.2 Quark and Quark *292*
8.5 Pair Annihilation in QCD *294*
8.6 Asymptotic Freedom *298*

9 Weak Interactions *307*
9.1 Charged Leptonic Weak Interactions *307*
9.2 Decay of the Muon *310*
9.3 Decay of the Neutron *315*
9.4 Decay of the Pion *321*
9.5 Charged Weak Interactions of Quarks *324*
9.6 Neutral Weak Interactions *329*
9.7 Electroweak Unification *338*
9.7.1 Chiral Fermion States *338*
9.7.2 Weak Isospin and Hypercharge *342*
9.7.3 Electroweak Mixing *345*

10 Gauge Theories *353*
10.1 Lagrangian Formulation of Classical Particle Mechanics *353*
10.2 Lagrangians in Relativistic Field Theory *354*
10.3 Local Gauge Invariance *358*

10.4 Yang–Mills Theory *361*
10.5 Chromodynamics *366*
10.6 Feynman Rules *369*
10.7 The Mass Term *372*
10.8 Spontaneous Symmetry-breaking *375*
10.9 The Higgs Mechanism *378*

11 Neutrino Oscillations *387*
11.1 The Solar Neutrino Problem *387*
11.2 Oscillations *390*
11.3 Confirmation *392*
11.4 Neutrino Masses *395*
11.5 The Mixing Matrix *397*

12 Afterword: What's Next? *401*
12.1 The Higgs Boson *401*
12.2 Grand Unification *405*
12.3 Matter/Antimatter Asymmetry *409*
12.4 Supersymmetry, Strings, Extra Dimensions *411*
12.4.1 Supersymmetry *411*
12.4.2 Strings *413*
12.5 Dark Matter/Dark Energy *414*
12.5.1 Dark Matter *414*
12.5.2 Dark Energy *416*
12.6 Conclusion *417*

A The Dirac Delta Function *423*

B Decay Rates and Cross Sections *429*
B.1 Decays *429*
B.1.1 Two-body Decays *429*
B.2 Cross Sections *430*
B.2.1 Two-body Scattering *430*

C Pauli and Dirac Matrices *433*
C.1 Pauli Matrices *433*
C.2 Dirac Matrices *434*

D Feynman Rules (Tree Level) *437*
D.1 External Lines *437*
D.2 Propagators *437*
D.3 Vertex Factors *438*

Index *441*

Preface to the First Edition

This introduction to the theory of elementary particles is intended primarily for advanced undergraduates who are majoring in physics. Most of my colleagues consider this subject inappropriate for such an audience – mathematically too sophisticated, phenomenologically too cluttered, insecure in its foundations, and uncertain in its future. Ten years ago I would have agreed. But in the last decade the dust has settled to an astonishing degree, and it is fair to say that elementary particle physics has come of age. Although we obviously have much more to learn, there now exists a coherent and unified theoretical structure that is simply too exciting and important to save for graduate school or to serve up in diluted qualitative form as a subunit of modern physics. I believe the time has come to integrate elementary particle physics into the standard undergraduate curriculum.

Unfortunately, the research literature in this field is clearly inaccessible to undergraduates, and although there are now several excellent graduate texts, these call for a strong preparation in advanced quantum mechanics, if not quantum field theory. At the other extreme, there are many fine popular books and a number of outstanding *Scientific American* articles. But very little has been written specifically for the undergraduate. This book is an effort to fill that need. It grew out of a one-semester elementary particles course I have taught from time to time at Reed College. The students typically had under their belts a semester of electromagnetism (at the level of Lorrain and Corson), a semester of quantum mechanics (at the level of Park), and a fairly strong background in special relativity.

In addition to its principal audience, I hope this book will be of use to beginning graduate students, either as a primary text, or as preparation for a more sophisticated treatment. With this in mind, and in the interest of greater completeness and flexibility, I have included more material here than one can comfortably cover in a single semester. (In my own courses I ask the students to read Chapters 1 and 2 on their own, and begin the lectures with Chapter 3. I skip Chapter 5 altogether, concentrate on Chapters 6 and 7, discuss the first two sections of Chapter 8, and then jump to Chapter 10.) To assist the reader (and the teacher) I begin each chapter with a brief indication of its purpose and content, its prerequisites, and its role in what follows.

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

This book was written while I was on sabbatical at the Stanford Linear Accelerator Center, and I would like to thank Professor Sidney Drell and the other members of the Theory Group for their hospitality.

DAVID GRIFFITHS
1986

Preface to the Second Edition

It is 20 years since the first edition of this book was published, and it is both gratifying and distressing to reflect that it remains, for the most part, reasonably up-to-date. There are, to be sure, some gross lacunae – the existence of the top quark had not been confirmed back then, and neutrinos were generally assumed (for no very good reason) to be massless. But the Standard Model, which is, in essence, the subject of the book, has proved to be astonishingly robust. This is a tribute to the theory, and at the same time an indictment of our collective imagination. I don't think there has been a comparable period in the history of elementary particle physics in which so little of a truly revolutionary nature has occurred. What about neutrino oscillations? Indeed: a fantastic story (I have added a chapter on the subject); and yet, this extraordinary phenomenon fits so comfortably into the Standard Model that one might almost say, in retrospect (of course), that it would have been more surprising if it had *not* been so. How about supersymmetry and string theory? Yes, but these must for the moment be regarded as speculations (I have added a chapter on contemporary theoretical developments). As far as solid experimental confirmation goes, the Standard Model (with neutrino masses and mixing) still rules.

In addition to the two new chapters already mentioned, I have brought the history up-to-date in Chapter 1, shortened Chapter 5, provided (I hope) a more compelling introduction to the Golden Rule in Chapter 6, and eliminated most of the old Chapter 8 on electromagnetic form factors and scaling (this was crucially important in interpreting the deep inelastic scattering experiments that put the quark model on a secure footing, but no one today doubts the existence of quarks, and the technical details are no longer so essential). What remains of Chapter 8 is now combined with the old Chapter 9 to make a new chapter on hadrons. Finally, I have prepared a complete solution manual (available free from the publisher, though only – I regret – to course instructors). Beyond this the changes are relatively minor.

Many people have sent me suggestions and corrections, or patiently answered my questions. I cannot thank everyone, but I would like to acknowledge some of those who were especially helpful: Guy Blaylock (UMass), John Boersma (Rochester), Carola Chinellato (Brazil), Eugene Commins (Berkeley), Mimi Gerstell (Cal Tech),

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

Nahmin Horwitz (Syracuse), Richard Kass (Ohio State), Janis McKenna (UBC), Jim Napolitano (RPI), Nic Nigro (Seattle), John Norbury (UW-Milwaukee), Jason Quinn (Notre Dame), Aaron Roodman (SLAC), Natthi Sharma (Eastern Michigan), Steve Wasserbeach (Haverford), and above all Pat Burchat (Stanford).

Part of this work was carried out while I was on sabbatical, at Stanford and SLAC, and I especially thank Patricia Burchat and Michael Peskin for making this possible.

David Griffiths
2008

Formulas and Constants

Particle Data

Mass in MeV/c^2, lifetime in seconds, charge in units of the proton charge.

Leptons (spin 1/2)

Generation	Flavor	Charge	Mass*	Lifetime	Principal Decays
first	e (electron)	-1	0.510999	∞	–
	ν_e (e neutrino)	0	0	∞	–
second	μ (muon)	-1	105.659	2.19703×10^{-6}	$e\nu_\mu\bar{\nu}_e$
	ν_μ (μ neutrino)	0	0	∞	–
third	τ (tau)	-1	1776.99	2.91×10^{-13}	$e\nu_\tau\bar{\nu}_e$, $\mu\nu_\tau\bar{\nu}_\mu$, $\pi^-\nu_\tau$
	ν_τ (τ neutrino)	0	0	∞	–

*Neutrino masses are extremely small, and for most purposes can be taken to be zero; for details see Chapter 11.

Quarks (spin 1/2)

Generation	Flavor	Charge	Mass*
first	d (down)	$-1/3$	7
	u (up)	2/3	3
second	s (strange)	$-1/3$	120
	c (charm)	2/3	1200
third	b (bottom)	$-1/3$	4300
	t (top)	2/3	174000

*Light quark masses are imprecise and speculative; for effective masses in mesons and baryons, see Chapter 5.

Mediators (spin 1)

Force	Mediator	Charge	Mass*	Lifetime	Principal Decays
Strong	g (8 gluons)	0	0	∞	–
Electromagnetic	γ (photon)	0	0	∞	–
Weak	$W^\pm$ (charged)	± 1	80,420	3.11×10^{-25}	$e^+\nu_e$, $\mu^+\nu_\mu$, $\tau^+\nu_\tau$, $cX \to$ hadrons
	Z^0 (neutral)	0	91,190	2.64×10^{-25}	e^+e^-, $\mu^+\mu^-$, $\tau^+\tau^-$, $q\bar{q} \to$ hadrons

Baryons (spin 1/2)

Baryon	Quark Content	Charge	Mass	Lifetime	Principal Decays
$N\begin{cases} p \\ n \end{cases}$	uud	1	938.272	∞	–
	udd	0	939.565	885.7	$pe\bar{\nu}_e$
Λ	uds	0	1115.68	2.63×10^{-10}	$p\pi^-, n\pi^0$
Σ^+	uus	1	1189.37	8.02×10^{-11}	$p\pi^0, n\pi^+$
Σ^0	uds	0	1192.64	7.4×10^{-20}	$\Lambda\gamma$
Σ^-	dds	−1	1197.45	1.48×10^{-10}	$n\pi^-$
Ξ^0	uss	0	1314.8	2.90×10^{-10}	$\Lambda\pi^0$
Ξ^-	dss	−1	1321.3	1.64×10^{-10}	$\Lambda\pi^-$
Λ_c^+	udc	1	2286.5	2.00×10^{-13}	$pK\pi$, $\Lambda\pi\pi$, $\Sigma\pi\pi$

Baryons (spin 3/2)

Baryon	Quark Content	Charge	Mass	Lifetime	Principal Decays
Δ	uuu, uud, udd, ddd	2,1,0,−1	1232	5.6×10^{-24}	$N\pi$
Σ^*	uus, uds, dds	1,0,−1	1385	1.8×10^{-23}	$\Lambda\pi$, $\Sigma\pi$
Ξ^*	uss, dss	0,−1	1533	6.9×10^{-23}	$\Xi\pi$
Ω^-	sss	−1	1672	8.2×10^{-11}	ΛK^-, $\Xi\pi$

Pseudoscalar Mesons (spin 0)

Meson	Quark Content	Charge	Mass	Lifetime	Principal Decays
$\pi^\pm$	$u\bar{d}$, $d\bar{u}$	1,−1	139.570	2.60×10^{-8}	$\mu\nu_\mu$
π^0	$(u\bar{u} - d\bar{d})/\sqrt{2}$	0	134.977	8.4×10^{-17}	$\gamma\gamma$
$K^\pm$	$u\bar{s}$, $s\bar{u}$	1,−1	493.68	1.24×10^{-8}	$\mu\nu_\mu$, $\pi\pi$, $\pi\pi\pi$
$K^0, \bar{K}^0$	$d\bar{s}$, $s\bar{d}$	0	497.65	$\begin{cases} K_S^0: 8.95 \times 10^{-11} \\ K_L^0: 5.11 \times 10^{-8} \end{cases}$	$\pi\pi$ $\pi e\nu_e$, $\pi\mu\nu_\mu$, $\pi\pi\pi$
η	$(u\bar{u} + d\bar{d} - 2s\bar{s})/\sqrt{6}$	0	547.51	5.1×10^{-19}	$\gamma\gamma$, $\pi\pi\pi$
η'	$(u\bar{u} + d\bar{d} + s\bar{s})/\sqrt{3}$	0	957.78	3.2×10^{-21}	$\eta\pi\pi$, $\rho\gamma$
$D^\pm$	$c\bar{d}$, $d\bar{c}$	1,−1	1869.3	1.04×10^{-12}	$K\pi\pi$, $K\mu\nu_\mu$, $Ke\nu_e$
$D^0, \bar{D}^0$	$c\bar{u}$, $u\bar{c}$	0	1864.5	4.1×10^{-13}	$K\pi\pi$, $Ke\nu_e$, $K\mu\nu_\mu$
$D_s^\pm$	$c\bar{s}$, $s\bar{c}$	1,−1	1968.2	5.0×10^{-13}	$\eta\rho$, $\phi\pi\pi$, $\phi\rho$
$B^\pm$	$u\bar{b}$, $b\bar{u}$	1,−1	5279.0	1.6×10^{-12}	$D^*\ell\nu_\ell$, $D\ell\nu_\ell$, $D^*\pi\pi\pi$
$B^0, \bar{B}^0$	$d\bar{b}$, $b\bar{d}$	0	5279.4	1.5×10^{-12}	$D^*\ell\nu_\ell$, $D\ell\nu_\ell$, $D^*\pi\pi$

Vector Mesons (spin 1)

Meson	Quark Content	Charge	Mass	Lifetime	Principal Decays
ρ	$u\bar{d}$, $(u\bar{u} - d\bar{d})/\sqrt{2}$, $d\bar{u}$	1,0,-1	775.5	4×10^{-24}	$\pi\pi$
K^*	$u\bar{s}$, $d\bar{s}$, $s\bar{d}$, $s\bar{u}$	1,0,-1	894	1×10^{-23}	$K\pi$
ω	$(u\bar{u} + d\bar{d})/\sqrt{2}$	0	782.6	8×10^{-23}	$\pi\pi\pi$, $\pi\gamma$
ψ	$c\bar{c}$	0	3097	7×10^{-21}	e^+e^-, $\mu^+\mu^-$, 5π, 7π
D^*	$c\bar{d}$, $c\bar{u}$, $u\bar{c}$, $d\bar{c}$	1,0,-1	2008	3×10^{-21}	$D\pi$, $D\gamma$
Υ	$b\bar{b}$	0	9460	1×10^{-20}	e^+e^-, $\mu^+\mu^-$, $\tau^+\tau^-$

Spin 1/2

Pauli Matrices:

$$\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\sigma_i \sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k, \quad (\mathbf{a}\cdot\boldsymbol{\sigma})(\mathbf{b}\cdot\boldsymbol{\sigma}) = \mathbf{a}\cdot\mathbf{b} + i\boldsymbol{\sigma}\cdot(\mathbf{a}\times\mathbf{b})$$

$$\sigma_i^\dagger = \sigma_i = \sigma_i^{-1}, \qquad e^{i\boldsymbol{\theta}\cdot\boldsymbol{\sigma}} = \cos\theta + i(\hat{\theta}\cdot\boldsymbol{\sigma})\sin\theta$$

Dirac Matrices:

$$\gamma^0 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma^i \equiv \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}, \quad \gamma^{0\dagger} = \gamma^0, \quad \gamma^{i\dagger} = -\gamma^i, \quad \gamma^0\gamma^{\mu\dagger}\gamma^0 = \gamma^\mu$$

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}, \quad g^{\mu\nu} = g_{\mu\nu} = \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{Bmatrix}$$

$$\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \{\gamma^\mu, \gamma^5\} = 0, \quad (\gamma^5)^2 = 1$$

(For product rules and trace theorems see Appendix C.)

Dirac Equation:

$$i\hbar\gamma^\mu\,\partial_\mu\psi - mc\,\psi = 0$$

$$(\not{p} - mc)u = 0, \quad (\not{p} + mc)v = 0, \quad \overline{u}(\not{p} - mc) = 0, \quad \overline{v}(\not{p} + mc) = 0$$

$$\overline{\psi} \equiv \psi^\dagger\gamma^0, \quad \overline{\Gamma} \equiv \gamma^0\Gamma^\dagger\gamma^0, \quad \not{a} \equiv a_\mu\gamma^\mu$$

Feynman Rules

	External Lines		Propagators
Spin 0:	Nothing		$\dfrac{i}{q^2 - (mc)^2}$
Spin 1/2:	Incoming particle:	u	$\dfrac{i(\not{q} + mc)}{q^2 - (mc)^2}$
	Incoming antiparticle:	$\overline{v}$	
	Outgoing particle:	$\overline{u}$	
	Outgoing antiparticle:	v	

$$\text{Spin 1:} \quad \begin{cases} \text{Incoming:} & \epsilon_\mu \\ \text{Outgoing:} & \epsilon_\mu^* \end{cases} \quad \begin{cases} \text{Massless:} & \dfrac{-ig_{\mu\nu}}{q^2} \\ \text{Massive:} & \dfrac{-i[g_{\mu\nu} - q_\mu q_\nu/(mc)^2]}{q^2 - (mc)^2} \end{cases}$$

(For vertex factors see Appendix D.)

Fundamental Constants

Planck's constant:	$\hbar$	$=$	1.05457×10^{-34} J s
		$=$	6.58212×10^{-22} MeV s
Speed of light:	c	$=$	2.99792×10^{8} m/s
Mass of electron:	m_e	$=$	9.10938×10^{-31}kg $= 0.510999$ MeV/c^2
Mass of proton:	m_p	$=$	1.67262×10^{-27}kg $= 938.272$ MeV/c^2
Electron charge (magnitude):	e	$=$	1.60218×10^{-19} C
		$=$	4.80320×10^{-10} esu
Fine structure constant:	α	$=$	$e^2/\hbar c = 1/137.036$
Bohr radius:	a	$=$	$\hbar^2/m_e e^2 = 5.29177 \times 10^{-11}$ m
Bohr energies:	E_n	$=$	$-m_e e^4/2\hbar^2 n^2 = -13.6057/n^2$ eV
Classical electron radius:	r_e	$=$	$e^2/m_e c^2 = 2.81794 \times 10^{-15}$ m
QED coupling constant:	g_e	$=$	$e\sqrt{4\pi/\hbar c} = 0.302822$
Weak coupling constants:	g_w	$=$	$g_e/\sin\theta_w = 0.6295$;
	g_z	$=$	$g_w/\cos\theta_w = 0.7180$
Weak mixing angle:	θ_w	$=$	$28.76°$ $(\sin^2\theta_w = 0.2314)$
Strong coupling constant:	g_s	$=$	1.214

Conversion Factors

1 Å	$=$	0.1 nm $= 10^{-10}$ m
1 fm	$=$	10^{-15} m
1 barn	$=$	10^{-28} m^2
1 eV	$=$	1.60218×10^{-19} J
1 MeV/c^2	$=$	1.78266×10^{-30} kg
1 Coulomb	$=$	2.99792×10^{-9} esu

Introduction

Elementary Particle Physics

Elementary particle physics addresses the question, 'What is matter made of?' at the most fundamental level – which is to say, on the smallest scale of size. It's a remarkable fact that matter at the subatomic level consists of tiny chunks, with vast empty spaces in between. Even more remarkable, these tiny chunks come in a small number of different types (electrons, protons, neutrons, pi mesons, neutrinos, and so on), which are then replicated in astronomical quantities to make all the 'stuff' around us. And these replicas are absolutely perfect copies – not just 'pretty similar', like two Fords coming off the same assembly line, but utterly *indistinguishable*. You can't stamp an identification number on an electron, or paint a spot on it – if you've seen one, you've seen them all. This quality of absolute identicalness has no analog in the macroscopic world. (In quantum mechanics it is reflected in the Pauli exclusion principle.) It enormously simplifies the task of elementary particle physics: we don't have to worry about big electrons and little ones, or new electrons and old ones – an electron is an electron is an electron. It didn't have to be so easy.

My first job, then, is to introduce you to the various kinds of elementary particles – the actors, if you will, in the drama. I could simply *list* them, and tell you their properties (mass, electric charge, spin, etc.), but I think it is better in this case to adopt a historical perspective, and explain how each particle first came on the scene. This will serve to endow them with character and personality, making them easier to remember and more interesting to watch. Moreover, some of the stories are delightful in their own right.

Once the particles have been introduced, in Chapter 1, the issue becomes, 'How do they interact with one another?' This question, directly or indirectly, will occupy us for the rest of the book. If you were dealing with two *macroscopic* objects, and you wanted to know how they interact, you would probably begin by holding them at various separation distances and measuring the force between them. That's how Coulomb determined the law of electrical repulsion between two charged pith balls, and how Cavendish measured the gravitational attraction of two lead weights. But you can't pick up a proton with tweezers or tie an electron onto the end of a piece of string; they're just too small. For practical reasons, therefore, we have to resort to

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

less direct means to probe the interactions of elementary particles. As it turns out, almost all of our experimental information comes from three sources: (1) scattering events, in which we fire one particle at another and record (for instance) the angle of deflection; (2) decays, in which a particle spontaneously disintegrates and we examine the debris; and (3) bound states, in which two or more particles stick together, and we study the properties of the composite object. Needless to say, determining the interaction law from such indirect evidence is not a trivial task. Ordinarily, the procedure is to *guess* a form for the interaction and compare the resulting theoretical predictions with the experimental data.

The formulation of such a guess ('model' is a more respectable term for it) is guided by certain general principles, in particular, special relativity and quantum mechanics. In the diagram below I have sketched out four realms of mechanics:

	Small →	
Fast ↓	Classical mechanics	Quantum mechanics
	Relativistic mechanics	Quantum field theory

The world of everyday life, of course, is governed by classical mechanics. But for objects that travel very fast (at speeds comparable to c), the classical rules are modified by special relativity, and for objects that are very small (comparable to the size of atoms, roughly speaking), classical mechanics is superseded by quantum mechanics. Finally, for things that are both fast *and* small, we require a theory that incorporates relativity and quantum principles: quantum field theory. Now, elementary particles *are* extremely small, of course, and typically they are also very fast. So, elementary particle physics naturally falls under the dominion of quantum field theory.

Please observe the distinction here between a *type of mechanics* and a *particular force law*. Newton's law of universal gravitation, for example, describes a specific interaction (gravity), whereas Newton's three laws of motion define a mechanical system (classical mechanics), which (within its jurisdiction) governs *all* interactions. The force law tells you what F *is*, in the case at hand; the mechanics tells you how to *use* F to determine the motion. The goal of elementary particle dynamics, then, is to guess a set of force laws which, within the context of quantum field theory, correctly describe particle behavior.

However, some general features of this behavior have nothing to do with the detailed form of the interactions. Instead they follow directly from relativity, from quantum mechanics, or from the combination of the two. For example, in relativity, energy and momentum are always conserved, but (rest) mass is not. Thus the decay $\Delta \to p + \pi$ is perfectly acceptable, even though the Δ weighs more than the sum of p plus π. Such a process would not be possible in classical mechanics, where mass is strictly conserved. Moreover, relativity allows for particles of zero (rest) mass – the very idea of a massless particle is nonsense in classical mechanics – and as we shall see, photons and gluons *are* massless.

In quantum mechanics a physical system is described by its *state*, s (represented by the wave function ψ_s in Schrödinger's formulation, or by the *ket* $|s\rangle$ in Dirac's theory). A physical process, such as scattering or decay, consists of a *transition* from one state to another. But in quantum mechanics the outcome is not uniquely determined by the initial conditions; all we can hope to calculate, in general, is the *probability* for a given transition to occur. This indeterminacy is reflected in the observed behavior of particles. For example, the charged pi meson ordinarily disintegrates into a muon plus a neutrino, but occasionally one will decay into an *electron* plus a neutrino. There's no difference in the original pi mesons; they're all identical. It is simply a fact of nature that a given particle can go either way.

Finally, the union of relativity and quantum mechanics brings certain extra dividends that neither one can offer by itself: the existence of antiparticles (with the same mass and lifetime as the particle itself, but opposite electric charge), a proof of the Pauli exclusion principle (which in nonrelativistic quantum mechanics is simply an ad hoc hypothesis), and the so-called *TCP theorem*. I'll tell you more about these later on; my purpose in mentioning them here is to emphasize that these are features of the mechanical system itself, not of the particular model. Short of a catastrophic revolution, they are untouchable. By the way, quantum field theory in all its glory is difficult and deep, but don't be alarmed: Feynman invented a beautiful and intuitively satisfying formulation that is not hard to learn; we'll come to that in Chapter 6. (The *derivation* of Feynman's rules from the underlying quantum field theory is a different matter, which can easily consume the better part of an advanced graduate course, but this need not concern us here.)

In the 1960s and 1970s a theory emerged that described all of the known elementary particle interactions, except gravity. (As far as we can tell, gravity is much too weak to play any significant role in ordinary particle processes.) This theory – or, more accurately, this collection of related theories, based on two families of elementary particles (quarks and leptons), and incorporating quantum electrodynamics, the Glashow–Weinberg–Salam theory of electroweak processes, and quantum chromodynamics – has come to be called the *Standard Model*. No one pretends that it is the final word on the subject, but at least we are now playing with a full deck of cards. Since 1978, when the Standard Model achieved the status of 'orthodoxy', it has met every experimental test. Moreover, it has an attractive aesthetic feature: all of the fundamental interactions derive from one general principle, the requirement of *local gauge invariance*. It seems certain that future developments will involve extensions of the Standard Model, not its repudiation. This book might be called an 'Introduction to the Standard Model'.

As that alternative title suggests, it is a book about elementary particle *theory*, with very little on experimental methods or instrumentation. These are important matters, and an argument can be made for integrating them into a text such as this, but they can also be distracting, and interfere with the clarity and elegance of the theory itself. I encourage you to read about the experimental aspects of the subject, and from time to time I will refer you to particularly accessible accounts. But for now, I'll confine myself to scandalously brief answers to the two most obvious experimental questions.

How Do You *Produce* Elementary Particles?

Electrons and protons are no problem; these are the stable constituents of ordinary matter. To produce electrons one simply heats up a piece of metal, and they come boiling off. If you want a *beam* of electrons, you just set up a positively charged plate nearby, to attract them over, and poke a small hole in it; the electrons that make it through the hole constitute the beam. Such an *electron gun* is the starting element in a television tube or an oscilloscope or an electron accelerator (Figure I.1).

To obtain protons you ionize hydrogen (in other words, strip off the electron). In fact, if you're using the protons as a *target*, you don't even need to bother about the electrons; they're so light that an energetic incident particle will knock them out of the way. Thus, a tank of hydrogen *is* essentially a tank of protons. For more exotic particles there are three main sources: cosmic rays, nuclear reactors, and particle accelerators.

- *Cosmic rays:* The earth is constantly bombarded with high-energy particles (principally protons) coming from outer space. What the source of these particles might be remains something of a mystery; at any rate, when they hit atoms in the upper atmosphere they produce showers of secondary particles (mostly muons and neutrinos, by the

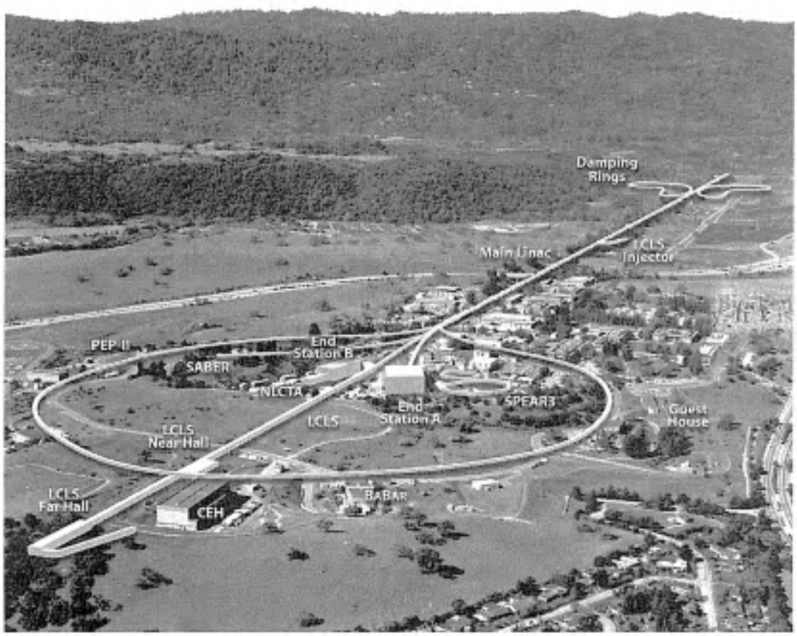


Fig. I.1 SLAC; the straight line is the accelerator itself. (Courtesy Stanford Linear Accelerator Center.)

time they reach ground level), which rain down on us all the time. As a source of elementary particles, cosmic rays have two virtues: they are free, and their energies can be enormous – far greater than we could possibly produce in the laboratory. But they have two major disadvantages: the rate at which they strike any detector of reasonable size is very low, and they are completely uncontrollable. So cosmic ray experiments call for patience and luck.

- *Nuclear reactors:* When a radioactive nucleus disintegrates, it may emit a variety of particles – neutrons, neutrinos, and what used to be called alpha rays (actually, alpha *particles*, which are bound states of two neutrons plus two protons), beta rays (actually, electrons or positrons), and gamma rays (actually, photons).
- *Particle accelerators:* You start with electrons or protons, accelerate them to high energy, and smash them into a target (Figure I.1). By skillful arrangements of absorbers and magnets, you can separate the particle species that you wish to study from the resulting debris. Nowadays it is possible in this way to generate intense secondary beams of positrons, muons, pions, kaons, *B*-mesons, antiprotons, and neutrinos, which in turn can be fired at another target. The stable particles – electrons, protons, positrons, and antiprotons – can even be fed into giant *storage rings* in which, guided by powerful magnets, they circulate at high speed for hours at a time, to be extracted and used at the required moment [1].

In general, the heavier the particle you want to produce, the higher must be the energy of the collision. That's why, historically, lightweight particles tend to be discovered first, and as time goes on, and accelerators become more powerful, heavier and heavier particles are found. It turns out that you gain enormously in *relative* energy if you collide two high-speed particles head-on, as opposed to firing one particle at a stationary target. (Of course, this calls for much better aim!) For this reason many contemporary experiments involve colliding beams from intersecting storage rings; if the particles miss on the first pass, they can try again the next time around. Indeed, with electrons and positrons (or protons and antiprotons) the *same ring* can be used, with the plus charges circulating in one direction and minus charges the other. Unfortunately, when a charged particle accelerates it radiates, thereby losing energy. In the case of circular motion (which, of course, involves acceleration) this is called *synchrotron radiation*, and it severely limits the efficiency of storage rings for energetic electrons (heavier particles with the same energy accelerate less, so synchrotron radiation is not such a problem for them). For this reason electron scattering experiments will increasingly turn to *linear* colliders, while storage rings will continue to be used for protons and heavier particles.

There is another reason why particle physicists are always pushing for higher energies: in general, the higher the energy of the collision, the closer the two particles come to one another. So if you want to study an interaction at very short range, you need very energetic particles. In quantum-mechanical terms, a particle of momentum p has an associated wavelength λ given by the de Broglie formula $\lambda = h/p$, where h is Planck's constant. At large wavelengths (low momenta) you can only hope to resolve relatively large structures; in order to examine something extremely small, you need comparably short wavelengths, and hence high momenta. If you like, consider this a manifestation of the uncertainty principle ($\Delta x \Delta p \geq h/4\pi$) – to make Δx small, Δp must be large. However you look at it, the conclusion is the same: to probe *small distances* you need *high energies*.

At present the most powerful accelerator in the world is the *Tevatron* at Fermilab (Figure I.2), with a maximum beam energy of almost 1 TeV. The tevatron (a proton–antiproton collider) began operation in 1983; its successor, the Superconducting Supercollider (SSC) was under construction in 1993 when the project was terminated by Congress. As a result, there has been a long period in which no fundamental progress was possible. This dry spell should end in 2008, when the Large Hadron Collider (LHC) at CERN starts taking data (Figure I.3). The LHC is designed to reach beam energies in excess of 7 TeV, and the hope is that this new terrain will include the Higgs particle, possibly supersymmetric particles, and – best of all – something completely unexpected [2]. It's not clear what comes after the LHC – most likely the proposed International Linear Collider (ILC). But, accelerators have become so huge (the SSC would have been 87 km in circumference) that there is not much room for expansion. Perhaps we are approaching the end

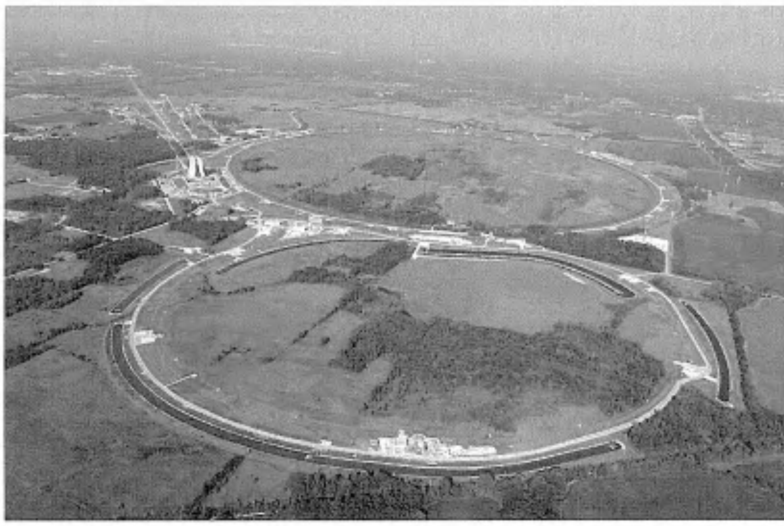


Fig. I.2 Fermilab; the large circle in the background is the Tevatron. (Courtesy Fermilab Visual Media Services.)

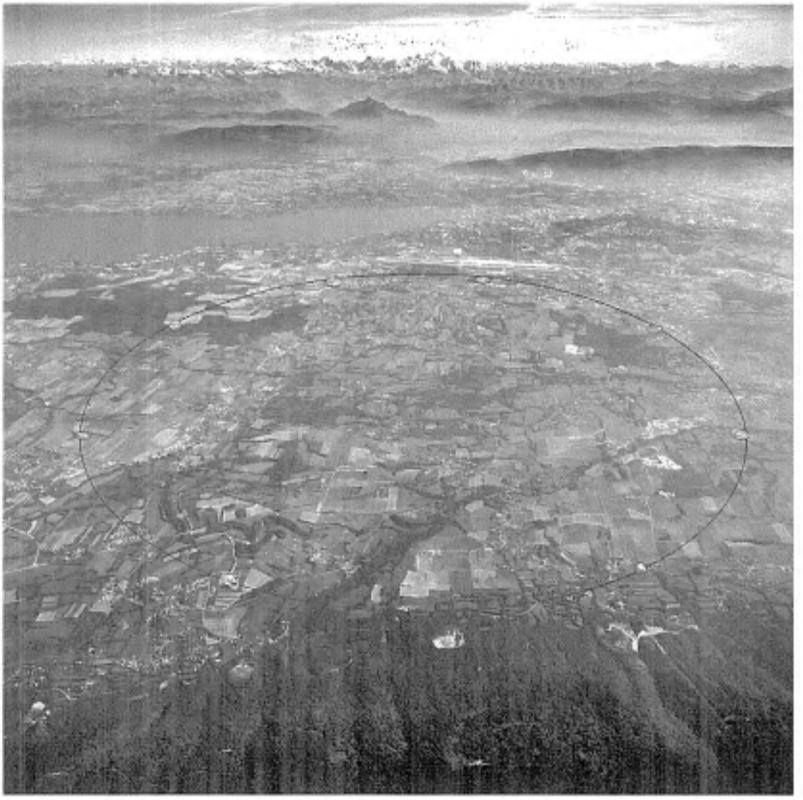


Fig. 1.3 CERN; the circle indicates the path of the LHC tunnel (formerly LEP) – Geneva and Mt Blanc are in the background. (Courtesy CERN.)

of the accelerator era, and particle physicists will have to turn to astrophysics and cosmology for information about higher energies. Or perhaps someone will have a clever new idea for squeezing energy onto an elementary particle.*

How Do You *Detect* Elementary Particles?

There are many kinds of particle detectors – Geiger counters, cloud chambers, bubble chambers, spark chambers, drift chambers, photographic emulsions, Čerenkov counters, scintillators, photomultipliers, and so on. Actually, a typical modern

* In macroscopic terms the amount of energy involved is not that great – after all, 1 TeV (10^{12} eV) is only 10^{-7} Joules; the problem is how to *deliver* that energy to a particle. No law of physics prevents you from doing so, but nobody has yet figured out a way to do it without gigantic (and expensive) machinery.

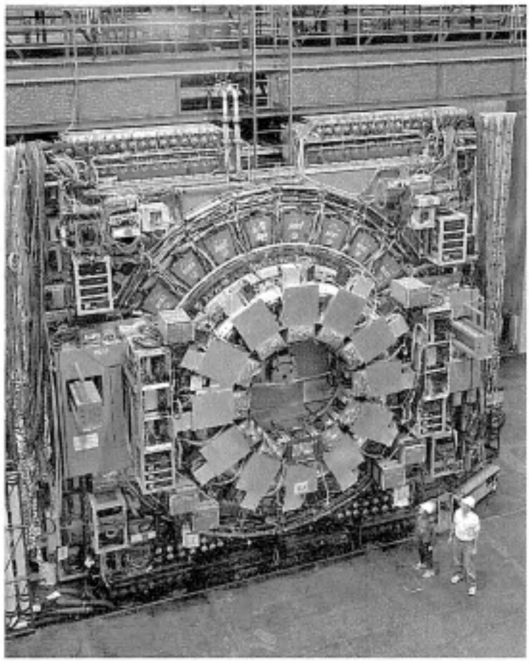


Fig. I.4 The CDF detector at Fermilab, where the top quark was discovered. (Courtesy Fermilab Visual Media Services.)

detector has whole arrays of these devices, wired up to a computer that tracks the particles and displays their trajectories on a television screen (Figure I.4). The details do not concern us, but there is one thing you should be aware of: most detection mechanisms rely on the fact that when high-energy charged particles pass through matter they ionize atoms along their path. The ions then act as 'seeds' in the formation of droplets (cloud chamber) or bubbles (bubble chamber) or sparks (spark chamber), as the case may be. But electrically *neutral* particles do not cause ionization, and they leave no tracks. For instance, if you look at the bubble chamber photograph in Figure 1.9, you will see that the five neutral particles are 'invisible'; their paths have been reconstructed by analyzing the tracks of the *charged* particles in the picture and invoking conservation of energy and momentum at each vertex. Notice also that most of the tracks in the picture are *curved* (actually, *all* of them are, to some extent; try holding a ruler up to one you think is straight). The bubble chamber was placed between the poles of a giant magnet; in a magnetic field B, a particle of charge q and momentum p will move in a circle of radius R given by the famous *cyclotron formula*: $R = pc/qB$, where c is the speed of light. The curvature of the track in a known magnetic field thus affords a very simple measure of the particle's momentum. Moreover, we can immediately tell the *sign* of the charge from the *direction* of the curve.

Units

Elementary particles are small, so for our purposes the *normal* mechanical units – grams, ergs, joules, and so on – are inconveniently large. Atomic physicists introduced the *electron volt* – the energy acquired by an electron when accelerated through a potential difference of 1 volt: 1 eV $= 1.6 \times 10^{-19}$ joules. For us the eV is inconveniently *small*, but we're stuck with it. Nuclear physicists use keV (10^3 eV); typical energies in particle physics are MeV (10^6 eV), GeV (10^9 eV), or even TeV (10^{12} eV). Momenta are measured in MeV/c (or GeV/c, or whatever), and masses in MeV/c^2. Thus the proton weighs 938 MeV/$c^2 = 1.67 \times 10^{-24}$ g.

Actually, particle theorists are lazy (or clever, depending on your point of view) – they seldom include the c's and $\hbar$'s ($\hbar \equiv h/2\pi$) in their formulas. You're just supposed to fit them in for yourself at the end, to make the dimensions come out right. As they say in the business, 'set $c = \hbar = 1$'. This amounts to working in units such that time is measured in centimeters and mass and energy in inverse centimeters; the unit of time is the time it takes light to travel 1 cm, and the unit of energy is the energy of a photon whose wavelength is 2π cm. Only at the end of the problem do we revert to conventional units. This makes everything look very elegant, but I thought it would be wiser in this book to keep all the c's and $\hbar$'s where they belong, so that you can check for dimensional consistency as you go along. (If this offends you, remember that it is easier for you to ignore an $\hbar$ you don't like than for someone else to conjure one up in just the right place.)

Finally, there is the question of what units to use for electric charge. In introductory physics courses most instructors favor the *SI* system, in which charge is measured in *coulombs*, and Coulomb's law reads

$$F = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r^2} \qquad \text{(SI)}$$

Most advanced work is done in the *Gaussian* system, in which charge is measured in *electrostatic units* (esu), and Coulomb's law is written

$$F = \frac{q_1 q_2}{r^2} \qquad \text{(G)}$$

But elementary particle physicists prefer the *Heaviside – Lorentz* system, in which Coulomb's law takes the form

$$F = \frac{1}{4\pi}\frac{q_1 q_2}{r^2} \qquad \text{(HL)}$$

The three units of charge are related as follows:

$$q_{\text{HL}} = \sqrt{4\pi}\, q_{\text{G}} = \frac{1}{\sqrt{\epsilon_0}}\, q_{\text{SI}}$$

In this book I shall use Gaussian units exclusively, in order to avoid unnecessary confusion in an already difficult subject. Whenever possible I will express results in terms of the *fine structure constant*

$$\alpha = \frac{e^2}{\hbar c} = \frac{1}{137.036}$$

where e is the charge of the electron in Gaussian units. Most elementary particle texts write this as $e^2/4\pi$, because they are measuring charge in Heaviside–Lorentz units and setting $c = \hbar = 1$; but everyone agrees that the *number* is 1/137.

Further reading

Since the early 1960s, the Particle Data Group at Berkeley has periodically issued a listing of the established particles and their properties. These are published every other year in *Reviews of Modern Physics* or *Journal of Physics G*, and summarized in a (free) booklet that can be ordered on the web at http:*pdg.lbl.gov*. In the early days this summary took the form of 'wallet cards', but by 2006 it had grown to a densely packed 315 pages. I shall refer to it as the *Particle Physics Booklet* (PPB). Every student of elementary particle physics must have a copy – don't leave home without it! The longer version, called the *Review of Particle Physics* (RPP) is the bible for professionals – the 2006 edition runs to 1231 pages, and it includes authoritative articles on every relevant subject, written by the world's leading experts [3]. If you want the definitive, up-to-date word on any particular topic, this is the place to go (it is also available on-line, at the Particle Data Group web site).

Particle physics is an enormous and rapidly changing subject. My aim in this book is to introduce you to some important ideas and methods, to give you a sense of what's out there to be learned, and perhaps to stimulate your appetite for more. If you want to read further in quantum field theory, I particularly recommend:

Bjorken, J. D. and Drell, S. D. (**1964**) *Relativistic Quantum Mechanics and Relativistic Quantum Fields*, McGraw-Hill, New York.

Itzykson, C. and Zuber, J.-B. (**1980**) *Quantum Field Theory*, McGraw-Hill, New York.

Peskin, M. E. and Schroeder, D. V. (**1995**) *An Introduction to Quantum Field Theory*, Perseus, Cambridge, M.A.

Ryder, L. H. (**1985**) *Quantum Field Theory*, Cambridge University Press, Cambridge, UK.

Sakurai, J. J. (**1967**) *Advanced Quantum Mechanics*, Addison-Wesley, Reading, M.A.

I warn you, however, that these are all difficult and advanced books.

For elementary particle physics itself, the following books (listed in order of increasing difficulty) are especially useful:

Close, F., Marten, M. and Sutton, C. (**1987**) *The Particle Explosion*, Oxford University Press, Oxford, UK.

Frauenfelder, H. and Henley, E. M. (**1991**) *Subatomic Physics*, 2nd edn, Prentice-Hall, Englewood Cliffs, N.J.

Gottfried, K. and Weisskopf, V. F. (**1984**) *Concepts of Particle Physics*, Oxford University Press, Oxford.

Perkins, D. H. (**2000**) *Introduction to High-Energy Physics*, 4th Ed, Cambridge University Press, Cambridge, UK.

Halzen, F. and Martin, A. D. (**1984**) *Quarks and Leptons*, John Wiley & Sons, Ltd, New York.

Roe, B. P. (**1996**) *Particle Physics at the New Millennium*, Springer, New York.

Aitchison, I. J. R. and Hey, A. J. G. (**2003**) *Gauge Theories in Particle Physics*, 3rd edn, Institute of Physics, Bristol, UK.

Seiden, A. (**2005**) *Particle Physics: A Comprehensive Introduction*, Addison-Wesley, San Francisco, C.A.

Quigg, C. (**1997**) *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions*, Addison-Wesley, Reading, M.A.

References

1 For a comprehensive bibliography see Chao, Q. W. (**2006**) *American Journal of Physics*, **74**, 855.

2 Smith, C. L. (July **2000**) *Scientific American*, 71; (a) Lederman, L. (**2007**) *Nature*, **448**, 310.

3 The current reference is Yao, W.-M. *et al.* (**2006**) *Journal of Physics*, **G 33**, 1. But because you will want to use the most recent versions, I will simply refer to them as *Particle Physics Booklet* and *Review of Particle Physics*, appending the year when it is relevant.

1

Historical Introduction to the Elementary Particles

This chapter is a kind of 'folk history' of elementary particle physics. Its purpose is to provide a sense of how the various particles were first discovered, and how they fit into the overall scheme of things. Along the way some of the fundamental ideas that dominate elementary particle theory are explained. This material should be read quickly, as background to the rest of the book. (As history, the picture presented here is certainly misleading, for it sticks closely to the main track, ignoring the false starts and blind alleys that accompany the development of any science. That's why I call it 'folk' history – it's the way particle physicists like to remember the subject – a succession of brilliant insights and heroic triumphs unmarred by foolish mistakes, confusion, and frustration. It wasn't really quite so easy.)

1.1
The Classical Era (1897–1932)

It is a little artificial to pinpoint such things, but I'd say that elementary particle physics was born in 1897, with J. J. Thomson's discovery of the electron [1]. (It is fashionable to carry the story all the way back to Democritus and the Greek atomists, but apart from a few suggestive words their metaphysical speculations have nothing in common with modern science, and although they may be of modest antiquarian interest, their genuine relevance is negligible.) Thomson knew that *cathode rays* emitted by a hot filament could be deflected by a magnet. This suggested that they carried electric charge; in fact, the direction of the curvature required that the charge be negative. It seemed, therefore, that these were not rays at all, but rather streams of particles. By passing the beam through crossed electric and magnetic fields, and adjusting the field strength until the net deflection was zero, Thomson was able to determine the velocity of the particles (about a tenth the speed of light) as well as their charge-to-mass ratio (Problem 1.1). This ratio turned out to be enormously greater than for any known ion, indicating either that the charge was extremely large or the mass was very small. Indirect evidence pointed to the second conclusion. Thomson called the particles *corpuscles*. Back in 1891, George Johnstone Stoney had introduced the term 'electron' for the fundamental unit of charge; later, that name was taken over for the particles themselves.

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

Thomson correctly surmised that these electrons were essential constituents of atoms; however, since atoms as a whole are electrically neutral and very much heavier than electrons, there immediately arose the problem of how the compensating plus charge – and the bulk of the mass – is distributed within an atom. Thomson himself imagined that the electrons were suspended in a heavy, positively charged paste, like (as he put it) the plums in a pudding. But Thomson's model was decisively repudiated by Rutherford's famous scattering experiment, which showed that the positive charge, and most of the mass, was concentrated in a tiny core, or *nucleus*, at the center of the atom. Rutherford demonstrated this by firing a beam of α particles (ionized helium atoms) into a thin sheet of gold foil (Figure 1.1). Had the gold atoms consisted of rather diffuse spheres, as Thomson supposed, then all of the α particles should have been deflected a bit, but none would have been deflected much – any more than a bullet is deflected much when it passes, say, through a bag of sawdust. What in *fact* occurred was that *most* of the α particles passed through the gold completely undisturbed, but a few of them bounced off at wild angles. Rutherford's conclusion was that the α particles had

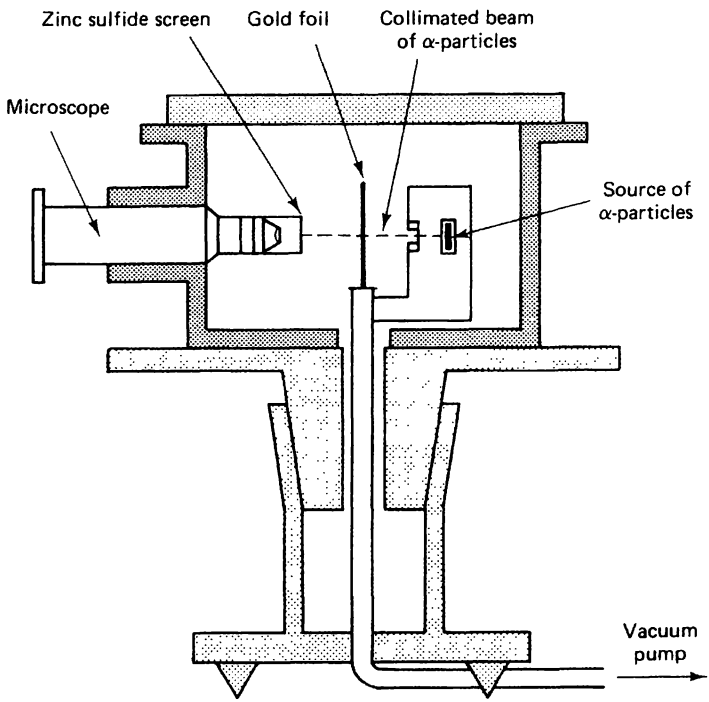


Fig. 1.1 Schematic diagram of the apparatus used in the Rutherford scattering experiment. Alpha particles scattered by the gold foil strike a fluorescent screen, giving off a flash of light, which is observed visually through a microscope.

encountered something very small, very hard, and very heavy. Evidently the positive charge, and virtually all of the mass, was concentrated at the center, occupying only a tiny fraction of the volume of the atom (the electrons are too light to play any role in the scattering; they are knocked right out of the way by the much heavier α particles).

The nucleus of the lightest atom (hydrogen) was given the name *proton* by Rutherford. In 1914 Niels Bohr proposed a model for hydrogen consisting of a single electron circling the proton, rather like a planet going around the sun, held in orbit by the mutual attraction of opposite charges. Using a primitive version of the quantum theory, Bohr was able to calculate the spectrum of hydrogen, and the agreement with experiment was nothing short of spectacular. It was natural then to suppose that the nuclei of heavier atoms were composed of two or more protons bound together, supporting a like number of orbiting electrons. Unfortunately, the next heavier atom (helium), although it does indeed carry two electrons, weighs *four* times as much as hydrogen, and lithium (three electrons) is *seven* times the weight of hydrogen, and so it goes. This dilemma was finally resolved in 1932 with Chadwick's discovery of the neutron – an electrically neutral twin to the proton. The helium nucleus, it turns out, contains two neutrons in addition to the two protons; lithium evidently includes four; and, in general, the heavier nuclei carry very roughly the same number of neutrons as protons. (The number of neutrons is in fact somewhat flexible – the same atom, chemically speaking, may come in several different *isotopes*, all with the same number of protons, but with varying numbers of neutrons.)

The discovery of the neutron put the final touch on what we might call the *classical period* in elementary particle physics. Never before (and I'm sorry to say never since) has physics offered so simple and satisfying an answer to the question, 'What is matter made of?' In 1932, it was all just protons, neutrons, and electrons. But already the seeds were planted for the three great ideas that were to dominate the *middle period* (1930–1960) in particle physics: Yukawa's meson, Dirac's positron, and Pauli's neutrino. Before we come to that, however, I must back up for a moment to introduce the photon.

1.2 The Photon (1900–1924)

In some respects, the photon is a very 'modern' particle, having more in common with the W and Z (which were not discovered until 1983) than with the classical trio. Moreover, it's hard to say exactly when or by whom the photon was really 'discovered', although the essential stages in the process are clear enough. The first contribution was made by Planck in 1900. Planck was attempting to explain the so-called *blackbody spectrum* for the electromagnetic radiation emitted by a hot object. Statistical mechanics, which had proved brilliantly successful in explaining other thermal processes, yielded nonsensical results when applied to electromagnetic fields. In particular, it led to the famous 'ultraviolet catastrophe', predicting that the total power radiated should be *infinite*. Planck found that he could escape the ultraviolet catastrophe – and fit the experimental curve – if he assumed

that electromagnetic radiation is *quantized*, coming in little 'packages' of energy

$$E = h\nu \tag{1.1}$$

where ν is the frequency of the radiation and h is a constant, which Planck adjusted to fit the data. The modern value of Planck's constant is

$$h = 6.626 \times 10^{-27} \text{ erg s} \tag{1.2}$$

Planck did not profess to know *why* the radiation was quantized; he assumed that it was due to a peculiarity in the emission process: for some reason a hot surface only gives off light* in little squirts.

Einstein, in 1905, put forward a far more radical view. He argued that quantization was a feature of the electromagnetic field itself, having nothing to do with the emission mechanism. With this new twist, Einstein adapted Planck's idea, and his formula, to explain the *photoelectric effect*: when electromagnetic radiation strikes a metal surface, electrons come popping out. Einstein suggested that an incoming light quantum hits an electron in the metal, giving up its energy ($h\nu$); the excited electron then breaks through the metal surface, losing in the process an energy w (the so-called *work function* of the material – an empirical constant that depends on the particular metal involved). The electron thus emerges with an energy

$$E \leq h\nu - w \tag{1.3}$$

(It may lose some energy before reaching the surface; that's the reason for the *in*equality.) Einstein's formula (Equation 1.3) is trivial to *derive*, but it carries an extraordinary implication: The maximum electron energy is *independent of the intensity of the light* and depends only on its *color* (frequency). To be sure, a more intense beam will knock out *more* electrons, but their *energies* will be the same.

Unlike Planck's theory, Einstein's met a hostile reception, and over the next 20 years he was to wage a lonely battle for the *light quantum* [2]. In saying that electromagnetic radiation is *by its nature* quantized, regardless of the emission mechanism, Einstein came dangerously close to resurrecting the discredited particle theory of light. Newton, of course, had introduced such a *corpuscular* model, but a major achievement of nineteenth-century physics was the decisive repudiation of Newton's idea in favor of the rival wave theory. No one was prepared to see that accomplishment called into question, even when the experiments came down on Einstein's side. In 1916 Millikan completed an exhaustive study of the photoelectric effect and was obliged to report that 'Einstein's photoelectric equation ... appears in every case to predict exactly the observed results. ... Yet the semicorpuscular theory by which Einstein arrived at his equation seems at present wholly untenable' [3].

* In this book the word *light* stands for *electromagnetic radiation*, whether or not it happens to fall in the visible region.

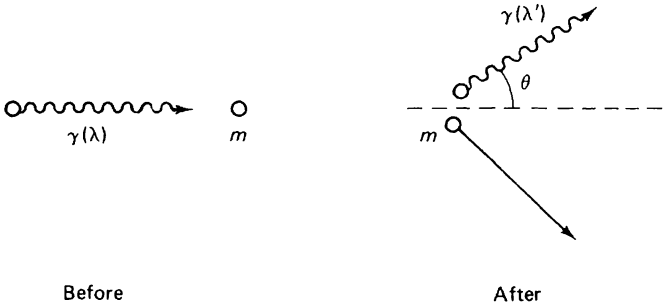


Fig. 1.2 Compton scattering. A photon of wavelength λ scatters off a particle, initially at rest, of mass m. The scattered photon carries wavelength λ' given by Equation 1.4.

What finally settled the issue was an experiment conducted by A. H. Compton in 1923. Compton found that the light scattered from a particle at rest is shifted in wavelength, according to the equation

$$\lambda' = \lambda + \lambda_c(1 - \cos\theta) \tag{1.4}$$

where λ is the incident wavelength, λ' is the scattered wavelength, θ is the scattering angle, and

$$\lambda_c = \mathrm{h}/mc \tag{1.5}$$

is the so-called *Compton wavelength* of the target particle (mass m). Now, this is *precisely* the formula you get (Problem 3.27) if you treat light as a particle of zero rest mass with energy given by Planck's equation, and apply the laws of conservation of (relativistic) energy and momentum – just as you would for an ordinary elastic collision (Figure 1.2). That clinched it; here was direct and incontrovertible experimental evidence that light behaves as a particle, on the subatomic scale. We call this particle the *photon* (a name suggested by the chemist Gilbert Lewis, in 1926); the symbol for a photon is γ (from *gamma ray*). How the particle nature of light on this level is to be reconciled with its well-established wave behavior on the macroscopic scale (exhibited in the phenomena of interference and diffraction) is a story I'll leave for books on quantum mechanics.

Although the photon initially *forced* itself on an unreceptive community of physicists, it eventually found a natural place in quantum field theory, and was to offer a whole new perspective on electromagnetic interactions. In classical electrodynamics, we attribute the electrical repulsion of two electrons, say, to the electric field surrounding them; each electron contributes to the field, and each one responds to the field. But in quantum field theory, the electric field is *quantized* (in the form of photons), and we may picture the interaction as consisting of a stream of photons passing back and forth between the two charges, each electron continually emitting photons and continually absorbing them. And the same goes for *any* noncontact

force: Where classically we interpret 'action at a distance' as 'mediated' by a *field*, we now say that it is mediated by an *exchange of particles* (the *quanta* of the field). In the case of electrodynamics, the mediator is the photon; for gravity, it is called the *graviton* (though a fully successful quantum theory of gravity has yet to be developed and it may well be centuries before anyone detects a graviton experimentally).

You will see later on how these ideas are implemented in practice, but for now I want to dispel one common misapprehension. When I say that every force is mediated by the exchange of particles, I am *not* speaking of a merely *kinematic* phenomenon. Two ice skaters throwing snowballs back and forth will of course move apart with the succession of recoils; they 'repel one another by exchange of snowballs', if you like. But that's *not* what is involved here. For one thing, this mechanism would have a hard time accounting for an *attractive* force. You might think of the mediating particles, rather, as 'messengers', and the message can just as well be 'come a little closer' as 'go away'.

I said earlier that in the 'classical' picture ordinary matter is made of atoms, in which electrons are held in orbit around a nucleus of protons and neutrons by the electrical attraction of opposite charges. We can now give this model a more sophisticated formulation by attributing the binding force to the exchange of photons between the electrons and the protons in the nucleus. However, for the purposes of atomic physics this is overkill, for in this context quantization of the electromagnetic field produces only minute effects (notably the Lamb shift and the anomalous magnetic moment of the electron). To excellent approximation we can pretend that the forces are given by Coulomb's law (together with various magnetic dipole couplings). The point is that in a bound state enormous numbers of photons are continually streaming back and forth, so that the 'lumpiness' of the field is effectively smoothed out, and classical electrodynamics is a suitable approximation to the truth. But in most elementary particle processes, such as the photoelectric effect or Compton scattering, *individual* photons are involved, and quantization can no longer be ignored.

1.3
Mesons (1934–1947)

Now there is one conspicuous problem to which the 'classical' model does not address itself at all: what holds the *nucleus* together? After all, the positively charged protons should repel one another violently, packed together as they are in such close proximity. Evidently there must be some other force, more powerful than the force of electrical repulsion, that binds the protons (and neutrons) together; physicists of that less imaginative age called it, simply, the *strong force*. But if there exists such a potent force in nature, why don't we notice it in everyday life? The *fact* is that virtually every force we experience directly, from the contraction of a muscle to the explosion of dynamite, is electromagnetic in origin; the only exception, outside a nuclear reactor or an atomic bomb, is gravity. The answer must be that, powerful though it is, the strong force is of very short *range*. (The range of a force is like the

arm's reach of a boxer – beyond that distance its influence falls off rapidly to zero. Gravitational and electromagnetic forces have *infinite* range, but the range of the strong force is about the size of the nucleus itself.)*

The first significant theory of the strong force was proposed by Yukawa in 1934 [4]. Yukawa assumed that the proton and neutron are attracted to one another by some sort of *field*, just as the electron is attracted to the nucleus by an electric field and the moon to the earth by a gravitational field. This field should properly be quantized, and Yukawa asked the question: what must be the properties of its *quantum* – the particle (analogous to the photon) whose exchange would account for the known features of the strong force? For example, the short range of the force indicated that the mediator would be rather heavy; Yukawa calculated that its mass should be nearly 300 times that of the electron, or about a sixth the mass of a proton (see Problem 1.2). Because it fell between the electron and the proton, Yukawa's particle came to be known as the *meson* (meaning 'middle-weight'). In the same spirit, the electron is called a *lepton* ('light-weight'), whereas the proton and neutron are *baryons* ('heavy-weight'). Now, Yukawa knew that no such particle had ever been observed in the laboratory, and he therefore assumed his theory was wrong. But at that time a number of systematic studies of cosmic rays were in progress, and by 1937 two separate groups (Anderson and Neddermeyer on the West Coast, and Street and Stevenson on the East) had identified particles matching Yukawa's description.† Indeed, the cosmic rays with which you are being bombarded every few seconds as you read this consist primarily of just such middle-weight particles.

For a while everything seemed to be in order. But as more detailed studies of the cosmic ray particles were undertaken, disturbing discrepancies began to appear. They had the wrong lifetime and they seemed to be significantly lighter than Yukawa had predicted; worse still, different mass measurements were not consistent with one another. In 1946 (after a period in which physicists were engaged in a less savory business) decisive experiments were carried out in Rome demonstrating that the cosmic ray particles interacted very weakly with atomic nuclei [5]. If this was really Yukawa's meson, the transmitter of the strong force, the interaction should have been dramatic. The puzzle was finally resolved in 1947, when Powell and his coworkers at Bristol [6] discovered that there are actually *two* middle-weight particles in cosmic rays, which they called π (or 'pion') and μ (or 'muon'). (Marshak reached the same conclusion simultaneously, on theoretical grounds [7].) The true Yukawa meson is the π; it is produced copiously in the upper atmosphere, but ordinarily disintegrates long before reaching the ground (see Problem 3.4). Powell's group exposed their photographic emulsions on mountain tops (see Figure 1.3). One of the decay products is the lighter (and longer lived) μ, and it is primarily muons that one observes at sea level. In the search for Yukawa's meson, then, the muon was simply an impostor, having nothing whatever to do

* This is a bit of an oversimplification. Typically, the forces go like $e^{-(r/a)}/r^2$, where a is the 'range'. For Coulomb's law and Newton's law of universal gravitation, $a = \infty$; for the strong force a is about 10^{-13} cm (1 fm).

† Actually, it was Robert Oppenheimer who drew the connection between these cosmic ray particles and Yukawa's meson.

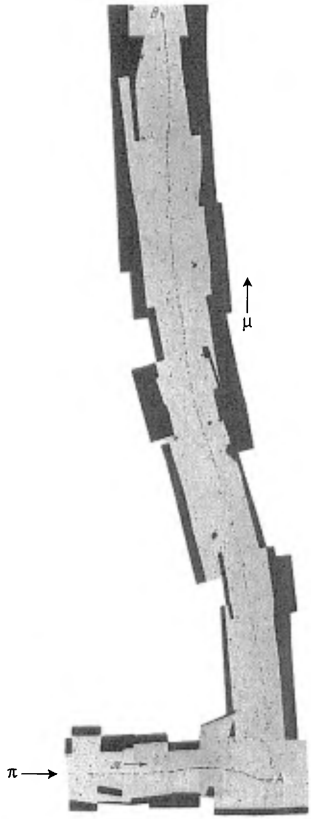


Fig. 1.3 One of Powell's earliest pictures showing the track of a pion in a photographic emulsion exposed to cosmic rays at high altitude. The pion (entering from the left) decays into a muon and a neutrino (the latter is electrically neutral, and leaves no track). (*Source*: Powell, C. F., Fowler, P. H. and Perkins, D. H. (1959) *The Study of Elementary Particles by the Photographic Method* Pergamon, New York. First published in (1947) *Nature* **159**, 694.)

with the strong interactions. In fact, it behaves in every way like a heavier version of the electron and properly belongs in the *lepton* family (though some people to this day call it the 'mu-meson' by force of habit).

1.4
Antiparticles (1930–1956)

Nonrelativistic quantum mechanics was completed in the astonishingly brief period 1923–1926, but the relativistic version proved to be a much thornier problem.

The first major achievement was Dirac's discovery, in 1927, of the equation that bears his name. The Dirac equation was supposed to describe free electrons with energy given by the relativistic formula $E^2 - \mathbf{p}^2c^2 = m^2c^4$. But it had a very troubling feature: for every positive-energy solution ($E = +\sqrt{\mathbf{p}^2c^2 + m^2c^4}$) it admitted a corresponding solution with *negative* energy ($E = -\sqrt{\mathbf{p}^2c^2 + m^2c^4}$). This meant that, given the natural tendency of every system to evolve in the direction of lower energy, the electron should 'runaway' to increasingly negative states, radiating off an infinite amount of energy in the process. To rescue his equation, Dirac proposed a resolution that made up in brilliance for what it lacked in plausibility: he postulated that the negative-energy states are all filled by an infinite 'sea' of electrons. Because this sea is always there, and perfectly uniform, it exerts no net force on anything, and we are not normally aware of it. Dirac then invoked the Pauli exclusion principle (which says that no two electrons can occupy the same state), to 'explain' why the electrons we *do* observe are confined to the positive-energy states. But if this is true, then what happens when we impart to one of the electrons in the 'sea' an energy sufficient to knock it into a positive-energy state? The *absence* of the 'expected' electron in the sea would be interpreted as a net positive charge in that location, and the absence of its expected negative energy would be seen as a net positive energy. Thus a 'hole in the sea' would function as an ordinary particle with *positive* energy and *positive* charge. Dirac at first hoped that these holes might be *protons*, but it was soon apparent that they had to carry the same mass as the electron itself – 2000 times too light to be a proton. No such particle was known at the time, and Dirac's theory appeared to be in trouble. What may have seemed a fatal defect in 1930, however, turned into a spectacular triumph in late 1931, with Anderson's discovery of the *positron* (Figure 1.4), a positively charged twin for the electron, with precisely the attributes Dirac required [8].

Still, many physicists were uncomfortable with the notion that we are awash in an infinite sea of invisible electrons, and in the 1940s Stuckelberg and Feynman provided a much simpler and more compelling interpretation of the negative-energy states. In the Feynman–Stuckelberg formulation, the negative-energy solutions are re-expressed as *positive*-energy states of a *different particle* (the positron); the electron and positron appear on an equal footing, and there is no need for Dirac's 'electron sea' or for its mysterious 'holes'. We'll see in Chapter 7 how this – the modern interpretation – works. Meantime, it turned out that the dualism in Dirac's equation is a profound and universal feature of quantum field theory: for *every* kind of particle there must exist a corresponding *antiparticle*, with the same mass but opposite electric charge. The positron, then, is the *antielectron*. (Actually, it is in principle completely arbitrary which one you call the 'particle' and which the 'antiparticle' – I could just as well have said that the electron is the antipositron. But since there are a lot of electrons around, and not so many positrons, we tend to think of electrons as 'matter' and positrons as 'antimatter'). The (negatively charged) antiproton was first observed experimentally at the Berkeley Bevatron in 1955, and the (neutral) antineutron was discovered at the same facility the following year [9].

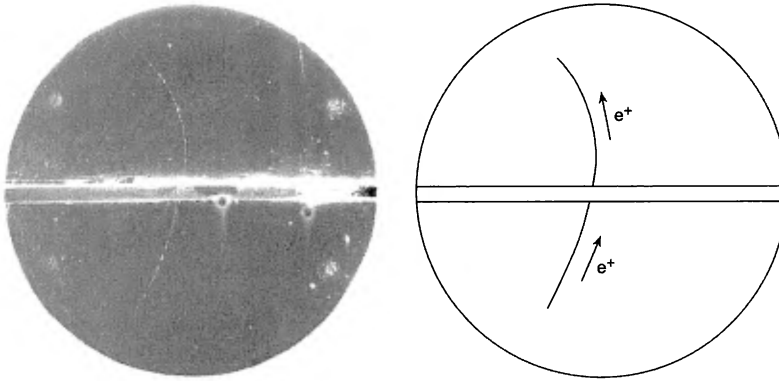


Fig. 1.4 The positron. In 1932, Anderson took this photograph of the track left in a cloud chamber by a cosmic ray particle. The chamber was placed in a magnetic field (pointing into the page), which caused the particle to travel in a curve. But was it a negative charge traveling downward or a positive charge traveling upward? In order to distinguish, Anderson had placed a lead plate across the center of the chamber (the thick horizontal line in the photograph). A particle passing through the plate slows down, and subsequently moves in a tighter circle. By inspection of the curves, it is clear that this particle traveled upward, and hence must have been positively charged. From the curvature of the track and from its texture, Anderson was able to show that the mass of the particle was close to that of the electron. (Photo courtesy California Institute of Technology.)

The standard notation for antiparticles is an overbar. For example, p denotes the proton and $\overline{p}$ the antiproton; n the neutron and $\overline{n}$ the antineutron. However, in some cases it is customary simply to specify the charge. Thus most people write e^+ for the positron (not $\overline{e}$) and μ^+ for the antimuon (not $\overline{\mu}$).* Some neutral particles are their *own* antiparticles. For example, the photon: $\overline{\gamma} \equiv \gamma$. In fact, you may have been wondering how the antineutron differs physically from the neutron, since both are uncharged. The answer is that neutrons carry other 'quantum numbers' besides charge (in particular, baryon number), which change sign for the antiparticle. Moreover, although its *net* charge is zero, the neutron *does* have a charge *structure* (positive at the center and near the surface, negative in between) and a magnetic dipole moment. These, too, have the opposite sign for $\overline{n}$.

There is a general principle in particle physics that goes under the name of *crossing symmetry*. Suppose that a reaction of the form

$$A + B \rightarrow C + D$$

is known to occur. Any of these particles can be 'crossed' over to the other side of the equation, provided it is turned into its antiparticle, and the resulting interaction

* But you must not *mix* conventions: $\overline{e}^+$ is ambiguous, like a double negative – the reader doesn't know if you mean the positron or the *anti*positron, (which is to say, the electron).

will also be allowed. For example,

$$A \to \overline{B} + C + D$$

$$A + \overline{C} \to \overline{B} + D$$

$$\overline{C} + \overline{D} \to \overline{A} + \overline{B}$$

In addition, the *reverse* reaction occurs: $C + D \to A + B$, but technically this derives from the principle of *detailed balance*, rather than from crossing symmetry. Indeed, as we shall see, the *calculations* involved in these various reactions are practically identical. We might almost regard them as different manifestations of the same fundamental process. However, there is one important *caveat* in all this: conservation of energy may veto a reaction that is otherwise permissible. For example, if A weighs less than the sum of B, C, and D, then the decay $A \to \overline{B} + C + D$ cannot occur; similarly, if A and C are light, whereas B and D are heavy, then the reaction $A + \overline{C} \to \overline{B} + D$ will not take place unless the initial kinetic energy exceeds a certain 'threshold' value. So perhaps I should say that the crossed (or reversed) reaction is *dynamically* permissible, but it may or may not be *kinematically* allowed. The power and beauty of crossing symmetry can scarcely be exaggerated. It tells us, for instance, that Compton scattering

$$\gamma + e^- \to \gamma + e^-$$

is 'really' the same process as pair annihilation

$$e^- + e^+ \to \gamma + \gamma$$

although in the laboratory they are completely different phenomena.

The union of special relativity and quantum mechanics, then, leads to a pleasing matter/antimatter symmetry. But this raises a disturbing question: how come *our* world is populated with protons, neutrons, and electrons, instead of antiprotons, antineutrons, and positrons? Matter and antimatter cannot coexist for long – if a particle meets its antiparticle, they annihilate. So maybe it's just a historical accident that in our corner of the universe there happened to be more matter than antimatter, and pair annihilation has vacuumed up all but a leftover residue of matter. If this is so, then presumably there are other regions of space in which antimatter predominates. Unfortunately, the astronomical evidence is pretty compelling that all of the observable universe is made of ordinary matter. In Chapter 12 we will explore some contemporary ideas about the 'matter–antimatter asymmetry'.

1.5 Neutrinos (1930–1962)

For the third strand in the story we return again to the year 1930 [10]. A problem had arisen in the study of nuclear beta decay. In beta decay, a radioactive nucleus A

is transformed into a slightly lighter nucleus B, with the emission of an electron:

$$A \rightarrow B + e^- \tag{1.6}$$

Conservation of charge requires that B carry one more unit of positive charge than A. (We now realize that the underlying process here is the conversion of a neutron, in A, into a proton, in B; but remember that in 1930 the neutron had not yet been discovered.) Thus the 'daughter' nucleus (B) lies one position farther along on the periodic table. There are many examples of beta decay: potassium goes to calcium ($^{40}_{19}\mathrm{K} \rightarrow ^{40}_{20}\mathrm{Ca}$), copper goes to zinc ($^{64}_{29}\mathrm{Cu} \rightarrow ^{64}_{30}\mathrm{Zn}$), tritium goes to helium ($^{3}_{1}\mathrm{H} \rightarrow ^{3}_{2}\mathrm{He}$), and so on.*

Now, it is a characteristic of two-body decays ($A \rightarrow B + C$) that the outgoing energies are kinematically determined, in the center-of-mass frame. Specifically, if the 'parent' nucleus (A) is at rest, so that B and e come out back-to-back with equal and opposite momenta, then conservation of energy dictates that the electron energy is (Problem 3.19)

$$E = \left(\frac{m_A^2 - m_B^2 + m_e^2}{2m_A} \right) c^2 \tag{1.7}$$

The point to notice is that E is *fixed* once the three masses are specified. But when the experiments are done, it is found that the emitted electrons vary considerably in energy; Equation 1.7 only determines the *maximum* electron energy for a particular beta decay process (see Figure 1.5).

This was a most disturbing result. Niels Bohr (not for the first time) was ready to abandon the law of conservation of energy.† Fortunately, Pauli took a more sober view, suggesting that another particle was emitted along with the electron, a silent accomplice that carries off the 'missing' energy. It had to be electrically neutral, to conserve charge (and also, of course, to explain why it left no track); Pauli proposed to call it the *neutron*. The whole idea was greeted with some skepticism, and in 1932 Chadwick preempted the name. But in the following year Fermi presented a theory of beta decay that incorporated Pauli's particle and proved so brilliantly successful that Pauli's suggestion had to be taken seriously. From the fact that the observed electron energies range up to the value given in Equation 1.7 it follows that the new particle must be extremely light; Fermi called it the *neutrino* ('little neutral one'). For reasons you'll see in a moment, we now call it the *anti*neutrino.

* The upper number is the *atomic weight* (the number of neutrons plus protons) and the lower number is the *atomic number* (the number of protons).

† It is interesting to note that Bohr was an outspoken critic of Einstein's light quantum (prior to 1924), that he mercilessly denounced Schrödinger's equation, discouraged Dirac's work on the relativistic electron theory (telling him, incorrectly, that Klein and Gordon had already succeeded), opposed Pauli's introduction of the neutrino, ridiculed Yukawa's theory of the meson, and disparaged Feynman's approach to quantum electrodynamics. Great scientists do not always have good judgment – especially when it concerns other people's work – but Bohr must hold the all-time record.

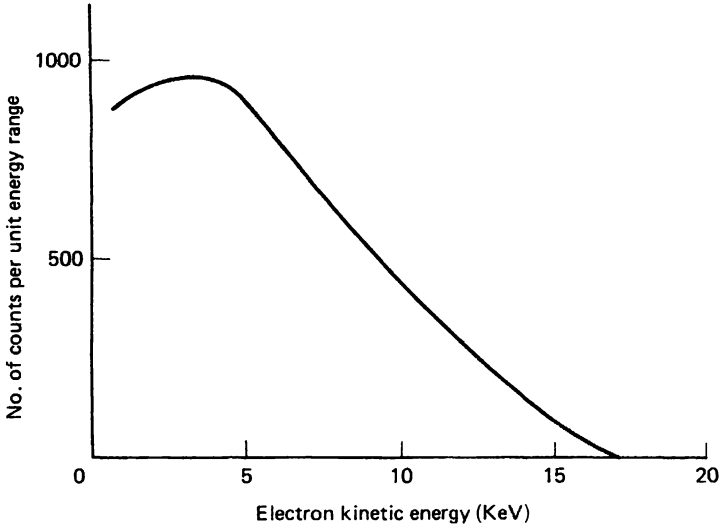


Fig. 1.5 The beta decay spectrum of tritium ($^3_1H \rightarrow ^3_2He$). (*Source*: Lewis, G. M. (1970) *Neutrinos*, Wykeham, London, p. 30.)

In modern terminology, then, the fundamental beta decay process is

$$n \rightarrow p^+ + e^- + \bar{\nu} \tag{1.8}$$

(neutron goes to proton plus electron plus antineutrino).

Now, you may have noticed something peculiar about Powell's picture of the disintegrating pion (Figure 1.3): the muon emerges at about 90° with respect to the original pion direction. (That's not the result of a *collision*, by the way; collisions with atoms in the emulsion account for the dither in the tracks, but they cannot produce an abrupt left turn.) What this kink indicates is that some *other* particle was produced in the decay of the pion, a particle that left no footprints in the emulsion, and hence must have been electrically neutral. It was natural (or at any rate *economical*) to suppose that this was again Pauli's neutrino:

$$\pi \rightarrow \mu + \nu \tag{1.9}$$

A few months after their first paper, Powell's group published an even more striking picture, in which the subsequent decay of the muon is also visible (Figure 1.6). By then muon decays had been studied for many years, and it was well established that the charged secondary is an electron. From the figure there is clearly a neutral product as well, and you might guess that it is another neutrino. However, this time it is actually *two* neutrinos:

$$\mu \rightarrow e + 2\nu \tag{1.10}$$

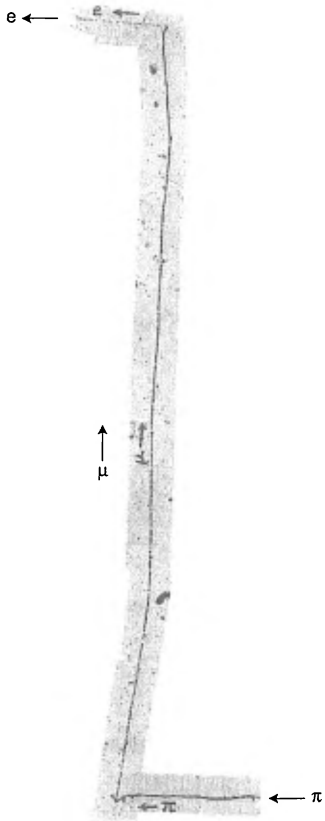


Fig. 1.6 Here, a pion decays into a muon (plus a neutrino); the muon subsequently decays into an electron (and two neutrinos). (*Source*: Powell, C. F., Fowler, P. H. and Perkins, D. H. (1959) *The Study of Elementary Particles by the Photographic Method* Pergamon, New York. First published in (1949) *Nature* **163**, 82.)

How do we know there are *two* of them? Same way as before: we repeat the experiment over and over, each time measuring the energy of the electron. If it always comes out the same, we know there are just two particles in the final state. But if it *varies*, then there must be (at least) three.* By 1949 it was clear that the

* Here, and in the original beta decay problem, conservation of angular momentum also requires a third outgoing particle, quite independently of energy conservation. But the spin assignments were not so clear in the early days, and for most people energy conservation was the compelling argument. In the interest of simplicity, I will keep angular momentum out of the story until Chapter 4.

electron energy in muon decay is *not* fixed, and the emission of two neutrinos was the accepted explanation. (By contrast, the *muon* energy in *pion* decay is perfectly constant, within experimental uncertainties, confirming that this is a genuine two-body decay.)

By 1950, then, there was compelling *theoretical* evidence for the existence of neutrinos, but there was still no direct *experimental* verification. A skeptic might have argued that the neutrino was nothing but a bookkeeping device – a purely hypothetical particle whose only function was to rescue the conservation laws. It left no tracks, and it didn't decay; in fact, no one had ever seen a neutrino *do anything*. The reason for this is that neutrinos interact extraordinarily weakly with matter; a neutrino of moderate energy could easily penetrate a thousand light years(!) of lead.* To have a chance of detecting one you need an extremely intense source. The decisive experiments were conducted at the Savannah River nuclear reactor in South Carolina, in the mid-1950s. Here Cowan and Reines set up a large tank of water and watched for the 'inverse' beta decay reaction

$$\bar{\nu} + p^+ \rightarrow n + e^+ \tag{1.11}$$

At their detector the antineutrino flux was calculated to be 5×10^{13} particles per square centimeter per second, but even at this fantastic intensity they could only hope for two or three events every hour. On the other hand, they developed an ingenious method for identifying the outgoing positron. Their results provided unambiguous confirmation of the neutrino's existence [11].

As I mentioned earlier, the particle produced in ordinary beta decay is actually an antineutrino, not a neutrino. Of course, since they're electrically neutral, you might ask – and many people did – whether there is any *difference* between a neutrino and an antineutrino. The neutral pion, as we shall see, is its *own* antiparticle; so too is the photon. On the other hand, the antineutron is definitely *not* the same as a neutron. So we're left in a bit of a quandary: *is* the neutrino the same as the antineutrino, and if not, what property distinguishes them? In the late 1950s, Davis and Harmer put this question to an experimental test [12]. From the positive results of Cowan and Reines, we know that the crossed reaction

$$\nu + n \rightarrow p^+ + e^- \tag{1.12}$$

must also occur, and at about the same rate. Davis looked for the analogous reaction using *anti*neutrinos:

$$\bar{\nu} + n \rightarrow p^+ + e^- \tag{1.13}$$

* That's a comforting realization when you learn that hundreds of billions of neutrinos per second pass through every square inch of your body, night and day, coming from the sun (they hit you from below, at night, having passed right through the earth).

He found that this reaction does *not* occur, and concluded that the neutrino and antineutrino are distinct particles.*

Davis's result was not unexpected. In fact, back in 1953 Konopinski and Mahmoud [13] had introduced a beautifully simple rule for determining which reactions – such as Equation 1.12 – will work, and which – like Equation 1.13 – will not. In effect,† they assigned a *lepton number* $L = +1$ to the electron, the muon, and the neutrino, and $L = -1$ to the positron, the positive muon, and the antineutrino (all other particles are given a lepton number of zero). They then proposed the *law of conservation of lepton number* (analogous to the law of conservation of charge): in any physical process, the sum of the lepton numbers before must equal the sum of the lepton numbers after. Thus the Cowan–Reines reaction (1.11) is allowed ($L = -1$ before and after), but the Davis reaction (1.13) is forbidden (on the left $L = -1$, on the right $L = +1$). It was in anticipation of this rule that I called the beta decay particle (Equation 1.8) an *anti*neutrino; likewise, the charged pion decays (Equation 1.9) should really be written

$$\begin{aligned} \pi^- &\to \mu^- + \bar{\nu} \\ \pi^+ &\to \mu^+ + \nu \end{aligned} \tag{1.14}$$

and the muon decays (Equation 1.10) are actually

$$\begin{aligned} \mu^- &\to e^- + \nu + \bar{\nu} \\ \mu^+ &\to e^+ + \nu + \bar{\nu} \end{aligned} \tag{1.15}$$

You might be wondering what *property* distinguishes the neutrino from the antineutrino. The cleanest answer is: *lepton number* – it's +1 for the neutrino and −1 for the antineutrino. These numbers are experimentally determinable, just as electric charge is, by watching how the particle in question interacts with others. (As we shall see, they also differ in their *helicity*: the neutrino is 'left-handed' whereas the antineutrino is 'right-handed'. But this is a technical matter best saved for later.)

There soon followed another curious twist to the neutrino story. Experimentally, the decay of a muon into an electron plus a *photon* is never observed:

$$\mu^- \not\to e^- + \gamma \tag{1.16}$$

and yet this process is consistent with conservation of charge and conservation of the lepton number. Now, a famous rule of thumb in particle physics (generally

* Actually, this conclusion is not as fireproof as it once seemed. It could be the spin state of the $\bar{\nu}$, rather than the fact that it is distinct from ν, that forbids reaction 1.13. Today, in fact, there are two viable models: *Dirac neutrinos*, which are distinct from their antiparticles, and *Majorana neutrinos*, for which ν and $\bar{\nu}$ are two states of the *same* particle. For most of this book, I shall assume we are dealing with Dirac neutrinos, but we'll return to the question in Chapter 11.

† Konopinski and Mahmoud [13] did not use this terminology, and they got the muon assignments wrong. But never mind, the essential idea was there.

attributed to Richard Feynman) declares that whatever is not expressly *forbidden* is *mandatory*. The absence of $\mu \to e + \gamma$ suggests a law of conservation of 'mu-ness', but then how are we to explain the observed decays $\mu \to e + \nu + \overline{\nu}$? The answer occurred to a number of people in the late 1950s and early 1960s [14]: suppose there are two different *kinds* of neutrino – one associated with the electron (ν_e) and one with the muon (ν_μ). If we assign a *muon number* $L_\mu = +1$ to μ^- and ν_μ, and $L_\mu = -1$ to μ^+ and $\overline{\nu}_\mu$, and at the same time an *electron number* $L_e = +1$ to e^- and ν_e, and $L_e = -1$ to e^+ and $\overline{\nu}_e$, and refine the conservation of lepton number into two separate laws – conservation of electron number and conservation of muon number – we can then account for *all* allowed and forbidden processes. Neutron beta decay becomes

$$n \to p^+ + e^- + \overline{\nu}_e \tag{1.17}$$

the pion decays are

$$\begin{aligned} \pi^- &\to \mu^- + \overline{\nu}_\mu \\ \pi^+ &\to \mu^+ + \nu_\mu \end{aligned} \tag{1.18}$$

and the muon decays take the form

$$\begin{aligned} \mu^- &\to e^- + \overline{\nu}_e + \nu_\mu \\ \mu^+ &\to e^+ + \nu_e + \overline{\nu}_\mu \end{aligned} \tag{1.19}$$

I said earlier that when pion decay was first analyzed it was 'natural' and 'economical' to assume that the outgoing neutral particle was the same as in beta decay, and that's quite true: it *was* natural and it *was* economical, but it was *wrong*.

The first experimental test of the two-neutrino hypothesis (and the separate conservation of electron and muon number) was conducted at Brookhaven in 1962 [15]. Using about 10^{14} antineutrinos from π^- decay, Lederman, Schwartz, Steinberger, and their collaborators identified 29 instances of the expected reaction

$$\overline{\nu}_\mu + p^+ \to \mu^+ + n \tag{1.20}$$

and no cases of the forbidden process

$$\overline{\nu}_\mu + p^+ \to e^+ + n \tag{1.21}$$

With only one kind of neutrino, the second reaction would be just as common as the first. (Incidentally, this experiment presented truly monumental shielding problems. Steel from a dismantled battleship was stacked up 44-feet thick, to make sure that nothing except neutrinos got through to the target.)

I mentioned earlier that neutrinos are extremely light – in fact, until fairly recently it was widely assumed (for no particularly good reason) that they are

Table 1.1 The lepton family, 1962–1976

	Lepton number	Electron number	Muon number
Leptons			
e^-	1	1	0
ν_e	1	1	0
μ^-	1	0	1
ν_μ	1	0	1
Antileptons			
e^+	−1	−1	0
$\bar{\nu}_e$	−1	−1	0
μ^+	−1	0	−1
$\bar{\nu}_\mu$	−1	0	−1

massless. This simplifies a lot of calculations, but we now know that it is not strictly true: neutrinos *have* mass, though we do not yet know what those masses are, except to reiterate that they are very small, even when compared to the electron's. What is more, over long distances neutrinos of one type can convert into neutrinos of another type (for example, electron neutrinos into muon neutrinos) – and back again, in a phenomenon known as *neutrino oscillation*. But this story belongs much later, and deserves a detailed treatment, so I'll save it for Chapter 11.

By 1962, then, the lepton family had grown to eight: the electron, the muon, their respective neutrinos, and the corresponding antiparticles (Table 1.1). The leptons are characterized by the fact that they do not participate in strong interactions. For the next 14 years things were pretty quiet, as far as the leptons go, so this is a good place to pause and catch up on the strongly interacting particles – the mesons and baryons, known collectively as the *hadrons*.

1.6 Strange Particles (1947–1960)

For a brief period in 1947, it was possible to believe that the major problems of elementary particle physics were solved. After a lengthy detour in pursuit of the muon, Yukawa's meson (the π) had finally been apprehended. Dirac's positron had been found, and Pauli's neutrino, although still at large (and, as we have seen, still capable of making mischief), was widely accepted. The role of the muon was something of a puzzle ('Who ordered *that?*' Rabi asked) – it seemed quite unnecessary in the overall scheme of things. On the whole, however, it looked in 1947 as though the job of elementary particle physics was essentially done.

But this comfortable state did not last long [16]. In December of that year, Rochester and Butler [17] published the cloud chamber photograph shown in Figure 1.8 Cosmic ray particles enter from the upper left and strike a lead plate,

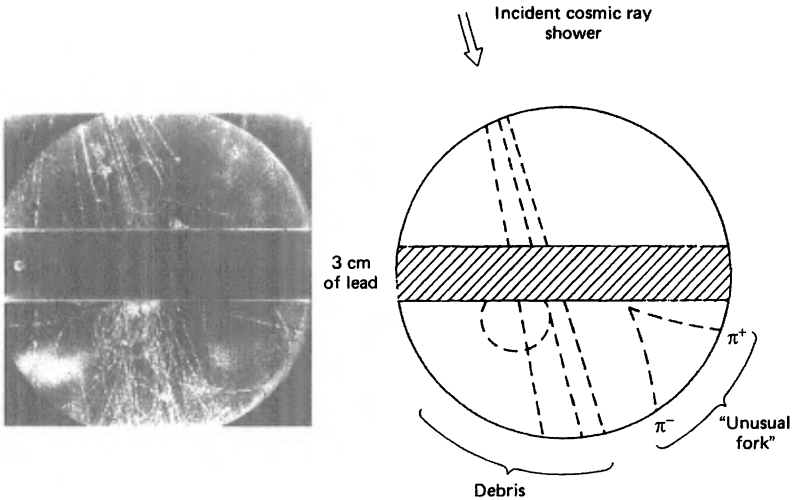


Fig. 1.7 The first strange particle. Cosmic rays strike a lead plate, producing a K^0, which subsequently decays into a pair of charged pions. (Photo courtesy of Prof. Rochester, G. D. (© 1947). *Nature*, **160**, 855. Copyright Macmillan Journals Limited.)

producing a neutral particle, whose presence is revealed when it decays into two charged secondaries, forming the upside-down 'V' in the lower right. Detailed analysis indicated that these charged particles are in fact a π^+ and a π^-. Here, then, was a new neutral particle with at least twice the mass of the pion; we call it the K^0 ('kaon'):

$$K^0 \rightarrow \pi^+ + \pi^- \tag{1.22}$$

In 1949 Brown and her collaborators published the photograph reproduced in Figure 1.8, showing the decay of a charged kaon:

$$K^+ \rightarrow \pi^+ + \pi^+ + \pi^- \tag{1.23}$$

(The K^0 was first known as the V^0 and later as the θ^0; the K^+ was originally called the τ^+. Their identification as neutral and charged versions of the same basic particle was not completely settled until 1956 – but that's another story, to which we shall return in Chapter 4.) The kaons behave in some respects like heavy pions, so the *meson* family was extended to include them. In due course, many more mesons were discovered – the η, the ϕ, the ω, the ρ's, and so on.

Meanwhile, in 1950 another neutral 'V' particle was found by Anderson's group at Cal Tech. The photographs were similar to Rochester's (Figure 1.7), but this time the products were a p^+ and a π^-. Evidently, this particle is substantially heavier

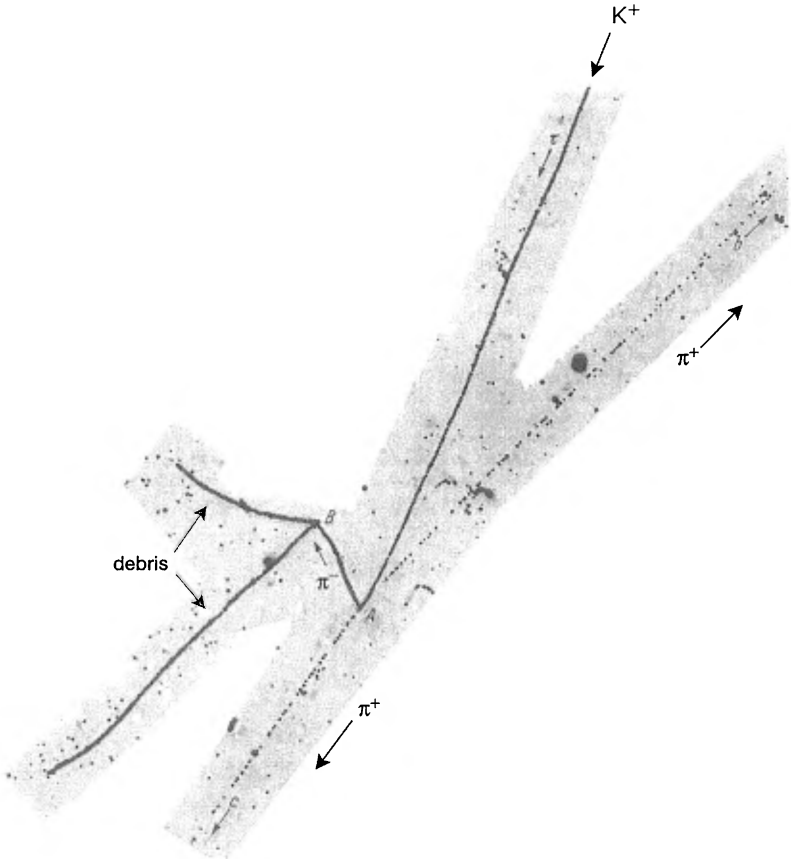


Fig. 1.8 K^+, entering from above, decays at A:$K^+ \rightarrow \pi^+ + \pi^+ + \pi^-$. (The π^- subsequently causes a nuclear disintegration at B.) (*Source*: Powell, C. F. Fowler, P. H. and Perkins, D. H. (1959) *The Study of Elementary Particles by the Photographic Method*, Pergamon, New York. First published in *Nature*, **163**, 82 (1949).)

than the proton; we call it the Λ:

$$\Lambda \rightarrow p^+ + \pi^- \tag{1.24}$$

The lambda belongs with the proton and the neutron in the *baryon* family. To appreciate this, we must go back for a moment to 1938. The question had arisen, 'Why is the proton stable?' Why, for example, doesn't it decay into a positron and a photon:

$$p^+ \rightarrow e^+ + \gamma \tag{1.25}$$

Needless to say, it would be unpleasant for us if this reaction were common (all atoms would disintegrate), and yet it does not violate any law known in 1938. (It does violate conservation of *lepton* number, but that law was not recognized, remember, until 1953.) Stückelberg [18] proposed to account for the stability of the proton by asserting a law of *conservation of baryon number*: assign to all baryons (which in 1938 meant the proton and the neutron) a 'baryon number' $A = +1$, and to the antibaryons ($\overline{p}$ and $\overline{n}$) $A = -1$; then the total baryon number is conserved in any physical process. Thus, neutron beta decay ($n \to p^+ + e^- + \overline{\nu}_e$) is allowed ($A = 1$ before and after), and so too is the reaction in which the antiproton was first observed:

$$p + p \to p + p + p + \overline{p} \tag{1.26}$$

($A = 2$ on both sides). But the proton, as the lightest baryon, has nowhere to go; conservation of baryon number guarantees its absolute stability.* If we are to retain the conservation of baryon number in the light of reaction (1.24), the lambda must be assigned to the baryon family. Over the next few years, many more heavy baryons were discovered – the Σ's, the Ξ's, the Δ's, and so on. By the way, unlike leptons and baryons, there is *no* conservation of mesons. In pion decay ($\pi^- \to \mu^- + \overline{\nu}_\mu$) a meson disappears, and in lambda decay ($\Lambda \to p^+ + \pi^-$) a meson is created.

It is some measure of the surprise with which these new heavy baryons and mesons were greeted that they came to be known collectively as 'strange' particles. In 1952, the first of the modern particle accelerators (the Brookhaven Cosmotron) began operating, and soon it was possible to produce strange particles in the laboratory (before this the only source had been cosmic rays) ... and with this the rate of proliferation increased. Willis Lamb began his Nobel Prize acceptance speech in 1955 with the following words [19]:

> *When the Nobel Prizes were first awarded in 1901, physicists knew something of just two objects which are now called "elementary particles": the electron and the proton. A deluge of other "elementary" particles appeared after 1930; neutron, neutrino, μ meson (sic), π meson, heavier mesons, and various hyperons. I have heard it said that "the finder of a new elementary particle used to be rewarded by a Nobel Prize, but such a discovery now ought to be punished by a $10,000 fine".*

Not only were the new particles unexpected; there is a more technical sense in which they seemed 'strange': they are *produced* copiously (on a time scale of about 10^{-23} seconds), but they *decay* relatively slowly (typically about 10^{-10} seconds). This suggested to Pais and others [20] that the mechanism involved in

* 'Grand unified theories' (GUTs) allow for a minute violation of baryon number conservation, and in these theories the proton is *not* absolutely stable (see Sections 2.6 and 12.2). As of 2007, no proton decay has been observed, and its lifetime is known to exceed 10^{29} years–which is pretty stable, when you consider that the age of the universe is about 10^{10} years.

their production is entirely different from that which governs their disintegration. In modern language, the strange particles are *produced* by the *strong* force (the same one that holds the nucleus together), but they *decay* by the *weak* force (the one that accounts for beta decay and all other neutrino processes). The details of Pais's scheme required that the strange particles be produced in *pairs* (so-called *associated production*). The experimental evidence for this was far from clear at that time, but in 1953 Gell-Mann [21] and Nishijima [22] found a beautifully simple and, as it developed, stunningly successful way to implement and improve Pais's idea. They assigned to each particle a new property (Gell-Mann called it 'strangeness') that (like charge, lepton number, and baryon number) is conserved in any strong interaction, but (*un*like those others) is *not* conserved in a weak interaction. In a pion–proton collision, for example, we might produce *two* strange particles:

$$\begin{aligned} \pi^- + p^+ &\to K^+ + \Sigma^- \\ &\to K^0 + \Sigma^0 \\ &\to K^0 + \Lambda \end{aligned} \tag{1.27}$$

Here, the K's carry strangeness $S = +1$, the Σ's and the Λ have $S = -1$, and the 'ordinary' particles – π, p, and n – have $S = 0$. But we never produce just *one* strange particle:

$$\begin{aligned} \pi^- + p^+ &\not\to \pi^+ + \Sigma^- \\ &\not\to \pi^0 + \Lambda \\ &\not\to K^0 + n \end{aligned} \tag{1.28}$$

On the other hand, when these particles *decay*, strangeness is *not* conserved:

$$\begin{aligned} \Lambda &\to p^+ + \pi^- \\ \Sigma^+ &\to p^+ + \pi^0 \\ &\to n + \pi^+ \end{aligned} \tag{1.29}$$

these are *weak* processes, which do not respect conservation of strangeness.

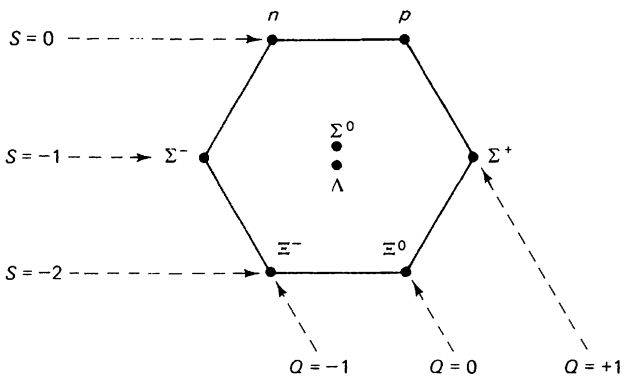
There is some arbitrariness in the assignment of strangeness numbers, obviously. We could just as well have given $S = +1$ to the Σ's and the Λ, and $S = -1$ to K^+ and K^0; in fact, in retrospect it would have been a little nicer that way. (In exactly the same sense, Benjamin Franklin's original convention for plus and minus charge was perfectly arbitrary at the time, and unfortunate in retrospect, since it made the current-carrying particle – the electron – negative.) The significant point is that there exists a consistent assignment of strangeness numbers to all the hadrons (baryons and mesons) that accounts for the observed strong processes and 'explains' why the others do not occur. (The leptons and the

photon don't experience strong forces at all, so strangeness does not apply to them.)

The garden that seemed so tidy in 1947 had grown into a jungle by 1960, and hadron physics could only be described as chaos. The plethora of strongly interacting particles was divided into two great families – the baryons and the mesons – and the members of each family were distinguished by charge, strangeness, and mass; but beyond that there was no rhyme or reason to it all. This predicament reminded many physicists of the situation in chemistry a century earlier, in the days before the periodic table, when scores of elements had been identified, but there was no underlying order or system. In 1960, the elementary particles awaited their own 'periodic table'.

1.7
The Eightfold Way (1961–1964)

The Mendeleev of elementary particle physics was Murray Gell-Mann, who introduced the so-called *Eightfold Way* in 1961 [23]. (Essentially the same scheme was proposed independently by Ne'eman.) The Eightfold Way arranged the baryons and mesons into weird geometrical patterns, according to their charge and strangeness. The eight lightest baryons fit into a hexagonal array, with two particles at the center:*

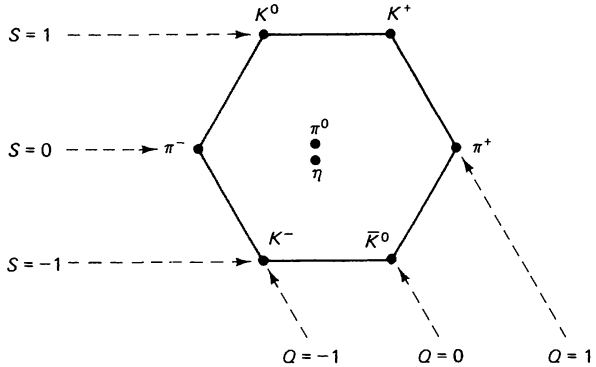


The baryon octet

This group is known as the *baryon octet*. Notice that particles of like charge lie along the downward-sloping *diagonal* lines: $Q = +1$ (in units of the proton charge) for the proton and the Σ^+; $Q = 0$ for the neutron, the Λ, the Σ^0, and the Ξ^0; $Q = -1$ for the Σ^- and the Ξ^-. *Horizontal* lines associate particles of like *strangeness*: $S = 0$ for the proton and neutron, $S = -1$ for the middle line, and $S = -2$ for the two Ξ's.

The eight lightest mesons fill a similar hexagonal pattern, forming the (*pseudoscalar*) *meson octet*:

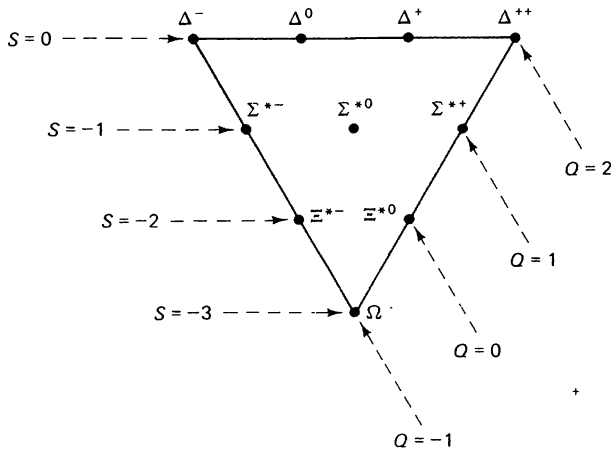
* The relative placement of the particles in the center is arbitrary, but in this book I shall always put the neutral member of the triplet (here the Σ^0) *above* the singlet (here the Λ).



The meson octet

Once again, diagonal lines determine charge and horizontal lines determine strangeness, but this time the top line has $S = 1$, the middle line $S = 0$, and the bottom line $S = -1$. (This discrepancy is again a historical accident; Gell-Mann could just as well have assigned $S = 1$ to the proton and neutron, $S = 0$ to the Σ's and the Λ, and $S = -1$ to the Ξ's. In 1953 he had no reason to prefer that choice, and it seemed most natural to give the familiar particles – proton, neutron, and pion – a strangeness of zero. After 1961, a new term – *hypercharge* – was introduced, which was equal to S for the mesons and to $S + 1$ for the baryons. But later developments revealed that strangeness was the better quantity after all, and the word 'hypercharge' has now been taken over for a quite different purpose.)

Hexagons were not the only figures allowed by the Eightfold Way; there was also, for example, a triangular array, incorporating 10 heavier baryons – the *baryon decuplet*:*



* In this book, for simplicity, I adhere to the old-fashioned notation in which the decuplet particles are designated Σ^* and Ξ^*; modern usage drops the star and puts the mass in parentheses: $\Sigma(1385)$ and $\Xi(1530)$.

Now, as Gell-Mann was fitting these particles into the decuplet, an absolutely lovely thing happened. Nine of the particles were known experimentally, but at that time the tenth particle (the one at the very bottom, with a charge of -1 and strangeness -3) was missing; no particle with these properties had ever been detected in the laboratory [24]. Gell-Mann boldly predicted that such a particle would be found, and told the experimentalists exactly how to produce it. Moreover, he calculated its mass (as you can for yourself, in Problem 1.6) and its lifetime (Problem 1.8) – and sure enough, in 1964 the famous *omega-minus* particle was discovered [25], precisely as Gell-Mann had predicted (see Figure 1.9).*

Since the discovery of the omega-minus (Ω^-), no one has seriously doubted that the Eightfold Way is correct. Over the next 10 years, every new hadron found a place in one of the Eightfold Way *supermultiplets*. Some of these are shown in Figure 1.10†. In addition to the baryon octet, decuplet, and so on, there exist of course an *anti*baryon octet, decuplet, etc, with opposite charge and opposite strangeness. However, in the case of the mesons, the antiparticles lie in the *same supermultiplet* as the corresponding particles, in the diametrically opposite positions. Thus the antiparticle of the pi-plus is the pi-minus, the anti-*K*-minus is the *K*-plus, and so on (the pi-zero and the eta are their *own* antiparticles).

Classification is the first stage in the development of any science. The Eightfold Way did more than merely classify the hadrons, but its real importance lies in the organizational structure it provided. I think it's fair to say that the Eightfold Way initiated the modern era in particle physics.

1.8 The Quark Model (1964)

But the very success of the Eightfold Way begs the question: *why* do the hadrons fit into these bizarre patterns? The periodic table had to wait many years for quantum mechanics and the Pauli exclusion principle to provide its explanation. An understanding of the Eightfold Way, however, came already in 1964, when Gell-Mann and Zweig independently proposed that all hadrons are in fact composed of even more elementary constituents, which Gell-Mann called *quarks* [26]. The

* A similar thing happened in the case of the periodic table. There were three famous 'holes' (missing elements) on Mendeleev's chart, and he predicted that new elements would be discovered to fill in the gaps. Like Gell-Mann, he confidently described their properties, and within 20 years all three – gallium, scandium, and germanium – were found.

† To be sure, there were occasional false alarms – particles that did not seem to fit Gell-Mann's scheme – but they always turned out to be experimental errors. Elementary particles have a way of appearing and then *dis*appearing. Of the 26 mesons listed on a standard table in 1963, 19 were later found to be spurious!

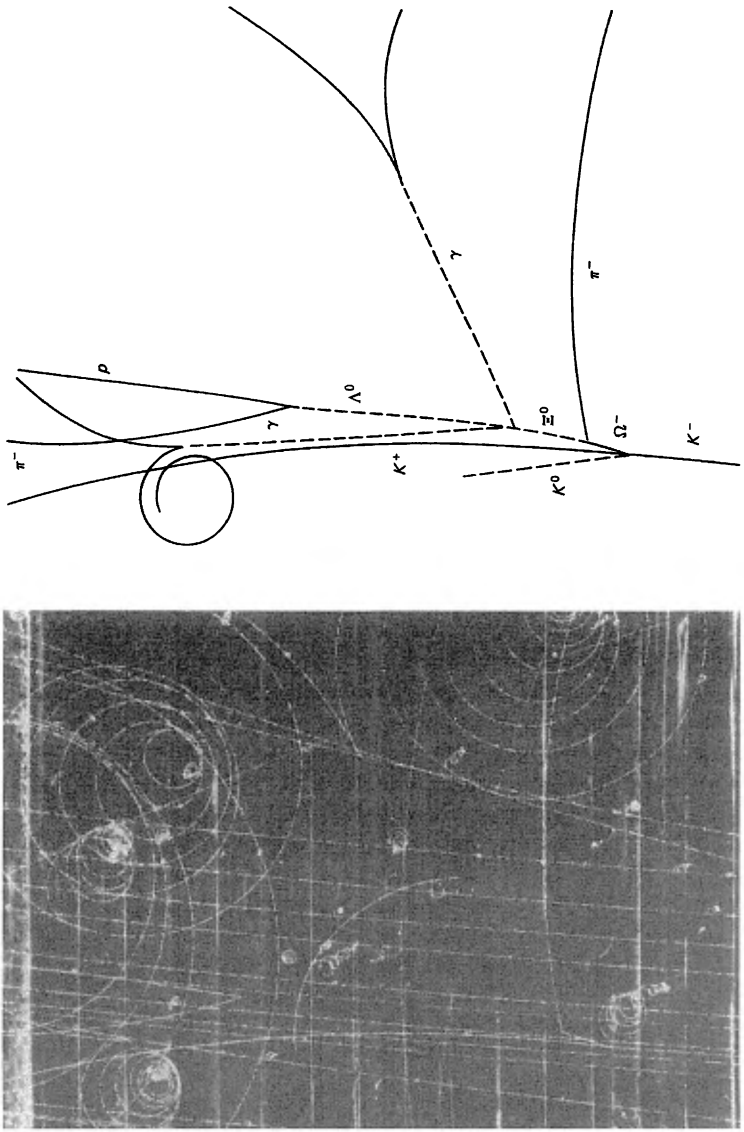


Fig. 1.9 The discovery of the Ω^-. The actual bubble chamber photograph is shown on the left; a line diagram of the relevant tracks is on the right. (Photo courtesy Brookhaven National Laboratory.)

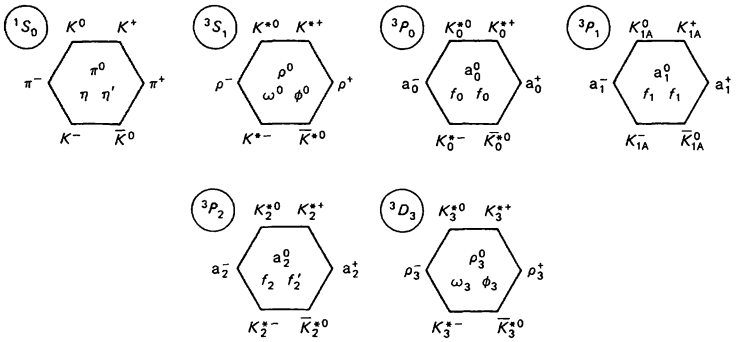
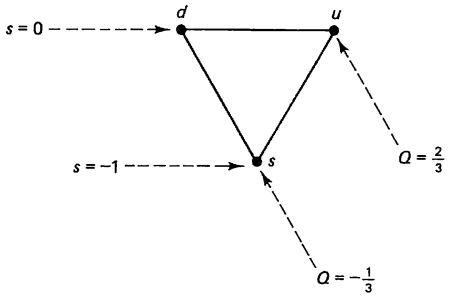


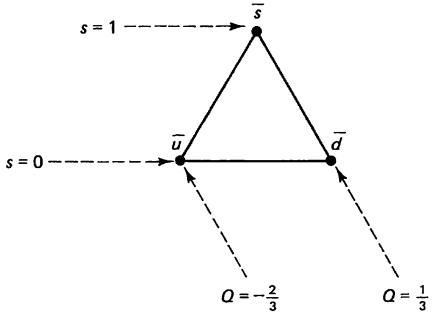
Fig. 1.10 Some meson nonets, labeled in spectroscopic notation (see Chapter 5). There are now at least 15 established nonets (though in some cases not all members have been discovered). For the baryons there are three complete octets (with spins 1/2, 3/2, and 5/2) and 10 others partly filled; there is only one complete decuplet, but 6 more are partly filled, and there are three known singlets.

quarks come in three types (or 'flavors'), forming a triangular 'Eightfold-Way' pattern:



The quarks

The u (for 'up') quark carries a charge of $\frac{2}{3}$ and a strangeness of zero; the d ('down') quark carries a charge of $-\frac{1}{3}$ and $S = 0$; the s (originally 'sideways', but now more commonly 'strange') quark carries a charge of $-\frac{1}{3}$ and $S = -1$. To each quark (q) there corresponds an *anti*quark ($\bar{q}$), with the opposite charge and strangeness:



The antiquarks

And there are two *composition rules*:

1. Every baryon is composed of three quarks (and every *anti*baryon is composed of three *anti*quarks).
2. Every meson is composed of a quark and an antiquark.

With this, it is a matter of elementary arithmetic to construct the baryon decuplet and the meson octet. All we need to do is list the combinations of three quarks (or quark–antiquark pairs) and add up their charge and strangeness:

The baryon decuplet

qqq	Q	S	Baryon
uuu	2	0	Δ^{++}
uud	1	0	Δ^{+}
udd	0	0	Δ^{0}
ddd	−1	0	Δ^{-}
uus	1	−1	Σ^{*+}
uds	0	−1	Σ^{*0}
dds	−1	−1	Σ^{*-}
uss	0	−2	Ξ^{*0}
dss	−1	−2	Ξ^{*-}
sss	−1	−3	Ω^{-}

Notice that there are 10 combinations of three quarks. Three u's, for instance, at $Q = \frac{2}{3}$ each, yield a total charge of +2 and a strangeness of zero. This is the Δ^{++} particle. Continuing down the table, we find all the members of the decuplet ending with the Ω^{-}, which is evidently made of three s quarks.

A similar enumeration of the quark–antiquark combinations yields the meson table:

The meson nonet

$q\bar{q}$	Q	S	Meson
$u\bar{u}$	0	0	π^{0}
$u\bar{d}$	1	0	π^{+}
$d\bar{u}$	−1	0	π^{-}
$d\bar{d}$	0	0	η
$u\bar{s}$	1	1	K^{+}
$d\bar{s}$	0	1	K^{0}
$s\bar{u}$	−1	−1	K^{-}
$s\bar{d}$	0	−1	$\bar{K}^{0}$
$s\bar{s}$	0	0	??

But wait! There are *nine* combinations here, and only eight particles in the meson octet. The quark model requires that there be a third meson (in addition to the π^0 and the η) with $Q = 0$ and $S = 0$. As it turns out, just such a particle had already been found experimentally – the η'. In the Eightfold Way, the η' had been classified as a *singlet*, all by itself. According to the quark model, it properly belongs with the other eight mesons to form the *meson nonet*. (Actually, since $u\bar{u}$, $d\bar{d}$, and $s\bar{s}$ all have $Q = 0$ and $S = 0$, it is not possible to say, on the basis of anything we have done so far, which is the π^0, which the η, and which the η'. But never mind, the point is that there are *three* mesons with $Q = S = 0$.) By the way, the *anti*mesons automatically fall in the same supermultiplet as the mesons: $u\bar{d}$ is the antiparticle of $d\bar{u}$, and vice versa.

You may have noticed that I avoided talking about the baryon *octet* – and it is far from obvious how we are going to get *eight* baryons by putting together three quarks. In truth, the procedure is perfectly straightforward, but it does call for some facility in handling spins, and I would rather save the details for Chapter 5. For now, I'll just tantalize you with the mysterious observation that if you take the decuplet and knock off the three corners (where the quarks are identical – *uuu*, *ddd*, and *sss*) and double the center (where all three are different – *uds*), you obtain precisely the eight states in the baryon octet. So the same set of quarks can account for the octet; it's just that some combinations do not appear at all, and one appears twice.

Indeed, *all* the Eightfold Way supermultiplets emerge naturally in quark model. Of course, the same combination of quarks can go to make a number of different particles: the delta-plus and the proton are both composed of two *u*'s and a *d*; the pi-plus and the rho-plus are both $u\bar{d}$, and so on. Just as the hydrogen atom (electron plus proton) has many different energy levels, a given collection of quarks can bind together in many different ways. But whereas the various energy levels in the electron/proton system are relatively close together (the spacings are typically several electron volts, in an atom whose rest energy is nearly 10^9 eV), so that we naturally think of them all as 'hydrogen', the energy spacings for different states of a bound quark system are very large, and we normally regard them as distinct particles. Thus we can, in principle, construct an infinite number of hadrons out of only three quarks. Notice, however, that *some* things are absolutely excluded in the quark model: for example, a baryon with $S = 1$ or $Q = -2$; no combination of the three quarks can produce these numbers (though they do occur for *anti*baryons). Nor can there be a *meson* with a charge of $+2$ (like the Δ^{++} baryon) or a strangeness of -3 (like the Ω^-). For a long time, there were major experimental searches for these so-called 'exotic' particles; their discovery would be devastating for the quark model, but none has ever been found (see Problem 1.11).

The quark model *does*, however, suffer from one profound embarrassment: in spite of the most diligent search, no one has ever seen an individual quark. Now, if a proton is really made out of three quarks, you'd think that if you hit one hard enough, the quarks ought to come popping out. Nor would they be hard to recognize, carrying as they do the unmistakable fingerprint of fractional charge – an ordinary Millikan oil drop experiment would clinch the identification. Moreover, at least one of the quarks should be absolutely stable; what could it decay

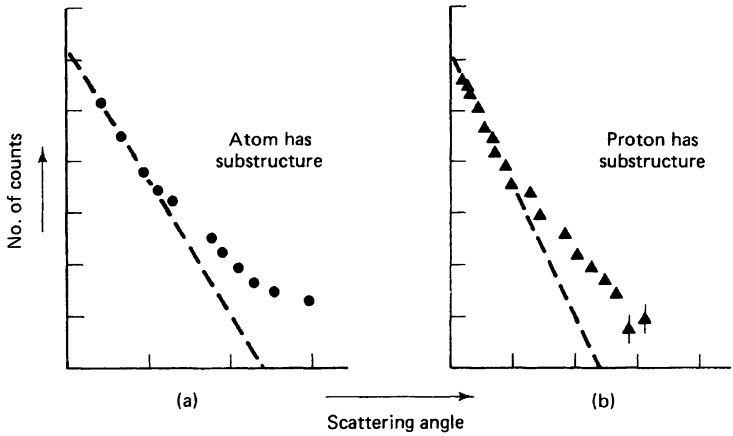


Fig. 1.11 (a) In Rutherford scattering, the number of particles deflected through large angles indicates that the atom has internal structure (a nucleus). (b) In deep inelastic scattering, the number of particles deflected through large angles indicates that the proton has internal structure (quarks). The dashed lines show what you would expect if the positive charge were uniformly distributed over the volume of (a) the atom, (b) the proton. (*Source*: Halzen, F. and Martin, A. D. (1984) *Quarks and Leptons*, John Wiley & Sons, New York, p. 17. Copyright © John Wiley & Sons, Inc. Reprinted by permission.)

into, since there is no lighter particle with fractional charge? So quarks ought to be *easy* to produce, *easy* to identify, and *easy* to store, and yet, no one has ever found one.

The failure of experiments to produce isolated quarks occasioned widespread skepticism about the quark model in the late 1960s and early 1970s. Those who clung to the model tried to conceal their disappointment by introducing the notion of *quark confinement*: perhaps, for reasons not yet understood, quarks are *absolutely confined* within baryons and mesons, so that no matter how hard you try, you cannot get them out. Of course, this doesn't explain anything, it just gives a name to our frustration. But it does pose sharply a critical theoretical question that is still not completely answered: what is the *mechanism* responsible for quark confinement? [27]

Even if all quarks are stuck inside hadrons, this does not mean they are inaccessible to experimental study. One can explore the interior of a proton in much the same way as Rutherford probed the inside of an atom – by firing things into it. Such experiments were carried out in the late 1960s using high-energy electrons at the Stanford Linear Accelerator Center (SLAC). They were repeated in the early 1970s using neutrino beams at CERN, and later still using protons. The results of these so-called 'deep inelastic scattering' experiments [28] were strikingly reminiscent of Rutherford's (Figure 1.11): *most* of the incident particles pass right through, whereas a small number bounce back sharply. This means that the charge of the proton is concentrated in small lumps, just as Rutherford's results indicated that the positive charge in an atom is concentrated at the nucleus [29]. However,

in the case of the proton the evidence suggests *three* lumps, instead of *one*. This is strong support for the quark model, obviously, but still not conclusive.

Finally, there was a theoretical objection to the quark model: it appears to violate the Pauli exclusion principle. In Pauli's original formulation, the exclusion principle states that no two electrons can occupy the same state. However, it was later realized that the same rule applies to all particles of half-integer spin (the proof of this is one of the most important achievements of quantum field theory). In particular, the exclusion principle should apply to quarks, which, as we shall see, must carry spin $\frac{1}{2}$. Now the Δ^{++}, for instance, is supposed to consist of three identical u quarks in the same state; it (and also the Δ^- and the Ω^-) appear to be inconsistent with the Pauli principle. In 1964, O. W. Greenberg proposed a way out of this dilemma [30]. He suggested that quarks not only come in three *flavors* (u, d, and s) but each of these also comes in three *colors* ('red', 'green', and 'blue', say). To make a baryon, we simply take one quark of each color; then the three u's in Δ^{++} are no longer identical (one's red, one's green, and one's blue). Since the exclusion principle only applies to *identical* particles, the problem evaporates.

The color hypothesis sounds like sleight of hand, and many people initially considered it the last gasp of the quark model. As it turned out, the introduction of color was extraordinarily fruitful [31]. I need hardly say that the term 'color' here has absolutely no connection with the ordinary meaning of the word. Redness, blueness, and greenness are simply *labels* used to denote three new properties that, in addition to charge and strangeness, the quarks possess. A *red* quark carries one unit of redness, zero blueness, and zero greenness; its antiparticle carries *minus* one unit of redness, and so on. We could just as well call these quantities X-ness, Y-ness, and Z-ness, for instance. However, the color terminology has one especially nice feature: it suggests a delightfully simple characterization of the particular quark combinations that are found in nature.

All naturally occurring particles are colorless.

By 'colorless' I mean that *either* the total amount of each color is zero *or* all three colors are present in equal amounts. (The latter case mimics the optical fact that light beams of three primary colors combine to make white.) This clever rule 'explains' (if that's the word for it) why you can't make a particle out of *two* quarks, or *four* quarks, and for that matter why *individual* quarks do not occur in nature. The only colorless combinations you can make are $q\bar{q}$ (the mesons), qqq (the baryons), and $\bar{q}\bar{q}\bar{q}$ (the antibaryons).*

* Of course, you can package together *combinations* of these – the deuteron, for example, is a *six* quark state (three u's and three d's). In 2003, there was a flurry of excitement over the apparent observation of four-quark 'mesons' (actually, $qq\bar{q}\bar{q}$) and pentaquark 'baryons' ($qqqq\bar{q}$). The latter now appear to have been statistical artifacts [32], but in at least one meson case (the so-called $X(3872)$ discovered at KEK in Japan), the four-quark interpretation seems to be holding up, though it is still not clear whether it is best thought of as a $D\bar{D}^*$ 'molecule' or as a meson in its own right [33].

1.9
The November Revolution and Its Aftermath (1974–1983 and 1995)

The decade from 1964 to 1974 was a barren time for elementary particle physics. The quark model, which had seemed so promising at the beginning, was in an uncomfortable state of limbo by the end. It had some striking successes: it neatly explained the Eightfold Way, and correctly predicted the lumpy structure of the proton. But it had two conspicuous defects: the experimental absence of free quarks and inconsistency with the Pauli principle. Those who liked the model papered over these failures with what seemed at the time to be rather transparent rationalizations: the idea of quark confinement and the color hypothesis. But I think it is safe to say that by 1974 most elementary particle physicists felt queasy, at best, about the quark model. The lumps inside the proton were called *partons*, and it was unfashionable to identify them explicitly with quarks.

Curiously enough, what rescued the quark model was not the discovery of free quarks, or an explanation of quark confinement, or confirmation of the color hypothesis, but something entirely different and (almost) [34] completely unexpected: the discovery of the psi meson. The ψ was first observed at Brookhaven by a group under C. C. Ting, in the summer of 1974. But Ting wanted to check his results before announcing them publicly, and the discovery remained an astonishingly well-kept secret until the weekend of November 10–11, when the new particle was discovered independently by Burton Richter's group at SLAC. The two teams then published simultaneously [35], Ting naming the particle J, and Richter calling it ψ. The J/ψ was an electrically neutral, extremely heavy meson – more than three times the weight of a proton (the original notion that mesons are 'middle-weight' and baryons 'heavy-weight' had long since gone by the boards). But what made this particle so unusual was its extraordinarily long lifetime, for the ψ lasted fully 10^{-20} seconds before disintegrating. Now, 10^{-20} seconds may not impress you as a particularly long time, but you must understand that the *typical* lifetimes for hadrons in this mass range are on the order of 10^{-23} seconds. So the ψ has a lifetime about a 1000 times longer than any comparable particle. It's as though someone came upon an isolated village in Peru or the Caucasus where people live to be 70 000 years old. That wouldn't just be some actuarial anomaly, it would be a sign of fundamentally new biology at work. And so it was with the ψ: its long lifetime, to those who understood, spoke of fundamentally new physics. For good reason, the events precipitated by the discovery of the ψ came to be known as the *November Revolution* [36].

In the months that followed, the true nature of the ψ meson was the subject of lively debate, but the explanation that won was provided by the quark model: the ψ is a bound state of a new (fourth) quark, the c (for *charm*) and its antiquark, $\psi = (c\bar{c})$. Actually, the idea of a fourth flavor, and even the whimsical name, had been introduced many years earlier by Bjorken and Glashow [37]. There was an

intriguing parallel between the leptons and the quarks:

$$\textit{Leptons}: e, \nu_e, \mu, \nu_\mu$$

$$\textit{Quarks}: d, u, s$$

If all mesons and baryons are made out of quarks, these two families are left as the *truly* fundamental particles. But why *four* leptons and only *three* quarks? Wouldn't it be nicer if there were four of each? Later, Glashow, Iliopoulos, and Maiani [38] offered more compelling technical reasons for wanting a fourth quark, but the simple idea of a parallel between quarks and leptons is another of those far-fetched speculations that turned out to have more substance than their authors could have imagined.

So when the ψ was discovered, the quark model was ready and waiting with an explanation. Moreover, it was an explanation pregnant with implications. For if a fourth quark exists, there should be all kinds of new baryons and mesons, carrying various amounts of charm. Some of these are shown in Figure 1.12; you can work out the possibilities for yourself (Problems 1.14 and 1.15). Notice that the ψ itself

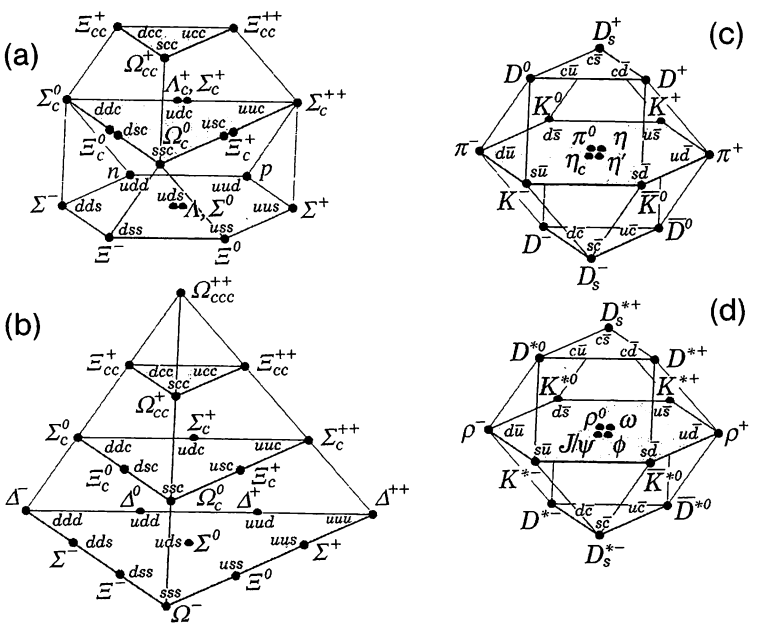
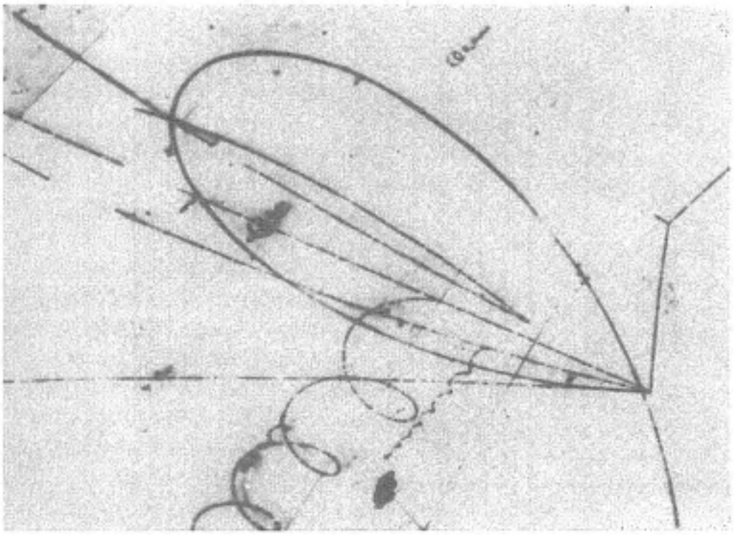


Fig. 1.12 Supermultiplets constructed using four-quark flavors: baryons (*a* and *b*) and mesons (*c* and *d*). (*Source: Review of Particle Physics.*)



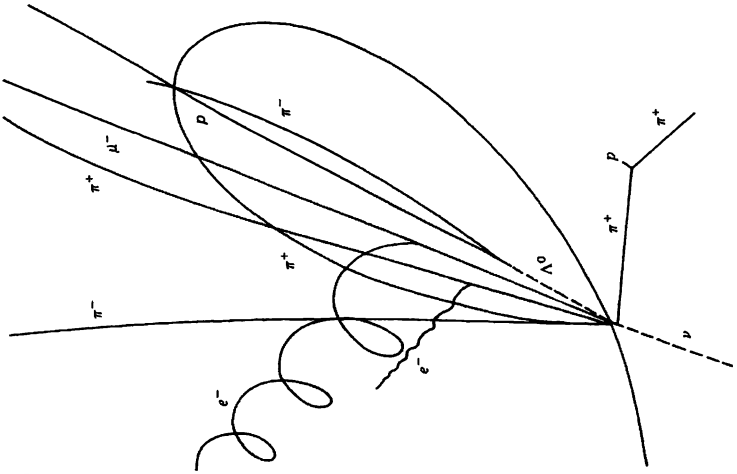


Fig. 1.13 The charmed baryon. The most probable interpretation of this event is $\nu_\mu + p \to \Lambda_c^+ + \mu^- + \pi^+ + \pi^-$. The charmed baryon decays ($\Lambda_c^+ \to \Lambda + \pi^+$) too soon to leave a track, but the subsequent decay of the Λ is clearly visible. (Photo courtesy of Samios, N. P. Brookhaven National Laboratory.)

carries no *net* charm, for if the c is assigned a charm of $+1$, then $\bar{c}$ will have a charm of -1; the charm of the ψ is, if you will, 'hidden'. To confirm the charm hypothesis, it was important to produce a particle with 'naked' (or 'bare') charm [39]. The first evidence for charmed baryons ($\Lambda_c^+ = udc$ and $\Sigma_c^{++} = uuc$) appeared already in 1975 (Figure 1.13) [40], followed later by $\Xi_c = usc$ and $\Omega_c = ssc$. (In 2002 there were hints of the first *doubly* charmed baryon at Fermilab.) The first charmed mesons ($D^0 = c\bar{u}$ and $D^+ = c\bar{d}$) were discovered in 1976 [41], followed by the charmed strange meson ($D_s^+ = c\bar{s}$) in 1977 [42]. With these discoveries, the interpretation of the ψ as $c\bar{c}$ was established beyond reasonable doubt. More important, the quark model itself was put back on its feet.

However, the story does not end there, for in 1975 a new *lepton* was discovered [43], spoiling Glashow's symmetry. This new particle (the tau) has its own neutrino, so we are up to six leptons, and only four quarks. But don't despair, because 2 years later a new heavy meson (the *upsilon*) was discovered [44], and quickly recognized as the carrier of a fifth quark, b (for *beauty*, or *bottom*, depending on your taste): $\Upsilon = b\bar{b}$. Immediately the search began for hadrons exhibiting 'naked beauty', or 'bare bottom.' (I'm sorry. *I* didn't invent this terminology. In a way, its silliness is a reminder of how wary people were of taking the quark model seriously, in the early days.) The first bottom baryon, $\Lambda_b^0 = udb$, was observed in the 1980's, and the second ($\Sigma_b^+ = uub$) in 2006; in 2007 the first baryon with a quark from all three generations was discovered ($\Xi_b^- = dsb$). The first bottom mesons ($\bar{B}^0 = b\bar{d}$ and $B^- = b\bar{u}$) were found in 1983 [45]. The $B^0/\bar{B}^0$ system has proven to be especially rich, and so-called 'B factories' are now operating at SLAC ('BaBar') and KEK ('Belle'). The *Particle Physics Booklet* also lists $B_s^0 = s\bar{b}$ and $B_c^+ = c\bar{b}$.

At this point, it didn't take a genius to predict that a sixth quark (t, for *truth*, of course, or *top*) would soon be found, restoring Glashow's symmetry with six quarks and six leptons. But the top quark turned out to be extraordinarily heavy and frustratingly elusive (at 174 GeV/c^2, it is over 40 times the weight of the bottom quark). Early searches for 'toponium' (a $t\bar{t}$ meson analogous to the ψ and Υ) were unsuccessful, both because the electron–positron colliders did not reach high enough energy and because, as we now realize, the top quark is simply too short-lived to form bound states – apparently there *are* no top baryons and mesons. The top quark's existence was not definitively established until 1995, when the Tevatron finally accumulated enough data to sustain strong indications from the previous year [46]. (The basic reaction is $u + \bar{u}$ (or $d + \bar{d}$) $\rightarrow t + \bar{t}$; the top and anti-top immediately decay, and it is by analyzing the decay products that one is able to infer their fleeting appearance.) Until the LHC begins operation, Fermilab will be the only accelerator in the world capable of producing top quarks.

1.10
Intermediate Vector Bosons (1983)

In his original theory of beta decay (1933), Fermi treated the process as a contact interaction, occurring at a single point, and therefore requiring no mediating

particle. As it happens, the weak force (which is responsible for beta decay) *is* of extremely short range, so that Fermi's model was not far from the truth, and yields excellent approximate results at low energies. However, it was widely recognized that this approach was bound to fail at high energies and would eventually have to be supplanted with a theory in which the interaction is mediated by the exchange of some particle. The mediator came to be known by the prosaic name *intermediate vector boson.* The challenge for theorists was to predict the properties of the intermediate vector boson, and for experimentalists, to produce one in the laboratory. You may recall that Yukawa, faced with the analogous problem for the strong force, was able to estimate the mass of the pion in terms of the range of the force, which he took to be roughly the same as the size of a nucleus. But we have no corresponding way to measure the range of the weak force; there are no 'weak bound states' whose size would inform us – the weak force is simply too feeble to bind particles together. For many years, predictions of the intermediate vector boson mass were little more than educated guesses (the 'education' coming largely from the failure of experiments at progressively higher energies to detect the particle). By 1962, it was known that the mass had to be at least half the proton mass; 10 years later the experimental lower limit had grown to 2.5 proton masses.

But it was not until the emergence of the electroweak theory of Glashow, Weinberg, and Salam that a really firm prediction of the mass became possible. In this theory, there are in fact *three* intermediate vector bosons, two of them charged ($W^{\pm}$) and one neutral (Z). Their masses were calculated to be [47]

$$M_W = 82 \pm 2\,\mathrm{GeV}/c^2, \quad M_Z = 92 \pm 2\,\mathrm{GeV}/c^2 \quad \text{(predicted)} \tag{1.30}$$

In the late 1970s, CERN began construction of a proton–antiproton collider designed specifically to produce these extremely heavy particles (bear in mind that the mass of the proton is 0.94 GeV/c^2, so we're talking about something nearly 100 times as heavy). In January 1983, the discovery of the W was reported by Carlo Rubbia's group [48], and 5 months later the same team announced discovery of the Z [49]. Their measured masses are

$$M_W = 80.403 \pm 0.029\,\mathrm{GeV}/c^2, \quad M_Z = 91.188 \pm 0.002\,\mathrm{GeV}/c^2 \quad \text{(measured)} \tag{1.31}$$

These experiments represent an extraordinary technical triumph [50], and they were of fundamental importance in confirming a crucial aspect of the Standard Model, to which the physics community was by that time heavily committed (and for which a Nobel Prize had already been awarded). Unlike the strange particles or the ψ, however, (but like the top quark a decade later) the intermediate vector bosons were long awaited and universally expected, so the general reaction was a sigh of relief, not shock or surprise.

1.11
The Standard Model (1978–?)

In the current view, then, all matter is made out of three kinds of elementary particles: leptons, quarks, and mediators. There are six leptons, classified according to their charge (Q), electron number (L_e), muon number (L_μ), and tau number (L_τ). They fall naturally into three *generations*:

Lepton classification

	l	Q	L_e	L_μ	L_r
First generation	e	-1	1	0	0
	ν_e	0	1	0	0
Second generation	μ	-1	0	1	0
	ν_μ	0	0	1	0
Third generation	τ	-1	0	0	1
	ν_τ	0	0	0	1

There are also six antileptons, with all the signs reversed. The positron, for example, carries a charge of $+1$ and an electron number -1. So there are really 12 leptons, all told.

Similarly, there are six 'flavors' of quarks, classified by charge, strangeness (S), charm (C), beauty (B), and truth (T). (For consistency, I suppose we should include 'upness', U, and 'downness', D, although these terms are seldom used. They are redundant, inasmuch as the only quark with $S = C = B = T = 0$ and $Q = \frac{2}{3}$, for instance, is the up quark, so it is not necessary to specify $U = 1$ and $D = 0$ as well.) The quarks, too, fall into three generations:

Quark classification

	q	Q	D	U	S	C	B	T
First generation	d	$-1/3$	-1	0	0	0	0	0
	u	$2/3$	0	1	0	0	0	0
Second generation	s	$-1/3$	0	0	-1	0	0	0
	c	$2/3$	0	0	0	1	0	0
Third generation	b	$-1/3$	0	0	0	0	-1	0
	t	$2/3$	0	0	0	0	0	1

Again, all signs would be reversed on the table of antiquarks. Meanwhile, each quark and antiquark comes in three colors, so there are 36 of them in all.

Finally, every interaction has its mediator – the photon for the electromagnetic force, two W's and a Z for the weak force, the graviton (presumably) for gravity ... but what about the strong force? In Yukawa's original theory the mediator of strong forces was the pion, but with the discovery of heavy mesons this simple picture could not stand; protons and neutrons could now exchange ρ's and η's and K's and ϕ's and all the rest of them. The quark model brought an even more radical revision: for if protons, neutrons, and mesons are complicated composite

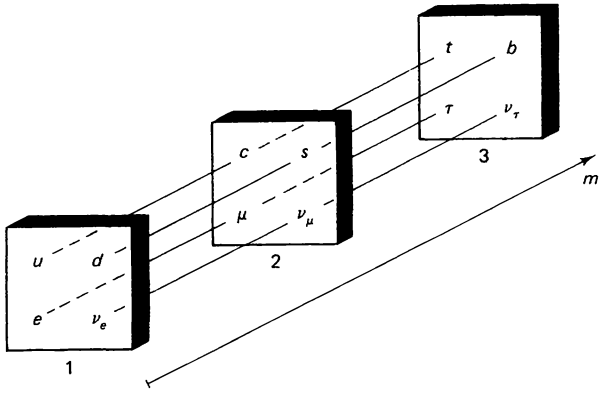


Fig. 1.14 The three generations of quarks and leptons, in order of increasing mass.

structures, there is no reason to believe their interaction *should* be simple. To study the strong force at the fundamental level, one should look, rather, at the interaction between individual quarks. So the question becomes: what particle is exchanged between two quarks, in a strong process? This mediator is called the *gluon*, and in the Standard Model there are eight of them. As we shall see, the gluons themselves carry color, and therefore (like the quarks) should not exist as isolated particles. We can hope to detect gluons only within hadrons or in colorless combinations with other gluons (*glueballs*). Nevertheless, there is substantial indirect experimental evidence for the existence of gluons: the deep inelastic scattering experiments showed that roughly half the momentum of a proton is carried by electrically neutral constituents, presumably gluons; the *jet* structure characteristic of inelastic scattering at high energies can be explained in terms of the disintegration of quarks and gluons in flight [51] and glueballs may conceivably have been observed [52].

This is all adding up to an embarrassingly large number of supposedly 'elementary' particles: 12 leptons, 36 quarks, 12 mediators (I won't count the graviton, since gravity is not included in the Standard Model). And, as we shall see later, the Glashow–Weinberg–Salam theory calls for at least one *Higgs* particle, so we have a minimum of 61 particles to contend with. Informed by our experience first with atoms and later with hadrons, many people have suggested that some, at least, of these 61 must be composites of more elementary subparticles (see Problem 1.18) [53]. Such speculations lie beyond the Standard Model and outside the scope of this book. Personally, I do not think the large number of 'elementary' particles in the Standard Model is by itself alarming, for they are tightly interrelated. The eight gluons, for example, are identical except for their colors, and the second and third generations mimic the first (Figure 1.14).

Still, it does seem odd that there should be *three* generations of quarks and leptons – after all, *ordinary* matter is made of up and down quarks (in the form of protons and neutrons) and electrons, all drawn from the first generation. Why are there two 'extra' generations; who needs 'em? It's a peculiar question, presuming a kind of purpose and efficiency on the part of the creator for which there is

little evidence ... but one can't help wondering. Actually, there is a surprising answer: as we shall see, the predominance of matter over antimatter admits a plausible accounting within the Standard Model, but only if there are (at least) three generations.

Of course, this begs the reverse question: why are there *only* three generations? Indeed, could there be more of them, which have not yet been discovered (presumably because they are too heavy to be made with existing machines)? As recently as 1988 [54], there were good reasons to anticipate a fourth generation, and perhaps even a fifth. But within a year that possibility was foreclosed by experiments at SLAC and CERN [55]. The Z^0 is (as Saddam would say) the 'mother of all particles', in the sense that it can decay (with a precisely calculable probability) into *any* quark/antiquark or lepton/antilepton pair ($e^- + e^+$, $u + \overline{u}$, $\nu_\mu + \overline{\nu}_\mu$, etc.), provided only that the particle's mass is less than half that of the Z^0 (else there wouldn't be enough energy to make the pair). So by measuring the lifetime of the Z^0 you can actually *count* the number of quarks and leptons with mass less than 45 GeV/c^2. The more there are, the shorter the lifetime of the Z^0, just as the more fatal diseases we are susceptible to the shorter our average lifespan becomes. The experiments show that the lifetime of the Z^0 is exactly what you would expect on the basis of the established three generations. Of course, the quarks (and conceivably even the charged lepton) in a putative fourth generation might be too heavy to affect the Z^0 lifetime, but it is hardly to be imagined that the fourth neutrino would suddenly jump to over 45 GeV/c^2. At any rate, what the experiments do unequivocally show is that the number of *light* neutrinos is 2.99 ± 0.06.

Although the Standard Model has survived unscathed for 30 years, it is certainly not the end of the story. There are many important issues that it simply does not address – it does not, for example, tell us how to calculate the quark and lepton masses.*

Quark and lepton masses (in MeV/c^2)

lepton	mass	quark	mass
ν_e	$<2 \times 10^{-6}$	u	2
ν_u	<0.2	d	5
ν_τ	<18	s	100
e	0.511	c	1200
μ	106	b	4200
τ	1777	t	174 000

In the Standard Model, these are simply empirical numbers, taken from experiment, but a mature theory, presumably, would *explain* them, just as we can for the atoms on the periodic table.† As we shall see, the Standard Model also takes as empirical input three angles and a phase in the Kobayashi–Maskawa matrix, and analogous numbers for the leptons, and the Weinberg angle describing electroweak

* There is substantial uncertainty in the light quark masses; I have rounded them off for the sake of clarity.

† Note, however, that the quark/lepton mass formula is going to look very strange, since it has to cover a range of at least 11 powers of 10, from the electron neutrino to the top quark.

mixing, and ... all told, there are over 20 arbitrary parameters in the Standard Model, and this is simply unacceptable in any 'final' theory [56].

On the experimental side, there remains much to be learned about neutrino oscillations (see Chapter 11), and CP violation (Chapter 12), but the most conspicuous missing link is the Higgs particle, which is necessary in the Standard Model to account for the masses of the W and Z (and perhaps all other particles as well). Like the top quark, the predicted mass of the Higgs has increased with time, as each new experiment failed to discover it. At this point, it is presumably beyond the range of any existing accelerator, and, since the cancellation of the SSC, the LHC is our best hope for finding this elusive particle.

Meanwhile, there are a number of theoretical speculations (supported as yet by no direct experimental evidence) that go beyond the Standard Model. There are the Grand Unified Theories (GUTs) that link the strong, electromagnetic, and weak interactions (Chapter 2); these are so widely accepted, at least in some form, as to be practically orthodox. Also very attractive to theorists is the idea of 'supersymmetry' (SUSY), which (among other things) would double the number of particles, associating with every fermion a boson, and vice versa. Thus the leptons would be joined by 'sleptons' ('selectrons', 'sneutrinos', etc.) and quarks by 'squarks'; the mediators would acquire twins (the 'photino', 'gluino', 'wino', and 'zino'). If subquarks or supersymmetric particles are discovered, this will be huge news, resetting the whole agenda for the next era in elementary particle physics. But except for several tantalizing false alarms [57], no evidence for either has yet appeared.

And then there is superstring theory, which since 1984 has captured the imagination of an entire generation of particle theorists. Superstrings promise not only to reconcile quantum mechanics and general relativity, and to eliminate the infinities that plague quantum field theory, but also to provide a unified 'theory of everything', from which all of elementary particle physics (including gravity) would emerge as an inescapable consequence. String theory has certainly enjoyed a brilliant and adventurous youth; it remains to be seen whether it can deliver on its extravagant ambition [58].

References

1 There are many good discussions of the history of elementary particle physics. My own favorite is the delightful little book by Yang, C. N. (1961) *Elementary Particles*, Princeton University Press, Princeton, N.J. More recent accounts are (a) Trefil's, J. S. (1980) *From Atoms to Quarks*, Scribners, New York; and (b) Close's, F. E. (1983) *The Cosmic Onion*, Heinemann Educational Books, London. The early days are treated well in (c) Keller's, A. (1983) *The Infancy of Atomic Physics*, Oxford University Press, Oxford; and (d) Weinberg's, S. (1983) *Subatomic Particles*, Scientific American Library, New York.For a fascinating and masterful account, see (e) Pais, A.(1986) *Inward Bound*, Clarendon Press, Oxford. For a comprehensive bibliography see (f) Hovis, R. C. and Kragh, H. (1991) *American Journal of Physics*, 59, 779.

2 The story is beautifully told by Pais, A. in his biography of Einstein, (1982) *Subtle is the Lord*, Clarendon Press, Oxford.

3 Millikan, R. A. (1916) *Physical Review*, **7**, 18. Quoted in (a) Pais, A. in his biography of Einstein, (1982) *Subtle is the Lord*, Clarendon Press, Oxford.

4 Heisenberg had suggested earlier that the deuteron is held together by exchange of *electrons*, in analogy with the hydrogen molecule ion (H_2^+). Yukawa was apparently the first to understand that a *new* force was involved, distinct both from electromagnetism and from the weak force responsible for beta decay. See Pais, A. (1986) *Inward Bound*, Clarendon Press, Oxford; and (a) Brown, L. M. and Rechenberg, H. (1988) *American Journal of Physics*, **56**, 982.

5 Conversi, M., Pancini, E. and Piccioni, O. (1947) *Physical Review*, **71**, 209.

6 Lattes, C. M. G. *et al.* (1947) *Nature*, **159**, 694; (1947) *Nature*, **160**, 453, 486.

7 Marshak, R. E. and Bethe, H. A. (1947) *Physical Review*, **72**, 506. Actually, Japanese physicists had come to a similar conclusion during the war; see (a) Pais, A. (1986) *Inward Bound*, Clarendon Press, Oxford, p. 453.

8 For an informal history of this discovery, see Anderson, C. D. (1961) *American Journal of Physics*, **29**, 825.

9 Chamberlain, O. *et al.* (1955) *Physical Review*, **100**, 947; (a) Cork, B. *et al.* (1956) *Physical Review*, **104**, 1193. See the articles by (b) Segrè, E. and Wiegand, C. E. (June 1956) *Scientific American*, 37 and by (c) Burbridge, G. and Hoyle, F. (April 1958), 34. Also (d) Kragh, H. (1989) *American Journal of Physics*, **57**, 1034.

10 The history of the neutrino is a fascinating story in its own right. See, for instance, Bernstein, J. (2004) *The Elusive Neutrino*, University Press of the Pacific, Honolulu; (a) Brown, L. M. (September 1978), *Physics Today*, 23; or (b) Morrison, P. (January 1956) *Scientific American*, 58. An extensive and useful bibliography on the neutrino is provided by (c) Lederman, L. M. (1970) *American Journal of Physics*, **38**, 129.

11 Reines, F. and Cowan, C. L. Jr., (1953) *Physical Review*, **92**, 8301; (a) Cowan, C. L. *et al.* (1956) *Science*, **124**, 103. Reines finally won the Nobel Prize for this work in 1995 (Cowan was dead by then).

12 Davis, R. and Harmer, D. S. (1959) *Bulletin of the American Physiological Society*, **4**, 217. See also (a) Cowan, C. L. Jr., and Reines, F. (1957) *Physical Review*, **106**, 825.

13 Konopinski, E. J. and Mahmoud, H. M. (1953) *Physical Review*, **92**, 1045.

14 Pontecorvo, B. (1960) *Soviet Physics JETP*, **37**, 1236; [translated from (1959) *Soviet Physics JETP*, **37**, 1751]; (a) Lee, T. D. (1960) Rochester Conference, New York, p. 567.

15 Danby, G. *et al.* (1962) *Physical Review Letters*, **9**, 36. See also (a) Lederman, L. (March 1963) *Scientific American*, 60.

16 Brown, L. M., Dresden, M. and Hoddeson, L. (November 1988), *Physics Today*, 56.

17 Rochester, G. D. and Butler, C. C. (1947) *Nature*, **160**, 855. See also G. D. Rochester's memoir in (a) Sekido, Y. and Elliot, H. (eds) (1985) *Early History of Cosmic Ray Studies*, Reidel, Dordrecht, p. 299.

18 Stückelberg himself did not use the term *baryon*, which was introduced by Pais, A. (1953) *Progress in Theoretical Physics*, **10**, 457.

19 *Les Prix Nobel 1955* The Nobel Foundation, Stockholm.

20 Pais, A. (1952) *Physical Review*, **86**, 663. The same (copious production, slow decay) could be said for the pion (and for that matter the neutron). But their decays produce neutrinos, and people were used to the idea that neutrino interactions are weak. What was new was a purely hadronic decay whose rate was characteristic of neutrino processes. For more on the history see (a) Pais, A. (1986) *Inward Bound*, Clarendon

Press, Oxford, p. 517; or (b) Brown, L. M., Dresden, M. and Hoddeson, L. (November **1988**), *Physics Today*, 60.

21 Gell-Mann, M. (**1953**) *Physical Review*, **92**, 883; (a) (**1956**) *Nuovo Cimento*, 4 (Suppl. 2), 848.

22 Nakano, T. and Nishijima, K. (**1953**) *Progress in Theoretical Physics*, **10**, 581.

23 The original papers are collected in Gell-Mann, M. and Ne'eman, Y. (**1964**) *The Eightfold Way*, Benjamin, New York.

24 Actually, there is a possibility that it was seen in a cosmic ray experiment in 1954 Eisenberg, Y. (**1954**) *Physical Review*, **96**, 541, but the identification was ambiguous.

25 Barnes, V. E. *et al.* (**1964**) *Physical Review Letters*, **12**, 204; (a) Gell-Mann, M. (**1953**) *Physical Review*, **92**, 883. See also (b) Fowler, W. B. and Samios, N. P. (October **1964**), *Scientific American*, 36.

26 An extensive bibliography on the quark model, and useful commentary, is given by Greenberg, O. W. (**1982**) *American Journal of Physics*, **50**, 1074. Many of the classic papers (including the original unpublished one by G. Zweig) are reprinted in (a) Lichtenberg, D. B. and Rosen, S. P. (eds) (**1980**) *Developments in the Quark Theory of Hadrons*, Hadronic Press, Nonantum.

27 Nambu, Y. (November **1976**) *Scientific American*, 48; (a) Johnson, K. (July **1979**) *Scientific American*, 112; (b) Rebbi, C. (February **1983**) *Scientific American*, 54.

28 Kendall, H. W. and Panofsky, W. K. H. (June **1971**) *Scientific American*, 61.

29 Jacob, M. and Landshoff, P. (March **1980**) *Scientific American*, 66.

30 Greenberg, O. W. (**1964**) *Physical Review Letters*, **13**, 598. Greenberg did not use this language; the terminology was introduced by (a) Lichtenberg, D. B. (**1970**) *Unitary Symmetry and Elementary Particles*, Academic Press, New York. See also (b) Glashow, S. L. (October **1975**) *Scientific American*, 38.

31 See Pais, A. (**1986**) *Inward Bound*, Clarendon Press, Oxford, pp. 562 and 602.

32 Trilling, G. (**2006**) *Review of Particle Physics*, p. 1019.

33 For status reports see Bauer, G. (**2006**) *International Journal of Modern Physics*, **A21**, 959; and (a) Swanson, E. S. (**2006**) Physics Reports, **429**, 243.

34 An exception was Iliopoulos, J. (**1974**) *At an International Conference of Particle Physicists*, London, in the summer of 1974, he remarked, 'I am ready to bet now a whole case [of wine] that the entire next Conference will be dominated by the discovery of charmed particles'.

35 Aubert, J. J. *et al.* (**1974**) *Physical Review Letters*, **33**, 1404; and (a) Augustin, J.-E. *et al.* (**1974**) *Physical Review Letters*, **33**, 1406.

36 A useful bibliography on this material, and reprints of the major articles, is given in Rosner, J. (ed) (**1981**) *New Physics*, published by the American Association of Physics Teachers, New York. The excitement of the November Revolution is captured in the SLAC publication; (a) (**1976**) *Beam Line*, **7** (11). See also the articles by (b) Drell, S. D. (June **1975**) *Scientific American*, 50, and (c) Glashow, S. L. and Trilling, G. (**2006**) RPP 2006, p. 1019.

37 Bjorken, B. J. and Glashow, S. L. (**1964**) *Physics Letters*, **11**, 255. In 1963 and 1964 there were many speculations about a possible fourth quark, for a variety of different reasons (see Pais, A. (**1986**) *Inward Bound*, Clarendon Press, Oxford, p. 601.

38 Glashow, S. L., Iliopoulos, J. and Maiani, L. (**1970**) *Physical Review*, **D2**, 1285.

39 Schwitters, R. F. (October **1977**) *Scientific American*, 56.

40 Cazzoli, E. G. *et al.* (**1975**) *Physical Review Letters*, **34**, 1125.

41 Goldhaber, G. *et al.* (1976) *Physical Review Letters*, **37**, 255; and (a) Peruzzi, I. (1976) *Physical Review Letters*, **37**, 569.
42 Brandelik, R. *et al.* *Physics Letters*, **B70**, 132.
43 Perl, M. *et al.* (1975) *Physical Review Letters*, **35**, 1489. See also (a) Perl, M. and Kirk, W. (March 1978) *Scientific American*, 50; and (b) Perl, M. (October 1997) *Physics Today*, 34. Perl was awarded the Nobel Prize in 1995 for the discovery of the τ.
44 Herb, S. W. *et al.* (1977) *Physical Review Letters*, **39**, 252. See also (a) Lederman, L. M. (October 1978) *Scientific American*, 72. It is an indication of how eager people were to find the fifth quark that the discoverers of the upsilon jumped the gun (b) (Hom, D. C. *et al.* (1976) *Physical Review Letters*, **36**, 1236), announcing a spurious particle now known fondly as the 'oops-Leon' (after Leon Lederman, head of the group).
45 Behrends, S. *et al.* (1983) *Physical Review Letters*, **50**, 881.
46 Abe, F. *et al.* (1995) *Physical Review Letters*, **74**, 2626; and (a) Abachi, S. *et al.* (1995) *Physical Review Letters*, **74**, 2632. See also (b) Liss, T. M. and Tipton, P. L. (September 1997), *Scientific American*, 54.
47 The formula for the W and Z masses was first obtained by Weinberg, S. (1967) *Physical Review Letters*, **19**, 1264. It involves a parameter θ_W whose value was unknown at that time, and all Weinberg could say for sure was that $M_W \geq 37$ GeV/c^2 and $M_Z \geq 75$ GeV/c^2. In the next 15 years θ_W was measured in a variety of experiments, and by 1982 the predictions had been refined, as indicated in Equation ((1.30)).
48 Arnison, G. *et al.* (1983) *Physics Letters*, **122B**, 103.
49 Arnison, G. *et al.* (1983) *Physics Letters*, **126B**, 398.
50 Cline, D. B. and Rubbia, C. (August 1980) *Physics Today*, 44; and (a) Cline, D. B., Rubbia, C. and van der Meer, S. (March 1982) *Scientific American*, 48. In 1984 Rubbia and van der Meer were awarded the Nobel Prize for this work. See (b) Sutton, C. (1984) *The Particle Connection*, Simon & Schuster, New York; and (c) Watkins, P. (1986) *Story of the W and Z*, Cambridge University Press, Cambridge.
51 Jacob, M. and Landshoff, P. (March 1980) *Scientific American*, 66.
52 Ishikawa, K. (November 1982), *Scientific American*, 142; (a) Sexton, J., Vaccarino, A. and Weingarten, D. (1995) *Physical Review Letters*, **75**, 4563.
53 See, for instance, the review article by Terezawa, H. (1984) *XXII International Conference on High-Energy Physics*, vol. I, Leipzig, p. 63; and (a) Harari, H. (April 1983) *Scientific American*, 56.
54 Cline, D. B. (August 1988) *Scientific American*, 60.
55 Abrams, G. S. *et al.* (1989) *Physical Review Letters*, **63**, 2173. See also (a) Feldman, G. J. and Steinberger, J. (February 1991), *Scientific American*, 70.
56 For a delightful discussion, see Cahn, R. N. (1996) *Reviews of Modern Physics*, **68**, 951.
57 On the search for subquarks, see Abe, F. *et al.* (1996) *Physical Review Letters*, **77**, 5336. For the evaporating evidence of supersymmetry in measurements of the anomalous magnetic moment of the muon, see (a) Schwarzschild, B. (February 2002) *Physics Today*, 18.
58 For a brilliant semi-popular account, see Greene, B. (1999) *The Elegant Universe*, W. W. Norton, New York. For an accessible introduction to the theory, see (a) Zwiebach, B. (2004) *A First Course in String Theory*, Cambridge, UK, Cambridge University Press.) For harsh critiques see (b) Smolin, L. (2006) *The Trouble with Physics*, Houghton Mifflin, Boston; and (c) Woit, P. (2006) *Not Even Wrong*, Perseus, New York.

Problems

1.1 If a charged particle is undeflected in passing through uniform crossed electric and magnetic fields **E** and **B** (mutually perpendicular and both perpendicular to the direction of motion), what is its velocity? If we now turn off the electric field, and the particle moves in an arc of radius R, what is its charge-to-mass ratio?

1.2 The mass of Yukawa's meson can be estimated as follows. When two protons in a nucleus exchange a meson (mass m), they must temporarily violate the conservation of energy by an amount mc^2 (the rest energy of the meson). The Heisenberg uncertainty principle says that you may 'borrow' an energy ΔE, provided you 'pay it back' in a time Δt given by $\Delta E\ \Delta t = \hbar/2$ (where $\hbar \equiv h/2\pi$). In this case, we need to borrow $\Delta E = mc^2$ long enough for the meson to make it from one proton to the other. It has to cross the nucleus (size r_0), and it travels, presumably, at some substantial fraction of the speed of light, so, roughly speaking, $\Delta t = r_0/c$. Putting all this together, we have

$$m = \frac{\hbar}{2r_0 c}$$

Using $r_0 = 10^{-13}$ cm (the size of a typical nucleus), calculate the mass of Yukawa's meson. Express your answer in MeV/c^2, and compare the observed mass of the pion. [*Comment:* If you find that argument compelling, I can only say that you're pretty gullible. Try it for an *atom*, and you'll conclude that the mass of the photon is about 7×10^{-30} g, which is nonsense. Nevertheless, it is a useful device for 'back-of-the-envelope' calculations, and it does very well for the pi meson. Unfortunately, many books present it as though it were a rigorous derivation, which it certainly is *not*. The uncertainty principle does *not* license violation of conservation of energy (nor does any such violation occur in this process; we shall see later on how this comes about). Moreover, it's an *inequality*, $\Delta E\ \Delta t \geq \hbar/2$, which at most could give you a *lower bound* on m. It *is* typically true that the *range* of a force is inversely proportional to the mass of the mediator, but the size of a bound state is not always a good measure of the range. (That's why the argument fails for the photon: the range of the electromagnetic force is infinite, but the size of an atom is not.) In general, when you hear a physicist invoke the uncertainty principle, keep a hand on your wallet.]

1.3 In the period before the discovery of the neutron, many people thought that the nucleus consisted of protons and *electrons*, with the atomic number equal to the excess number of protons. Beta decay seemed to support this idea – after all, electrons come popping out; doesn't that imply that there were electrons inside? Use the position–momentum uncertainty relation, $\Delta x\ \Delta p \geq \hbar/2$, to estimate the minimum momentum of an electron confined to a nucleus (radius 10^{-13} cm). From the relativistic energy–momentum relation, $E^2 - \mathbf{p}^2c^2 = m^2c^4$, determine the corresponding energy and compare it with that of an electron emitted in, say, the beta decay of tritium (Figure 1.5). (This result convinced some people that the beta decay electron could *not* have been rattling around inside the nucleus, but must be produced in the disintegration itself.)

1.4 The *Gell-Mann/Okubo mass formula* relates the masses of members of the baryon octet (ignoring small differences between p and n; Σ^+, Σ^0, and Σ^-; and Ξ^0 and Ξ^-):

$$2(m_N + m_\Xi) = 3m_\Lambda + m_\Sigma$$

Using this formula, together with the known masses of the *nucleon* N (use the average of p and n), Σ (again, use the average), and Ξ (ditto), 'predict' the mass of the Λ. How close do you come to the observed value?

1.5 The same formula applies to the mesons (with $\Sigma \to \pi$, $\Lambda \to \eta$, etc.), except that in this case, for reasons that remain something of a mystery, you must use the *squares* of the masses. Use this to 'predict' the mass of the η. How close do you come?

1.6 The mass formula for decuplets is much simpler – equal spacing between the rows:

$$m_\Delta - m_{\Sigma^*} = m_{\Sigma^*} - m_{\Xi^*} = m_{\Xi^*} - m_\Omega$$

Use this formula (as Gell-Mann did) to predict the mass of the Ω^-. (Use the average of the first two spacings to estimate the third.) How close is your prediction to the observed value?

1.7 (a) Members of the baryon decuplet typically decay after 10^{-23}seconds into a lighter baryon (from the baryon octet) and a meson (from the pseudo-scalar meson octet). Thus, for example, $\Delta^{++} \to p^+ + \pi^+$. List all decay modes of this form for the Δ^-, Σ^{*+}, and Ξ^{*-}. Remember that these decays must conserve charge and strangeness (they are *strong* interactions).

(b) In any decay, there must be sufficient mass in the original particle to cover the masses of the decay products. (There may be *more* than enough; the extra will be 'soaked up' in the form of kinetic energy in the final state.) Check each of the decays you proposed in part (a) to see which ones meet this criterion. The others are kinematically forbidden.

1.8 (a) Analyze the possible decay modes of the Ω^-, just as you did in Problem 1.7 for the Δ, Σ^*, and Ξ^*. See the problem? Gell-Mann predicted that the Ω^- would be 'metastable' (i.e. much longer lived than the other members of the decuplet), for precisely this reason. (The Ω^- *does* in fact decay, but by the much slower *weak* interaction, which does not conserve strangeness.)

(b) From the bubble chamber photograph (Figure 1.9), measure the length of the Ω^- track, and use this to estimate the lifetime of the Ω^-. (Of course, you don't know how fast it was going, but it's a safe bet that the speed was less than the velocity of light; let's say it was going about $0.1c$. Also, you don't know if the reproduction has enlarged or shrunk the scale, but never mind: this is quibbling over factors of 2, or 5, or maybe even 10. The important point is that the lifetime is many orders of magnitude longer than the 10^{-23} seconds characteristic of all other members of the decuplet).

1.9 Check the *Coleman–Glashow relation* [*Phys. Rev.* **B134**, 671 (1964)]:

$$\Sigma^+ - \Sigma^- = p - n + \Xi^0 - \Xi^-$$

(the particle names stand for their masses).

1.10 Look up the table of 'known' mesons compiled by Roos, M. (1963) *Reviews of Modern Physics*, **35**, 314, and compare the current *Particle Physics Booklet* to determine which of the 1963 mesons have stood the test of time. (Some of the names have been changed, so you will have to work from other properties, such as mass, charge, strangeness, etc.)

1.11 Of the spurious particles you identified in Problem 1.10, which are 'exotic' (i.e., inconsistent with the quark model)? How many of the surviving mesons are exotic?

1.12 How many different *meson* combinations can you make with 1, 2, 3, 4, 5, or 6 different quark flavors? What's the general formula for n flavors?

1.13 How many different *baryon* combinations can you make with 1, 2, 3, 4, 5, or 6 different quark flavors? What's the general formula for n flavors?

1.14 Using four quarks (u, d, s, and c), construct a table of all the possible baryon species. How many combinations carry a charm of $+1$? How many carry charm $+2$, and $+3$?

1.15 Same as Problem 1.14, but this time for *mesons*.

1.16 Assuming the top quark is too short-lived to form bound states ('truthful' mesons and baryons), list the 15 distinct meson combinations $q\bar{q}$ (not counting antiparticles) and the 35 distinct baryon combinations qqq. From the *Particle Physics Booklet* and/or other sources, determine which of these have been found experimentally. Give their name, mass, and year of discovery (just the lightest one, in each case). Thus, for instance, one baryon entry would be

$$sss : \Omega^-, \ 1672 \text{ MeV}/c^2, \quad 1964.$$

All hadrons are (presumably) various excitations of these 50 quark combinations.

1.17 A. De Rujula, H. Georgi, and S. L. Glashow [*Physical Review*, **D12**, 147 (1975)] estimated the so-called *constituent quark masses** to be: $m_u = m_d = 336$ MeV/c^2, $m_s = 540$ MeV/c^2, and $m_c = 1500$ MeV/c^2 (the bottom quark is about 4500 MeV/c^2). If they are right, the average binding energy for members of the baryon octet is -62 MeV. If they all had *exactly* this binding energy, what would their masses be? Compare the *actual* values and give the percent error. (Don't try this on the other supermultiplets, however. There really is no reason to suppose that the binding energy is the same for all members of the group. The problem of hadron masses is a thorny issue, to which we shall return in Chapter 5.)

1.18 Shupe, M. (1979) [*Physics Letters*, **86B**, 87] proposed that all quarks and leptons are composed of two even more elementary constituents: c (with charge $-1/3$) and n (with charge zero) – and their respective antiparticles, $\bar{c}$ and $\bar{n}$. You're allowed to combine them in groups of three particles or three antiparticles (ccn, for example, or $\bar{n}\bar{n}\bar{n}$). Construct all of the eight quarks and leptons in the first generation in this manner. (The other generations are supposed to be excited states.) Notice that each of the *quark* states admits three possible permutations (ccn, cnc, ncc, for example) – these correspond to the three colors. Mediators can be constructed from three particles plus three antiparticles. $W^\pm$, Z^0, and γ involve three *like* particles and three like antiparticles ($W^- = ccc\bar{n}\bar{n}\bar{n}$, for instance). Construct W^+, Z^0, and γ in this way. Gluons involve mixed combinations ($ccn\bar{c}\bar{c}\bar{n}$, for instance). How many possibilities are there in all? Can you think of any way to reduce this down to eight?

1.19 Your roommate is a chemistry major. She knows all about protons, neutrons, and electrons, and she sees them in action every day in the laboratory. But she is skeptical when you tell her about positrons, muons, neutrinos, pions, quarks, and intermediate vector bosons. Explain to her why none of these plays any direct role in chemistry. (For instance, in the case of the muon a reasonable answer might be 'They are unstable, and last only a millionth of a second before disintegrating.')

* For reasons we will come to in due course, the effective mass of a quark bound inside a hadron is not the same as the 'bare' mass of the 'free' quark.

2

Elementary Particle Dynamics

This chapter introduces the fundamental forces by which elementary particles interact, and the ***Feynman diagrams*** *we use to represent these interactions. The treatment is entirely qualitative and can be read quickly to get a sense of the 'lay of the land'. The quantitative details will come in Chapters 6 through 9.*

2.1
The Four Forces

As far as we know, there are just four fundamental forces in nature: *strong, electromagnetic, weak,* and *gravitational.* They are listed in the following table in order of decreasing strength:*

Force	Strength	Theory	Mediator
Strong	10	Chromodynamics	Gluon
Electromagnetic	10^{-2}	Electrodynamics	Photon
Weak	10^{-13}	Flavordynamics	W and Z
Gravitational	10^{-42}	Geometrodynamics	Graviton

To each of these forces there belongs a physical theory. The classical theory of gravity is, of course, Newton's law of universal gravitation. Its relativistic generalization is Einstein's general theory of relativity ('geometrodynamics' would be a better term). A completely satisfactory quantum theory of gravity has yet to be worked out; for the moment, most people assume that gravity is simply too weak to play a significant role in elementary particle physics. The physical theory that describes electromagnetic forces is called *electrodynamics.* It was given its classical formulation

* The 'strength' of a force is an intrinsically ambiguous notion – after all, it depends on the nature of the source and on how far away you are. So the numbers in this table should not be taken too literally, and (especially in the case of the weak force) you will see quite different figures quoted elsewhere.

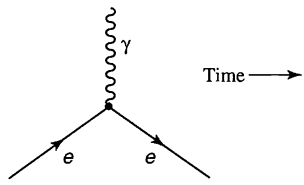
Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

by Maxwell over one hundred years ago. Maxwell's theory was already consistent with special relativity (for which it was, in fact, the main inspiration). The quantum theory of electrodynamics was perfected by Tomonaga, Feynman, and Schwinger in the 1940s. The weak forces, which account for nuclear beta decay (and also, as we have seen, the decay of the pion, the muon, and many of the strange particles), were unknown to classical physics; their theoretical description was given a relativistic quantum formulation right from the start. The first theory of the weak forces was presented by Fermi in 1933; it was refined by Lee and Yang, Feynman and Gell-Mann, and many others, in the 1950s, and put into its present form by Glashow, Weinberg, and Salam, in the 1960s. For reasons that will appear in due course, the theory of weak interactions is sometimes called *flavordynamics* [1]; in this book, I refer to it simply as the Glashow–Weinberg–Salam (GWS) theory. (The GWS model treats weak and electromagnetic interactions as different manifestations of a single *electroweak* force, and in this sense the four forces reduce to three.) As for the strong forces, beyond the pioneering work of Yukawa in 1934 there really *was* no theory until the emergence of chromodynamics in the 1970s.

Each of these forces is mediated by the exchange of a particle. The gravitational forces are mediated by the *graviton*, electromagnetic forces are mediated by the *photon*, strong forces by the *gluon*, and weak forces by the *intermediate vector bosons*, W and Z. These mediators transmit the force between one quark or lepton and another. In principle, the force of impact between a bat and a baseball is nothing but the combined interaction of the quarks and leptons in one with the quarks and leptons in the other. More to the point, the strong force between two protons, say, which Yukawa took to be a fundamental and irreducible process, must be regarded as a complicated interaction of six quarks. This is clearly not the place to look for simplicity. Rather, we must begin by analyzing the force between one truly elementary particle and another. In this chapter, I will show you qualitatively how each of the relevant forces acts on individual quarks and leptons. Subsequent chapters develop the machinery needed to make the theory quantitative.

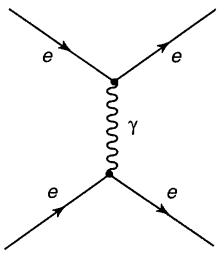
2.2 Quantum Electrodynamics (QED)

Quantum electrodynamics (QED) is the oldest, the simplest, and the most successful of the dynamical theories; the others are self-consciously modeled on it. So I'll begin with a description of QED. *All electromagnetic phenomena are ultimately reducible to the following elementary process*:



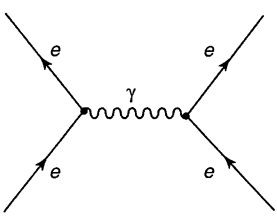
In these figures time flows *horizontally*, to the right, so this diagram reads: a charged particle, e, enters, emits (or absorbs) a photon, γ, and exits. For the sake of argument, I'll assume that the charged particle is an electron; it could just as well be a quark, or any lepton except a neutrino (the latter is neutral, of course, and does not experience an electromagnetic force).

To describe more complicated processes, we simply combine two or more replicas of this *primitive vertex*. Imagine that you have a bag full of 'tinker toy' models of the primitive vertex, made out of flexible plastic. You can snap them together, photon-to-photon or electron-to-electron (but in the latter case you must preserve the direction of the arrows). Consider, for example, the following:



Here, two electrons enter, a photon passes between them (I need not say which one emits the photon and which one absorbs it; the diagram represents *both* orderings), and the two exit.* This diagram, then, describes the interaction between two electrons; in the classical theory, we would call it the Coulomb repulsion of like charges. In QED, this process is called *Møller scattering*; we say that the interaction is 'mediated by the exchange of a photon', for reasons that should now be apparent.

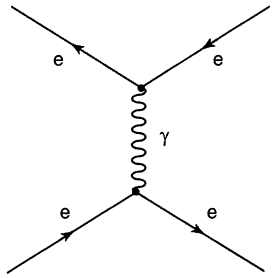
You're allowed to twist these 'Feynman diagrams' around into any topological configuration you like – for example, we could stand the previous picture on its side:



A particle line running 'backward in time' (an arrow pointing toward the left) is interpreted as the corresponding *antiparticle* going *forward* (the photon is its own antiparticle, that's why I didn't need an arrow on the photon line). In this process

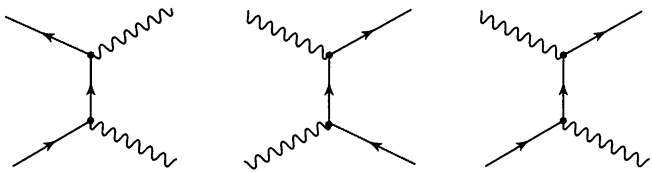
* In reading a Feynman diagram it sometimes helps to picture a vertical line that sweeps along to the right, representing the passage of time. In the beginning (far left) it intersects two electron lines, in the middle it encounters the exchanged photon, and at the end (far right) there are again just two electrons.

an electron and a positron* annihilate to form a photon, which in turn produces a new electron–positron pair. An electron and a positron went *in*, an electron and a positron came *out* (not the *same* ones, but then, since all electrons are identical, it hardly matters). This represents the interaction of two opposite charges: their Coulomb attraction. In QED, this process is called *Bhabha scattering*. Actually, there is a quite different diagram which also describes Bhabha scattering:

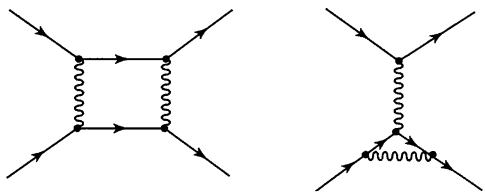


As we shall see, both diagrams must be included in the analysis.

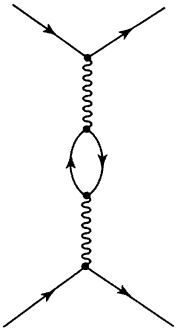
Using just two vertices we can also construct the following diagrams, describing, respectively, pair annihilation, $e^- + e^+ \to \gamma + \gamma$; pair production, $\gamma + \gamma \to e^- + e^+$; and Compton scattering, $e^- + \gamma \to e^- + \gamma$:

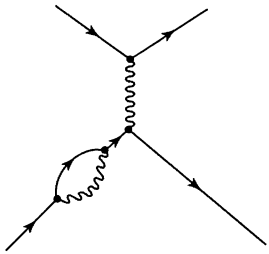


Notice that Bhabha and Møller scattering are related by crossing symmetry (Section 1.4), as are the three processes shown here. In terms of Feynman diagrams, crossing symmetry corresponds to twisting or rotating the figure. If we allow more vertices (just reach in the bag and pull out a few more tinker toys), the possibilities rapidly proliferate; for example, with four vertices we obtain, among others, the following diagrams:



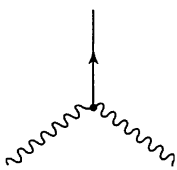
* Some authors would label the upper left and lower right lines in this diagram with $\bar{e}$, to remind you that it's an antiparticle. I think this is dangerous notation. The arrow already tells you it's the antiparticle, and a literal reading would suggest that it is an antiparticle going backwards in time ... which would be a *particle*. I prefer to label all lines with the *particle* symbol, and let the arrow tell you whether it is in fact the antiparticle.



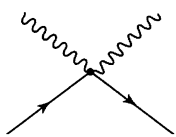


In each of these figures two electrons went in and two electrons came out. They too describe the repulsion of like charges (Møller scattering). The 'innards' of the diagram are irrelevant as far as the observed process is concerned. Internal lines (those which begin and end within the diagram) represent particles that are not observed – indeed, that *cannot* be observed without entirely changing the process. We call them *virtual particles*. Only the *external* lines (those that enter or leave the diagram) represent 'real' (observable) particles. The external lines, then, tell you what physical process is occurring; the internal lines describe the *mechanism* involved.

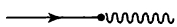
At the purely qualitative level this is such a childishly simple game that there's a serious danger you will inadvertently embellish the rules. If you find yourself drawing a Feynman diagram that contains the vertex



for example, or



or you snap a photon line onto an electron line

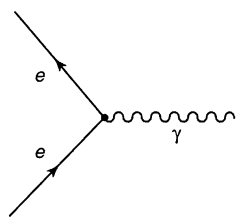


you have made a mistake – the bag contains no such tinker toys, and the snaps just don't work when you try to hook a photon to an electron. Your diagram might conceivably describe some *other* interaction, but it's not electrodynamics.

Feynman diagrams are purely symbolic; they do *not* represent particle trajectories (as you might see them in, say, a bubble chamber photograph). The horizontal dimension is *time*, but vertical spacing does *not* correspond to physical separation. For instance, in Bhabha scattering the electron and positron are *attracted*, not repelled (as the diverging lines might seem to suggest). All that the diagram says is: 'Once there was an electron and a positron; they exchanged a photon; then there was an electron and a positron again'.

Quantitatively, each Feynman diagram stands for a particular *number*, which can be calculated using the so-called *Feynman rules* (you'll learn how to do this in Chapter 6). Suppose you want to analyze a certain physical process (say, Møller scattering). First you draw all the diagrams that have the appropriate external lines (the one with two vertices, all the ones with four vertices, and so on), then you evaluate the contribution of each diagram, using the Feynman rules, and add it all up. The *sum total* of all Feynman diagrams with the given external lines represents the actual physical process. Of course, there's a wee problem here: there are infinitely many Feynman diagrams for any particular reaction! Fortunately, each vertex within a diagram introduces a factor of $\alpha = e^2/\hbar c = 1/137$, the *fine structure constant*. Because this is such a small number, diagrams with more and more vertices contribute less and less to the final result, and, depending on the accuracy you need, may be ignored. In fact, in QED it is rare to see a calculation that includes diagrams with more than four vertices. The answers are only approximate, to be sure, but when the approximation is valid to six significant digits, only the most fastidious are likely to complain.

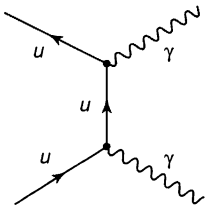
The Feynman rules enforce conservation of energy and momentum at each vertex, and hence for the diagram as a whole. It follows that the primitive QED vertex *by itself* does not represent a possible physical process. We can draw the diagram, but calculation would assign to it the number *zero*. The reason is purely kinematical: $e^- \to e^- + \gamma$ would violate conservation of energy. (In the center-of-mass frame the electron is initially at rest, so its energy is mc^2. It cannot decay into a photon plus a recoiling electron because the latter alone would require an energy greater than mc^2.) Nor, for instance, is $e^- + e^+ \to \gamma$ kinematically possible, although it is easy enough to draw the diagram:



In the center-of-mass system the electron and positron enter symmetrically with equal and opposite velocities, so the total momentum before the collision is obviously zero. But the *final* momentum *cannot* be zero, since photons always

travel at the speed of light; an electron–positron pair can annihilate to make *two* photons, but not *one*. *Within a larger diagram*, however, these figures are perfectly acceptable, because, although energy and momentum must be conserved at each vertex, *a virtual particle does not carry the same mass* as the corresponding free particle. In fact, a virtual particle can have *any* mass.* In the business, we say that virtual particles do not lie on their *mass shell*. *External* lines, by contrast, represent *real* particles, and these do carry the 'correct' mass.†

I have been assuming that the charged particle in question is an electron,‡ but it could just as well be a muon, say, or a quark. What would you make of the following diagram?



Here a $u/\bar{u}$ pair annihilates, producing two photons (*one* photon, remember, is kinematically forbidden). Because of quark confinement you're not going to witness this as a scattering experiment, but what if the quarks were bound together in the form of a meson – a π^0, for example? This diagram would represent the 'decay' of the π^0: $\pi^0 \to \gamma + \gamma$. I put the word in quotes, because in a deeper sense this is not a decay at all – it's just ordinary old pair annihilation, in which the original pair happen to be bound together as a meson. This explains why the π^0 has a lifetime 9 orders of magnitude smaller than its charged siblings ($\pi^\pm$) – it decays by an electromagnetic process, whereas the others have to await the weak interactions, which are much slower.

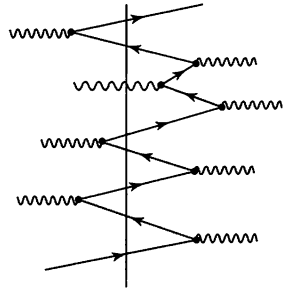
I cannot resist telling you an amusing fable, but you must promise not to take it too seriously. Feynman claimed that his advisor (J. A. Wheeler) once offered the

* In special relativity, the energy E, momentum, $\mathbf{p}$, and mass m of a free particle are related by the equation $E^2 - \mathbf{p}^2c^2 = m^2c^4$. But for a *virtual* particle $E^2 - \mathbf{p}^2c^2$ can take on *any* value. Many authors interpret this to mean that virtual processes violate conservation of energy (see Problem 1.2). Personally, I consider this misleading, at best. Energy is *always* conserved.

† Actually, the *physical* distinction between real and virtual particles is not quite as sharp as I have implied. If a photon is emitted on Alpha Centauri, and absorbed in your eye, it is technically a virtual photon, I suppose. However, in general, the farther a virtual particle is from its mass shell the shorter it lives, so a photon from a distant star would have to be extremely close to its 'correct' mass – it would have to be *almost* 'real'. As a calculational matter, you would get essentially the same answer if you treated the process as two separate events (emission of a real photon by star, followed by absorption of a real photon by eye). You might say that a real particle is a virtual particle that lasts long enough that we don't care to inquire how it was produced, or how it is eventually absorbed.

‡ In practice, the term 'quantum electrodynamics' is usually taken to mean the interaction of electrons, positrons, and photons, unless otherwise specified.

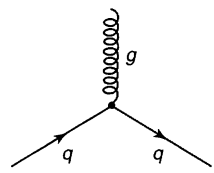
following explanation for why all electrons are identical: there's only *one* of 'em! It's riding along on a diagram of the form



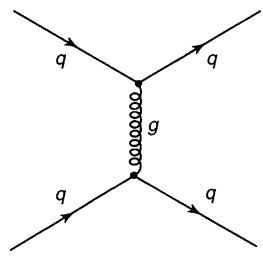
At a given instant (the vertical line) the electron is present (on this segment) four times as a particle and three times as an antiparticle – but it's all the same electron. Of course, this does imply that the number of positrons in the universe should equal the number of electrons (give or take one), but apart from that it's kind of cute.

2.3
Quantum Chromodynamics (QCD)

In chromodynamics, *color* plays the role of charge, and the fundamental process (analogous to $e \to e + \gamma$) is quark $\to$ quark plus gluon ($q \to q + g$):*



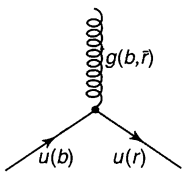
As before, we combine two or more such 'primitive vertices' to represent more complicated processes. For example, the force between two quarks (which is responsible in the first instance for binding quarks together to make hadrons, and indirectly for holding the neutrons and protons together to form a nucleus) is described in lowest order by the diagram:



* Since leptons do not carry color, they do not participate in the strong interactions.

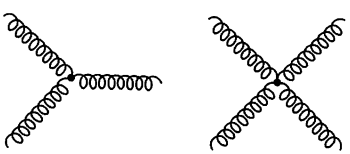
We say that the force between two quarks is 'mediated' by the exchange of gluons.

At this level chromodynamics is very similar to electrodynamics. However, there are also important differences, most conspicuously the fact that whereas there is only one kind of electric charge (it can be positive or negative, to be sure, but a *single number* suffices to characterize the charge of a particle), there are *three* kinds of color (red, green, and blue). In the fundamental process $q \to q + g$, the color of the quark (but not its flavor) may change. For example, a blue up-quark may convert into a red up-quark. Since color (like charge) is always conserved, this means that the gluon must carry away the difference – in this instance, one unit of blueness and *minus* one unit of redness:



Gluons, then, are 'bicolored', carrying one positive unit of color and one negative unit. There are evidently $3 \times 3 = 9$ possibilities here, and you might expect there to be nine kinds of gluons. For technical reasons, which we'll come to in Chapter 8, there are actually only eight.

Since the gluons *themselves* carry color (unlike the photon, which is electrically neutral), they couple directly to other gluons, and hence in addition to the fundamental quark–gluon vertex, we also have primitive gluon–gluon vertices; in fact, two kinds: three-gluon vertices and four-gluon vertices:



This direct gluon–gluon coupling makes chromodynamics a lot more complicated than electrodynamics, but also far richer, allowing, for instance, the possibility of glueballs (bound states of interacting gluons, with no quarks in sight).

Another difference between chromodynamics and electrodynamics is the size of the *coupling constant*. Remember that each vertex in QED introduces a factor of $\alpha = 1/137$, and the smallness of this number means that we need only consider Feynman diagrams with a small number of vertices. Experimentally, the corresponding coupling constant for the strong forces, α_s – as determined, say, from the force between two protons – is greater than 1, and the *bigness* of this number has plagued particle physics for decades. Instead of contributing less and less, the more complex diagrams contribute *more and more*, and Feynman's procedure, which worked so well in QED, is apparently doomed. One of the great

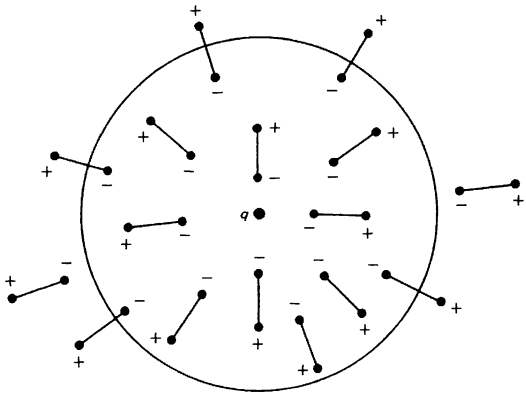


Fig. 2.1 Screening of a charge q by a dielectric medium.

triumphs of quantum chromodynamics (QCD) was the discovery that in this theory the number that plays the role of coupling 'constant' is in fact not constant at all, but depends on the separation distance between the interacting particles (we call it a 'running' coupling constant). Although at the relatively *large* distances characteristic of nuclear physics it is big, at very short distances (less than the size of a proton) it becomes quite small. This phenomenon is known as *asymptotic freedom* [2]; it means that within a proton or a pion, say, the quarks rattle around without interacting much. Just such behavior was found experimentally in the deep inelastic scattering experiments. From a theoretical point of view, the discovery of asymptotic freedom rescued the Feynman calculus as a legitimate tool for QCD, in the high-energy regime.

Even in electrodynamics, the effective coupling depends somewhat on how far you are from the source. This can be understood qualitatively as follows. Picture first a positive point charge q embedded in a dielectric medium (i.e. a substance whose molecules become polarized in the presence of an electric field). The negative end of each molecular dipole will be attracted toward q, and the positive end repelled away, as shown in Figure 2.1 As a result, the particle acquires a 'halo' of negative charge that partially cancels its field. In the presence of the dielectric, then, the *effective* charge of any particle is somewhat reduced:

$$q_{\text{eff}} = q/\epsilon \tag{2.1}$$

(The factor ϵ by which the field is reduced is called the *dielectric constant* of the material; it is a measure of the ease with which the substance can be polarized [3].) Of course, if you are closer than the nearest molecule, then there is no such screening, and you 'see' the full charge q. Thus if you were to make a graph of the

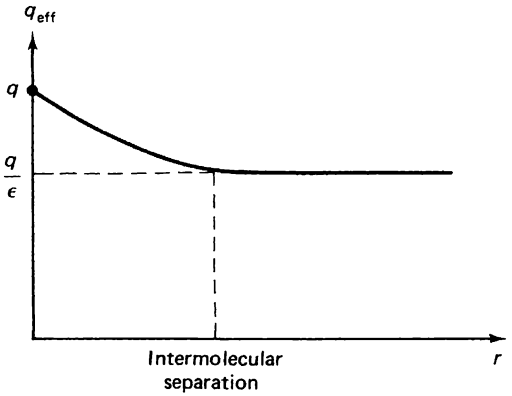
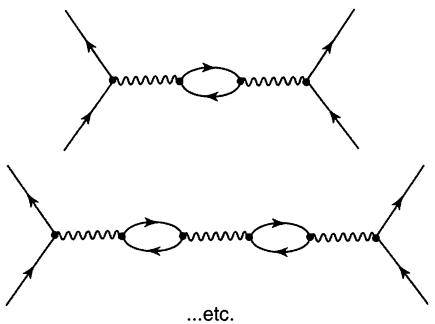


Fig. 2.2 Effective charge as a function of distance.

effective charge, as a function of distance, it would look something like Figure 2.2 The effective charge *increases* at very small distances.

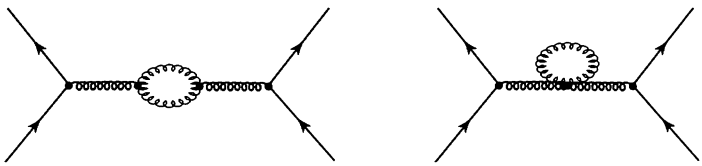
Now, it so happens that in *quantum* electrodynamics the vacuum itself behaves like a dielectric; it sprouts positron–electron pairs, as shown in Feynman diagrams such as these:



The virtual electron in each 'bubble' is attracted toward q, and the virtual positron is repelled away; the resulting *vacuum polarization* partially screens the charge and reduces its field. Once again, however, if you get *too* close to q, the screening disappears. What plays the role of the 'intermolecular spacing' in this case is the Compton wavelength of the electron, $\lambda_c = h/mc = 2.43 \times 10^{-10}$ cm. For distances smaller than this the effective charge increases, just as it did in Figure 2.2. Notice that the *unscreened* ('close-up') charge, which you might regard as the 'true' charge of the particle, is *not* what we measure in any ordinary experiment,

since we are seldom working at such minute separation distances.* What we have always called 'the charge of the electron' is actually the fully screened *effective* charge.

So much for electrodynamics. The same thing happens in QCD, but with an important added ingredient. Not only do we have the quark–quark–gluon vertex (which, by itself, would again lead to an increasing coupling strength at short distances), but now there are *also* the direct gluon–gluon vertices. So in addition to the diagrams analogous to vacuum polarization in QED, we must now also include gluon loops, such as these:



It is not clear a priori what influence these diagrams will have on the story [4]; as it turns out, their effect is the *opposite*: There occurs a kind of competition between the quark polarization diagrams (which drive α_s *up* at short distances) and gluon polarization (which drives it *down*). Since the former depends on the number of quarks in the theory (hence on the number of *flavors*, f), whereas the latter depends on the number of gluons (hence on the number of *colors*, n), the winner in the competition depends on the relative number of flavors and colors. The critical parameter turns out to be

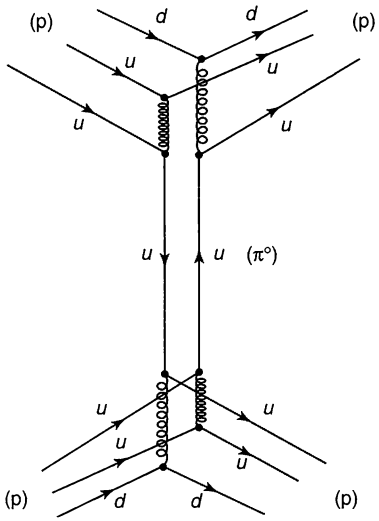
$$a \equiv 2f - 11n \tag{2.2}$$

If this number is *positive*, then, as in QED, the effective coupling *increases* at short distances; if it is *negative*, the coupling *decreases*. In the Standard Model, $f = 6$ and $n = 3$, so $a = -21$, and the QCD coupling decreases at short distances. This is the origin of asymptotic freedom.

The final distinction between QED and QCD is that whereas many particles carry electric charge, no naturally occurring particles carry color. Quarks are confined in colorless packages of two (mesons) and three (baryons). As a consequence, the processes we actually observe in the laboratory are necessarily indirect and complicated manifestations of chromodynamics. It is as though our only access to electrodynamics came from the van der Waals forces between neutral molecules. For example, the (strong) force between two protons involves (among many others)

* An exception is the Lamb shift – a tiny perturbation in the spectrum of hydrogen – in which the influence of vacuum polarization (or rather, its *absence* at short distances) is clearly discernible.

the following diagram:



You will recognize here the remnants of Yukawa's pion-exchange model, but the entire process is enormously more complex than Yukawa ever imagined.

If QCD is correct, it must contain the explanation for quark confinement; that is, it must be possible to *prove*, as a consequence of this theory, that quarks can only exist in the form of colorless combinations. Presumably this proof will take the form of a demonstration that the potential energy increases without limit as the quarks are pulled farther and farther apart, so that it would require an infinite energy (or at any rate, enough to create new quark–antiquark pairs) to separate them completely (see Figure 2.3). So far, no one has provided a conclusive proof that QCD implies confinement (see, however, Reference 27 in Chapter 1). The difficulty is that confinement involves the *long-range* behavior of the quark–quark interaction, but this is precisely the regime in which the Feynman calculus fails.*

2.4 Weak Interactions

There is no particular name for the 'stuff' that produces weak forces, in the sense that electric charge produces electromagnetic forces and color produces strong

* There are strong indications that a 'phase transition' occurs at extremely high densities – three or four times that of an atomic nucleus – leading to deconfinement and the so-called *quark–gluon plasma*. Thus, free quarks may have existed in the first moments after the Big Bang, and efforts are underway to recreate similar conditions (on a smaller scale!) in the laboratory, using the Relativistic Heavy Ion Collider (RHIC) at Brookhaven (5).

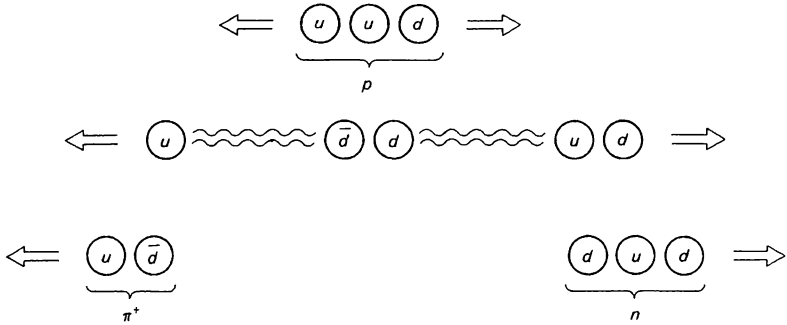
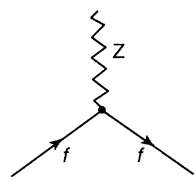


Fig. 2.3 A possible scenario for quark confinement: as we pull a u quark out of the proton, a pair of quarks is created, and instead of a free quark, we are left with a pion and a neutron.

forces. Some people call it 'weak charge'. Whatever word you use, all quarks and all leptons carry it [6]. (Leptons have no color, so they do not participate in the strong interactions; neutrinos have no charge, so they experience no electromagnetic forces; but *all* of them join in the weak interactions.) There are two kinds of weak interactions: *charged* (mediated by the Ws) and *neutral* (mediated by the Z). The neutral weak interactions are much simpler, so I'll start with them.*

2.4.1 Neutral

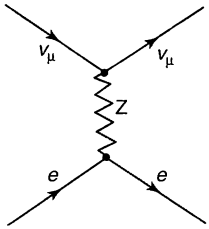
The fundamental neutral vertex is:†



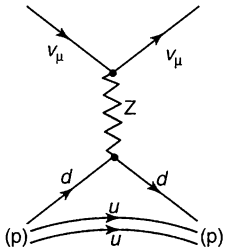
where f can be any lepton or any quark. The Z mediates such processes as neutrino–electron scattering ($\nu_\mu + e^- \to \nu_\mu + e^-$):

* Although charged weak interactions were known right from the start (beta decay is the classic example), the theoretical possibility of neutral weak processes was not appreciated until 1958. The GWS model includes neutral weak interactions as essential ingredients, and their existence was first confirmed in neutrino scattering experiments at CERN, in 1973 (7).

† It is traditional to use a wavy line for the photon, and a springy line for the gluon, but there is no consistency in the literature for the weak mediators. I'm going to use a jagged line, but this is not standard notation. (I'll use a solid line for spin-1/2 particles, which is standard, and a dashed line for spin 0, which is not.)

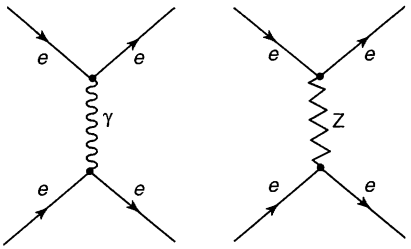


and neutrino–proton scattering ($\nu_\mu + p \to \nu_\mu + p$)



(in the latter case, two 'spectator' quarks go along for the ride, bound to the d by strong forces – gluon exchange – that, for simplicity, we do not draw).*

Notice that any process mediated by the photon could also be mediated by the Z – for example, electron–electron scattering:



Presumably there is a minute correction to Coulomb's law attributable to the second diagram, but the photon-mediated process overwhelmingly dominates. Experiments at DESY (in Hamburg) studied the reaction $e^- + e^+ \to \mu^- + \mu^+$ at very high energy and found unmistakable evidence of a contribution from the Z [8]. In atomic physics, neutral weak contamination of electromagnetic processes can sometimes be teased out by exploiting the fact that weak interactions carry a unique fingerprint: they violate conservation of parity (mirror symmetry) [9].

* There are also, of course, diagrams in which the Z couples to one of the u quarks.

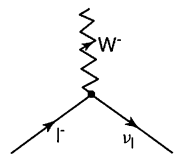
Still, to observe a *pure* neutral weak interaction one has to resort to neutrino scattering, in which there is no competing electromagnetic mechanism – and neutrino experiments are notoriously difficult.

2.4.2 Charged

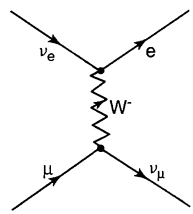
The primitive vertices for strong, electromagnetic, and neutral weak interactions all share the feature that the *same* quark or lepton comes out as went in – accompanied, of course, by a gluon, photon, or Z, as the case may be. Well ... OK: in QCD the *color* of the quark may change, but the *flavor* never does. The charged weak interactions are the only ones that change flavor, and in this sense they are the only ones capable of causing a 'true' decay (as opposed to a mere repackaging of the quarks, or a hidden pair production or annihilation). I'll begin with the charged weak interactions of leptons.*

2.4.2.1 Leptons

The fundamental charged vertex looks like this:



A negative lepton (it could be e^-, μ^-, or τ^-) converts into the corresponding neutrino, with emission of a W^- (or absorption of a W^+): $l^- \to \nu_l + W^-$.† As always, we combine the primitive vertices to generate more complicated reactions. For example, the process $\mu^- + \nu_e \to e^- + \nu_\mu$ would be represented by the diagram:

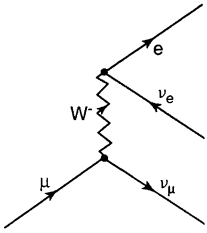


Such a neutrino–muon scattering event would be hard to set up in the laboratory, but with a slight twist essentially the same diagram describes the decay of the

* The discovery of neutrino oscillations will force some modifications in this picture, but we do not know yet exactly what form they will take (perhaps they will bring the theory into line with the quarks), so for the moment I will stick to the simple (pre-oscillation) story.

† This implies, of course, that the crossed reaction $l^+ \to \bar{\nu}_l + W^+$ is also allowed.

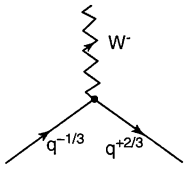
muon, $\mu^- \to e^- + \nu_\mu + \overline{\nu}_e$:



This is the cleanest of all charged weak interactions; we'll study it in detail in Chapter 10.*

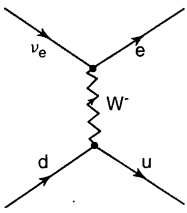
2.4.3 Quarks

Notice that the leptonic weak vertices connect members of the *same generation*: e^- converts to ν_e (with emission of W^-), or $\mu^- \to \mu^-$ (emitting a Z), but e^- never goes to μ^- nor μ^- to ν_e. In this way, the theory enforces the conservation of electron number, muon number, and tau number. It is tempting to suppose that the same rule applies to the quarks, so that the fundamental charged vertex is:



A quark with charge $-\frac{1}{3}$ (which is to say: d, s, or b) converts into the corresponding quark with charge $+\frac{2}{3}$ (u, c, or t, respectively), with the emission of a W^-. The outgoing quark carries the same color as the ingoing one, but a different flavor.†

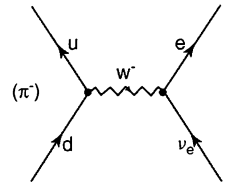
The far end of the W line can couple to leptons (a 'semileptonic' process), or to other quarks (a purely hadronic process). The most important semileptonic process is $d + \nu_e \to u + e$:



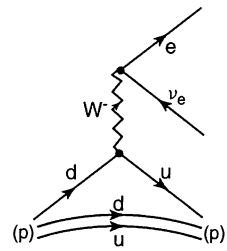
* Technically, it is only the lowest-order contribution to muon decay, but in weak interaction theory one almost never needs to consider higher-order corrections.

† It's not that the W^- carries off the 'missing' flavor – the W's *have* no flavor; flavor is simply *not conserved* in the charged weak interactions.

Because of quark confinement, this process would never occur in nature as it stands. However, turned on its side, and with the $\overline{u}$ and d bound together (by the strong force), this diagram represents a possible decay of the pion, $\pi^- \to e^- + \overline{\nu}_e$:

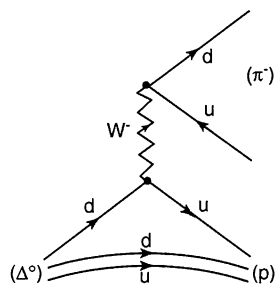


(For reasons to be discussed later, the more common decay is actually $\pi^- \to \mu^- + \overline{\nu}_\mu$, but the diagram is the same, with e replaced by μ.) Moreover, essentially the same diagram accounts for the beta decay of the neutron ($n \to p^+ + e^- + \overline{\nu}_e$):



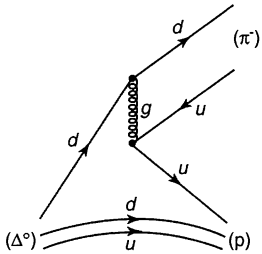
Thus, apart from strong interaction contamination (in the form of the spectator quarks), the decay of the neutron is identical in structure to the decay of the muon, and closely related to the decay of the pion. In the days before the quark model, these appeared to be three very different processes.

Eliminating the electron–neutrino vertex in favor of a second quark vertex we obtain a purely hadronic weak interaction, $\Delta^0 \to p^+ + \pi^-$:*



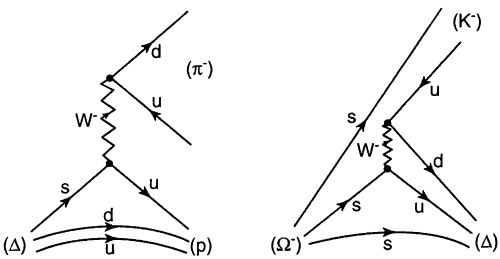
* The Δ^0 has the same quark content as the neutron, but this decay is not possible for neutrons because they are not heavy enough to make a proton and a pion.

Actually, this particular decay also proceeds by the strong interaction:



the weak mechanism is an immeasurably small contribution. We'll see more realistic examples of nonleptonic weak interactions in a moment.

So far, it's all pretty simple: the quarks mimic the leptons – the only difference is that the strong force (to which, remember, the leptons are immune) complicates the picture with spectators and what-not that have nothing to do with the basic weak process. Sad to say, this story is a little *too* simple. For as long as the fundamental quark vertex is allowed to operate only *within each generation*, we can never hope to account for strangeness-changing weak interactions, such as the decay of the lambda ($\Lambda \to p^+ + \pi^-$) or the omega-minus ($\Omega^- \to \Lambda + K^-$), which involves the conversion of a strange quark into an up quark:



The solution to this dilemma was suggested by Cabibbo in 1963, perfected by Glashow, Illiopoulos, and Maiani (GIM) in 1970, and extended to three generations by Kobayashi and Maskawa (KM) in 1973.* The essential idea is that the quark generations are 'skewed,' for the purposes of weak interactions.†

* The Cabibbo/GIM/KM mechanism will be discussed more fully in Chapter 9.

† Technically, this applies to the neutral as well as the charged weak interactions. But in the former case it doesn't matter, and I have tried to keep the story as clear as possible by avoiding the issue at that stage. Historically, when there were only three quarks known it was a puzzle why (experimentally) there were no strangeness-changing neutral weak interactions. The GIM mechanism introduced a fourth quark (four years *before* the November Revolution), and a 2 × 2 'KM matrix', to provide for a miraculous cancellation, the net effect of which (in the neutral case) was the same as if we had never 'skewed' the quarks in the first place.

Instead of

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix} \tag{2.3}$$

the weak force couples the pairs

$$\begin{pmatrix} u \\ d' \end{pmatrix}, \begin{pmatrix} c \\ s' \end{pmatrix}, \begin{pmatrix} t \\ b' \end{pmatrix} \tag{2.4}$$

where d', s', and b' are *linear combinations* of the physical quarks d, s, and b:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.5}$$

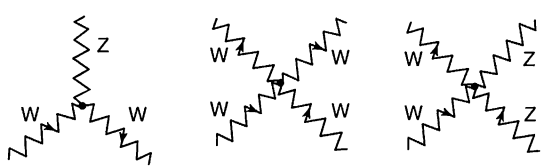
If this 3×3 *Kobayashi–Maskawa matrix* were the *unit* matrix, then d', s', and b' would be the same as d, s, and b, and no 'cross-generational' transitions could occur. 'Upness-plus-downness' would be absolutely conserved (just as the electron number is); 'strangeness-plus-charm' would be conserved (like muon number); and so would 'topness-plus-bottomness' (like tau number). But it's *not* the unit matrix (although it's pretty close); experimentally, the magnitudes of the matrix elements are [10]

$$\begin{pmatrix} 0.974 & 0.227 & 0.004 \\ 0.227 & 0.973 & 0.042 \\ 0.008 & 0.042 & 0.999 \end{pmatrix} \tag{2.6}$$

V_{ud} measures the coupling of u to d, V_{us} the coupling of u to s, and so on. The fact that the latter is nonzero is what permits strangeness-changing processes, such as the decay of the Λ and the Ω^-, to occur.*

2.4.4
Weak and Electromagnetic Couplings of W and Z

There are also direct couplings of W and Z to one another, in GWS theory (just as there are direct gluon–gluon couplings in QCD):



* Neutrino oscillations involve cross-generational couplings in the lepton sector, so it may be that we will have a 'KM matrix' for the leptons as well. See Chapter 11.

Moreover, because the W is charged, it couples to the photon:

Although these interactions are critical for the internal consistency of the theory, they are of limited practical importance at this stage (see Problem 2.6).

2.5 Decays and Conservation Laws

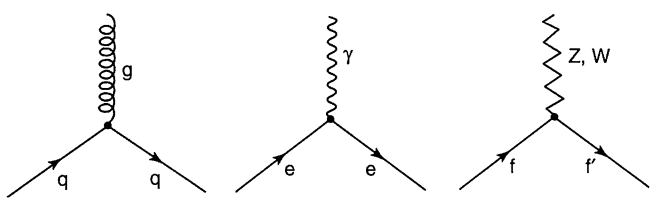
One of the most striking general properties of elementary particles is their tendency to disintegrate; we might almost call it a universal principle that *every particle decays into lighter particles, unless prevented from doing so by some conservation law*. The photon is stable (having zero mass, there is nothing lighter for it to decay into); the electron is stable (it's the lightest charged particle, so conservation of charge prevents its decay); the proton is presumably stable (it's the lightest baryon, and the conservation of baryon number saves it – likewise conservation of lepton number protects the lightest of the neutrinos). By the same token, the positron, the antiproton, and the lightest antineutrino are stable. But *most* particles spontaneously disintegrate – even the neutron, although it becomes stable in the protective environment of many atomic nuclei. In practice, our world is populated mainly by protons, neutrons, electrons, photons, and neutrinos; more exotic things are created now and then (by collisions) but they don't last long. Each unstable species has a characteristic mean lifetime,* τ: for the muon it's 2.2×10^{-6} sec; for the π^+ it's 2.6×10^{-8} sec; for the π^0 it's 8.3×10^{-17} sec. In fact, most particles exhibit several different decay *modes*; 64% of all K^+'s, for example, decay into $\mu^+ + \nu_\mu$, but 21% go to $\pi^+ + \pi^0$, 6% to $\pi^+ + \pi^+ + \pi^-$, 5% to $(e^+ + \nu_e + \pi^0)$, and so on. One of the goals of elementary particle theory is to calculate these lifetimes and *branching ratios*.

A given decay is governed by one of the three fundamental forces: $\Delta^{++} \to p^+ + \pi^+$, for example, is a strong decay; $\pi^0 \to \gamma + \gamma$ is electromagnetic; and $\Sigma^- \to n + e^- + \overline{\nu}_e$ is weak. How can you tell? Well, if a photon comes out, the process is certainly electromagnetic, and if a neutrino emerges, the process is certainly weak. But if neither a photon nor a neutrino is present, it's a little harder to say. For example, $\Sigma^- \to n + \pi^-$ is weak, but $\Delta^- \to n + \pi^-$ is strong. I'll show you in a moment how to figure that out, but first I want to mention the most dramatic *experimental* difference between strong, electromagnetic, and weak decays: a typical

* The *lifetime* τ is related to the *half-life* $t_{1/2}$ by the formula $t_{1/2} = (\ln 2)\tau = 0.693\tau$. The half-life is the time it takes for half the particles in a large sample to disintegrate (see Section 6.1).

strong decay involves a lifetime around 10^{-23} sec, a typical electromagnetic decay takes about 10^{-16} sec, and weak decay times range from around 10^{-13} sec (for the τ) up to 15 min (for the neutron). For a given type of interaction, the decay generally proceeds more rapidly the larger the mass difference between the original particle and the decay products, just as a ball rolls faster down a steeper hill.* It is this kinematic effect that accounts for the enormous range in weak interaction lifetimes. In particular, the proton and electron together are *so* close to the neutron's mass that the decay $n \to p^+ + e^- + \bar{\nu}_e$ *barely* makes it at all, and the lifetime of the neutron is greater by far than that of any other unstable particle. Experimentally, though, there is a vast separation in lifetime between strong and electromagnetic decays (a factor of about 10 million), and again between electromagnetic and weak decays (a factor of at least a thousand). Indeed, particle physicists are so used to thinking in terms of 10^{-23} sec as the 'normal' unit of time that the handbooks generally classify anything with a lifetime greater than 10^{-17} sec or so as a 'stable' particle!†

Now, what about the conservation laws which, as I say, permit certain reactions and forbid others? To begin with there are the purely kinematic conservation laws – conservation of energy and momentum (which we shall study in Chapter 3) and conservation of angular momentum (which comes in Chapter 4). The fact that a particle cannot spontaneously decay into particles heavier than itself is actually a consequence of conservation of energy (although it may seem so 'obvious' as to require no explanation at all). The kinematic conservation laws apply to *all* interactions – strong, electromagnetic, weak, and for that matter anything else that may come along in the future – since they derive from special relativity itself. However, our concern right now is with the *dynamical* conservation laws that follow from the structure of the fundamental vertices:



* There are exceptions: $\pi^+ \to \mu^+ + \nu_\mu$, for example, is shorter by a factor of 10^4 than $\pi^+ \to e^+ + \nu_e$, but such cases cry out for some special explanation.

† Incidentally, 10^{-23} sec is about the time it takes light to cross a proton (diameter $\sim 10^{-15}$ m). You obviously cannot determine the lifetime of such a particle with a stop-watch, or even by measuring the length of its track (as you did for the Ω^- in Problem 1.8(b)) – it doesn't move far enough to *leave* a track. Instead, you make a histogram of *mass* measurements, and invoke the uncertainty principle: $\Delta E \Delta t \geq \hbar/2$. Here $\Delta E = (\Delta m)c^2$, and $\Delta t = \tau$, so

$$\tau \geq \frac{\hbar}{2(\Delta m)c^2}$$

Thus the *spread in mass* is a measure of the particle's lifetime. (Technically it's only a *lower bound* on τ, but for such short-lived particles we are presumably right up against the uncertainty limit [11]).

Since all physical processes are obtained by sticking these together in elaborate combinations, anything that is conserved at each vertex must be conserved for the reactions as a whole. So, what do we have?

1. *Charge*: All three interactions, of course, conserve electric charge. In the case of the weak interactions, the lepton (or quark) that comes out may not have the same charge as the one that went in, but if so, the difference is carried away by the W.
2. *Color*: The electromagnetic and weak interactions do not affect color. At a strong vertex the quark color does change, but the difference is carried off by the gluon. (The direct gluon–gluon couplings also conserve color.) However, since naturally occurring particles are always colorless, the observable manifestation of color conservation is pretty trivial: zero in, zero out.
3. *Baryon number*: In all the primitive vertices, if a quark goes *in*, a quark comes *out*, so the total number of quarks present is a constant. In this arithmetic we count *anti*quarks as *negative*, so that, for example, at the vertex $q + \bar{q} \rightarrow g$ the quark number is zero before and zero after. Of course, we never see individual quarks, only baryons (with quark number 3), antibaryons (quark number -3), and mesons (quark number zero). So, in practice, it is more convenient to speak of the conservation of baryon number (1 for baryons, -1 for antibaryons, and 0 for everything else). The baryon number is just $\frac{1}{3}$ the quark number. Notice that there is no analogous conservation of *meson number*; since mesons carry zero quark number, a given collision or decay can produce as many mesons as it likes, consistent with conservation of energy.
4. *Lepton number*: The strong forces do not touch leptons at all; in an electromagnetic interaction the same particle comes out (accompanied by a photon) as went in; and in the weak interactions if a lepton goes in, a lepton comes out (not necessarily the *same* one, this time). So, lepton number is absolutely conserved. Until recently there appeared to be no cross-generation mixing among the leptons, so electron number, muon number, and tau number were all separately conserved. This remains true in most cases, but neutrino oscillations indicate that it is not absolute.*

* There would be a similar conservation of generation type for quarks (upness-plus-downness, strangeness-plus-charm, and beauty-plus-truth), but here the intergenerational mixing has been obvious for decades. Still, because the off-diagonal elements in the KM matrix are relatively small, cross-generational decays tend to be suppressed, and processes that require *two* such crossings are extremely rare – hence an old rule that 'forbids' decays with $\Delta S = 2$.

5. *Flavor*: What about quark *flavor?* Flavor is conserved at a strong or electromagnetic vertex, but *not* at a weak vertex, where an up-quark may turn into a down quark or a strange quark, with nothing at all picking up the lost upness or supplying the 'gained' downness or strangeness. Because the weak forces *are* so weak, we say that the various flavors are *approximately* conserved. In fact, as you may remember, it was precisely this approximate conservation that led Gell-Mann to introduce the notion of strangeness in the first place. He 'explained' the fact that strange particles are always produced in pairs:

$$\pi^-(d\bar{u}) + p^+(uud) \rightarrow K^+(u\bar{s}) + \Sigma^-(dds) \tag{2.7}$$

for instance, but

$$\pi^-(d\bar{u}) + p^+(uud) \not\rightarrow \pi^+(u\bar{d}) + \Sigma^-(dds) \tag{2.8}$$

by arguing that the latter violates conservation of strangeness. (Actually, this *is* a possible *weak* interaction, but it will never be seen in the laboratory, because it must compete against enormously more probable strong processes that do conserve strangeness.) In *decays*, however, the nonconservation of strangeness is very conspicuous, because for many particles this is the only way they *can* decay; there *is* no competition from strong or electromagnetic processes. The Λ, for instance, is the lightest strange baryon; if it is to decay, it must go to n (or p) plus something. But the lightest strange meson is the K, and n (or p) plus K weighs substantially more than the Λ. If the Λ decays at all (and it *does* as we know: $\Lambda \rightarrow p^+ + \pi^-$ 64% of the time, and $\Lambda \rightarrow n + \pi^0$ 36% of the time), then strangeness cannot be conserved, and the reaction must proceed by the weak interaction. By contrast, the Δ^0 (with a strangeness of zero) can go to $p^+ + \pi^-$ or $n + \pi^0$ by the *strong* interaction, and its lifetime is accordingly much shorter.
6. *The OZI rule*: Finally, I must tell you about one very peculiar case that has been on my conscience since Chapter 1. I have in mind the decay of the ψ meson, which, you will recall, is a bound state of the charmed quark and its antiquark: $\psi = c\bar{c}$. The ψ has an anomalously long lifetime ($\sim 10^{-20}$ sec); the question is, why? It has nothing to do with

conservation of charm; the net charm of the ψ is zero. The ψ lifetime is short enough so that the decay is clearly due to the strong interactions. But why is it a thousand times slower than a strong decay 'ought' to be? The explanation (if you can call it that) goes back to an old observation by Okubo, Zweig, and Iizuka, known as the 'OZI rule'. These authors [12] were puzzled by the fact that the ϕ meson (whose quark content, $s\bar{s}$, makes it the strange analog to the ψ) decays much more often into two K's than into three π's (the *two* pion decay is forbidden for other reasons, which we will come to in Chapter 4), in spite of the fact that the three-pion decay is energetically favored (the mass of two K's is 990 MeV/c^2; three π's weigh only 415 MeV/c^2). In Figure 2.4, we see that the three-pion diagram can be cut in two by *snipping only gluon lines.*
The OZI rule states that such processes are 'suppressed'. Not absolutely forbidden, mind you, for the decay $\phi \to 3\pi$ *does* in fact occur, but far *less likely* than one would otherwise have supposed. The OZI rule is related to asymptotic freedom, in the following sense: in an OZI-suppressed diagram the gluons must be 'hard' (high energy), since they carry the energy necessary to make the hadrons into which they fragment. But asymptotic freedom says that gluons couple weakly at high energies (short ranges). By contrast, in OZI-allowed processes the gluons are typically 'soft' (low energy), and in this regime the coupling is strong. Qualitatively, at

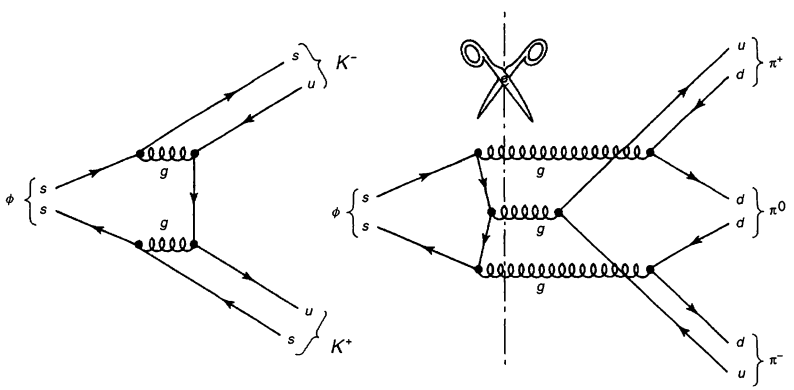


Fig. 2.4 The OZI rule: if the diagram can be cut in two by slicing only gluon lines (and not cutting any external lines), the process is suppressed.

least, this accounts for the OZI rule. (The quantitative details will have to await a more complete understanding of QCD.) But what does all this have to do with the ψ? Well, presumably the same rule applies, suppressing $\psi \to 3\pi$, and leaving the decay into two charmed D mesons (analogs to the K, but with the charmed quarks in place of the strange quarks) as the favored route. Only there's a new twist in the ψ system, for the D's turn out to be too heavy: a pair of D's weighs more than the ψ. So the decay $\psi \to D^+ + D^-$ (or $D^0 + \overline{D}^0$) is *kinematically* forbidden, while $\psi \to 3\pi$ is *OZI suppressed*, and it is to this happy combination that the ψ owes its unusual longevity.

2.6 Unification Schemes

At one time, electricity and magnetism were two distinct subjects, the one dealing with pith balls, batteries, and lightning; the other with lodestones, bar magnets, and the North Pole. But in 1820 Oersted noticed that an electric current could deflect a magnetic compass needle, and 10 years later Faraday discovered that a moving magnet could generate an electric current in a nearby loop of wire. By the time Maxwell put the whole theory together in its final form, electricity and magnetism were properly regarded as two aspects of a single subject: electromagnetism.

Einstein dreamed of going a step further, combining gravity with electrodynamics in a single *unified field theory*. Although this program was not successful, a similar vision inspired Glashow, Weinberg, and Salam to join the *weak* and electromagnetic forces. Their theory starts out with four massless mediators, but, as it develops, three of them acquire mass (by the so-called *Higgs mechanism*), becoming the W's and the Z, while one remains massless: the photon. Although experimentally a reaction mediated by W or Z is quite different from one mediated by the γ, they are both manifestations of a single *electroweak* interaction. The relative *weakness* of the weak force is attributable to the enormous mass of the intermediate vector bosons; its *intrinsic* strength is in fact somewhat *greater* than that of the electromagnetic force, as we shall see in Chapter 9.

Beginning in the early 1970s, many people have worked on the obvious next step: combining the strong force (chromodynamics) with the electroweak force (GWS). Several different schemes for implementing this *grand unification* are now on the table, and although it is too soon to draw any definitive conclusions, the basic idea is widely accepted. You will recall that the strong coupling constant α_s *decreases* at short distances (which is to say, for very high-energy collisions). So too does the weak coupling α_w, but at a slower rate. Meanwhile, the electromagnetic coupling constant, α_e, which is the smallest of the three, *increases*. Could it be that they all converge to a common limiting value, at extremely high energy (Figure 2.5)? Such

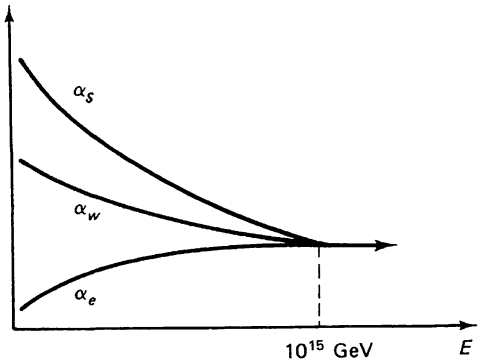


Fig. 2.5 Evolution of the three fundamental coupling constants.

is the promise of the grand unified theories (GUTs). Indeed, from the functional form of the running coupling constants it is possible to estimate the energy at which this unification occurs: around 10^{15} GeV. This is, of course, astronomically higher than any currently accessible energy (remember, the mass of the Z is $90\,\mathrm{GeV}/c^2$). Nevertheless, it is an exciting idea, for it means that the observed difference in strength among the three interactions is an 'accident' resulting from the fact that we are obliged to work at low energies, where the unity of the forces is obscured. If we could just get in close enough to see the 'true' strong, electric, and weak charges, without any of the screening effects of vacuum polarization, we would find that they are all equal. How nice!

Another prediction of the GUTs is that the proton is unstable, although its half-life is fantastically long (at least 10^{19} times the age of the universe). It has often been remarked that conservation of charge and color are in a sense more 'fundamental' than the conservation of baryon number and lepton number, because charge is the 'source' for electrodynamics, and color for chromodynamics. If these quantities were not conserved, QED and QCD would have to be completely reformulated. But baryon number and lepton number do not function as sources for any interaction, and their conservation has no deep dynamical significance. In the grand unified theories new interactions are contemplated, permitting decays such as

$$p^+ \to e^+ + \pi^0 \quad \text{or} \quad p^+ \to \bar{\nu}_\mu + \pi^+ \tag{2.9}$$

in which baryon number and lepton number change. Several major experiments have searched for these rare proton decays, but so far the results are negative [13].

If grand unification works, all of elementary particle physics will be reduced to the action of a single force. The final step, then, will be to bring in gravity, vindicating at last Einstein's dream, with the ultimate unification. At this point superstring theory is the most promising approach.* Stay tuned!

* See Section 12.2 for more on grand unification, and Section 12.4 for supersymmetry and superstrings.

References

1 Consistent etymology would call for *geusidynamics*, from the Greek word for 'flavor'; see Gaillard, M. (April 1981) *Physics Today*, 74. M. Gaillard suggests *asthenodynamics*, from the Greek word for *weak*.

2 Gross, Politzer, and Wilcek won the 2004 Nobel Prize for the discovery of asymptotic freedom. For an accessible account see Gross, D. J. (January 1987) *Physics Today*, 39.

3 See, for example, Purcell, E. M. (1985) *Electricity and Magnetism*, 2nd edn, McGraw-Hill, New York, Sec. 10.1.

4 Quigg, C. (April 1985) *Scientific American*, 84, gives a qualitative interpretation of the effect of gluon polarization.

5 For a status report, see S. Aronson and T. Ludlam (2005) Hunting the Quark Gluon Plasma, BNL-73847 Brookhaven National Laboratory, Brookhaven.

6 The classic papers on weak interaction theory up to 1960 are collected in Kabir, P. K. (ed) (1963) *The Development of Weak Interaction Theory*, Gordon & Breach, New York. A similar collection covering the modern era is contained in (a) Lai, C. H. (ed) (1981) *Gauge Theory of Weak and Electromagnetic Interactions*, World Scientific, Singapore. See also (b) Commins, E. D. and Bucksbaum, P. H. (1983) *Weak Interactions of Leptons and Quarks*, Cambridge University Press, Cambridge, UK.

7 Hasert, F. J. *et al.* (1973) *Physics Letters*, **46B**, 138; and *Nuclear Physics*, **B73**, 1. See also (a) Cline, D. B., Mann, A. K. and Rubbia, C. (December 1974) *Scientific American*, 108.

8 Wu, S. L. (1984) *Physics Reports*, **107**, 59), Section 5.6. See also (a) Bouchiat, M. -A. and Pottier, L. (June 1984) *Scientific American*, 100.

9 See Levi, B. G. (April 1997) *Physics Today*, 17.

10 The numbers are from the *Particle Physics Booklet*, (2006).

11 See Gillespie, D. T. (1973) *A Quantum Mechanics Primer*, International Textbook Co., London, p. 78, for a careful justification of this procedure.

12 Okubo, S. (1963) *Physics Letters*, **5**, 165; (a)Zweig, G. (1964) CERN Preprints TH 401 and TH 412; (b)Iizuka, J. (1966) *Progress in Physics Suppl.*, **37**, 21.

13 Weinberg, S. (June 1981) *Scientific American*, 64; (a) LoSecco, J. M., Reines, F. and Sinclair, D. (June 1985) *Scientific American*, 54. The best current limits on the proton lifetime come from Super-Kamiokande; see (b)Shiozawa, M. *et al.* (1998) *Physical Review Letters*, **81**, 3319.

Problems

2.1 Calculate the ratio of the gravitational attraction to the electrical repulsion between two stationary electrons. (Do I need to tell you how far apart they are?)

2.2 Sketch the lowest-order Feynman diagram representing *Delbruck scattering*: $\gamma + \gamma \to \gamma + \gamma$. (This process, the scattering of light by light, has no analog in classical electrodynamics.)

2.3 Draw all the fourth-order (four vertex) diagrams for Compton scattering. (There are 17 of them; disconnected diagrams don't count.)

2.4 Determine the mass of the virtual photon in each of the lowest-order diagrams for Bhabha scattering (assume the electron and positron are at rest). What is its velocity? (Note that these answers would be impossible for *real* photons.)

2.5 (a) Which decay do you think would be more likely,

$$\Xi^- \to \Lambda + \pi^- \quad \text{or} \quad \Xi^- \to n + \pi^-$$

Explain your answer, and confirm it by looking up the experimental data.

(b) Which decay of the $D^0(c\bar{u})$ meson is most likely,

$$D^0 \to K^- + \pi^+, \quad D^0 \to \pi^- + \pi^+, \quad \text{or} \quad D^0 \to K^+ + \pi^-$$

Which is *least* likely? Draw the Feynman diagrams, explain your answer and check the experimental data. (One of the successful predictions of the Cabibbo/GIM/KM model was that charmed mesons should decay preferentially into strange mesons, even though *energetically* the 2π mode is favored.)

(c) How about the 'beautiful' (B) mesons? Should they go to the D's, K's, or π's?

2.6 Draw all the lowest-order diagrams contributing to the process $e^+ + e^- \to W^+ + W^-$. (One of them involves the direct coupling of Z to W's and another the coupling of γ to W's, so when LEP (the electron–positron collider at CERN) achieved sufficient energy to make two W's, in 1996, these exotic processes could be studied experimentally. See B. Schwarzschild, *Physics Today* (September 1996), p. 21.)

2.7 Examine the following processes, and state for each one whether it is *possible* or *impossible*, according to the Standard Model (which does not include GUTs, with their potential violation of the conservation of lepton number and baryon number). In the former case, state which interaction is responsible – strong, electromagnetic, or weak; in the latter case, cite a conservation law that prevents it from occurring.* (Following the usual custom, I will not indicate the charge when it is unambiguous, thus γ, Λ, and n are neutral; p is positive, e is negative; etc.)

(a) $p + \bar{p} \to \pi^+ + \pi^0$
(b) $\eta \to \gamma + \gamma$
(c) $\Sigma^0 \to \Lambda + \pi^0$
(d) $\Sigma^- \to n + \pi^-$
(e) $e^+ + e^- \to \mu^+ + \mu^-$
(f) $\mu^- \to e^- + \bar{\nu}_e$
(g) $\Delta^+ \to p + \pi^0$
(h) $\bar{\nu}_e + p \to n + e^+$
(i) $e + p \to \nu_e + \pi^0$
(j) $p + p \to \Sigma^+ + n + K^0 + \pi^+ + \pi^0$
(k) $p \to e^+ + \gamma$
(l) $p + p \to p + p + p + \bar{p}$
(m) $n + \bar{n} \to \pi^+ + \pi^- + \pi^0$
(n) $\pi^+ + n \to \pi^- + p$
(o) $K^- \to \pi^- + \pi^0$
(p) $\Sigma^+ + n \to \Sigma^- + p$
(q) $\Sigma^0 \to \Lambda + \gamma$
(r) $\Xi^- \to \Lambda + \pi^-$
(s) $\Xi^0 \to p + \pi^-$
(t) $\pi^- + p \to \Lambda + K^0$
(u) $\pi^0 \to \gamma + \gamma$
(v) $\Sigma^- \to n + e + \bar{\nu}_e$

2.8 Some decays involve two (or even all three) different forces. Draw possible Feynman diagrams for the following processes:

(a) $\mu \to e + e + e^+ + \nu_\mu + \bar{\nu}_e$

(b) $\Sigma^+ \to p + \gamma$

What interactions are involved? (Both these decays have been observed, by the way.)

* *Note:* A *collision* is never *kinematically* forbidden. If you claim, for example, that reaction (e) is forbidden by conservation of energy (because the electron weighs less than the muon), you are at least half wrong – it can (and does) occur, as long as the electrons have enough kinetic energy to make up the difference. But don't try to play this game for *decays* – a single particle cannot decay into heavier secondaries no matter *what* its kinetic energy is, as you can easily see by examining the process in the rest frame of the decaying particle.

2.9 The Υ meson, $b\bar{b}$, is the bottom-quark analog to the ψ, $c\bar{c}$. Its mass is 9460 MeV/c^2, and its lifetime is 1.5×10^{-20} sec. From this information, what can you say about the mass of the B meson, $u\bar{b}$? (The observed mass is 5280 MeV/c^2.)

2.10 The ψ' meson, at 3686 MeV/c^2, has the same quark content as the ψ (viz. $c\bar{c}$). Its principal decay mode is $\psi' \to \psi + \pi^+ + \pi^-$. Is this a strong interaction? Is it OZI suppressed? What lifetime would you expect for the ψ'? (The observed value is 3×10^{-21} sec.)

2.11 Figure 1.9 shows the first confirmed production of an Ω^-, in a hydrogen bubble chamber. The incident K^- evidently hit a stationary particle X, producing a K^0, a K^+, and the Ω^-. (a) What was the charge of the X? What was its strangeness? What particle do you suppose it was? (b) Follow each line in the right-hand diagram, listing every reaction as you go along; also specify what kind of interaction – (strong, electromagnetic, or weak – was responsible. (In case the diagram is unclear, the two photons are supposed to come from the same point. Incidentally, while $\gamma \to e^- + e^+$ is impossible in vacuum (it doesn't conserve momentum), it *does* occur in the vicinity of a nucleus – the nucleus soaks up the 'missing' momentum. The reaction is really $\gamma + p \to e^- + e^+ + p$, but the p leaves no track, because it is so heavy that it scarcely moves; the electron and positron carry off the photon's energy, and the proton simply acts as a passive momentum 'sink'.)

2.12 The W^- was discovered in 1983 at CERN, using proton/antiproton scattering:

$$p + \bar{p} \to W^- + X$$

where X represents one or more particles. What is the most likely X, for this process? Draw a Feynman diagram for your reaction, and explain why your X is more probable than the various alternatives.

3

Relativistic Kinematics

In this chapter, I summarize the basic principles, notation, and terminology of relativistic kinematics. This is material you must know cold *in order to understand Chapters 6 through 10 (it is not needed for Chapters 4 and 5, however, and if you prefer you can read them first). Although the treatment is reasonably self-contained, I do assume that you have encountered special relativity before – if not, you should pause here and read the appropriate chapter in any introductory physics text before proceeding. If you are already quite familiar with relativity, this chapter will be an easy review – but read through it anyway because some of the notation may be new to you.*

3.1
Lorentz Transformations

According to the special theory of relativity [1], the laws of physics apply just as well in a reference system moving at constant velocity as they do in one at rest. An embarrassing implication of this is that there's no way of telling which system (if any) *is* at rest, and hence there is no way of knowing what 'the' velocity of any other system might be. So perhaps I had better start over. *Ahem.*

According to the special theory of relativity [1], the laws of physics are equally valid in all *inertial* reference systems. An inertial system is one in which Newton's first law (the law of inertia) is obeyed: objects keep moving in straight lines at constant speeds unless acted upon by some force.* It's easy to see that any two inertial systems must be moving at constant velocity with respect to one another, and conversely, that any system moving at constant velocity with respect to an inertial system is itself inertial.

Imagine, then, that we have two inertial frames, S and S', with S' moving at uniform velocity $\mathbf{v}$ (magnitude v) with respect to S (S, then, is moving at velocity $-\mathbf{v}$ with respect to S'). We may as well lay out our coordinates in such a way that the motion is along the common x/x' axis (Figure 3.1), and set the master clocks at the origin in each system so that both read zero at the instant the two coincide (that is, $t = t' = 0$ when $x = x' = 0$). Suppose, now, that some event occurs at position

* If you are wondering whether a freely falling system in a uniform gravitational field is 'inertial', you know more than is good for you. Let's just keep gravity out of it.

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

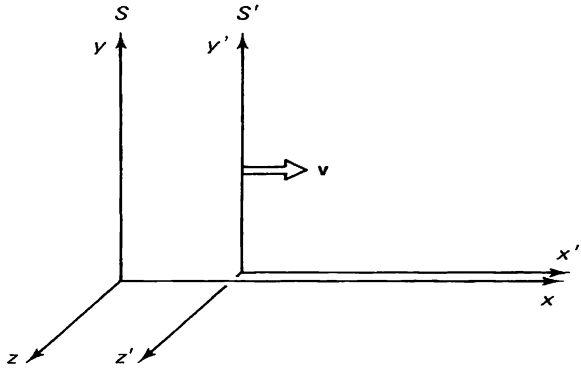


Fig. 3.1 The inertial systems S and S'.

(x, y, z) and time t in S. *Question*: What are the space-time coordinates (x', y', z') and t' of this *same event* in S'? The answer is provided by the Lorentz transformations:

$$\begin{aligned} &\text{i. } x' = \gamma(x - vt) \\ &\text{ii. } y' = y \\ &\text{iii. } z' = z \\ &\text{iv. } t' = \gamma\left(t - \frac{v}{c^2}x\right) \end{aligned} \tag{3.1}$$

where

$$\gamma \equiv \frac{1}{\sqrt{1 - v^2/c^2}} \tag{3.2}$$

The inverse transformations, which take us back from S' to S, are obtained by simply changing the sign of v (see Problem 3.1):

$$\begin{aligned} &\text{i}'. \ x = \gamma(x' + vt') \\ &\text{ii}'. \ y = y' \\ &\text{iii}'. \ z = z' \\ &\text{iv}'. \ t = \gamma\left(t' + \frac{v}{c^2}x'\right) \end{aligned} \tag{3.3}$$

The Lorentz transformations have a number of immediate consequences, of which I mention briefly the most important:

1. *The relativity of simultaneity:* If two events occur at the same time in S, but at different locations, then they do *not* occur at the same time in S'. Specifically, if $t_A = t_B$, then

$$t'_A = t'_B + \frac{\gamma v}{c^2}(x_B - x_A) \tag{3.4}$$

(see Problem 3.2). Events that are simultaneous in one inertial system, then, are not simultaneous in others.

2. *Lorentz contraction:* Suppose a stick lies on the x' axis, at rest in S'. Say one end is at the origin ($x' = 0$) and the other is at L' (so its length in S' is L'). What is its length as measured in S? Since the stick is *moving* with respect to S, we must be careful to record the positions of its two ends at the same instant, say $t = 0$. At that moment, the left end is at $x = 0$ and the right end, according to Equation (i), is at $x = L'/\gamma$. Thus the length of the stick is $L = L'/\gamma$, in S. Notice that γ is always greater than or equal to 1. It follows that a *moving object is shortened* by a factor of γ, as compared with its length in the system in which it is at rest. Notice that Lorentz contraction only applies to lengths *along the direction of motion*; perpendicular dimensions are not affected.
3. *Time dilation:* Suppose the clock at the origin in S' ticks off an interval T'; for simplicity, say it runs from $t' = 0$ to $t' = T'$. How long is this period as measured in S? Well, it begins when $t = 0$, and it ends when $t' = T'$ at $x' = 0$, so (according to Equation (iv′)) $t = \gamma T'$. Evidently the clocks in S tick off a *longer* interval, $T = \gamma T'$, by that same factor of γ; or, put it the other way around: *moving clocks run slow.*

   Unlike Lorentz contraction, which is only indirectly relevant to elementary particle physics, time dilation is a commonplace in the laboratory. For, in a sense, every unstable particle has a built-in clock: whatever it is that tells the particle when its time is up. And these internal clocks do indeed run slow when the particle is moving. That is to say, a moving particle lasts longer (by a factor of γ) than it would at rest.* (The tabulated lifetimes are, of course, for particles at rest.) In fact, the cosmic ray muons produced in the upper atmosphere would never make it to ground level were it not for time dilation (see Problem 3.4).
4. *Velocity addition:* Suppose a particle is moving in the x direction at speed u', with respect to S'. What is its speed, u, with respect to S? Well, it travels a distance $\Delta x = \gamma(\Delta x' + v\,\Delta t')$ in a time $\Delta t = \gamma[\Delta t' + (v/c^2)\Delta x']$, so

$$\frac{\Delta x}{\Delta t} = \frac{\Delta x' + v\,\Delta t'}{\Delta t' + (v/c^2)\Delta x'} = \frac{(\Delta x'/\Delta t') + v}{1 + (v/c^2)(\Delta x'/\Delta t')}.$$

* Actually, the disintegration of an individual particle is a random process; when we speak of a 'lifetime' we really mean the *average* lifetime of that particle type. When I say that a moving particle lasts longer, I really mean that the *average* lifetime of a *group* of moving particles is longer.

But $\Delta x/\Delta t = u$, and $\Delta x'/\Delta t' = u'$, so

$$u = \frac{u' + v}{1 + (u'v/c^2)} \tag{3.5}$$

Notice that if $u' = c$, then $u = c$ also: *the speed of light is the same in all inertial systems.*

It can sometimes be confusing to figure out in a particular context, which numbers should be primed and what signs attach to the velocities, so I personally remember three rules: moving sticks are short (by a factor of γ), moving clocks are slow (by a factor of γ) – so put the γ (which, remember, is *greater* than 1) on whichever side of the equation you need to achieve these results, – and

$$v_{AC} = \frac{v_{AB} + v_{BC}}{1 + (v_{AB}v_{BC}/c^2)} \tag{3.6}$$

where v_{AB} (for instance) means the velocity of A with respect to B. The numerator is the classical result (the so-called 'Galilean velocity addition rule'); the denominator is Einstein's correction – it is very close to 1 unless the velocities are close to c.

3.2 Four-vectors

It is convenient at this point to introduce some simplifying notation. We define the *position-time four-vector* x^μ, $\mu = 0, 1, 2, 3$, as follows:

$$x^0 = ct, \qquad x^1 = x, \qquad x^2 = y, \qquad x^3 = z \tag{3.7}$$

In terms of x^μ, the Lorentz transformations take on a more symmetrical appearance:

$$\begin{aligned} x^{0'} &= \gamma(x^0 - \beta x^1) \\ x^{1'} &= \gamma(x^1 - \beta x^0) \\ x^{2'} &= x^2 \\ x^{3'} &= x^3 \end{aligned} \tag{3.8}$$

where

$$\beta \equiv \frac{v}{c} \tag{3.9}$$

More compactly:

$$x^{\mu'} = \sum_{\nu=0}^{3} \Lambda^\mu_\nu x^\nu \qquad (\mu = 0, 1, 2, 3) \tag{3.10}$$

The coefficients Λ^{μ}_{ν} may be regarded as the elements of a matrix Λ:

$$\Lambda = \begin{bmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.11}$$

(i.e. $\Lambda^0_0 = \Lambda^1_1 = \gamma$; $\Lambda^1_0 = \Lambda^0_1 = -\gamma\beta$; $\Lambda^2_2 = \Lambda^3_3 = 1$; and all the rest are zero). To avoid writing lots of Σ's, we shall follow Einstein's 'summation convention', which says that repeated Greek indices (one as subscript, one as superscript) are to be summed from 0 to 3. Thus Equation 3.10 becomes, finally,*

$$x^{\mu'} = \Lambda^{\mu}_{\nu} x^{\nu} \tag{3.12}$$

A special virtue of this tidy notation is that the same form describes Lorentz transformations that are *not* along the x direction; in fact, the S and S' axes need not even be parallel; the Λ matrix is more complicated, naturally, but Equation 3.12 still holds. (On the other hand, there is no real loss of generality in using Equation 3.11, since we are always free to *choose* parallel axes, and to align the x axis along the direction of v.)

Although the individual coordinates of an event change, in accordance with Equation 3.12, when we go from S to S', there is a particular *combination* of them that remains the same (Problem 3.8):

$$I \equiv (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = (x^{0'})^2 - (x^{1'})^2 - (x^{2'})^2 - (x^{3'})^2 \tag{3.13}$$

Such a quantity, which has the same value in any inertial system, is called an *invariant*. (In the same sense, the quantity $r^2 = x^2 + y^2 + z^2$ is invariant under *rotations*.) Now, I would *like* to write this invariant in the form of a sum: $\Sigma^3_{\mu=0} x^{\mu} x^{\mu}$, but unfortunately there are those three irritating minus signs. To keep track of them, we introduce the *metric*, $g_{\mu\nu}$, whose components can be displayed as a matrix $\mathbf{g}$:

$$\mathbf{g} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{3.14}$$

* In an expression such as this the Greek letter used for the summation index, ν, is of course completely arbitrary. The same goes for the 'hanging' index μ, although it must match on the two sides of the equation. Thus Equation (3.12) could just as well be written $x^{\kappa'} = \Lambda^{\kappa}_{\lambda} x^{\lambda}$. Either expression stands for the set of four equations:

$$x^{0'} = \Lambda^0_0 x^0 + \Lambda^0_1 x^1 + \Lambda^0_2 x^2 + \Lambda^0_3 x^3$$
$$x^{1'} = \Lambda^1_0 x^0 + \Lambda^1_1 x^1 + \Lambda^1_2 x^2 + \Lambda^1_3 x^3$$
$$x^{2'} = \Lambda^2_0 x^0 + \Lambda^2_1 x^1 + \Lambda^2_2 x^2 + \Lambda^2_3 x^3$$
$$x^{3'} = \Lambda^3_0 x^0 + \Lambda^3_1 x^1 + \Lambda^3_2 x^2 + \Lambda^3_3 x^3$$

(i.e. $g_{00} = 1$; $g_{11} = g_{22} = g_{33} = -1$; all the rest are zero).* With the help of $g_{\mu\nu}$, the invariant I can be written as a double sum:

$$I = \sum_{\mu=0}^{3}\sum_{\nu=0}^{3} g_{\mu\nu}x^{\mu}x^{\nu} = g_{\mu\nu}x^{\mu}x^{\nu} \tag{3.15}$$

Carrying things a step further, we define the *covariant* four-vector x_{μ} (index *down*) as follows:

$$x_{\mu} \equiv g_{\mu\nu}x^{\nu} \tag{3.16}$$

(i.e. $x_0 = x^0$, $x_1 = -x^1$, $x_2 = -x^2$, $x_3 = -x^3$). To emphasize the distinction we call the 'original' four-vector x^{μ} (index *up*) a *contravariant* four-vector. The invariant I can then be written in its cleanest form:

$$I = x_{\mu}x^{\mu} \tag{3.17}$$

(or, equivalently, $x^{\mu}x_{\mu}$). All this will no doubt seem like monstrous notational overkill, just to keep track of three pesky minus signs, but it's actually very simple, once you get used to it. (What's more, it generalizes nicely to non-Cartesian coordinate systems and to the curved spaces encountered in general relativity, though neither of these is relevant to us here.)

The position-time four-vector x^{μ} is the archetype for all *four-vectors*. We define a four-vector, a^{μ}, as a four-component object that transforms in the same way x^{μ} does when we go from one inertial system to another, to wit:

$$a^{\mu'} = \Lambda^{\mu}_{\nu}a^{\nu} \tag{3.18}$$

with the same coefficients Λ^{μ}_{ν}. To each such (contravariant) four-vector we associate a *co*variant four-vector a_{μ}, obtained by simply changing the signs of the spatial components, or, more formally

$$a_{\mu} = g_{\mu\nu}a^{\nu} \tag{3.19}$$

Of course, we can go back from covariant to contravariant by reversing the signs again:

$$a^{\mu} = g^{\mu\nu}a_{\nu} \tag{3.20}$$

where $g^{\mu\nu}$ are technically the elements in the matrix $\mathbf{g}^{-1}$ (however, since our metric is its own inverse, $g^{\mu\nu}$ is the same as $g_{\mu\nu}$). Given any two four-vectors, a^{μ} and b^{μ}, the quantity

$$a^{\mu}b_{\mu} = a_{\mu}b^{\mu} = a^0b^0 - a^1b^1 - a^2b^2 - a^3b^3 \tag{3.21}$$

* I should warn you that some physicists define the metric with the opposite signs $(-1, 1, 1, 1)$. It doesn't *matter* much – if I is an invariant, so too is $-I$. But it does mean that you must be on the lookout for unfamiliar signs. Fortunately, most *particle* physicists nowadays use the convention in Equation 3.14.

is invariant (the same number in any inertial system). We shall refer to it as the *scalar product* of a and b; it is the four-dimensional analog to the dot product of two three-vectors (there is no four-vector analog to the cross product).*

If you get tired of writing indices, feel free to use the dot notation:

$$a \cdot b \equiv a_\mu b^\mu \tag{3.22}$$

However, you will then need a way to distinguish this four-dimensional scalar product from the ordinary dot product of two three-vectors. The best way is to be scrupulously careful to put an arrow over all three-vectors (except perhaps the velocity, **v**, which, since it is not part of a four-vector, is not subject to ambiguity). In this book, I use boldface for three-vectors. Thus

$$a \cdot b = a^0 b^0 - \mathbf{a} \cdot \mathbf{b} \tag{3.23}$$

We also use the notation a^2 for the scalar product of a^μ with itself:†

$$a^2 \equiv a \cdot a = (a^0)^2 - \mathbf{a}^2 \tag{3.24}$$

Notice, however, that a^2 need not be positive. Indeed, we can classify all four-vectors according to the sign of a^2:

$$\begin{array}{ll} \text{If } a^2 > 0, & a^\mu \text{ is called } \textit{timelike} \\ \text{If } a^2 < 0, & a^\mu \text{ is called } \textit{spacelike} \\ \text{If } a^2 = 0, & a^\mu \text{ is called } \textit{lightlike} \end{array} \tag{3.25}$$

From *vectors* it is a short step to *tensors*: a second-rank tensor, $s^{\mu\nu}$, carries two indices, has $4^2 = 16$ components, and transforms with *two* factors of Λ:

$$s^{\mu\nu\prime} = \Lambda^\mu_\kappa \Lambda^\nu_\sigma s^{\kappa\sigma} \tag{3.26}$$

a third-rank tensor, $t^{\mu\nu\lambda}$, has three indices, $4^3 = 64$ components, and transforms with *three* factors of Λ:

$$t^{\mu\nu\lambda\prime} = \Lambda^\mu_\kappa \Lambda^\nu_\sigma \Lambda^\lambda_\tau t^{\kappa\sigma\tau} \tag{3.27}$$

* The closest thing is $(a^\mu b^\nu - a^\nu b^\mu)$, but this is a second-rank *tensor*, not a four-vector (see below).

† On its face, this is dangerously ambiguous notation, since a^2 could also be the second spatial component of a^μ. But in practice we so seldom refer to individual components that this causes no problems (if you really *mean* the component, better say so explicitly). More serious is the potential confusion between a^2 and the square of the magnitude of the three-vector *part* of a^μ. I personally write the latter in bold face, to avoid any possible misunderstanding: $\mathbf{a}^2 \equiv \mathbf{a} \cdot \mathbf{a}$. This is not standard notation, however, and if you prefer some other device, that's fine. But I do urge you to find a clear way to distinguish a^2 from $\mathbf{a}^2$, or you are asking for real trouble down the road.

and so on. In this hierarchy, a vector is a tensor of rank one, and a scalar (invariant) is a tensor of rank zero. We construct covariant and 'mixed' tensors by lowering indices (at the cost of a minus sign for each spatial index), for example

$$s^{\mu}{}_{\nu} = g_{\nu\lambda} s^{\mu\lambda}; \quad s_{\mu\nu} = g_{\mu\kappa} g_{\nu\lambda} s^{\kappa\lambda} \tag{3.28}$$

and so on. Notice that the product of two tensors is itself a tensor: $(a^{\mu} b^{\nu})$ is a tensor of second rank; $(a^{\mu} t^{\nu\lambda\sigma})$ is a tensor of fourth rank; and so on. Finally, we can obtain from any tensor of rank $n+2$ a 'contracted' tensor of rank n, by summing like upper and lower indices. Thus $s^{\mu}{}_{\mu}$ is a scalar; $t_{\nu}{}^{\mu\nu}$ is a vector; $a_{\mu} t^{\mu\nu\lambda}$ is a second-rank tensor.

3.3 Energy and Momentum

Suppose you're driving down the highway, and pretend for the sake of argument that you're going at close to the speed of light. You might want to keep an eye on two different 'times': if you're worried about making an appointment in San Francisco, you should check the stationary clocks posted now and then along the side of the road. But if you're wondering when would be an appropriate time to stop for a bite to eat, it would be more sensible to look at the watch on your wrist. For, according to relativity, the moving clock (in this case, your watch) is running slow (relative to the 'stationary' clocks on the ground), and so too is your heart rate, your metabolism, your speech and thought, *every*thing. Specifically, while the 'ground' time advances by an infinitesimal amount dt, your own (or *proper*) time advances by the smaller amount $d\tau$:

$$d\tau = \frac{dt}{\gamma} \tag{3.29}$$

At normal driving speeds, of course, γ is so close to 1 that dt and $d\tau$ are essentially identical, but in elementary particle physics the distinction between laboratory time (read off the clock on the wall) and particle time (as it would appear on the particle's watch) is crucial. Although we can always get from one to the other, using Equation 3.29, in practice it is usually most convenient to work with proper time, *because τ is invariant – all* observers can read the particle's watch, and at any given moment they must all agree on what it says, even though their *own* clocks may differ from it and from one another.

When we speak of the 'velocity' of a particle (with respect to the laboratory), we mean, of course, the distance it travels (measured in the lab frame) divided by the time it takes (measured on the lab clock):

$$\mathbf{v} = \frac{d\mathbf{x}}{dt} \tag{3.30}$$

But in view of what has just been said, it is also useful to introduce the *proper velocity*, η, which is the distance traveled (again, measured in the lab frame) divided

by the *proper* time:*

$$\boldsymbol{\eta} \equiv \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\tau} \tag{3.31}$$

According to Equation 3.29, the two velocities are related by a factor of γ:

$$\boldsymbol{\eta} = \gamma \mathbf{v} \tag{3.32}$$

However, $\boldsymbol{\eta}$ is much easier to work with, for if we want to go from the lab system, S, to a moving system, S', *both the numerator and the denominator in Equation 3.30 must be transformed* – leading to the cumbersome velocity addition rule Equation 3.5 – whereas in Equation 3.31 *only the numerator transforms*; $d\tau$, as we have seen, is invariant. In fact, proper velocity is part of a four-vector:

$$\eta^{\mu} = \frac{\mathrm{d}x^{\mu}}{\mathrm{d}\tau} \tag{3.33}$$

whose zeroth component is

$$\eta^{0} = \frac{\mathrm{d}x^{0}}{\mathrm{d}\tau} = \frac{\mathrm{d}(ct)}{(1/\gamma)\,\mathrm{d}t} = \gamma c \tag{3.34}$$

Thus

$$\eta^{\mu} = \gamma(c, v_x, v_y, v_z) \tag{3.35}$$

Incidentally, $\eta_{\mu}\eta^{\mu}$ should be invariant, and it *is*:

$$\eta_{\mu}\eta^{\mu} = \gamma^2(c^2 - v_x^2 - v_y^2 - v_z^2) = \gamma^2 c^2(1 - v^2/c^2) = c^2 \tag{3.36}$$

They don't make 'em more invariant than *that!*

Classically, momentum is mass times velocity. We would like to carry this over in relativity, but the question arises: which velocity should we use – *ordinary* velocity or *proper* velocity? Classical considerations offer no clue, for the two are equal in the nonrelativistic limit. In a sense, it's just a matter of definition, but there is a subtle and compelling reason why ordinary velocity would be a *bad* choice, whereas proper velocity is a *good* choice. The point is this: if we defined momentum as $m\mathbf{v}$, then the law of conservation of momentum would be inconsistent with the principle of relativity (if it held in one inertial system, it would *not* hold in other inertial

* Proper velocity is a hybrid quantity, in the sense that distance is measured in the *lab* frame, whereas time is measured in the *particle* frame. Some people object to the adjective 'proper' in this context, holding that this should be reserved for quantities measured entirely in the particle frame. Of course, in its *own* frame the particle never moves at all – its velocity is zero. If my terminology disturbs you, call η the 'four-velocity'. I should add that although proper velocity is the more convenient quantity to *calculate* with, ordinary velocity is still the more *natural* quantity from the point of view of an observer watching a particle fly past.

systems). But, if we define momentum as $m\boldsymbol{\eta}$, then conservation of momentum *is* consistent with the principle of relativity (if it holds in one inertial system, it automatically holds in all inertial systems). I'll let you prove this for yourself in Problem 3.12. Mind you, this doesn't guarantee that momentum *is* conserved – that's a matter for *experiments* to decide. But it *does* say that if we're hoping to extend momentum conservation to the relativistic domain, we had better *not* define momentum as $m\mathbf{v}$, whereas $m\boldsymbol{\eta}$ is perfectly acceptable.

That's a tricky argument, and if you didn't follow it, try reading that last paragraph again. The upshot is that in relativity, momentum is defined as mass times *proper* velocity:

$$\mathbf{p} \equiv m\boldsymbol{\eta} \tag{3.37}$$

Since proper velocity is part of a four-vector, the same goes for momentum:

$$p^\mu = m\eta^\mu \tag{3.38}$$

The spatial components of p^μ constitute the (relativistic) momentum three-vector:

$$\mathbf{p} = \gamma m\mathbf{v} = \frac{m\mathbf{v}}{\sqrt{1 - v^2/c^2}} \tag{3.39}$$

Meanwhile, the 'temporal' component is

$$p^0 = \gamma mc \tag{3.40}$$

For reasons that will appear in a moment, we define the *relativistic energy*, E, as

$$E \equiv \gamma mc^2 = \frac{mc^2}{\sqrt{1 - v^2/c^2}} \tag{3.41}$$

The zeroth component of p^μ, then, is E/c. Thus, energy and momentum together make up a four-vector – the *energy–momentum four-vector* (or *four-momentum*)

$$p^\mu = \left(\frac{E}{c}, p_x, p_y, p_z\right) \tag{3.42}$$

Incidentally, from Equations 3.36 and 3.38 we have

$$p_\mu p^\mu = \frac{E^2}{c^2} - \mathbf{p}^2 = m^2c^2 \tag{3.43}$$

which, again, is manifestly invariant.

The relativistic momentum (Equation 3.37) reduces to the classical expression in the nonrelativistic regime ($v \ll c$), but the same cannot be said for relativistic energy (Equation 3.41). To see how this quantity comes to be called 'energy,' we

expand the radical in a Taylor series:

$$E = mc^2\left(1 + \tfrac{1}{2}\frac{v^2}{c^2} + \frac{3}{8}\frac{v^4}{c^4} + \cdots\right) = mc^2 + \tfrac{1}{2}mv^2 + \frac{3}{8}m\frac{v^4}{c^2} + \cdots \tag{3.44}$$

Notice that the *second* term here corresponds to the classical kinetic energy, while the leading term (mc^2) is a constant. Now you may recall that in classical mechanics only *changes* in energy are physically significant – you can add a constant with impunity. In this sense, the relativistic formula *is* consistent with the classical one, in the limit $v \ll c$ where the higher terms in the expansion are negligible. The constant term, which survives even when $v = 0$, is called the *rest energy*;

$$R \equiv mc^2 \tag{3.45}$$

the remainder, which is energy attributable to the motion of the particle, is the *relativistic kinetic energy*:*

$$T \equiv mc^2(\gamma - 1) = \frac{1}{2}mv^2 + \frac{3}{8}m\frac{v^4}{c^2} + \cdots \tag{3.46}$$

In classical mechanics, there is no such thing as a massless particle; its momentum ($m\mathbf{v}$) would be zero, its kinetic energy ($\frac{1}{2}mv^2$) would be zero, it could sustain no force, since $F = ma$, and hence (by Newton's third law) it could not exert a force on anything *else* – it would be a dynamical ghost. At first glance you might suppose that the same would be true in relativity, but a careful inspection of the formulas

$$\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - v^2/c^2}}, \qquad E = \frac{mc^2}{\sqrt{1 - v^2/c^2}} \tag{3.47}$$

reveals a loophole: when $m = 0$, the numerators are zero, but if $v = c$, the denominators *also* vanish, and these equations are indeterminate (0/0). So it is just possible that *we could allow $m = 0$, provided the particle always travels at the speed of light.* In this case, Equation 3.47 will not serve to define E and $\mathbf{p}$; nevertheless, Equation 3.43 still holds:

$$v = c, \qquad E = |\mathbf{p}|c \qquad \text{(for massless particles)} \tag{3.48}$$

Personally, I would regard this 'argument' as a joke, were it not for the fact that massless particles (photons) are known to exist in nature, they *do* travel at the speed of light, and their energy and momentum *are* related by Equation 3.48. So we have

* Notice that I have never mentioned 'relativistic mass' in all this. It is a superfluous quantity that serves no useful function. In case you encounter it, the definition is $m_{rel} \equiv \gamma m$; it has died out because it differs from E only by a factor of c^2. Whatever can be said about m_{rel} could just as well be said about E. For instance, the 'conservation of relativistic mass' is nothing but conservation of *energy*, with a factor of c^2 divided out.

to take the loophole seriously. You may well ask: if Equation 3.47 doesn't define $\mathbf{p}$ and E, what *does* determine the momentum and energy of a massless particle? Not the mass (that's zero by assumption); not the speed (that's always c). How, then, *does* a photon with an energy of 2 eV differ from a photon with an energy of 3 eV? Relativity offers no answer to this question, but curiously enough *quantum* mechanics *does*, in the form of Planck's formula:

$$E = h\nu \tag{3.49}$$

It is the *frequency* of the photon that determines its energy and momentum: the 2-eV photon is *red*, and the 3-eV photon is *purple* !

3.4 Collisions

So far, relativistic energy and momentum are nothing but *definitions*; the *physics* resides in the empirical fact that these quantities are *conserved*. In relativity, as in classical mechanics, the cleanest application of the conservation laws is to *collisions*. Imagine first a classical collision, in which object A hits object B (perhaps they are both carts on an air table), producing objects C and D (Figure 3.2). Of course, C and D might be the same as A and B; but we may as well allow that some paint (or whatever) rubs off A onto B, so that the final masses are not the same as the original ones. (We *do* assume, however, that A, B, C, and D are the only actors in the drama; if some wreckage, W, is left at the scene, then we would be talking about a more complicated process: $A + B \rightarrow C + D + W$.) By its nature, a collision is something that happens so fast that no *external* force, such as gravity, or friction with the track, has an appreciable influence. Classically, mass and momentum are always conserved in such a process; kinetic energy may or may not be conserved.

3.4.1 Classical Collisions

1. Mass is conserved: $m_A + m_B = m_C + m_D$.
2. Momentum is conserved: $\mathbf{p}_A + \mathbf{p}_B = \mathbf{p}_C + \mathbf{p}_D$.
3. Kinetic energy may or may not be conserved.

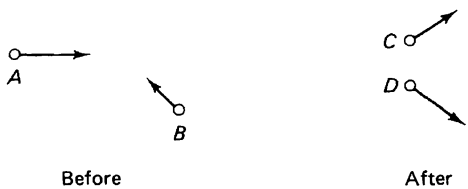


Fig. 3.2 A collision in which $A + B \rightarrow C + D$.

I like to distinguish three types of collisions: 'sticky' ones, in which the kinetic energy *decreases* (typically, it is converted into heat); 'explosive' ones, in which the kinetic energy *increases* (for example, suppose A has a compressed spring on its front bumper, and the catch is released in the course of the collision so that spring energy is converted into kinetic energy); and *elastic* ones, in which the kinetic energy is conserved.

(a) *Sticky* (kinetic energy decreases): $T_A + T_B > T_C + T_D$.
(b) *Explosive* (kinetic energy increases): $T_A + T_B < T_C + T_D$.
(c) *Elastic* (kinetic energy conserved): $T_A + T_B = T_C + T_D$.

In the extreme case of type (a), the two particles stick together, and there is really only *one* final object: $A + B \rightarrow C$. In the extreme case of type (b), a single object breaks in two: $A \rightarrow C + D$ (in the language of particle physics, A *decays* into $C + D$).

3.4.2 Relativistic Collisions

In a relativistic collision, *energy and momentum are always conserved.* In other words, all four components of the energy–momentum four-vector are conserved. As in the classical case, *kinetic* energy may or may not be conserved.

1. Energy is conserved: $E_A + E_B = E_C + E_D$.
2. Momentum is conserved: $\mathbf{p}_A + \mathbf{p}_B = \mathbf{p}_C + \mathbf{p}_D$.
3. Kinetic energy may or may not be conserved.

(The first two can be combined into a single expression: $p_A^\mu + p_B^\mu = p_C^\mu + p_D^\mu$.)

Again, we can classify collisions as sticky, explosive, or elastic, depending on whether the kinetic energy decreases, increases, or remains the same. Since the *total* energy (rest plus kinetic) is *always* conserved, it follows that rest energy (and hence also mass) increases in a sticky collision, decreases in an explosive collision, and is unchanged in an elastic collision.

(a) *Sticky* (kinetic energy decreases): rest energy and mass increase.
(b) *Explosive* (kinetic energy increases): rest energy and mass decrease.
(c) *Elastic* (kinetic energy is conserved): rest energy and mass are conserved.

Please note: *except in elastic collisions, mass is not conserved.** For example, in the decay $\pi^0 \rightarrow \gamma + \gamma$ the initial mass was 135 MeV/c^2, but the final mass is *zero.*

* In the old terminology, we would say that *relativistic* mass is conserved, but *rest* mass is not.

Here *rest* energy was converted into *kinetic* energy (or, in the absurd language of the popular press, infuriating to anyone with the slightest respect for dimensional consistency, 'mass was converted into energy'). Conversely, if mass *is* conserved, then the collision was elastic. In elementary particle physics, there is only one way this ever happens: the *same particles* come out as went in*– electron–proton scattering ($e + p \to e + p$), for example.

In spite of a certain structural similarity between the classical and relativistic analyses, there is a striking difference in the interpretation of inelastic collisions. In the classical case, we say that energy is converted from kinetic form to some 'internal' form (thermal energy, spring energy, etc.), or vice versa. In the relativistic analysis, we say that it goes from kinetic energy to *rest* energy or vice versa. How can these possibly be consistent? After all, relativistic mechanics is supposed to reduce to classical mechanics in the limit $v \ll c$. The answer is that all 'internal' forms of energy are reflected in the rest energy of an object. A hot potato *weighs more* than a cold potato; a compressed spring *weighs more* than a relaxed spring. On the macroscopic scale, rest energies are enormously greater than internal energies, so these mass differences are utterly negligible in everyday life, and very small even at the atomic level. Only in nuclear and particle physics are typical internal energies comparable to typical rest energies. Nevertheless, in principle, whenever you weigh an object, you are measuring not only the rest energies (masses) of its constituent parts, but all of their kinetic and interaction energies as well.

3.5 Examples and Applications

Solving problems in relativistic kinematics is as much an art as a science. Although the *physics* involved is minimal – nothing but conservation of energy and conservation of momentum – the *algebra* can be formidable. Whether a given problem takes two lines or seven pages depends a lot on how skillful and experienced you are at manipulating the tools and the tricks of the trade. I now propose to work a few examples, pointing out as I go along some of the labor-saving devices that are available to you [2].

Example 3.1 Two lumps of clay, each of mass m, collide head-on at $\frac{3}{5}c$ (Figure 3.3). They stick together. *Question*: What is the mass M of the final composite lump?

Solution: Conservation of energy says $E_1 + E_2 = E_M$. Conservation of momentum says $\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{p}_M$. In this case, conservation of momentum is trivial: $\mathbf{p}_1 = -\mathbf{p}_2$, so the final lump is at rest (which was obvious from the start). The initial energies are

* In principle, if there existed two distinct pairs of particles (A, B and C, D) that happened to add up to the same total mass, then I suppose the reaction $A + B \to C + D$ might be considered 'elastic', but in reality there *are* no such coincidences, so to a particle physicist the word 'elastic' has come to *mean* that the same particles come out as went in.

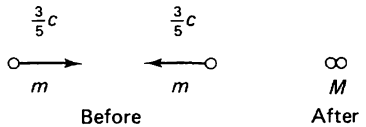


Fig. 3.3 Sticky collision of two equal masses (Example 3.1).

equal, so conservation of energy yields

$$Mc^2 = 2E_m = \frac{2mc^2}{\sqrt{1-(3/5)^2}} = \frac{5}{4}(2mc^2)$$

Conclusion: $M = \frac{5}{2}m$. Notice that this is *greater* than the sum of the initial masses; in sticky collisions kinetic energy is converted into rest energy, so the mass increases.

Example 3.2 A particle of mass M, initially at rest, decays into two pieces, each of mass m (Figure 3.4). *Question*: What is the speed of each piece as it flies off?

Solution: This is, of course, the reverse of the process in Example 3.1 Conservation of momentum just says that the two lumps fly off in opposite directions at equal speeds. Conservation of energy requires that

$$M = \frac{2m}{\sqrt{1-v^2/c^2}}, \quad \text{so} \quad v = c\sqrt{1-(2m/M)^2}$$

This answer makes no sense unless M exceeds $2m$: there has to be at least enough rest energy available to cover the rest energies in the final state (any *extra* is fine; it can be soaked up in the form of kinetic energy). We say that $M = 2m$ is the *threshold* for the process $M \to 2m$ to occur. The deuteron, for example, is below the threshold for decay into proton plus neutron ($m_d = 1875.6$ Mev$/c^2$; $m_p + m_n = 1877.9$ MeV$/c^2$), and therefore is stable. A deuteron can be *pulled* apart, but only by pumping enough energy into the system to make up the difference. (If it puzzles you that a bound state of p and n should weigh *less* than the sum of its parts, the point is that the binding energy of the deuteron – which, like all internal energy, is reflected in its rest mass – is *negative*. Indeed, for *any* stable bound state the binding energy must be negative; if the composite particle weighs *more* than the sum of its constituents, it will spontaneously disintegrate.)

Example 3.3 A pion at rest decays into a muon plus a neutrino (Figure 3.5). *Question*: What is the speed of the muon?

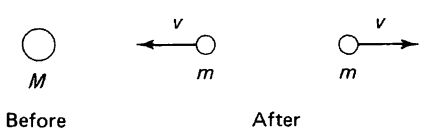


Fig. 3.4 A particle decays into two equal pieces. (Example 3.2).

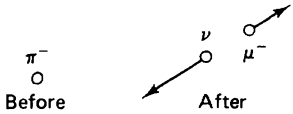


Fig. 3.5 Decay of the charged pion (Example 3.3).

Solution: Conservation of energy requires $E_\pi = E_\mu + E_\nu$. Conservation of momentum gives $\mathbf{p}_\pi = \mathbf{p}_\mu + \mathbf{p}_\nu$; but $\mathbf{p}_\pi = 0$, so $\mathbf{p}_\mu = -\mathbf{p}_\nu$. Thus the muon and the neutrino fly off back-to-back, with equal and opposite momenta. To proceed, we need a formula relating the energy of a particle to its momentum; Equation 3.43 does the job.*

> *Suggestion 1. To get the energy of a particle, when you know its momentum (or vice versa), use the invariant*

$$E^2 - \mathbf{p}^2c^2 = m^2c^4 \tag{3.50}$$

In the present case, then:

$$E_\pi = m_\pi c^2$$
$$E_\mu = c\sqrt{m_\mu^2c^2 + \mathbf{p}_\mu^2}$$
$$E_\nu = |\mathbf{p}_\nu|c = |\mathbf{p}_\mu|c$$

Putting these into the equation for conservation of energy, we have

$$m_\pi c^2 = c\sqrt{m_\mu^2c^2 + \mathbf{p}_\mu^2} + |\mathbf{p}_\mu|c$$

or

$$(m_\pi c - |\mathbf{p}_\mu|)^2 = m_\mu^2c^2 + \mathbf{p}_\mu^2$$

Solving for $|\mathbf{p}_\mu|$,

$$|\mathbf{p}_\mu| = \frac{m_\pi^2 - m_\mu^2}{2m_\pi}c$$

Meanwhile, the energy of the muon (from Equation 3.50) is

$$E_\mu = \frac{m_\pi^2 + m_\mu^2}{2m_\pi}c^2$$

* You might be inclined to solve Equation 3.39 for the velocity, and plug the result into Equation 3.41, but that would be a very poor strategy. In general, velocity is a bad parameter to work with, in relativity. Better to use Equation 3.43, which takes you *directly* back and forth between E and $\mathbf{p}$.

Once we know the energy and momentum of a particle, it is easy to find its velocity. If $E = \gamma mc^2$ and $\mathbf{p} = \gamma m\mathbf{v}$, dividing gives

$$\mathbf{p}/E = \mathbf{v}/c^2$$

Suggestion 2. If you know the energy and momentum of a particle, and you want to determine its velocity, use

$$\mathbf{v} = \mathbf{p}c^2/E \tag{3.51}$$

So the answer to our problem is

$$v_\mu = \frac{m_\pi^2 - m_\mu^2}{m_\pi^2 + m_\mu^2}c$$

Putting in the actual masses, I get $v_\mu = 0.271c$.

There is nothing wrong with that calculation; it was a straightforward and systematic exploitation of the conservation laws. But I want to show you now a faster way to get the energy and momentum of the muon, by using four-vector notation. (I should put a superscript μ on all the four-vectors, but I don't want you to confuse the space-time index μ with the particle identifier μ, so here, and often in the future, I will suppress the space-time indices, and use a dot to indicate the scalar product.) Conservation of energy and momentum requires

$$p_\pi = p_\mu + p_\nu, \quad \text{or} \quad p_\nu = p_\pi - p_\mu$$

Taking the scalar product of each side with itself, we obtain

$$p_\nu^2 = p_\pi^2 + p_\mu^2 - 2p_\pi \cdot p_\mu$$

But

$$p_\nu^2 = 0; \;\; p_\pi^2 = m_\pi^2 c^2, \;\; p_\mu^2 = m_\mu^2 c^2; \;\; \text{and} \;\; p_\pi \cdot p_\mu = \frac{E_\pi}{c}\frac{E_\mu}{c} = m_\pi E_\mu$$

Therefore

$$0 = m_\pi^2 c^2 + m_\mu^2 c^2 - 2m_\pi E_\mu$$

from which E_μ follows immediately.

By the same token

$$p_\mu = p_\pi - p_\nu$$

Squaring yields

$$m_\mu^2 c^2 = m_\pi^2 c^2 - 2m_\pi E_\nu$$

But $E_\nu = |\mathbf{p}_\nu|c = |\mathbf{p}_\mu|c$, so

$$2m_\pi |\mathbf{p}_\mu| = (m_\pi^2 - m_\mu^2)c$$

which gives us $|\mathbf{p}_\mu|$. In this case, the problem was simple enough that the savings afforded by four-vector notation are meager, but in more complicated problems the benefits can be enormous.

> *Suggestion 3. Use four-vector notation, and exploit the invariant dot product. Remember that $p^2 = m^2c^2$ (Equation 3.43) for any (real) particle.*

One reason why the use of invariants is so powerful in this business is that we are free to evaluate them in any inertial system we like. Frequently, the laboratory frame is not the simplest one to work with. In a typical scattering experiment, for instance, a beam of particles is fired at a stationary target. The reaction under study might be, say, $p + p \rightarrow$ whatever, but in the laboratory the situation is asymmetrical, since one proton is moving and the other is at rest. Kinematically, the process is much simpler when viewed from a system in which the two protons approach one another with equal speeds. We call this the *center-of-momentum* (CM) frame, because in this system the total (three-vector) momentum is *zero*.

Example 3.4 The Bevatron at Berkeley was built with the idea of producing antiprotons, by the reaction $p + p \rightarrow p + p + p + \bar{p}$. That is, a high-energy proton strikes a proton at rest, creating (in addition to the original particles) a proton–antiproton pair. *Question*: What is the threshold energy for this reaction (i.e. the minimum energy of the incident proton)?

Solution: In the laboratory the process looks like Figure 3.6a; in the CM frame, it looks like Figure 3.6b. Now, what is the condition for threshold? Answer: Just barely enough incident energy to create the two extra particles. In the lab frame, it is hard to see how we would formulate this condition, but in the CM it is easy: *all four final particles must be at rest*, with nothing 'wasted' in the form of kinetic energy. (We can't have that in the *lab* frame, of course, since conservation of momentum requires that there be some residual motion.)

Let p^μ_{TOT} be the *total* energy–momentum four-vector in the lab; it is conserved, so it doesn't matter whether we evaluate it before or after the collision. We'll do it *before*:

$$p^\mu_{\text{TOT}} = \left(\frac{E + mc^2}{c}, |\mathbf{p}|, 0, 0\right)$$

where E and $\mathbf{p}$ are the energy and momentum of the incident proton, and m is the proton mass. Let $p^\mu_{\text{TOT}}{}'$ be the total energy–momentum four-vector in the CM. Again, we can evaluate it before or after the collision; this time we'll do it *after*:

$$p^\mu_{\text{TOT}}{}' = (4mc, 0, 0, 0)$$

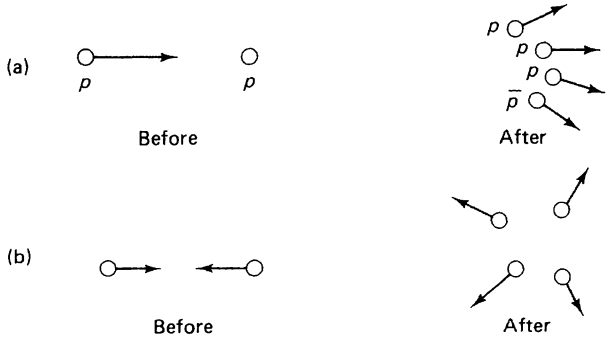


Fig. 3.6 $p + p \rightarrow p + p + p + \bar{p}$. (a) In the lab frame; (b) in the CM frame.

since (at threshold) all four particles are at rest. Now $p^{\mu}_{\text{TOT}} \neq p^{\mu}_{\text{TOT}}{}'$, obviously, but the invariant products $p_{\mu\text{TOT}}p^{\mu}_{\text{TOT}}$ and $p'_{\mu\text{TOT}}p^{\mu}_{\text{TOT}}{}'$ *are* equal:

$$\left(\frac{E}{c} + mc\right)^2 - \mathbf{p}^2 = (4mc)^2$$

Using the standard invariant (Equation 3.50) to eliminate $\mathbf{p}^2$, and solving for E, we find

$$E = 7mc^2$$

Evidently, the incident proton must carry a kinetic energy at least six times its rest energy, for this process to occur. (And in fact the first antiprotons were discovered when the machine reached about 6000 MeV.)

This is perhaps a good place to emphasize the distinction between a *conserved* quantity and an *invariant* quantity. Energy is *conserved* – the same value *after* the collision as *before* – but it is not invariant. Mass is *invariant* – the same in all inertial systems – but it is not conserved. Some quantities are both invariant and conserved (e.g. electric charge); many are neither (speed, for instance). As Example 3.4 indicates, the clever exploitation of conserved and invariant quantities can save you a lot of messy algebra. It also demonstrates that some problems are easier to analyze in the CM system, whereas others may be simpler in the lab frame.

> *Suggestion 4. If a problem seems cumbersome in the lab frame, try analyzing it in the CM system.*

Even if you're dealing with something more complicated than a collision of two identical particles, the CM (in which $\mathbf{p}_{\text{TOT}} = 0$) is still a useful reference frame, for in this system conservation of momentum is trivial: zero before, zero after. But you might wonder whether there *is* always a CM frame. In other words, given a swarm of particles with masses $m_1, m_2, m_3, \ldots$, and velocities $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots$, does there necessarily exist an inertial system in which their total (three-vector) momentum

is zero? The answer is *yes*; I will prove it by finding the velocity of that frame and demonstrating that this velocity is less than c. The total energy and momentum in the lab frame (S) are

$$E_{\text{TOT}} = \sum_i \gamma_i m_i c^2; \quad \mathbf{p}_{\text{TOT}} = \sum_i \gamma_i m_i \mathbf{v}_i \tag{3.52}$$

Since p^μ_{TOT} is a four-vector, we can use the Lorentz transformations to get the momentum in system S', moving in the direction of $\mathbf{p}_{\text{TOT}}$ with speed v

$$|\mathbf{p}'_{\text{TOT}}| = \gamma \left(|\mathbf{p}_{\text{TOT}}| - \beta \frac{E_{\text{TOT}}}{c} \right)$$

In particular, this momentum is *zero* if v is chosen such that

$$\frac{v}{c} = \frac{|\mathbf{p}_{\text{TOT}}|c}{E_{\text{TOT}}} = \frac{|\sum \gamma_i m_i \mathbf{v}_i|}{\sum \gamma_i m_i c}$$

Now, the length of the sum of three-vectors cannot exceed the sum of their lengths (this geometrically evident fact is known as the *triangle inequality*), so

$$\frac{v}{c} \leq \frac{\sum \gamma_i m_i (v_i/c)}{\sum \gamma_i m_i}$$

and since $v_i < c$, we can be sure that $v < c$.* Thus the CM system always exists, and its velocity relative to the lab frame is given by

$$\mathbf{v}_{\text{CM}} = \frac{\mathbf{p}_{\text{TOT}} c^2}{E_{\text{TOT}}} \tag{3.53}$$

It seems odd, looking back at the answer to Example 3.4, that it takes an incident kinetic energy *six times* the proton rest energy to produce a $p/\bar{p}$ pair. After all, we're only creating $2mc^2$ of new rest energy. This example illustrates the inefficiency of scattering off a stationary target; conservation of momentum forces you to waste a lot of energy as kinetic energy in the final state. Suppose we could have fired the two protons at one another, making the laboratory itself the CM system. Then it would suffice to give each proton a kinetic energy of only mc^2, one-sixth of what the stationary-target experiment requires. This realization led, in the early 1970s, to the development of colliding-beam machines (see Figure 3.7). Today, virtually every new machine in high-energy physics is a collider.

Example 3.5 Suppose two identical particles, each with mass m and kinetic energy T, collide head-on. *Question:* What is their *relative* kinetic energy, T' (i.e. the kinetic energy of one in the rest system of the other)?

* I am tacitly assuming that at least one of the particles is massive. If *all* of them are massless, we *may* obtain $v = c$, in which case there is no CM system. For example, there is no CM frame for a single photon.

Fig. 3.7 Two experimental arrangements: (a) Colliding beams; (b) fixed target.

Solution: There are many ways to do this one. A quick method is to write down the total four-momentum in the CM and in the lab

$$p^{\mu}_{\text{TOT}} = \left(\frac{2E}{c}, 0\right), \quad p^{\mu}_{\text{TOT}}{}' = \left(\frac{E' + mc^2}{c}, \mathbf{p}'\right)$$

set $(p_{\text{TOT}})^2 = (p_{\text{TOT}}')^2$:

$$\left(\frac{2E}{c}\right)^2 = \left(\frac{E' + mc^2}{c}\right)^2 - \mathbf{p}'^2$$

use Equation 3.50 to eliminate $\mathbf{p}'$

$$2E^2 = mc^2(E' + mc^2)$$

and express the answer in terms of $T = E - mc^2$ and $T' = E' - mc^2$

$$T' = 4T\left(1 + \frac{T}{2mc^2}\right) \tag{3.54}$$

The *classical* answer would have been $T' = 4T$, to which this reduces when $T \ll mc^2$. (In the rest system of B, A has, classically, twice the velocity, and hence four times as much kinetic energy, as in the CM.) Now, a factor of 4 is *some* benefit, to be sure, but the *relativistic* gain can be greater by far. Colliding electrons with a laboratory kinetic energy of 1 GeV, for example, would have a *relative* kinetic energy of 4000 GeV!

References

1 There are many excellent textbooks on special relativity. I especially recommend Smith, J. H. (1996) *Introduction to Special Relativity*, Dover, New York. For a fascinating (but unorthodox) approach, see (a) Taylor, E. F. and Wheeler, J. A. (1992) *Spacetime Physics*, 2nd edn. Freeman, San Francisco, C.A.

2 If you want to go into this much more deeply, the standard reference is Hagedorn, R. (1964) *Relativistic Kinematics*, Benjamin, New York.

Problems

3.1 Solve Equation 3.1 for x, y, z, t in terms of x', y', z', t', and check that you recover Equation 3.3.

3.2 (a) Derive Equation 3.4.

(b) According to clocks on the ground (system S), streetlights A and B (situated 4 km apart) were both turned on at precisely 8:00 P.M. Which one went on first according to an observer on a train (system S'), which moves from A toward B at $\frac{3}{5}$ the speed of light? How much later (in seconds) did the other light go on? *Note*: As always in relativity, we are talking here about what S' *observed, after* correcting for the time it took the light to reach her, *not* what she actually *saw* (which would depend on where she was located on the train).

3.3 (a) How do *volumes* transform? (If a container has volume V' in its own rest frame, S', what is its volume as measured by an observer in S, with respect to which it is moving at speed v?)

(b) How do *densities* transform? (If a container holds ρ' molecules per unit volume in its own rest frame, S', how many molecules per unit volume does it carry in S?)

3.4 Cosmic ray muons are produced high in the atmosphere (at 8000 m, say) and travel toward the earth at very nearly the speed of light (0.998 c, say).

(a) Given the lifetime of the muon (2.2×10^{-6} sec), how far would it go before disintegrating, according to prerelativistic physics? Would the muons make it to ground level?

(b) Now answer the same question using *relativistic* physics. (Because of time dilation, the muons last longer, so they travel farther.)

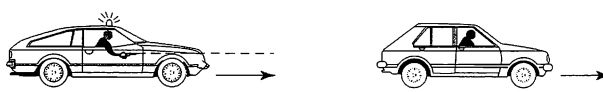
(c) Pions are also produced in the upper atmosphere. In fact, the sequence is proton (from outer space) hits proton (in atmosphere) $\rightarrow p + p +$ pions. The pions then decay into muons: $\pi^- \rightarrow \mu^- + \bar{\nu}_\mu$; $\pi^+ \rightarrow \mu^+ + \nu_\mu$. But, the lifetime of the pion is much shorter (2.6×10^{-8} s). Assuming the pions have the same speed (0.998 c), will they reach ground level?

3.5 Half the muons in a monoenergetic beam decay in the first 600 m. How fast are they going?

3.6 As the outlaws escape in their getaway car, which goes $\frac{3}{4}c$, the cop fires a bullet from the pursuit car, which only goes $\frac{1}{2}c$. The muzzle velocity (speed relative to gun) of the bullet is $\frac{1}{3}c$. Does the bullet reach its target

(a) According to prerelativistic physics?

(b) According to relativity?



3.7 Find the matrix M that inverts Equation 3.12: $x^\mu = M^\mu_\nu x^{\nu\prime}$ (use Equation 3.3). Show that M is the matrix inverse of Λ: $\Lambda M = 1$.

3.8 Show that the quantity I (in Equation 3.13) is invariant under Lorentz transformations (Equation 3.8).

3.9 Given two four-vectors, $a^\mu = (3, 4, 1, 2)$ and $b^\mu = (5, 0, 3, 4)$, find: a_μ, b_μ, a^2, b^2, $\mathrm{a} \cdot \mathrm{b}$, a^2, b^2, and $a \cdot b$. Characterize a^μ and b^μ as timelike, spacelike, or lightlike.

3.10 A second-rank tensor is called *symmetric* if it is unchanged when you switch the indices ($s^{\nu\mu} = s^{\mu\nu}$); it is *antisymmetric* if it changes sign ($a^{\nu\mu} = -a^{\mu\nu}$).

(a) How many independent elements are there in a symmetric tensor? (Since $s^{12} = s^{21}$, these would count as only *one* independent element.)

(b) How many independent elements are there in an antisymmetric tensor?

(c) Show that symmetry is preserved by Lorentz transformations – that is, if $s^{\mu\nu}$ is symmetric, so too is $s^{\mu\nu\prime}$. What about *anti*symmetry?

(d) If $s^{\mu\nu}$ is symmetric, show that $s_{\mu\nu}$ is also symmetric. If $a^{\mu\nu}$ is antisymmetric, show that $a_{\mu\nu}$ is antisymmetric.

(e) If $s^{\mu\nu}$ is symmetric and $a^{\mu\nu}$ is antisymmetric, show that $s^{\mu\nu}a_{\mu\nu} = 0$.

(f) Show that any second-rank tensor ($t^{\mu\nu}$) can be written as the sum of an antisymmetric part ($a^{\mu\nu}$) and a symmetric part ($s^{\mu\nu}$): $t^{\mu\nu} = a^{\mu\nu} + s^{\mu\nu}$. Construct $s^{\mu\nu}$ and $a^{\mu\nu}$ explicitly, given $t^{\mu\nu}$.

3.11 A particle is traveling at $\frac{3}{5}c$ in the x direction. Determine its proper velocity, η^μ (all four components).

3.12 Consider a collision in which particle A (with 4-momentum p_A^μ) hits particle B (4-momentum p_B^μ), producing particles C (p_C^μ) and D (p_D^μ). Assume the (relativistic) energy and momentum are conserved in systems S ($p_A^\mu + p_B^\mu = p_C^\mu + p_D^\mu$). Using the Lorentz transformations (Eq. 3.12), show that energy and momentum are also conserved in S'.

3.13 Is p^μ timelike, spacelike, or lightlike for a (real) particle of mass m? How about a massless particle? How about a virtual particle?

3.14 How much more does a hot potato weigh than a cold one (in kg)?

3.15 A pion traveling at speed v decays into a muon and a neutrino, $\pi^- \to \mu^- + \bar{\nu}_\mu$. If the neutrino emerges at 90° to the original pion direction, at what angle does the μ come off? [*Answer*: $\tan\theta = (1 - m_\mu^2/m_\pi^2)/(2\beta\gamma^2)$.]

3.16 Particle A (energy E) hits particle B (at rest), producing particles $C_1, C_2, \ldots$: $A + B \to C_1 + C_2 + \cdots + C_n$. Calculate the threshold (i.e. minimum E) for this reaction, in terms of the various particle masses.

$$\left[\textit{Answer}: \ \frac{M^2 - m_A^2 - m_B^2}{2m_B}\,c^2, \text{ where } M \equiv m_1 + m_2 + \ldots + m_n\right]$$

3.17 Use the result of Problem 3.16 to find the threshold energies for the following reactions, assuming the target proton is stationary:*

(a) $p + p \to p + p + \pi^0$

(b) $p + p \to p + p + \pi^+ + \pi^-$

(c) $\pi^- + p \to p + \bar{p} + \mathrm{n}$

(d) $\pi^- + p \to K^0 + \Sigma^0$

(e) $p + p \to p + \Sigma^+ + K^0$

3.18 The first man-made Ω^- (Fig. 1.9) was created by firing a high-energy proton at a stationary hydrogen atom to produce a K^+/K^- pair: $p + p \to p + p + K^+ + K^-$; the K^- in turn hit another stationary proton, $K^- + p \to \Omega^- + K^0 + K^+$. What minimum kinetic energy is required (for the incident proton), to make an Ω^- in this way? (Gell-Mann must have done this calculation to see whether the experiment would be feasible.)

Beware: The *Particle Physics Booklet* (and most other sources) list particle 'masses' in MeV. For example, the mass of the muon is quoted as 105.658 MeV. What they *mean*, of course, is the *rest energy* of the muon: $m_\mu c^2 = 105.658$ MeV – or, what is the same thing, $m_\mu = 105.658\ \mathrm{MeV}/c^2$. It is safest to convert formulas from mass to rest energy *before* plugging in any numbers. In this case, for example, multiply top and bottom by c^2, to get $E_{min} = [(Mc^2)^2 - (m_Ac^2)^2 - (m_Bc^2)^2]/2(m_Bc^2)$.

3.19 Particle A, at rest, decays into particles B and C ($A \to B + C$).

(a) Find the energy of the outgoing particles, in terms of the various masses.

$$\left[\textit{Answer:}\ E_B = \frac{m_A^2 + m_B^2 - m_C^2}{2m_A} c^2\right]$$

(b) Find the magnitudes of the outgoing momenta.

$$\left[\begin{array}{l} \textit{Answer:}\ |\mathbf{p}_B| = |\mathbf{p}_C| = \dfrac{\sqrt{\lambda(m_A^2, m_B^2, m_C^2)}}{2m_A c,} \\ \\ \text{where } \lambda \text{ is the so-called } \textit{triangle function}: \\ \\ \lambda(x, y, z) \equiv x^2 + y^2 + z^2 - 2xy - 2xz - 2yz. \end{array}\right]$$

(c) Note that λ factors: $\lambda(a^2, b^2, c^2) = (a + b + c)(a + b - c)(a - b + c)(a - b - c)$. Thus $|\mathbf{p}_B|$ goes to zero when $m_A = m_B + m_C$, and runs imaginary if $m_A < (m_B + m_C)$. Explain.

3.20 Use the result of Problem 3.19 to find the CM energy of each decay product in the following reactions (see footnote to Problem 3.17):

(a) $\pi^- \to \mu^- + \overline{\nu}_\mu$
(b) $\pi^0 \to \gamma + \gamma$
(c) $K^+ \to \pi^+ + \pi^0$
(d) $\Lambda \to p + \pi^-$
(e) $\Omega^- \to \Lambda + K^-$

3.21 A pion at rest decays into a muon and a neutrino ($\pi^- \to \mu^- + \overline{\nu}_\mu$). On the average, how far will the muon travel (in vacuum) before disintegrating? [*Answer*: $d = [(m_\pi^2 - m_\mu^2)/(2m_\pi m_\mu)]c\tau = 186$ m.]

3.22 Particle A, at rest, decays into three or more particles: $A \to B + C + D + \cdots$.

(a) Determine the maximum and minimum energies that B can have in such a decay, in terms of the various masses.

(b) Find the maximum and minimum electron energies in muon decay, $\mu^- \to e^- + \overline{\nu}_e + \nu_\mu$.

3.23 (a) A particle traveling at speed u approaches an identical particle at rest. What is the speed (v) of each particle in the CM frame? (*Classically*, of course, it would just be $u/2$.)
[*Answer*: $(c^2/u)\left(1 - \sqrt{1 - u^2/c^2}\right)$]

(b) Find $\gamma \equiv 1/\sqrt{1 - v^2/c^2}$ in terms of $\gamma' \equiv 1/\sqrt{1 - u^2/c^2}$.
[*Answer*: $\sqrt{(\gamma' + 1)/2}$]

(c) Use your result in part (b) to express the kinetic energy of each particle in the CM frame, and thus re-derive Equation 3.54

3.24 In reactions of the type $A + B \to A + C_1 + C_2 + \cdots$ (in which particle A scatters off particle B, producing $C_1, C_2, \ldots$), there is another inertial frame, in addition to the lab (B at rest) and the CM ($\mathbf{p}_{TOT} = 0$), which is sometimes useful. It is called the *Breit*, or 'brick wall' frame, and it is the system in which A recoils with its momentum reversed ($\mathbf{p}_{after} = -\mathbf{p}_{before}$), as though it had bounced off a brick wall. Take the case of elastic scattering ($A + B \to A + B$); if particle A carries energy E, and scatters at an angle θ, in the CM, what is its energy in the Breit frame? Find the velocity of the Breit frame (magnitude and direction) relative to the CM.

3.25 In a two-body scattering event, $A + B \rightarrow C + D$, it is convenient to introduce the *Mandelstam variables*

$$s \equiv (p_A + p_B)^2/c^2$$

$$t \equiv (p_A - p_C)^2/c^2$$

$$u = (p_A - p_D)^2/c^2$$

(a) Show that $s + t + u = m_A^2 + m_B^2 + m_C^2 + m_D^2$.
The *theoretical* virtue of the Mandelstam variables is that they are Lorentz invariants, with the same value in any inertial system. *Experimentally*, though, the more accessible parameters are energies and scattering angles.

(b) Find the CM energy of A, in terms of s, t, u and the masses. [*Answer*: $E_A^{\text{CM}} = (s + m_A^2 - m_B^2)c^2/2\sqrt{s}$.]

(c) Find the Lab (B at rest) energy of A. [*Answer*: $E_A^{\text{lab}} = (s - m_A^2 - m_B^2)c^2/2m_B$.]

(d) Find the total CM energy ($E_{TOT} = E_A + E_B = E_C + E_D$). [*Answer*: $E_{\text{TOT}}^{\text{CM}} = \sqrt{s}c^2$.]

3.26 For elastic scattering of identical particles, $A + A \rightarrow A + A$, show that the Mandelstam variables (Problem 3.25) become

$$s = 4(\mathbf{p}^2 + m^2c^2)/c^2$$

$$t = -2\mathbf{p}^2(1 - \cos\theta)/c^2$$

$$u = -2\mathbf{p}^2(1 + \cos\theta)/c^2$$

where $\mathbf{p}$ is the CM momentum of the incident particle, and θ is the scattering angle.

3.27 Work out the kinematics of Compton scattering: a photon of wavelength λ collides elastically with a charged particle of mass m. If the photon scatters at angle θ, find its outgoing wavelength, λ'. [*Answer*: $\lambda' = \lambda + (h/mc)(1 - \cos\theta)$.]

4

Symmetries

Symmetries play an important role in elementary particle physics, in part because of their relation to conservation laws and in part because they permit one to make some progress when a complete dynamical theory is not yet available. The first section of this chapter contains some general remarks about the mathematical description of symmetry (group theory) and the relation between symmetry and conservation laws (Noether's theorem). We then take up the case of rotational symmetry and its relation to angular momentum and spin. This leads in turn to the 'internal' symmetries – isospin, SU(3), and flavor SU(6). Finally, we consider 'discrete' symmetries – parity, charge conjugation, and time reversal. Except for the theory of spin (Section 4.2) – which will be used extensively in later chapters – and the material on parity in Section 4.1 – which is useful background for Chapter 9 – this chapter can be studied as superficially (or as deeply) as the reader wishes. I recommend a quick pass at this stage and a return to specific sections later, if warranted. Some knowledge of matrix theory is presupposed; readers familiar with quantum mechanics will find the sections on angular momentum an easy review (those unacquainted with quantum mechanics may find them hopelessly obscure, in which case they should study the relevant chapter of an introductory quantum text). Group theory is touched on here in a scandalously cursory fashion (my main purpose is to introduce some standard terminology); a serious student of elementary particle physics should plan eventually to study this subject in far greater detail.

4.1
Symmetries, Groups, and Conservation Laws

Take a look at the graph in Figure 4.1. I won't tell you what the functional form of $f(x)$ is, but this much is clear: It's an *odd* function, $f(-x) = -f(x)$. (If you don't believe me, trace the curve, rotate the tracing by 180°, and check that it perfectly fits the original.) From this it follows, for instance, that

$$[f(-x)]^6 = [f(x)]^6, \qquad \int_{-3}^{+3} f(x)\,\mathrm{d}x = 0,$$

$$\left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{+2} = \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{-2}, \qquad \int_{-7}^{+7} [f(x)]^2\,\mathrm{d}x = 2\int_{0}^{+7} [f(x)]^2\,\mathrm{d}x, \tag{4.1}$$

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

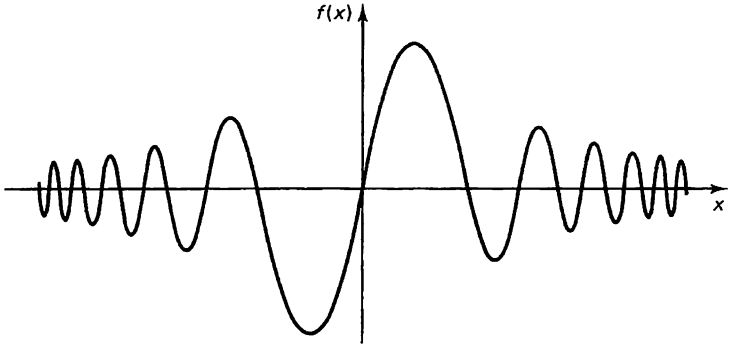


Fig. 4.1 An odd function.

that no cosines appear in the Fourier expansion of $f(x)$, and that its Taylor series contains only odd powers of x. In fact, you can deduce quite a lot about $f(x)$, even though you don't know its functional form, just from the observation that it has a particular *symmetry* – oddness, in this case. In physics, intuition or a general principle often suggests symmetries in a problem, and their systematic exploitation can be an extremely powerful tool.*

The most striking examples of symmetry in physics are, I suppose, crystals. But we're not so much interested here in static symmetries of *shape* as in dynamical symmetries of *motion*. The Greeks apparently believed that the symmetries of nature should be directly reflected in the motion of objects: stars must move in circles because those are the most symmetrical trajectories. Of course, planets do not, and that was embarrassing (it was not the last time that naive intuitions about symmetry ran into trouble with experiment). Newton recognized that fundamental symmetries are revealed not in the motions of individual objects, but in the *set of all possible* motions – symmetries are manifest in the *equations of motion* rather than in particular solutions to those equations. Newton's law of universal gravitation, for instance, exhibits spherical symmetry (the force is the same in all directions), and yet planetary orbits are elliptical. Thus the underlying symmetry of the system is only indirectly revealed to us; indeed, you might wonder how we would ever have discovered it from the observed planetary trajectories, if we didn't have a pretty strong hunch that the gravitational field of the sun 'ought' to be spherically symmetrical.

It was not until 1917 that the dynamical implications of symmetry were completely understood. In that year, Emmy Noether published her famous theorem

* In some respects, the appeal to symmetry is characteristic of an *incomplete* theory. For example, if we somehow discovered the explicit form of $f(x)$, say, $f(x) = e^{-x^2}\sin(x^3)$, then the theorems in Equation 4.1 would lose their luster. Why bother with *partial* information when we can have it *all*? But even in a mature theory, symmetry considerations often lead to deeper understanding and calculational simplification. For instance, if you're called upon to integrate $f(x)$ from -3 to $+3$, it pays to notice that $f(x)$ is odd, even if you *do* know its functional form.

Table 4.1 Symmetries and conservation laws.

Symmetry		Conservation law
Translation in time	↔	Energy
Translation in space	↔	Momentum
Rotation	↔	Angular momentum
Gauge transformation	↔	Charge

relating symmetries and conservation laws:

Noether's Theorem: Symmetries ↔ Conservation laws

Every symmetry of nature yields a conservation law; conversely, every conservation law reflects an underlying symmetry. For example, the laws of physics are symmetrical with respect to *translations in time* (they work the same today as they did yesterday). Noether's theorem relates this invariance to conservation of energy. If a system is invariant under translations in space, then momentum is conserved; if it is symmetrical under rotations about a point, then angular momentum is conserved. Similarly, the invariance of electrodynamics under gauge transformations leads to conservation of charge (we call this an *internal* symmetry, in contrast to the space-time symmetries). I'm not going to *prove* Noether's theorem; the details are not terribly enlightening [1]. The important thing is the profound and beautiful *idea* that symmetries are associated with conservation laws (see Table 4.1).

I have been speaking rather casually about symmetries, and I cited some examples; but what precisely *is* a symmetry? It is an operation you can perform (at least conceptually) on a system that leaves it *invariant* – that carries it into a configuration indistinguishable from the original one. In the case of the function in Figure 4.1, changing the sign of the argument, $x \to -x$, and multiplying the whole thing by -1, $f(x) \to -f(-x)$, is a symmetry operation. For a meatier example, consider the equilateral triangle (Figure 4.2). It is carried into itself by a clockwise rotation through 120° (R_+), and by a counterclockwise rotation through 120° (R_-), by flipping it about the vertical axis a (R_a), or around the axis through b (R_b), or c (R_c). Is that all? Well, doing *nothing (I)* obviously leaves it invariant, so this too is a symmetry operation, albeit a pretty trivial one. And then we could combine operations – for example, rotate clockwise through 240°. But that's the same as rotating *counter* clockwise by 120° (i.e. $R_+^2 = R_-$). As it turns out, we have already identified all the distinct symmetry operations on the equilateral triangle (see Problem 4.1).

The set of all symmetry operations (on a particular system) has the following properties:

1. *Closure*: If R_i and R_j are in the set, then the *product*, R_iR_j – meaning: first perform R_j, then perform R_i* – is also in the set; that is, there exists some R_k such that $R_iR_j = R_k$.

* Note the 'backwards' ordering. Think of the symmetry operations as acting on a system to their right: $R_iR_j(\Delta) = R_i[R_j(\Delta)]$; R_j acts first, and then R_i acts on the result.

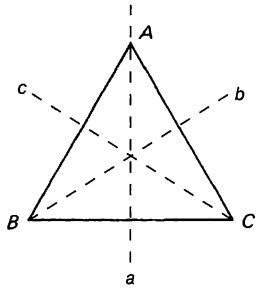


Fig. 4.2 Symmetries of the equilateral triangle.

2. *Identity*: There is an element I such that $IR_i = R_iI = R_i$ for all elements R_i.
3. *Inverse*: For every element R_i there is an *inverse*, R_i^{-1}, such that $R_iR_i^{-1} = R_i^{-1}R_i = I$.
4. *Associativity*: $R_i(R_jR_k) = (R_iR_j)R_k$.

These are the defining properties of a mathematical *group*. Indeed, group theory may be regarded as the systematic study of symmetries. Note that group elements need not *commute*: $R_iR_j \neq R_jR_i$, in general. If all the elements *do* commute, the group is called *Abelian*. Translations in space and time form Abelian groups; rotations (in three dimensions) do *not* [2]. Groups can be *finite* (like the triangle group, which has just six elements) or *infinite* (for example, the set of integers, with addition playing the role of group 'multiplication'). We shall encounter *continuous* groups (such as the group of all rotations in a plane), in which the elements depend on one or more continuous parameters* (the angle of rotation, in this case), and *discrete* groups, in which the elements may be labeled by an index that takes on only integer values (all finite groups are, of course, discrete).

As it turns out, most of the groups of interest in physics can be formulated as groups of *matrices*. The Lorentz group, for instance, consists of the set of 4×4 Λ matrices introduced in Chapter 3. In elementary particle physics, the most common groups are of the type mathematicians call $U(n)$: the collection of all unitary $n \times n$ matrices (see Table 4.2). (A *unitary matrix* is one whose inverse is equal to its transpose conjugate: $U^{-1} = \tilde{U}^*$.) If we restrict ourselves further to unitary matrices with determinant 1, the group is called $SU(n)$. (The S stands for 'special', which just means 'determinant 1'.) If we limit ourselves to *real* unitary matrices, the group is $O(n)$. (O stands for 'orthogonal'; an orthogonal matrix is one whose inverse is equal to its transpose: $O^{-1} = \tilde{O}$.) Finally, the group of real, orthogonal, $n \times n$ matrices of determinant 1 is $SO(n)$; $SO(n)$ may be thought of as the group of all *rotations* in a space of n dimensions. Thus, $SO(3)$ describes the

* If this dependence takes the form of an *analytic* function, it is called a *Lie* group. All of the continuous groups one encounters in physics are Lie groups [3].

Table 4.2 Important symmetry groups.

Group name	Dimension	Matrices in group
$U(n)$	$n \times n$	unitary ($\tilde{U}^* U = 1$)
$SU(n)$	$n \times n$	unitary, determinant 1
$O(n)$	$n \times n$	orthogonal ($\tilde{O}O = 1$)
$SO(n)$	$n \times n$	orthogonal, determinant 1

rotational symmetry of our world, a symmetry that is related by Noether's theorem to the conservation of angular momentum. Indeed, the entire quantum theory of angular momentum is really closet group theory. It so happens that $SO(3)$ is almost identical in mathematical structure to $SU(2)$, which is the most important *internal* symmetry in elementary particle physics. So the theory of angular momentum, to which we turn next, will actually serve us twice.

One final thing. Every group G can be *represented* by a group of *matrices*: for every group element a there is a corresponding matrix M_a, and the correspondence respects group multiplication, in the sense that if $ab = c$, then $M_a M_b = M_c$. A representation need not be 'faithful': there may be many distinct group elements represented by the same matrix. (Mathematically, the group of matrices is *homo*morphic, but not necessarily *iso*morphic, to G.) Indeed, there is a trivial case, in which we represent every element by the 1×1 *unit* matrix (which is to say, the number 1). If G is a group of matrices, such as $SU(6)$ or $O(18)$, then it is (obviously) a representation of *itself* – we call it the *fundamental* representation. But there will, in general, be many other representations, by matrices of various dimensions. For example, $SU(2)$ has representations of dimension 1 (the trivial one), 2 (the matrices themselves), 3, 4, 5, and in fact *every* positive integer. A major problem in group theory is the characterization of all the representations of a given group.

Of course, you can always construct a new representation by combining two old ones, thus

$$M_a = \begin{pmatrix} \boxed{M_a^{(1)}} & \text{(zeros)} \\ \text{(zeros)} & \boxed{M_a^{(2)}} \end{pmatrix}$$

But we don't count this separately; when we list the representations of a group, we are talking about the so-called *irreducible* representations, which *cannot* be decomposed into block-diagonal form. Actually, you have already encountered several examples of group representations, probably without realizing it: an ordinary scalar belongs to the one-dimensional representation of the rotation group, $SO(3)$, and a vector belongs to the three-dimensional representation; four-vectors belong to the four-dimensional representation of the Lorentz group; and the curious geometrical arrangements of Gell-Mann's Eightfold Way correspond to irreducible representations of the group $SU(3)$.

4.2 Angular Momentum

The earth, in its motion, carries two kinds of angular momentum: *orbital* angular momentum, rmv, associated with its annual revolution around the sun, and *spin* angular momentum, $I\omega$, associated with its daily rotation about the north–south axis. The same goes for the electron in a hydrogen atom: it too carries both orbital and spin angular momentum. In the macroscopic case, the distinction is not terribly profound; after all, the spin angular momentum of the earth is nothing but the sum total of the 'orbital' angular momenta of all the rocks and dirt clods that make it up, in their daily 'orbit' around the axis. In the case of the electron this interpretation is not open to us: the electron, as far as we know, is a true point particle; its spin angular momentum is not attributable to constituent parts revolving about an axis, but is simply an intrinsic property of the particle itself (see Problem 4.8).

Classically, we are free to measure all three components of the orbital angular momentum vector, $\mathbf{L} = \mathbf{r} \times m\mathbf{v}$, to any desired precision, and these components can assume any values whatever. In quantum mechanics, however, it is *impossible in principle* to measure all three components simultaneously; a measurement of L_x, say, inevitably alters the value of L_y, by an unpredictable amount. The best we can do is to measure the *magnitude* of $\mathbf{L}$, (or rather, its *square*: $L^2 = \mathbf{L} \cdot \mathbf{L}$) together with *one component* (which we customarily take to be the z component, L_z). Furthermore, these measurements can only return certain 'allowed' values.* Specifically, a (competent) measurement of L^2 always yields a number of the form

$$l(l+1)\hbar^2 \tag{4.2}$$

where l is a nonnegative integer:

$$l = 0, 1, 2, 3, \ldots \tag{4.3}$$

For a given value of l, a measurement of L_z always gives a result of the form

$$m_l\hbar \tag{4.4}$$

where m_l is an integer in the range $-l$ to $+l$:

$$m_l = -l, -l+1, \ldots, -1, 0, +1, \ldots, l-1, l \tag{4.5}$$

$(2l+1)$ possibilities. Figure 4.3 may help you to visualize the situation. Here $l = 2$, so the magnitude of $\mathbf{L}$ is $\sqrt{6}\hbar = 2.45\ \hbar$; L_z can assume the values $2\hbar$, $\hbar$, 0, $-\hbar$, or

* I am not going to *prove* the quantization rules for angular momentum, and if this material is new to you, I suggest that you consult a textbook on quantum mechanics. All I propose to do here is summarize the essential results we will need in what follows.

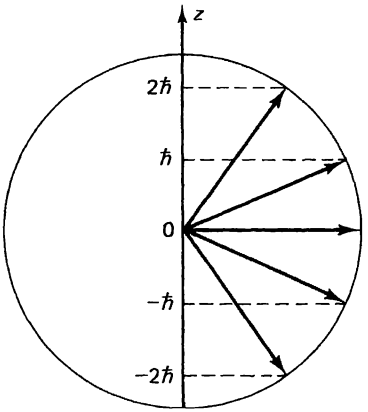


Fig. 4.3 Possible orientations of the angular momentum vector for $l = 2$.

$-2\hbar$. Notice that the angular momentum vector cannot be oriented purely in the z direction.

The same goes for spin angular momentum: a measurement of $S^2 = \mathbf{S} \cdot \mathbf{S}$ can only return values of the form

$$s(s+1)\hbar^2 \tag{4.6}$$

In the case of spin, however, the quantum number s can be a *half*-integer as well as an integer:

$$s = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, \tfrac{5}{2}, \ldots \tag{4.7}$$

For a given value of s, a measurement of S_z must yield an answer of the form

$$m_s\hbar \tag{4.8}$$

where m_s is an integer or half-integer (whichever s is) in the range $-s$ to s:

$$m_s = -s, -s+1, \ldots, s-1, s \tag{4.9}$$

$(2s+1)$ possibilities.

Now, a given particle can be given any *orbital* angular momentum l you like, but for each type of particle, the value of s is *fixed*. Every pion or kaon, for example, has $s = 0$; every electron, proton, neutron, and quark carries $s = \frac{1}{2}$; for the ρ, the ψ, the photon, and the gluon, $s = 1$; for the Δ's and the Ω^-, $s = \frac{3}{2}$; and so on. We call s the 'spin' of the particle. Particles with half-integer spin are *fermions* – all

Table 4.3 Classification of particles by spin.

Bosons (integer spin)		Fermions (half-integer spin)		
Spin 0	**Spin 1**	**Spin $\frac{1}{2}$**	**Spin $\frac{3}{2}$**	
–	Mediators	Quarks/Leptons	–	← Elementary
Pseudoscalar mesons	Vector mesons	Baryon octet	Baryon decuplet	← Composite

baryons, leptons, and quarks are fermions; particles with integer spin are *bosons* – all mesons and mediators are bosons (see Table 4.3).*

4.2.1
Addition of Angular Momenta

Angular momentum states are represented by 'kets': $|l\ m_l\rangle$ or $|s\ m_s\rangle$. Thus, if I say the electron in a hydrogen atom occupies the orbital state $|3\ -1\rangle$ and the spin state $|\frac{1}{2}\ \frac{1}{2}\rangle$, I am telling you that $l = 3$, $m_l = -1$, $s = \frac{1}{2}$ (which is unnecessary, of course; if it is an electron, s *must* be $\frac{1}{2}$), and $m_s = \frac{1}{2}$. Now, it may happen that we are not interested in the spin and orbital angular momenta *separately*, but rather in the *total* angular momentum, $\mathbf{L} + \mathbf{S}$. (In the presence of *coupling* between $\mathbf{L}$ and $\mathbf{S}$ – tidal, if it's the earth–sun system; magnetic, for the electron–proton system – it is the *sum*, and not $\mathbf{L}$ and $\mathbf{S}$ individually, that will be conserved.) Or perhaps we are studying the two quarks that go to make a ψ meson; in this case, as we shall see, the *orbital* angular momentum is zero, but we are confronted with the problem of combining the two quark spins to get the total spin of the ψ: $\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2$. In either case, the question arises: *how do we add two angular momenta*†

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2 \tag{4.10}$$

Classically, of course, we just add the components. But in quantum mechanics we do not have access to all three components; we are obliged to work with *one* component and the *magnitude*. So the question becomes: if we combine states $|j_1 m_1\rangle$ and $|j_2 m_2\rangle$, what *total* angular momentum state(s) $|jm\rangle$ do we get? The z components still add, naturally, so

$$m = m_1 + m_2 \tag{4.11}$$

but the magnitudes do *not*; it all depends on the relative orientation of $\mathbf{J}_1$ and $\mathbf{J}_2$ (Figure 4.4). If they are parallel the magnitudes add, but if they are antiparallel the

* The terms 'fermion' and 'boson' refer to the rules for constructing composite wavefunctions for identical particles: boson wave functions must be *symmetric* under interchange of any two particles, fermion wavefunctions are *anti*symmetric. This leads to the Pauli exclusion principle for fermions, and to profound differences in the statistical mechanics of the two particle types. The 'connection between spin and statistics' (all fermions have half-integer spin and all bosons have integer spin) is a deep theorem in quantum field theory.

† I'll use the letter $\mathbf{J}$ for generic angular momentum – it could be orbital ($\mathbf{L}$), spin ($\mathbf{S}$), or some combined quantity.

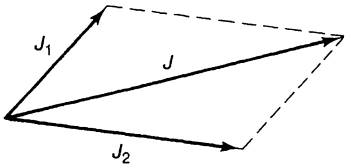


Fig. 4.4 Addition of angular momenta.

magnitudes subtract; in general, the magnitude of the vector sum is somewhere between these extremes. As it turns out, we get *every j from* $(j_1 + j_2)$ down to $|j_1 - j_2|$, in integer steps [4]:

$$j = |j_1 - j_2|, |j_1 - j_2| + 1, \ldots, (j_1 + j_2) - 1, (j_1 + j_2) \tag{4.12}$$

For instance, a particle of spin 1 in an orbital state $l = 3$ could have total angular momentum $j = 4$ (i.e. $J^2 = 20\hbar^2$), or $j = 3$ ($J^2 = 12\hbar^2$), or $j = 2$ ($J^2 = 6\hbar^2$).

Example 4.1 A quark and an antiquark are bound together, in a state of zero orbital angular momentum, to form a meson. *Question:* What are the possible values of the meson's spin?

Solution: Quarks (and therefore also antiquarks) carry spin $\frac{1}{2}$, so we can get $\frac{1}{2} + \frac{1}{2} = 1$ or $\frac{1}{2} - \frac{1}{2} = 0$. The spin-0 combination gives us the 'pseudoscalar' mesons (π's, K's, η, η') – 'scalar' means spin 0, 'pseudo-' will be explained shortly. The spin-1 combination gives the 'vector' mesons (ρ's, K^*'s, ϕ, ω) – 'vector' means spin 1.

To add *three* angular momenta, we combine two of them first, using Equation 4.12, and then add on the third. Thus, if we allow the quarks in Example 4.1 an orbital angular momentum $l > 0$, we get mesons with spin $l + 1$, l, and $l - 1$. Because the orbital quantum number has to be an integer, all mesons carry integer spin (they are bosons). By the same token, all baryons (made up of *three* quarks) must have half-integer spin (they are fermions).

Example 4.2 Suppose you combine three quarks in a state of zero orbital angular momentum. *Question:* What are the possible spins of the resulting baryon?

Solution: From *two* quarks, each spin $\frac{1}{2}$, we get a total angular momentum of $\frac{1}{2} + \frac{1}{2} = 1$ or $\frac{1}{2} - \frac{1}{2} = 0$. Adding in the third quark yields $1 + \frac{1}{2} = \frac{3}{2}$ or $1 - \frac{1}{2} = \frac{1}{2}$ (when the first two add to 1), and $0 + \frac{1}{2} = \frac{1}{2}$ (when the first two add to zero). Thus the baryon can have a spin of $\frac{3}{2}$ or $\frac{1}{2}$ (and the latter can be achieved in two different ways). In practice, $s = \frac{3}{2}$ is the decuplet, $s = \frac{1}{2}$ is the octet, and evidently, the quark model would allow for another family with $s = \frac{1}{2}$. (If we permit the quarks to revolve around one another, throwing in some *orbital* angular momentum, the number of possibilities increases accordingly – but the total will always be a half-integer).

Well, Equation 4.12 tells us what *total* angular momenta j we can obtain by combining j_1 and j_2, but occasionally we require the explicit decomposition of

$2 \times 1/2$												
			5/2									
			+5/2									
			5/2	3/2								
+2	1/2	1	3/2	+3/2								
	+2	−1/2	1/5	4/5	5/2	3/2						
	+1	+1/2	4/5	−1/5	+1/2	+1/2						
			1	−1/2	2/5	3/5	5/2	3/2				
			0	+1/2	3/5	−2/5	−1/2	−1/2				
					0	−1/2	3/5	2/5	5/2	3/2		
					−1	+1/2	2/5	−3/5	−3/2	−3/2		
							−1	−1/2	4/5	1/5	5/2	
							−2	+1/2	1/5	−4/5	−5/2	
									−2	−1/2	1	

Fig. 4.5 Clebsch–Gordan coefficients for $j_1 = 2$, $j_2 = \frac{1}{2}$. (A square root sign over each number is implied.)

$|j_1 m_1\rangle\,|j_2 m_2\rangle$ into specific states of total angular momentum, $|jm\rangle$:

$$|j_1 m_1\rangle|j_2 m_2\rangle = \sum_{j=|j_1-j_2|}^{(j_1+j_2)} C^{jj_1j_2}_{mm_1m_2}|jm\rangle, \quad \text{with } m = m_1 + m_2 \tag{4.13}$$

The numbers $C^{jj_1j_2}_{mm_1m_2}$ are known as *Clebsch–Gordan coefficients*. A book on advanced quantum mechanics will explain how to calculate them. In practice, we normally look them up in a table. (There is one in the *Particle Physics Booklet*, and the case $j_1 = 2$, $j_2 = \frac{1}{2}$ is reproduced in Figure 4.5) The Clebsch–Gordan coefficients tell you the probability of getting $j(j+1)\hbar^2$, for any particular allowed j, if we measure J^2 on a system consisting of two angular momentum states $|j_1 m_1\rangle$ and $|j_2 m_2\rangle$: the probability is the *square* of the corresponding Clebsch–Gordan coefficient.

Example 4.3 The electron in a hydrogen atom occupies the orbital state $|2\ -1\rangle$ and the spin state $|\frac{1}{2}\ \frac{1}{2}\rangle$. *Question:* If we measure J^2, what values might we get, and what is the probability of each?

Solution: The possible values of j are $l + s = 2 + \frac{1}{2} = \frac{5}{2}$ and $l - s = 2 - \frac{1}{2} = \frac{3}{2}$. The z components add: $m = -1 + \frac{1}{2} = -\frac{1}{2}$. We go to the Clebsch–Gordan table (Figure 4.5) labeled $2 \times \frac{1}{2}$, which indicates that we are combining $j_1 = 2$ with $j_2 = \frac{1}{2}$, and look for the horizontal row, labeled -1, $\frac{1}{2}$; these are the values of m_1 and m_2. Reading off the two entries, we find $|2\ -1\rangle|\frac{1}{2}\ \frac{1}{2}\rangle = \sqrt{\frac{2}{5}}\ |\frac{5}{2}\ -\frac{1}{2}\rangle - \sqrt{\frac{3}{5}}\ |\frac{3}{2}\ -\frac{1}{2}\rangle$. So the probability of getting $j = \frac{5}{2}$ is $\frac{2}{5}$, and the probability of getting $j = \frac{3}{2}$ is $\frac{3}{5}$. Notice that the probabilities add to 1, as, of course, they *must*.

Example 4.4 We know from Example 4.1 that two spin-$\frac{1}{2}$ states combine to give spin 1 and spin 0. *Problem:* Find the explicit Clebsch–Gordan decomposition for these states.

Solution: Consulting the $\frac{1}{2} \times \frac{1}{2}$ table, we find

$$\begin{aligned}
|\tfrac{1}{2}\,\tfrac{1}{2}\rangle|\tfrac{1}{2}\,\tfrac{1}{2}\rangle &= |1\,1\rangle \\
|\tfrac{1}{2}\,\tfrac{1}{2}\rangle|\tfrac{1}{2}\,-\tfrac{1}{2}\rangle &= (\tfrac{1}{\sqrt{2}})|1\,0\rangle + (\tfrac{1}{\sqrt{2}})|0\,0\rangle \\
|\tfrac{1}{2}\,-\tfrac{1}{2}\rangle|\tfrac{1}{2}\,\tfrac{1}{2}\rangle &= (\tfrac{1}{\sqrt{2}})|1\,0\rangle - (\tfrac{1}{\sqrt{2}})|0\,0\rangle \\
|\tfrac{1}{2}\,-\tfrac{1}{2}\rangle|\tfrac{1}{2}\,-\tfrac{1}{2}\rangle &= |1\,-1\rangle
\end{aligned} \tag{4.14}$$

Thus the three spin 1 states are

$$\left.\begin{aligned}
|1\,1\rangle &= |\tfrac{1}{2}\,\tfrac{1}{2}\rangle|\tfrac{1}{2}\,\tfrac{1}{2}\rangle \\
|1\,0\rangle &= (\tfrac{1}{\sqrt{2}})\left[|\tfrac{1}{2}\,\tfrac{1}{2}\rangle|\tfrac{1}{2}\,-\tfrac{1}{2}\rangle + |\tfrac{1}{2}\,-\tfrac{1}{2}\rangle|\tfrac{1}{2}\,\tfrac{1}{2}\rangle\right] \\
|1\,-1\rangle &= |\tfrac{1}{2}\,-\tfrac{1}{2}\rangle|\tfrac{1}{2}\,-\tfrac{1}{2}\rangle
\end{aligned}\right\} \tag{4.15}$$

whereas the spin 0 state is

$$|0\,0\rangle = (1/\sqrt{2})\left[|\tfrac{1}{2}\,\tfrac{1}{2}\rangle|\tfrac{1}{2}\,-\tfrac{1}{2}\rangle - |\tfrac{1}{2}\,-\tfrac{1}{2}\rangle|\tfrac{1}{2}\,\tfrac{1}{2}\rangle\right] \tag{4.16}$$

By the way, Equations (4.15) and (4.16) can be read directly off the Clebsch–Gordan table; the coefficients work both directions:

$$|jm\rangle = \sum_{j_1 j_2} C^{jj_1j_2}_{mm_1m_2}|j_1 m_1\rangle|j_2 m_2\rangle \tag{4.17}$$

This time we read down the columns, instead of along the rows. The spin-1 combination is called the 'triplet', for obvious reasons, and spin 0 is called the 'singlet'. For future reference, notice that the triplet is *symmetric* under interchange of the particles, $1 \leftrightarrow 2$, whereas the singlet is antisymmetric (that is, it changes sign). Incidentally, in a singlet state the spins are oppositely aligned (antiparallel); however, it is *not* the case that in a triplet state the spins are necessarily parallel; they are for $m = 1$ and $m = -1$, but not for $m = 0$.

4.2.2 Spin $\frac{1}{2}$

The most important spin system is $s = \frac{1}{2}$; the proton, neutron, electron, all quarks, and all leptons carry spin $\frac{1}{2}$. Furthermore, once you understand the formalism for $s = \frac{1}{2}$, any other case is a relatively simple matter to work out. So I will pause here to develop the theory of spin $\frac{1}{2}$ in some detail.

A particle with spin $\frac{1}{2}$ can have $m_s = \frac{1}{2}$ ('spin up'), or $m_s = -\frac{1}{2}$ ('spin down'). Informally, we represent these two states by arrows: $\uparrow$ and $\downarrow$. But a better notation is afforded by two-component column vectors, or *spinors*:

$$|\tfrac{1}{2}\,\tfrac{1}{2}\rangle = \begin{pmatrix}1\\0\end{pmatrix}, \quad |\tfrac{1}{2}\,-\tfrac{1}{2}\rangle = \begin{pmatrix}0\\1\end{pmatrix} \tag{4.18}$$

It is often said that a particle of spin $\frac{1}{2}$ can only exist in one or the other of these two states, but that is quite false. The most general state of a spin-$\frac{1}{2}$ particle is the *linear combination*

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4.19}$$

where α and β are two complex numbers. It is true that a *measurement* of S_z can only return the value $+\frac{1}{2}\hbar$ or $-\frac{1}{2}\hbar$, but the first outcome, say, does not prove that the particle was in the state $\uparrow$ prior to the measurement. In the general case (Equation 4.19), $|\alpha|^2$ is the *probability* that a measurement of S_z would yield the value $+\frac{1}{2}\hbar$, and $|\beta|^2$ is the probability of getting $-\frac{1}{2}\hbar$. Since these are the only allowed results, it follows that

$$|\alpha|^2 + |\beta|^2 = 1 \tag{4.20}$$

Apart from this 'normalization' condition, there is no *a priori* constraint on the numbers α and β.

Suppose now that we are to measure S_x or S_y on a particle in the generic state given by Equation 4.19 What results might we get, and what is the probability of each? Symmetry dictates that the allowed values be $\pm\frac{1}{2}\hbar$ – after all, it's perfectly arbitrary which direction we choose to call z in the first place. But determining the probabilities is not so simple. To each component of $\mathbf{S}$ we associate a 2×2 matrix:*

$$\hat{S}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{S}_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{S}_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.21}$$

The eigenvalues of $\hat{S}_x$ are $\pm\frac{\hbar}{2}$, and corresponding normalized eigenvectors are†

$$\chi_\pm = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \pm\frac{1}{\sqrt{2}} \end{pmatrix} \tag{4.22}$$

* Again, the *derivation* of these matrices will be found in any quantum-mechanics text. My purpose here is to show you *how* angular momentum is handled in particle physics, not to explain *why* it is done this way.

† A nonzero column matrix

$$\chi = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

is called an *eigenvector* of a given $n \times n$ matrix M if

$$M\chi = \lambda\chi$$

for some number λ (the *eigenvalue*). Notice that any *multiple* of χ is still an eigenvector, with the same eigenvalue.

(see Problem 4.15). An arbitrary spinor $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$ can be written as a linear combination of these eigenvectors:

$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = a\begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix} + b\begin{pmatrix}\frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}}\end{pmatrix} \tag{4.23}$$

where

$$a = \left(\frac{1}{\sqrt{2}}\right)[\alpha+\beta]; \quad b = \left(\frac{1}{\sqrt{2}}\right)[\alpha-\beta] \tag{4.24}$$

The probability that a measurement of S_x will yield the value $\frac{1}{2}\hbar$ is $|a|^2$; the probability of getting $-\frac{1}{2}\hbar$ is $|b|^2$. Evidently, $|a|^2+|b|^2=1$ (see Problem 4.16).

The general procedure, of which this was a particular instance, is as follows:

1. Construct the matrix, $\hat{A}$, representing the observable A in question.
2. The allowed values of A are the eigenvalues of $\hat{A}$.
3. Write the state of the system as a linear combination of eigenvectors of $\hat{A}$; the absolute square of the coefficient of the ith eigenvector is the probability that a measurement of A would yield the ith eigenvalue.

Example 4.5 Suppose we measure $(S_x)^2$ on a particle in the state $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$. *Question:* What values might we get, and what is the probability of each?

Solution. The matrix representing $(S_x)^2$ is the square of the matrix representing S_x:

$$\hat{S}_x^2 = \frac{\hbar^2}{4}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \tag{4.25}$$

Since

$$\frac{\hbar^2}{4}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \frac{\hbar^2}{4}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$$

every spinor is an eigenvector of $\hat{S}_x^2$, with eigenvalue $\frac{\hbar^2}{4}$. Thus we would be *certain* to get $\frac{\hbar^2}{4}$ (probability 1). The same goes for $\hat{S}_y^2$ and $\hat{S}_z^2$, so every spinor is an eigenstate of $\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2$, with eigenvalue $\frac{3\hbar^2}{4}$. This should come as no surprise – in general, for spin s we must have $S^2 = s(s+1)\hbar^2$. ▒▒

For mathematical purposes, the factor of $\frac{\hbar}{2}$ in Equation 4.21 is ugly, and it is customary to introduce the *Pauli spin matrices*:

$$\sigma_x = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \quad \sigma_y = \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}, \quad \sigma_z = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \tag{4.26}$$

so that $\hat{\mathbf{S}} = (\frac{\hbar}{2})\sigma$. The Pauli matrices have many interesting properties, some of which are explored in Problems 4.19 and 4.20. We shall encounter them repeatedly in the course of this book.

In a sense, spinors (two-component objects) occupy an intermediate position between scalars (one component) and vectors (three components). Now, when you rotate your coordinate axes, the components of a vector change, in a prescribed manner (see Problem 4.6), and we might inquire how the components of a *spinor* transform, under the same circumstances. The answer [5] is provided by the following rule:

$$\begin{pmatrix} \alpha' \\ \beta' \end{pmatrix} = U(\boldsymbol{\theta}) \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{4.27}$$

where $U(\boldsymbol{\theta})$ is the 2×2 matrix

$$U(\boldsymbol{\theta}) = \mathrm{e}^{-i(\boldsymbol{\theta}\cdot\boldsymbol{\sigma})/2} \tag{4.28}$$

The vector $\boldsymbol{\theta}$ points along the axis of rotation, and its magnitude is the angle of rotation, in the right-hand sense, about that axis. Notice that the exponent here is itself a matrix! An expression of this form is to be interpreted as shorthand for the power series:

$$\mathrm{e}^A \equiv 1 + A + \tfrac{1}{2}A^2 + \tfrac{1}{3!}A^3 + \cdots \tag{4.29}$$

(see Problem 4.21).* As you can check for yourself (Problem 4.22), $U(\boldsymbol{\theta})$ is a unitary matrix of determinant 1; in fact, the set of all such rotation matrices constitutes the group $SU(2)$. Thus spin-$\frac{1}{2}$ particles transform under rotations according to the two-dimensional representation of $SU(2)$. Similarly, particles of spin 1, described by *vectors*, belong to the three-dimensional representation of $SU(2)$; spin-$\frac{3}{2}$ particles, described by a four-component object, transform under the four-dimensional representation of $SU(2)$; and so on. (The construction of these higher-dimensional representations is explored in Problem 4.23.) You're probably wondering what $SU(2)$ has to do with rotations; well, as I mentioned earlier, $SU(2)$ is essentially† the same group as $SO(3)$, the group of rotations in three dimensions. Particles of different spin, then, belong to different *representations* of the rotation group.

* *Beware*: For matrices it is *not* the case that $\mathrm{e}^A\mathrm{e}^B = \mathrm{e}^{A+B}$, in general. You might want to check this by using the matrices in Problem 4.21. However, the usual rule *does* apply if A and B *commute* (i.e. if $AB = BA$).

† There is actually a subtle distinction between $SU(2)$ and $SO(3)$. According to Problem 4.21, the matrix U for rotation through an angle of 2π is -1; a spinor *changes sign* under such a rotation. And yet, *geometrically*, a rotation through 2π is equivalent to no rotation at all. $SU(2)$ is a kind of 'doubled' version of $SO(3)$, in which you don't come back to the beginning until you have turned through 720°. In this sense, spinor representations of $SU(2)$ are not 'true' representations of the rotation group, and that's why they do not appear in classical physics. In quantum mechanics only the *square* of the wave function carries physical significance, and in the squaring the minus sign goes away.

4.3 Flavor Symmetries

There's an extraordinary thing about the neutron, which Heisenberg observed shortly after its discovery in 1932: apart from the obvious fact that it carries no charge, it is almost identical to the proton. In particular, their masses are astonishingly close, $m_p = 938.28\ \mathrm{MeV}/c^2, m_n = 939.57\ \mathrm{MeV}/c^2$. Heisenberg [6] proposed that we regard them as two 'states' of a *single* particle, the *nucleon*. Even the small difference in mass might be attributed to the fact that the proton is charged, since the energy stored in its electric field contributes, according to Einstein's formula $E = mc^2$, to its inertia. (Unfortunately, this argument suggests that the *proton* should be the heavier of the two, which is not only untrue, but would be disastrous for the stability of matter. More on this in a moment.) If we could somehow 'turn off' all electric charge, the proton and neutron would, according to Heisenberg, be indistinguishable. Or, to put it more prosaically, the *strong* forces experienced by protons and neutrons are identical.

To implement Heisenberg's idea, we write the nucleon as a two-component column matrix

$$N = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{4.30}$$

with

$$p = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad n = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4.31}$$

This is nothing but notation, of course, but it is notation seductively reminiscent of the spinors we encountered in the theory of angular momentum. By direct analogy with spin, **S**, we are led to introduce *isospin*, **I**.* However, **I** is not a vector in *ordinary* space, with components along the coordinate directions x, y, and z, but rather in an abstract 'isospin space', with components we will call I_1, I_2, and I_3. On this understanding, we may borrow the entire apparatus of angular momentum, as developed earlier in the chapter. The nucleon carries isospin $\frac{1}{2}$, and the third component, I_3, has the eigenvalues† $+\frac{1}{2}$ (the proton) and $-\frac{1}{2}$ (the neutron):

$$p = |\tfrac{1}{2}\ \tfrac{1}{2}\rangle, \qquad n = |\tfrac{1}{2}\ -\tfrac{1}{2}\rangle \tag{4.32}$$

The proton is 'isospin up'; the neutron is 'isospin down'.

This is still just notation; the *physics* comes in Heisenberg's proposition that the *strong interactions are invariant under rotations in isospin space*, just as, for example, electrical forces are invariant under rotations in ordinary configuration space. We

* The word derives from the misleading older term *isotopic spin* (introduced by Wigner in 1937). Nuclear physicists use the (better) word *isobaric* spin.

† There is no factor of $\hbar$ in this case; isospin is dimensionless.

call this an *internal symmetry*, because it has nothing to do with space and time, but rather with the relations between different particles. A rotation through 180° about axis number 1 in isospin space converts protons into neutrons, and vice versa. If the strong force is invariant under rotations in isospin space, it follows, by Noether's theorem, that *isospin is conserved in all strong interactions*, just as angular momentum is conserved in processes with rotational invariance in ordinary space.*

In the language of group theory, Heisenberg asserted that the strong interactions are invariant under an internal symmetry group $SU(2)$, and the nucleons belong to the two-dimensional representation (isospin $\frac{1}{2}$). In 1932, this was a bold suggestion; today the evidence is all around us, most conspicuously in the 'multiplet' structure of the hadrons. Recall the Eightfold Way diagrams in Chapter 1: the horizontal rows all display exactly the feature that caught Heisenberg's eye in the case of the nucleons; they have very similar masses but different charges. To each of these multiplets, we assign a particular isospin I, and to each member of the multiplet, we assign a particular I_3. For the pions, $I = 1$:

$$\pi^+ = |1\,1\rangle, \quad \pi^0 = |1\,0\rangle, \quad \pi^- = |1\,-1\rangle \tag{4.33}$$

for the Λ, $I = 0$:

$$\Lambda = |0\,0\rangle \tag{4.34}$$

for the Δ's, $I = \frac{3}{2}$:

$$\Delta^{++} = |\tfrac{3}{2}\,\tfrac{3}{2}\rangle, \quad \Delta^+ = |\tfrac{3}{2}\,\tfrac{1}{2}\rangle, \quad \Delta^0 = |\tfrac{3}{2}\,-\tfrac{1}{2}\rangle, \quad \Delta^- = |\tfrac{3}{2}\,-\tfrac{3}{2}\rangle \tag{4.35}$$

and so on. To determine the isospin of a multiplet, just count the number of particles it contains; since I_3 ranges from $-I$ to $+I$, in integer steps, the number of particles in the multiplet is $2I + 1$:

$$\text{multiplicity} = 2I + 1 \tag{4.36}$$

The third component of isospin, I_3, is related to the charge, Q, of the particle. We assign the maximum value, $I_3 = I$, to the member of the multiplet with the highest charge, and fill in the rest in order of decreasing Q. For the 'pre-1974' hadrons – those composed of u, d, and s quarks only – the explicit relation between Q and I_3 is the *Gell-Mann–Nishijima formula*:

$$Q = I_3 + \tfrac{1}{2}(A + S) \tag{4.37}$$

* It is tempting to overstate the so-called 'charge independence' of the strong forces (the fact that they are the same for protons as for neutrons). It does *not* say that you will get the same result if you substitute an *individual* proton for a neutron, only if you interchange *all* protons and neutrons. (For example, there exists a bound state of the proton and the neutron – the deuteron – but there is *no* bound state of two protons or two neutrons.) Indeed, any such assertion would be incompatible with the Pauli exclusion principle, since a proton and a neutron can be in the same quantum state, but two neutrons (or two protons) cannot.

where A is the baryon number and S is the strangeness.* Originally, this equation was a purely empirical observation, but in the context of the quark model it follows simply from the isospin assignments for quarks: u and d form a 'doublet' (like the proton and the neutron):

$$u = |\tfrac{1}{2}\,\tfrac{1}{2}\rangle, \quad d = |\tfrac{1}{2}\,-\tfrac{1}{2}\rangle \tag{4.38}$$

and all the other flavors carry isospin zero† (see Problems 4.25 and 4.26).

But classification is not all that isospin does for us. It also has important *dynamical* implications. For example, suppose we have *two* nucleons. From the rules for addition of angular momenta we know that the combination gives a total isospin of 1 or 0. Specifically (using Example 4.4), we obtain a symmetric isotriplet:

$$\begin{aligned} |1\,1\rangle &= pp \\ |1\,0\rangle &= \left(\tfrac{1}{\sqrt{2}}\right)(pn + np) \\ |1\,-1\rangle &= nn \end{aligned} \tag{4.39}$$

and an antisymmetric isosinglet:

$$|0\,0\rangle = \left(\tfrac{1}{\sqrt{2}}\right)(pn - np) \tag{4.40}$$

Experimentally, the neutron and proton form a single bound state, the deuteron (d); there is no bound state of two protons or of two neutrons. Thus the deuteron must be an isosinglet. If it were a triplet, all three states would have to occur, since they differ only by a rotation in isospin space. Evidently, there is a strong attraction in the $I = 0$ channel, but not in the $I = 1$ channel. Presumably, the potential describing the interaction between two nucleons contains a term of the form $\mathbf{I}^{(1)} \cdot \mathbf{I}^{(2)}$, which takes the value $\frac{1}{4}$ in the triplet configuration and $-\frac{3}{4}$ in the singlet (see Problem 4.27).

Isospin invariance has implications, too, for nucleon–nucleon scattering. Consider the processes

$$\begin{aligned} &\text{(a)}\ p + p \to d + \pi^+ \\ &\text{(b)}\ p + n \to d + \pi^0 \\ &\text{(c)}\ n + n \to d + \pi^- \end{aligned} \tag{4.41}$$

Since the deuteron carries $I = 0$, the isospin states on the right are $|1\ 1\rangle$, $|1\ 0\rangle$, and $|1{-}1\rangle$, respectively, whereas those on the left are $pp = |11\rangle$, $nn = |1\ -1\rangle$,

* Since Q, A, and S are all conserved by the electromagnetic forces, it follows that I_3 is also conserved. However, the other two components (I_1 and I_2), and hence also I itself, are *not* conserved in electromagnetic interactions. For example, in the decay $\pi^0 \to \gamma + \gamma$, I goes from 1 to 0. As for the weak interactions, they don't even conserve S, so I_3 is not conserved in weak processes (for example, $\Lambda \to p + \pi^-$ takes $I_3 = 0$ to $I_3 = -\frac{1}{2}$).

† Since isospin pertains only to the strong forces, it is not a relevant quantity for leptons. For consistency, all leptons and mediators are assigned isospin zero.

and $pn = (\frac{1}{\sqrt{2}})(|1\ 0\rangle + |0\ 0\rangle)$.* Only the $I = 1$ combination contributes (since the final state in each case is pure $I = 1$, and isospin is conserved), so the *scattering amplitudes* are in the ratio

$$\mathscr{M}_a : \mathscr{M}_b : \mathscr{M}_c = 1 : \left(\frac{1}{\sqrt{2}}\right) : 1 \tag{4.42}$$

As we shall see,† the *cross section*, σ, goes like the absolute square of the amplitude; thus

$$\sigma_a : \sigma_b : \sigma_c = 2 : 1 : 2 \tag{4.43}$$

Process (c) would be hard to set up in the laboratory, but (a) and (b) have been measured, and (when corrections are made for electromagnetic effects) they are found to be in the predicted ratio [7].

As a final example, let's consider pion–nucleon scattering, $\pi N \to \pi N$. There are six elastic processes:

$$\begin{aligned}
&(a)\ \pi^+ + p \to \pi^+ + p && (b)\ \pi^0 + p \to \pi^0 + p \\
&(c)\ \pi^- + p \to \pi^- + p && (d)\ \pi^+ + n \to \pi^+ + n \\
&(e)\ \pi^0 + n \to \pi^0 + n && (f)\ \pi^- + n \to \pi^- + n
\end{aligned} \tag{4.44}$$

and four charge-exchange processes:

$$\begin{aligned}
&(g)\ \pi^+ + n \to \pi^0 + p && (h)\ \pi^0 + p \to \pi^+ + n \\
&(i)\ \pi^0 + n \to \pi^- + p && (j)\ \pi^- + p \to \pi^0 + n
\end{aligned} \tag{4.45}$$

Since the pion carries $I = 1$, and the nucleon $I = \frac{1}{2}$, the total isospin can be $\frac{3}{2}$ or $\frac{1}{2}$. So there are just two distinct amplitudes here: $\mathscr{M}_3$, for $I = \frac{3}{2}$, and $\mathscr{M}_1$, for $I = \frac{1}{2}$. From the Clebsch–Gordan tables we find the following decompositions:

$$\begin{aligned}
\pi^+ + p\text{:}\quad & |1\,1\rangle|\tfrac{1}{2}\,\tfrac{1}{2}\rangle && = |\tfrac{3}{2}\,\tfrac{3}{2}\rangle \\
\pi^0 + p\text{:}\quad & |1\,0\rangle|\tfrac{1}{2}\,\tfrac{1}{2}\rangle && = \sqrt{\tfrac{2}{3}}\,|\tfrac{3}{2}\,\tfrac{1}{2}\rangle - \left(\tfrac{1}{\sqrt{3}}\right)|\tfrac{1}{2}\,\tfrac{1}{2}\rangle \\
\pi^- + p\text{:}\quad & |1\,-1\rangle|\tfrac{1}{2}\,\tfrac{1}{2}\rangle && = \left(\tfrac{1}{\sqrt{3}}\right)|\tfrac{3}{2}\,-\tfrac{1}{2}\rangle - \sqrt{\tfrac{2}{3}}\,|\tfrac{1}{2}\,-\tfrac{1}{2}\rangle \\
\pi^+ + n\text{:}\quad & |1\,1\rangle|\tfrac{1}{2}\,-\tfrac{1}{2}\rangle && = \left(\tfrac{1}{\sqrt{3}}\right)|\tfrac{3}{2}\,\tfrac{1}{2}\rangle + \sqrt{\tfrac{2}{3}}\,|\tfrac{1}{2}\,\tfrac{1}{2}\rangle \\
\pi^0 + n\text{:}\quad & |1\,0\rangle|\tfrac{1}{2}\,-\tfrac{1}{2}\rangle && = \sqrt{\tfrac{2}{3}}\,|\tfrac{3}{2}\,-\tfrac{1}{2}\rangle + \left(\tfrac{1}{\sqrt{3}}\right)|\tfrac{1}{2}\,-\tfrac{1}{2}\rangle \\
\pi^- + n\text{:}\quad & |1\,-1\rangle|\tfrac{1}{2}\,-\tfrac{1}{2}\rangle && = |\tfrac{3}{2}\,-\tfrac{3}{2}\rangle
\end{aligned} \tag{4.46}$$

* Add Equations 4.39 and 4.40

† The theory of scattering amplitudes and cross sections will be developed in Chapter 6. In this and the following paragraph, I anticipate later results, but I hope it is clear from the context how the calculation proceeds. If you wish, skip these two paragraphs for now.

Reactions (a) and (f) are pure $I = \frac{3}{2}$:

$$\mathscr{M}_a = \mathscr{M}_f = \mathscr{M}_3 \tag{4.47}$$

The others are all mixtures; for example,

$$\mathscr{M}_c = \tfrac{1}{3}\mathscr{M}_3 + \tfrac{2}{3}\mathscr{M}_1, \qquad \mathscr{M}_j = \left(\tfrac{\sqrt{2}}{3}\right)\mathscr{M}_3 - \left(\tfrac{\sqrt{2}}{3}\right)\mathscr{M}_1 \tag{4.48}$$

(I'll let you work out the rest, see Problem 4.28). The cross sections, then, stand in the ratio

$$\sigma_a : \sigma_c : \sigma_j = 9|\mathscr{M}_3|^2 : |\mathscr{M}_3 + 2\mathscr{M}_1|^2 : 2|\mathscr{M}_3 - \mathscr{M}_1|^2 \tag{4.49}$$

At a CM energy of 1232 MeV, there occurs a famous and dramatic bump in pion–nucleon scattering, first discovered by Fermi *et al.* in 1951 [8]; here the pion and nucleon join to form a short-lived 'resonance' state – the Δ. We know the Δ carries $I = \frac{3}{2}$, so we expect that at this energy $\mathscr{M}_3 \gg \mathscr{M}_1$, and hence

$$\sigma_a : \sigma_c : \sigma_j = 9 : 1 : 2 \tag{4.50}$$

Experimentally, it is easier to measure the *total* cross sections, so (c) and (j) are combined:

$$\frac{\sigma_{\text{tot}}(\pi^+ + p)}{\sigma_{\text{tot}}(\pi^- + p)} = 3 \tag{4.51}$$

As you can see in Figure 4.6, this prediction is well satisfied by the data.

In the late 1950s history repeated itself. Just as in 1932 the proton and neutron were seen to form a pair, it was now increasingly clear that the nucleons, the Λ, the Σ's, and the Ξ's together, constituted a natural grouping within the baryon family. They all carry spin $\frac{1}{2}$, and their masses are similar. It is true that the latter range from 940 MeV/c^2, for the nucleons, up to 1320 MeV/c^2, for the Ξ's, so it would be stretching things a bit to argue that they are all different states of one particle, as Heisenberg had suggested for the proton and neutron. Nevertheless, it was tempting to regard these eight baryons as a *supermultiplet*, and this presumably meant that they belonged in the same representation of some enlarged symmetry group, in which the $SU(2)$ of isospin would be incorporated as a *subgroup*. The critical question became: what is the larger group? (The 'Eight Baryon Problem', as it was called, was not always phrased this way; at the time, most physicists were surprisingly ignorant of group theory. Gell-Mann worked out most of the formalism he needed from scratch, and only later learned that it was well known to mathematicians.) The Eightfold Way was Gell-Mann's solution to the Eight Baryon Problem. The symmetry group is $SU(3)$; the octets constitute eight-dimensional representations of $SU(3)$, the decuplet a 10-dimensional representation, and so on. One thing that made this case more difficult than Heisenberg's was that no naturally

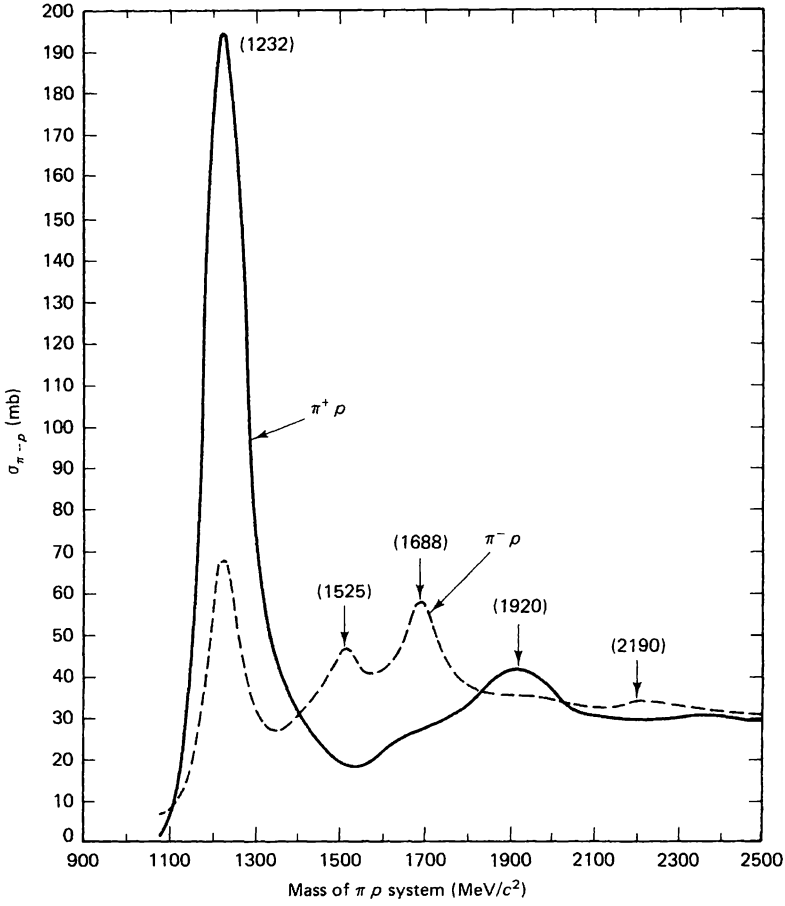


Fig. 4.6 Total cross sections for π^+p (solid line) and π^-p (dashed line) scattering. (*Source*: Gasiorowicz, S. (1966) *Elementary Particle Physics*, John Wiley & Sons, New York, p. 294. Reprinted by permission of John Wiley and Sons, Inc.)

occurring particles fall into the fundamental (three-dimensional) representation of $SU(3)$, as the nucleons, and later the K's, the Ξ's, and so on, do for $SU(2)$. This role was reserved for the quarks: u, d, and s together form a three-dimensional representation of $SU(3)$, which breaks down into an isodoublet (u, d) and an isosinglet (s) under $SU(2)$.

Of course, when the charmed quark came along, the flavor symmetry group of the strong interactions expanded once again – this time to $SU(4)$ (some $SU(4)$ supermultiplets are shown in Figure 1.13). But things barely paused there before the arrival of the bottom quark, taking us to $SU(5)$, and finally the top quark, $SU(6)$.

Table 4.4 Quark masses (MeV/c^2)

Quark flavor	Bare mass	Effective mass
u	2	336
d	5	340
s	95	486
c	1300	1550
b	4200	4730
t	174 000	177 000

Warning: These numbers are somewhat speculative and model dependent [12].

However, there is an important *caveat* in this neat hierarchy: isospin, $SU(2)$, is a very 'good' symmetry; the members of an isospin multiplet differ in mass by at most 2 or 3%, which is about the level at which electromagnetic corrections would be expected.* But the Eightfold Way, $SU(3)$, is a badly 'broken' symmetry; mass splittings within the baryon octet are around 40%. The symmetry breaking is even worse when we include charm; the $\Lambda_c^+(udc)$ weighs more than twice the $\Lambda(uds)$, although they are in the same $SU(4)$ supermultiplet. It is worse still with bottom, and absolutely terrible with top, which doesn't form bound states at all.

Why is isospin such a good symmetry, the Eightfold Way fair, and flavor $SU(6)$ so poor? The Standard Model blames it all on the quark masses. Now, the theory of quark masses is a slippery business, given the fact that they are not accessible to direct experimental measurement. Various arguments [9] suggest that the u and d quarks are *intrinsically* very light, about 10 times the mass of the electron. However, within the confines of a hadron, their *effective* mass is much greater. The precise value, in fact, depends on the context; it tends to be a little higher in baryons than in mesons (more on this in Chapter 5). In somewhat the same way, the effective inertia of a spoon is greater when you're stirring honey than when you're stirring tea, and in either case it exceeds the true mass of the spoon. Generally speaking, the effective mass of a quark in a hadron is about 350 MeV/c^2 greater than its bare mass [10] (see Table 4.4). Compared to this, the quite different *bare* masses of up and down quarks are practically irrelevant; they *function* as though they had identical masses. But the s quark is distinctly heavier, and the c, b, and t quarks are widely separated. Apart from the differences in quark masses, the strong interactions treat all flavors equally. Thus isospin is a good symmetry because the effective u and d masses are so nearly equal (which is to say, on a more fundamental level, because their *bare* masses are so small); the Eightfold Way is a *fair* symmetry because the effective mass of the strange quark is not *too* far from that of the u and d. But

* Indeed, it used to be thought that isospin was an *exact* symmetry of the strong interactions, and *all* of the symmetry breaking was attributable to electromagnetic contamination. The fact that the n–p mass splitting is in the wrong direction to be purely electromagnetic was troubling, however, and we now believe that $SU(2)$ is only an *approximate* symmetry of the strong interactions.

the heavy quarks are so far apart that their flavor symmetry is severely broken. Of course, this 'explanation' raises two further questions: (i) *Why* does the binding of quarks into hadrons increase their effective mass by about 350 MeV/c^2? The answer presumably lies within QCD, but the details are not fully understood [11]. (ii) Why do the *bare* quarks have the particular masses they do? Is there some pattern here? To this question, the Standard Model offers no answer; the six bare quark masses, and also the six lepton masses, are simply input parameters, for now, and it is the business of theories *beyond* the Standard Model to say where they come from.

4.4 Discrete Symmetries

4.4.1 Parity

Prior to 1956, it was taken for granted that the laws of physics are ambidextrous; that is, the mirror image of any physical process also represents a perfectly possible physical process [13]. To be sure, we drive on the right (at least, Americans do) and our hearts are on the left, but these are obviously historical or evolutionary accidents; it could just as well have been the other way around. Indeed, most physicists held the mirror symmetry (or 'parity invariance') of the laws of nature to be self-evident. But in 1956, Lee and Yang [14] were led to wonder (for reasons we will come back to at the end of this section) whether there had been any *experimental* test of this assumption. Searching the literature, they were surprised to discover that although there was ample evidence for parity invariance in strong and electromagnetic processes, there was no confirmation in the case of *weak* interactions. They proposed a test, which was carried out later that year by Wu [15], to settle the issue. In this famous experiment, radioactive cobalt 60 nuclei were carefully aligned, so that their spins pointed in, say, the z direction (Figure 4.7). Cobalt 60 undergoes beta decay ($^{60}\text{Co} \rightarrow {}^{60}\text{Ni} + e + \overline{\nu}_e$), and Wu recorded the direction of the emitted electrons. What she found was that most of them came out in the 'southerly' direction, *opposite* to the nuclear spin.

That's all there was to it. But that simple observation had astonishing implications. For suppose we examine the mirror image of that same process (Figure 4.8). The image nucleus rotates in the opposite direction; its spin points *downward*. And yet, the electrons (in the mirror) still came off downward. In the mirror, then, the electrons are emitted preferentially in the *same* direction as the nuclear spin. Here, then, is a physical process whose mirror image does *not* occur in nature; evidently parity is *not* an invariance of the weak interactions. If it *were*, the electrons in Wu's experiment would have to come out in equal numbers, 'north' and 'south', but they *don't*.

The overthrow of parity had a profound effect on physicists – devastating to some, exhilarating to others [16]. The violation is not a small effect; as we shall

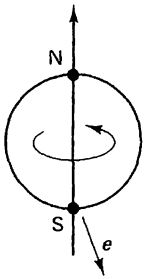


Fig. 4.7 In the beta decay of cobalt 60, most electrons are emitted in the direction opposite to the nuclear spin.

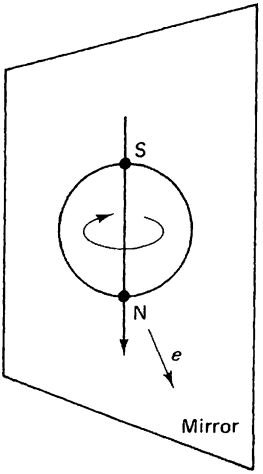


Fig. 4.8 Mirror image of Figure 4.7: Most electrons are emitted parallel to the nuclear spin.

see in Chapter 9, it is in fact 'maximal'. Nor is it limited to beta decay in cobalt; once you *look* for it, parity violation is practically the *signature* of the weak force. It is most dramatically revealed in the behavior of the neutrino. In the theory of angular momentum, the axis of quantization is, by convention, the z axis. Of course, the orientation of the z axis is completely up to us, but if we are dealing with a particle traveling through the laboratory at velocity $\mathbf{v}$, a natural choice suggests itself: why not pick the *direction of motion* as the z axis? The value of m_s/s for this axis is called the *helicity* of the particle. Thus a particle of spin $\frac{1}{2}$ can have a helicity of $+1(m_s = \frac{1}{2})$ or $-1(m_s = -\frac{1}{2})$; we call the former 'right-handed' and the latter 'left-handed.'* The difference is not terribly profound, however, because it is not Lorentz-invariant. Suppose I have a right-handed electron going to the right (Figure 4.9a), and someone else looks at it from an inertial system traveling to the right at a speed *greater* than v. From *his* perspective, the electron is going to the *left* (Figure 4.9b); but it is still spinning the same way, so this observer will say it's a *left*-handed electron. In other words, you can convert a right-handed electron into a left-handed one simply by changing your frame of reference. That's what I mean, when I say the distinction is not Lorentz-invariant.

But what if we applied that same reasoning to a *neutrino* – taken, for the moment, to be *massless*, so it travels at the speed of light, and hence there *is* no observer traveling faster? It is *impossible* to 'reverse the direction of motion' of a (massless) neutrino by getting into a faster-moving reference system, and therefore the helicity

* In Chapter 9, I shall introduce a technical distinction between 'handedness' and *helicity*, but for the moment I will use the terms interchangeably.

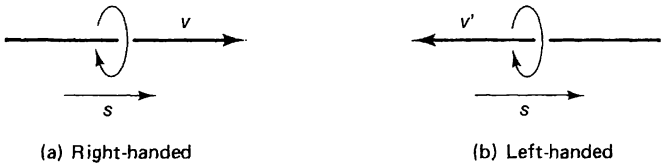


Fig. 4.9 Helicity. In (a) the spin and velocity are parallel (helicity +1); in (b) they are antiparallel (helicity −1).

of a neutrino (or any other massless particle)* is Lorentz-invariant – a fixed and fundamental property, which is *not* an artifact of the observer's reference frame. It becomes an important *experimental* matter to determine the helicity of a given neutrino. Until the mid-fifties, everyone assumed that half of all neutrinos would be left-handed, and half right-handed, just like photons. What they in fact *discovered* was that

NEUTRINOS ARE LEFT-HANDED;
ANTINEUTRINOS ARE RIGHT-HANDED.

Of course, it's tough to measure the helicity of a neutrino directly; they're hard enough to detect at *all*. There is, however, a relatively easy indirect method, using the decay of the pion: $\pi^- \to \mu^- + \bar{\nu}_\mu$. If the pion is at rest, the muon and the antineutrino come out back to back (Figure 4.10). Moreover, since the pion has spin 0, the muon and the antineutrino spins must be oppositely aligned.† Therefore, if the antineutrino is right-handed, the muon must be right-handed too (in the pion rest frame) – and this is precisely what is found experimentally [17]. Measurement of the *muon* helicity, then, enables us to determine the *antineutrino* helicity. By the same token, in π^+ decay, the antimuon is always left-handed, and this indicates that the neutrino is left-handed. By contrast, consider the decay of the *neutral* pion, $\pi^0 \to \gamma + \gamma$. Once again, in any given decay the two photons must have the same helicity. But this is an *electromagnetic* process, which respects parity, and thus, on the average, we get just as many right-handed photon pairs as left-handed pairs. Not so for neutrinos; they only interact weakly, and every one is left-handed; the

Fig. 4.10 Decay of π^- at rest.

* For massless particles, only the maximal value of $|m_s|$ occurs. For example, the photon can have $m_s = +1$ or $m_s = -1$, but not $m_s = 0$. So the helicity of a massless particle is always ± 1. In the case of the photon, these represent states of left- and right-circular polarization. The absence of $m_s = 0$ corresponds to the absence of longitudinal polarization in classical optics.

† The orbital angular momentum (if there is any) points perpendicular to the outgoing velocities, so it does not affect this argument.

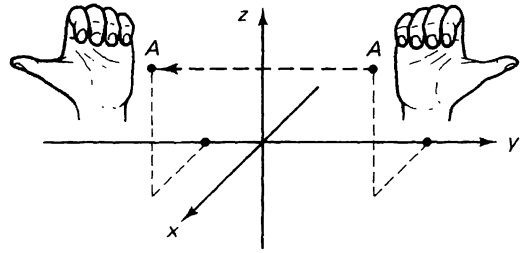
mirror image of a neutrino does not exist.* That is about the starkest violation of mirror symmetry you could ask for.†

In spite of its violation in weak processes, parity invariance remains a valid symmetry of the strong and electromagnetic interactions. It is useful, therefore, to develop some formalism and terminology. First, a minor technical point: Instead of *reflections*, which oblige us to choose arbitrarily the plane of the 'mirror', we will talk about *inversions*, in which every point is carried through the origin to the diametrically opposite location (Figure 4.11). Both transformations turn a right hand into a left hand; in fact, an inversion is nothing but a reflection followed by a rotation (180° about the y axis, in the figure). Thus in the cases of interest (which *also* possess rotational symmetry) it is a matter of indifference which one is used. Let P denote inversion; we call it the 'parity operator'. If the system in question is a *right* hand, P turns it into an upside-down and backward *left* hand (Figure 4.11b). When applied to a vector, $\mathbf{a}$, P produces a vector pointing in the opposite direction: $P(\mathbf{a}) = -\mathbf{a}$. How about the cross product of two vectors: $\mathbf{c} = \mathbf{a} \times \mathbf{b}$? Well, if P changes the sign of $\mathbf{a}$ and of $\mathbf{b}$, then $\mathbf{c}$ itself does *not* change sign: $P(\mathbf{c}) = \mathbf{c}$. Very strange! Evidently, there are two kinds of vectors – 'ordinary' ones, which change sign under the parity transformation, and this other type, of which the cross product is the classic example, which do *not*. We call the former 'polar' vectors, when the distinction must be drawn, and the latter 'pseudo' (or 'axial') vectors. Notice that the cross product of a polar vector with a *pseudo* vector would be a *polar* vector.

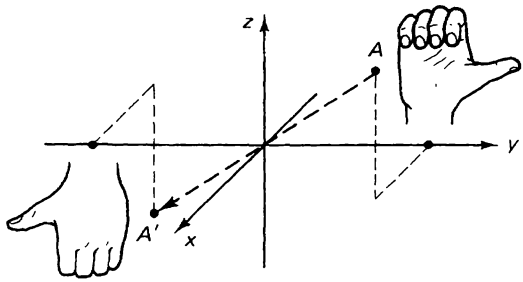
You have encountered pseudovectors before, though probably without using this language; angular momentum is one, and so is the magnetic field. In a theory with parity invariance, you must never add a vector to a pseudovector. Consider, for example, in the Lorentz force law: $\mathbf{F} = q[\mathbf{E} + (\mathbf{v} \times \mathbf{B})/c]$; $\mathbf{v}$ is a vector, and $\mathbf{B}$ is a pseudovector, so $\mathbf{v} \times \mathbf{B}$ is a vector, and it is legal to add it to $\mathbf{E}$. But $\mathbf{B}$ itself could never

* This is too strong a statement. There could, I suppose, be right-handed neutrinos around, but they do not interact with ordinary matter by any mechanism presently known. In fact, since we now know that neutrinos have a small but nonzero mass, right-handed neutrinos *must* exist. None of this, however, alters the fact that when a π^- decays, the emerging μ^- is right-handed in the *CM* frame and that by itself destroys mirror symmetry.
By the way, back in 1929, shortly after the publication of Dirac's equation, Weyl presented a beautifully simple theory of massless particles of spin $\frac{1}{2}$, which had the feature that they carried a fixed 'handedness'. At the time, Weyl's theory aroused limited interest, since there were no massless particles known, except for the photon, which carries spin 1. When Pauli introduced the neutrino, in 1931, you might suppose that he would dust off Weyl's theory and put it to use. He did not. Pauli rejected Weyl's theory out of hand, on the ground that it violated mirror symmetry. He lived to regret this mistake, and in 1957, Weyl's theory was triumphantly vindicated.

† It may occur to you, as it did to many physicists at the time, that if we simultaneously convert all particles into their antiparticles, then a kind of mirror symmetry is restored; the image of $\pi^- \to \mu^- + \bar{\nu}_\mu$ (with a right-handed antineutrino) becomes $\pi^+ \to \mu^+ + \nu_\mu$ (with a left-handed neutrino), which is perfectly okay. This realization was some comfort, until 1964, when it, too, was shown to fail. More on this in the following sections.



(a) Reflection (in the x-z plane)
$(x, y, z) \to (x, -y, z)$



(b) Inversion $(x, y, z) \to (-x, -y, -z)$

Fig. 4.11 Reflections and inversions.

be added to **E**. As we shall see, it is precisely the addition of a vector to a pseudovector in the theory of weak interactions that leads to the breakdown of parity.

Finally, the *dot* product of two polar vectors does not change sign under P, but the dot product of a polar vector and a pseudovector (or the triple product of three vectors: $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$) *does* change sign. So there are two kinds of *scalars*, too: the 'ordinary' kind, which don't change sign, and 'pseudoscalars,' which do. All this is summarized in Table 4.5.*

If you apply the parity operator *twice*, of course, you're right back where you started:

$$P^2 = I \tag{4.52}$$

(The parity group, then, consists of just two elements: I and P.) It follows that the eigenvalues of P are ± 1 (Problem 4.34). For example, scalars and pseudovectors have eigenvalue $+1$, whereas vectors and pseudoscalars have eigenvalue -1. The hadrons are eigenstates of P and can be classified according to their eigenvalue,

* The terminology extends very simply to special relativity: $a^\mu = (a^0, \mathbf{a})$ is called a pseudovector if its *spatial* components constitute a pseudovector $P(\mathbf{a}) = \mathbf{a}$; p is a pseudoscalar if it goes into minus itself under spatial inversions $P(p) = -p$.

Table 4.5 Scalars and vectors under parity

Scalar	:	$P(s) = s$
Pseudoscalar	:	$P(p) = -p$
Vector (or polar vector)	:	$P(\mathbf{v}) = -\mathbf{v}$
Pseudovector (or axial vector)	:	$P(\mathbf{a}) = \mathbf{a}$

just as they are classified by spin, charge, isospin, strangeness, and so on. According to quantum field theory, the parity of a fermion (half-integer spin) must be *opposite* to that of the corresponding antiparticle, while the parity of a boson (integer spin) is the *same* as its antiparticle. We take the *quarks* to have *positive* intrinsic parity, so the antiquarks are negative.* The parity of a composite system in its ground state is the *product* of the parities of its constituents (we say that parity is a 'multiplicative' quantum number, in contrast to charge, strangeness, and so on, which are 'additive').† Thus the baryon octet and decuplet have positive parity, $(+1)^3$, whereas the pseudoscalar and vector meson nonets have negative parity, $(-1)(+1)$. (The prefix 'pseudo' tells you the parity of the particles.) For an excited state (of two particles) there is an extra factor of $(-1)^l$, where l is the orbital angular momentum [18]. Thus, in general, the mesons carry a parity of $(-1)^{l+1}$ (see Table 4.6). Meanwhile, the *photon* is a *vector* particle (it is represented by the vector potential A^μ); its spin is 1 and its intrinsic parity is -1.

The mirror symmetry of strong and electromagnetic interactions means that parity is conserved in all such processes. Originally, everyone took it for granted that the same goes for the weak interactions as well. But a disturbing paradox arose in the early fifties, known as the 'tau–theta puzzle'. Two strange mesons, called at the time τ and θ, appeared to be identical in every respect – same mass, same spin (zero), same charge, and so on – except that one of them decayed into two pions and the other into three pions, states of opposite parity:

$$\begin{aligned} \theta^+ &\to \pi^+ + \pi^0 && (P = (-1)^2 = +1) \\ \tau^+ &\to \left\{ \begin{matrix} \pi^+ + \pi^0 + \pi^0 \\ \pi^+ + \pi^+ + \pi^- \end{matrix} \right\} && (P = (-1)^3 = -1) \end{aligned} \tag{4.53}$$

* This choice is completely arbitrary; we could just as well do it the other way around. Indeed, in principle we could assign positive parity to some quark flavors and negative to others. This would lead to a different set of hadronic parities, but the *conservation* of parity would still hold. The rule stated here is obviously the *simplest*, and it leads to the conventional assignments.

† There is less to this distinction than meets the eye; in a sense, it results from a notational anomaly. Scrupulous consistency would require that we write the parity operator in exponential form, $P = e^{i\pi K}$, with the operator K playing a role analogous to, say, spin (Equation 4.28). The eigenvalues of K would be 0 and 1, corresponding to $+1$ and -1 for P, and multiplication of parities would correspond to addition of K.

Table 4.6 Quantum numbers of some meson nonets

Orbital angular momentum	Net spin	J^{PC}	Observed Nonet $I=1$	$I=\frac{1}{2}$	$I=0$	Average mass (MeV/c^2)
$l=0$	$s=0$	0^{-+}	π	K	η, η'	400
	$s=1$	1^{--}	ρ	K^*	ϕ, ω	900
$l=1$	$s=0$	1^{+-}	b_1	K_{1_B}	h_1, h_1	1200
	$s=1$	0^{++}	a_0	K_0^*	f_0, f_0	1100
	$s=1$	1^{++}	a_1	K_{1_A}	f_1, f_1	1300
	$s=1$	2^{++}	a_2	K_2^*	f'_2, f_2	1400

It seemed peculiar that two otherwise identical particles should carry different parity. The alternative, suggested by Lee and Yang in 1956 was that τ and θ are really the *same particle* (now known as the K^+), and parity is simply not conserved in one of the decays. This idea prompted their search for evidence of parity invariance in the weak interactions and, when they found none, to their proposal for an experimental test.

4.4.2 Charge Conjugation

Classical electrodynamics is invariant under a change in the sign of all electric charges; the potentials and fields reverse their signs, but there is a compensating charge factor in the Lorentz law, so the forces still come out the same. In elementary particle physics, we introduce an operation that generalizes this notion of 'changing the sign of the charge'. It is called *charge conjugation, C*, and it converts each particle into its antiparticle:

$$C|p\rangle = |\bar{p}\rangle \tag{4.54}$$

'Charge conjugation' is something of a misnomer, for C can be applied to a neutral particle, such as the neutron (yielding an antineutron), and it changes the sign of *all* the 'internal' quantum numbers – charge, baryon number, lepton number, strangeness, charm, beauty, truth – while leaving mass, energy, momentum, and spin untouched.

As with P, application of C *twice* brings us back to the original state:

$$C^2 = I \tag{4.55}$$

and hence the eigenvalues of C are ± 1. Unlike P, however, most of the particles in nature are clearly *not* eigenstates of C. For if $|p\rangle$ is an eigenstate of C, it follows that

$$C|p\rangle = \pm|p\rangle = |\bar{p}\rangle \tag{4.56}$$

so $|p\rangle$ and $|\bar{p}\rangle$ differ at most by a sign, which means that they represent the same physical state. Thus, *only those particles that are their own antiparticles can be eigenstates of C*. This leaves us the photon, as well as all those mesons that lie at the center of their Eightfold-Way diagrams: π^0, η, η', ρ^0, ϕ, ω, ψ, and so on. Because the photon is the quantum of the electromagnetic field, which changes sign under C, it makes sense that the photon's 'charge conjugation number' is -1. It can be shown [19] that a system consisting of a spin-$\frac{1}{2}$ particle and its antiparticle, in a configuration with orbital angular momentum l and total spin s, constitutes an eigenstate of C with eigenvalue $(-1)^{l+s}$. According to the quark model, the mesons in question are of precisely this form: for the pseudoscalars, $l = 0$ and $s = 0$, so $C = +1$; for the vectors, $l = 0$ and $s = 1$, so $C = -1$. (Often, as in Table 4.6, C is listed as though it were a valid quantum number for the entire supermultiplet; in *fact* it pertains only to the central members.)

Charge conjugation is a multiplicative quantum number, and, like parity, it is conserved in the strong and electromagnetic interactions. Thus, for example, the π^0 decays into two photons:

$$\pi^0 \to \gamma + \gamma \tag{4.57}$$

(for n photons $C = (-1)^n$, so in this case $C = +1$ before and after), but it cannot decay into three photons. Similarly, the ω goes to $\pi^0 + \gamma$, but never to $\pi^0 + 2\gamma$. In the strong interactions, charge conjugation invariance requires, for example, that the energy distributions of the charged pions in the reaction

$$p + \bar{p} \to \pi^+ + \pi^- + \pi^0 \tag{4.58}$$

should (on average) be identical [20]. On the other hand, charge conjugation is *not* a symmetry of the weak interactions: when applied to a neutrino (left-handed, remember), C gives a left-handed *anti*neutrino, which does not occur. So the charge-conjugated version of any process involving neutrinos is *not* a possible physical process. And purely hadronic weak interactions also show violations of C as well as P.

Because so few particles are eigenstates of C, its direct application in elementary particle physics is rather limited. Its power can be somewhat extended, if we confine our attention to the strong interactions, by combining it with an appropriate isospin transformation. Rotation by 180° about the number 2 axis in isospin space* will carry I_3 into $-I_3$, converting, for instance, a π^+ into a π^-. If we then apply the charge conjugation operator, we come back to π^+. Thus the charged pions are eigenstates of this combined operator, even though they are not eigenstates of C alone. For some reason the product transformation is called 'G-parity':

$$G = CR_2, \quad \text{where} \quad R_2 = e^{i\pi I_2} \tag{4.59}$$

* Some authors use the number 1 axis. Obviously, any axis in the 1–2 plane will do the job.

All mesons that carry no strangeness (or charm, beauty, or truth) are eigenstates of G;* for a multiplet of isospin I the eigenvalue is given (Problem 4.36) by

$$G = (-1)^I C \tag{4.60}$$

where C is the charge conjugation number of the neutral member. For a single pion, $G = -1$, and for a state with n pions

$$G = (-1)^n \tag{4.61}$$

This is a very handy result, for it tells you how many pions can be emitted in a particular decay. For example, the ρ mesons, with $I = 1$, $C = -1$, and hence $G = +1$, can go to *two* pions, but not to three, whereas the ϕ, the ω, and the ψ (all $I = 0$) can go to *three*, but not to two.

4.4.3
CP

As we have seen, the weak interactions are not invariant under the parity transformation P; the cleanest evidence for this is the fact that the antimuon emitted in pion decay

$$\pi^+ \to \mu^+ + \nu_\mu \tag{4.62}$$

always comes out left-handed. Nor are the weak interactions invariant under C, for the charge-conjugated version of this reaction would be

$$\pi^- \to \mu^- + \bar{\nu}_\mu \tag{4.63}$$

with a left-handed muon, whereas in fact the muon always comes out *right*-handed. However, if we *combine* the two operations we're back in business: CP turns the left-handed antimuon into a *right*-handed muon, which is exactly what we observe in nature. Many people who had been shocked by the fall of parity were consoled by this realization; perhaps, it was the combined operation that our intuition had been talking about all along – maybe what we should have *meant* by the 'mirror image' of a right-handed electron was a left-handed *positron*.† If we had defined parity from the start to be what we now call CP, the trauma of parity violation might have been avoided (or at least postponed). It is too late to change the terminology

* K^+, for example, is *not* an eigenstate of G, for R_2 takes it to K^0, and C takes that to $\overline{K}^0$. The idea could be extended to the K's, by using an appropriate $SU(3)$ transformation in place of R_2, but since $SU(3)$ is not a very good symmetry of the strong forces, there is little percentage in doing so.

† Incidentally, we could perfectly well take electric charge to be a pseudoscalar in classical electrodynamics; $\mathbf{E}$ becomes a pseudovector and $\mathbf{B}$ a vector, but the results are all the same. It is really a matter of taste whether you say the mirror image of a plus charge is positive or negative. But it seems simplest to say the charge does *not* change, and this is the standard convention.

now, but at least this helps to appease our visceral sense that the world 'ought' to be left–right symmetric.

4.4.3.1 Neutral Kaons

CP invariance has bizarre implications for the neutral K mesons, as was first pointed out in a classic paper by Gell-Mann and Pais [21]. They noted that the K^0, with strangeness $+1$, can turn into its antiparticle $\overline{K}^0$, strangeness -1

$$K^0 \rightleftharpoons \overline{K}^0 \tag{4.64}$$

through a second-order weak interaction we now represent by the 'box' diagrams in Figure 4.12.* As a result, the particles we normally observe in the laboratory are not K^0 and $\overline{K}^0$, but rather some linear combination of the two. In particular, we can form eigenstates of CP, as follows. Because the K's are pseudoscalars

$$P|K^0\rangle = -|K^0\rangle, \quad P|\overline{K}^0\rangle = -|\overline{K}^0\rangle \tag{4.65}$$

On the other hand, from Equation 4.54

$$C|K^0\rangle = |\overline{K}^0\rangle, \quad C|\overline{K}^0\rangle = |K^0\rangle \tag{4.66}$$

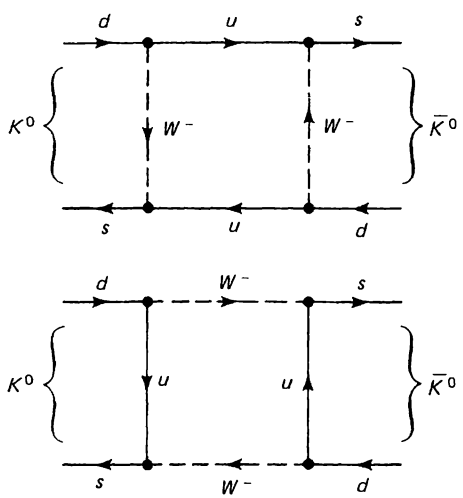


Fig. 4.12 Feynman diagrams contributing to $K^0 \rightleftharpoons \overline{K}^0$. (There are others, including those with one or both u quarks replaced by c or t.)

* The possibility of such an interconversion is almost unique to the neutral kaon system; among the 'stable' hadrons the only other candidates are $D^0/\overline{D}^0$, $B^0/\overline{B}^0$, and $B_s^0/\overline{B}_s^0$ (Problem 4.38).

Accordingly,

$$CP|K^0\rangle = -|\overline{K}^0\rangle, \quad CP|\overline{K}^0\rangle = -|K^0\rangle \tag{4.67}$$

and hence the (normalized) eigenstates of CP are

$$|K_1\rangle = \left(\frac{1}{\sqrt{2}}\right)(|K^0\rangle - |\overline{K}^0\rangle) \text{ and } |K_2\rangle = \left(\frac{1}{\sqrt{2}}\right)(|K^0\rangle + |\overline{K}^0\rangle) \tag{4.68}$$

with

$$CP|K_1\rangle = |K_1\rangle \quad \text{and} \quad CP|K_2\rangle = -|K_2\rangle \tag{4.69}$$

Assuming CP is conserved in the weak interactions, K_1 can only decay into a state with $CP = +1$, whereas K_2 must go to a state with $CP = -1$. Typically, neutral kaons decay into two or three pions. But we have already seen that the two-pion configuration carries a parity of +1, and the three-pion system has $P = -1$ (Equation 4.53); both have $C = +1$. *Conclusion:* K_1 decays into two pions; K_2 decays into three pions (never two):*

$$K_1 \to 2\pi, \quad K_2 \to 3\pi \tag{4.70}$$

Now, the 2π decay is much faster, because the energy released is greater. So if we start with a beam of K^0's

$$|K^0\rangle = \left(\frac{1}{\sqrt{2}}\right)(|K_1\rangle + |K_2\rangle) \tag{4.71}$$

the K_1 component will quickly decay away, and down the line we shall have a beam of pure K_2's. Near the source, we should see a lot of 2π events, but farther along we expect only 3π decays.

Well ... that's a lot to swallow. As Cronin put it, in a delightful memoir [22]:

> *So these gentlemen, Gell-Mann and Pais, predicted that in addition to the short-lived K mesons, there should be long-lived K mesons. They did it beautifully, elegantly and simply. I think theirs is a paper one should read sometime just for its pure beauty of reasoning. It was published in the Physical Review in 1955. A very lovely thing! You get shivers up and down your spine, especially when you find you understand it. At the time, many of the most distinguished theoreticians thought this prediction was really baloney.*

* Actually, with the right combination of orbital angular momenta, it is possible to construct a $CP = +1$ state of the $\pi^+\pi^-\pi^0$ system, but while this might allow K_1 to decay (rarely) into 3π, it does not alter the critical fact that K_2 cannot go to 2π.

But it wasn't baloney, and in 1956, Lederman and his collaborators discovered the K_2 meson at Brookhaven [23]. Experimentally, the two lifetimes are

$$\tau_1 = 0.895 \times 10^{-10} \text{ sec}$$
$$\tau_2 = 5.11 \times 10^{-8} \text{ sec} \tag{4.72}$$

so the K_1's are mostly gone after a few centimeters, whereas the K_2's can travel many meters. Notice that K_1 and K_2 are not antiparticles of one another, like K^0 and $\overline{K}^0$; rather, each is its own antiparticle ($C = -1$ for K_1 and $C = +1$ for K_2). They differ ever-so-slightly in mass; experiments give [24]

$$m_2 - m_1 = 3.48 \times 10^{-6} \text{ eV/c}^2 \tag{4.73}$$

The neutral kaon system adds a subtle twist to the old question, 'What is a particle?' Kaons are typically *produced* by the strong interactions, in eigenstates of strangeness (K^0 and $\overline{K}^0$), but they *decay* by the *weak* interactions, as eigenstates of CP (K_1 and K_2). Which, then, is the 'real' particle? If we hold that a 'particle' must have a unique lifetime, then the 'true' particles are K_1 and K_2.* But we need not be so dogmatic. In practice, it is sometimes more convenient to use one set, and sometimes, the other. The situation is in many ways analogous to polarized light. Linear polarization can be regarded as a superposition of left-circular polarization and right-circular polarization. If you imagine a medium that preferentially absorbs right-circularly polarized light, and shine on it a linearly polarized beam, it will become progressively more left-circularly polarized as it passes through the material, just as a K^0 beam turns into a K_2 beam. But whether you choose to analyze the process in terms of states of linear or circular polarization is largely a matter of taste.

4.4.3.2 *CP* Violation

The neutral kaons provide a perfect experimental system for testing CP invariance. By using a long enough beam, we can produce an arbitrarily pure sample of the long-lived species. If at this point, we observe a 2π decay, we shall know that CP has been violated. Such an experiment was reported by Cronin and Fitch in 1964.[25] At the end of a beam 57 feet long, they counted 45 two pion events in a total of 22,700 decays. That's a tiny fraction (roughly 1 in 500), but unmistakable evidence of CP violation. Evidently, the long-lived neutral kaon is *not* a perfect eigenstate of CP after all, but contains a small admixture of K_1:

$$|K_L\rangle = \frac{1}{\sqrt{1+|\epsilon|^2}}(|K_2\rangle + \epsilon|K_1\rangle) \tag{4.74}$$

* This, incidentally, was the position advocated by Gell-Mann and Pais.

The coefficient ϵ is a measure of nature's departure from perfect CP invariance;* experimentally, its magnitude is about 2.24×10^{-3}.

Although the effect is small, CP violation poses a far deeper problem than parity ever did. The nonconservation of parity was quickly incorporated into the theory of weak interactions (in fact, *part* of the 'new' theory – Weyl's equation for the neutrino – had been waiting in the wings for many years). Parity violation was easier to handle precisely because it was such a *large* effect: *all* neutrinos are left-handed, not just 50.01% of them. Parity is, in this sense, *maximally* violated, in the weak interactions. By contrast, CP violation is a small effect by any measure. Within the Standard Model, it can be accommodated by including an empirical phase factor (δ) in the Cabibbo–Kobayashi–Maskawa (CKM) matrix, provided that there are (at least) three generations of quarks. Indeed, it was this realization that led Kobayashi and Maskawa to propose a third generation of quarks in 1973, before even charm was discovered. [27]

The Fitch–Cronin experiment destroyed the last hope for any form of exact mirror symmetry in nature. And subsequent study of the semileptonic decays of K_L revealed even more dramatic evidence of CP violation. Although 32% of all K_L's decay by the 3π mode we have discussed, 41% go to

$$\text{(a)}\ \pi^+ + e^- + \overline{\nu}_e \qquad \text{or} \qquad \text{(b)}\ \pi^- + e^+ + \nu_e \tag{4.75}$$

Notice that CP takes (a) into (b), so if CP were conserved, and K_L were a pure eigenstate, (a) and (b) would be equally probable. But experiments show [28] that K_L decays more often into a positron than into an electron, by a fractional amount 3.3×10^{-3}. Here, for the first time, is a process that makes an absolute distinction between matter and antimatter, and provides an unambiguous, convention-free definition of positive charge: *it is the charge carried by the lepton preferentially produced in the decay of the long-lived neutral K meson.* The fact that CP violation permits unequal treatment of particles and antiparticles suggests that it may be responsible for the dominance of matter over antimatter in the universe. [29] We will explore this further in Chapter 12.

For almost 40 years, the decay of K_L was the only context in which CP violation was observed in the laboratory. In 1981, Carter and Sanda pointed out that the violation should also occur with the neutral B mesons. [30] To explore this possibility, 'B-factories' were constructed at SLAC and KEK (in Japan), designed specifically to produce enormous numbers of $B^0/\overline{B}^0$ pairs [31]. By 2001, their detectors ('BaBar' and 'Belle', respectively) had recorded incontrovertible evidence of CP violation in neutral B decays. [32]† Unlike the kaon system, where CP violation is a tiny effect in relatively common decays (such as Equation 4.75), for the B's it tends to be a large effect in extremely rare decays. For example, the branching ratio

* This is not the only route by which K_L can decay to 2π; in the Standard Model, there is also a small 'direct' CP violation that does not involve $K^0 \leftrightarrow \overline{K}^0$ mixing, but is associated instead with the so-called 'penguin' diagrams (Problem 4.40). Direct CP violation in $K_L \to 2\pi$ was confirmed in 1999 [26].

† 'Direct' CP violation in neutral B decays was confirmed by both labs in 2004 [33].

for $B^0 \to K^+ + \pi^-$ is only 1.82×10^{-5}, but this decay is 13% more common than its CP 'mirror image' $\overline{B}^0 \to K^- + \pi^+$. So far, this is the only other system in which CP violation has been detected.*

4.4.4
Time Reversal and the *TCP* Theorem

Suppose we made a movie of some physical process, say, an elastic collision of two billiard balls. If we ran the movie backward, would it depict a possible physical process, or would the viewer be able to say with certainty 'No, no, that's impossible; the film must be running in reverse'? In the case of classical elastic collisions, the 'time-reversed' process *is* perfectly possible. (To be sure, if we put a *lot* of billiard balls in the picture, the backward version might be highly *improbable*; we would be *surprised* to see the balls gather themselves together into a perfect triangle, with a single cue ball rolling away, and we would strongly *suspect* that the film had been reversed. But that's just because we know it would be extraordinarily difficult to set up the necessary starting conditions, such that all the balls would roll together at just the right speeds and in just the right directions. Thus the initial conditions may give us a clue to the 'arrow of time', but the laws governing the collisions themselves work just as well forward as backward.) Until fairly recently, it was taken for granted that all elementary particle interactions share this time-reversal invariance. But with the downfall of parity, it was natural to wonder whether time reversal was really so sacred.[36]

As it turns out, time reversal is a lot harder to test than P or C. In the first place, whereas many particles are eigenstates of P, and some are eigenstates of C, none is an eigenstate of T (the 'time-reversal operator', which runs the movie backward).† So we cannot check the 'conservation of T' simply by multiplying numbers, the way we can for P and C. The most direct test would be to take a particular reaction (say, $n + p \to d + \gamma$), and run it in reverse ($d + \gamma \to n + p$). For corresponding conditions of momentum, energy, and spin, the reaction rate should be the same in either direction. (This is called the 'principle of detailed balance', and it follows directly from time-reversal invariance.) Such tests work fine for the strong and electromagnetic interactions, and a variety of processes have been checked. The results have always been negative (no evidence of T violation), but this is hardly surprising.

* There is some evidence for $B_s^0/\overline{B}_s^0$ mixing [34], and more recently $D^0/\overline{D}^0$ mixing [35], but as yet no evidence of CP violation in either case. Because the b quark – like the s quark – cannot decay without crossing a generation boundary, the B mesons – like the Ks – tend to be relatively long-lived (10^{-12} s). The c quark, by contrast, can go to an s without crossing a boundary, and that makes the D mesons short-lived (10^{-15} s). This is one reason the B system is a more promising place to look for CP violation, even though D's are easier to produce.

† A particle can be identical to its mirror image, and, if it's neutral, to its own antiparticle, but it can't be identical to itself-going-backward-in-time (at least, not if anything ever happens to it).

On the basis of our experience with P and C, we expect to see a failure of time reversal in the weak interactions, if anywhere. Unfortunately, inverse-reaction experiments are tough to do in the weak interactions. Take, for instance, the typical weak decay $\Lambda \to p^+ + \pi^-$. The inverse reaction would be $p^+ + \pi^- \to \Lambda$, but we are never going to see such a process, because the strong interaction of the proton and the pion will totally swamp the feeble weak interaction. To avoid strong and electromagnetic contamination, we might go to a neutrino process. But it is notoriously difficult to do accurate measurements on neutrinos, and here we are presumably looking for a very small effect. In practice, therefore, the critical tests of T invariance involve careful measurements of quantities that should be precisely zero if T is a perfect symmetry. The classic example is a static electric dipole moment of an elementary particle.* Probably, the most sensitive experimental tests are the upper limits on the electric dipole moments of the neutron [37] and the electron:[38]

$$d_n < (6 \times 10^{-26}\ \text{cm})\, e, \quad d_e < (1.6 \times 10^{-27}\ \text{cm})\, e \tag{4.76}$$

where e is the charge of the proton; *no* experiment has shown direct evidence of T violation.

Nevertheless, there is a compelling reason to believe that time reversal cannot be a perfect symmetry of nature. It comes from the so-called TCP theorem, one of the deepest results of quantum field theory [39]. Based only on the most general assumptions – Lorentz invariance, quantum mechanics, and the idea that interactions are represented by fields – the TCP theorem states that the combined operation of time reversal, charge conjugation, and parity (in any order) is an exact symmetry of *any* interaction. It is simply *impossible to construct* a quantum field theory in which the product TCP is not conserved. If, as the Fitch–Cronin experiment demonstrated, CP is violated, there must be a compensating violation of T. Of course, like any assertion of impossibility, the TCP theorem may just be a measure of our lack of imagination; it must be tested in the laboratory, and that is one reason it is so important to look for independent evidence of T violation. But the TCP theorem has other implications that are also subject to experimental verification: if the theorem is correct, every particle must have precisely the same mass and lifetime as its antiparticle.† Measurements have been made on a number of particle–antiparticle pairs; the most sensitive test to date is the $K^0 - \overline{K}^0$ mass

* For an elementary particle, the dipole moment **d**, would have to point along the axis of the spin, s; there is no other direction available. But **d** is a vector, whereas s is a pseudovector, so a nonzero dipole moment would imply violation of P. Similarly, s changes sign under time reversal, but **d** does not, so a nonzero **d** would also (and more interestingly) mean violation of T. For further details, see Ramsey, ref. [32].

† This would also follow from C invariance. However, since we know that the latter is violated, it is significant that the equality of masses and lifetimes (also magnetic moments, incidentally, although they have opposite *signs*) follows from the far weaker assumption of TCP symmetry.

difference, which, as a fraction of the K^0 mass, is known to be less than 10^{-18}. So the *TCP* theorem is on extremely firm ground theoretically, and it is relatively secure experimentally. Indeed, as one prominent theorist has put it, if a departure is ever found, 'all hell breaks loose'.

References

1 See, for example, Halzen, F. and Martin, A. D. (1984) *Quarks and Leptons*, John Wiley & Sons, New York. Section 14.2. For a more complete (if somewhat dated) discussion; see (a) Hill, E. L. (1951) *Reviews of Modern Physics*, 23, 253.

2 If this fact is not familiar, you might want to look at Halliday, D., Resnick, R. and Walker, J. (2001) *Fundamentals of Physics*, 6th edn, John Wiley & Sons, New York, Section 11.3.

3 For accessible treatments of group theory see Tinkham, M. (2003) *Group Theory and Quantum Mechanics*, Dover, New York; (a) Lipkin, H. J. (2002) *Lie Groups for Pedestrians*, Dover, New York.

4 See, for example, Merzbacher, E. (1998) *Quantum Mechanics*, 3rd edn, John Wiley & Sons, New York, Chapter 17, Section 5.

5 See Tinkham, M. (2003) *Group Theory and Quantum Mechanics*, Dover, New York; (a) Lipkin, H. J. (2002) *Lie Groups for Pedestrians*, Dover, New York.90 Chapter 16, Sections 3 and 4.

6 Heisenberg, W. (1932) *Zeitschrift Fur Physik*, 77, 1. An English translation of this classic paper appears in (a) Brink, D. M. (1965) *Nuclear Forces*, Pergamon, Elmsford, NY.

7 Fliagin, V. B. et al. (1959) *Soviet Physics, JETP*, 35 (8), 592.

8 Anderson, H. L., Fermi, E., Long, E. A. and Nagle, D. E. (1952) *Physical Review*, 85, 936.

9 De Rujula, A., Georgi, M. and Glashow, S. L. (1975) *Physical Review*, D12, 147; (a) Weinberg, S. (1977) *Transactions of the New York Academy Sciences, Series II*, 38, 185; (b) Gasiorowicz, S. and Rosner, J. L. (1981) *American Journal of Physics*, 49, 962; (c) Gasser, J. and Leutwyler, H. (1982) *Physics Report*, 87, 78.

10 Holstein, B. R. (1995) *American Journal of Physics*, 63, 14.

11 A qualitatively plausible mechanism is suggested by the 'MIT Bag Model'. Free quarks of mass m, confined within a spherical shell of radius R, are found to have an effective mass $m_{\text{eff}} = \sqrt{m^2 + (\hbar x/Rc)^2}$, where x is a dimensionless number around 2.5. Using the radius of the proton (say, 1.5×10^{-13} cm) for R, we obtain $m_{\text{eff}} = 330$ MeV/c^2 for the up and down quarks. See Close, F. E. (1979) *An Introduction to Quarks and Partons*, Academic, London, Section 18.1.

12 Bare masses are taken from the *Particle Physics Booklet* (2006). The light quark effective masses are somewhat lower in mesons and higher in baryons; the best-fit values depend on the context. See Tables 5.3, 5.5, and 5.6.

13 For an illuminating introduction to parity see Feynman, R. P., Leighton, R. B. and Sands, M. (1965) *The Feynman Lectures on Physics*, vol. III, Addison-Wesley, Reading, MA, p. 17–22.

14 Lee, T. D. and Yang, C. N. (1956) *Physical Review*, 104, 254.

15 Wu, C. S. et al. (1957) *Physical Review*, 105, 1413. In the interest of clarity I am ignoring the formidable technical difficulties involved in this experiment. To keep the cobalt nuclei aligned, the sample had to be maintained at a temperature of less than 1° K for 10 minutes. Small wonder that no earlier experiments had stumbled on evidence of parity violation.

16 See, for example, Pauli's letter to Weisskopf in Pauli, W. (1964) *Collected Scientific Papers*, vol. **I**, (eds R. Kronig and V. F. Weisskopf), Wiley-Interscience, New York, p. xii; (a) Morrison, P. (April 1957) *Scientific American*, 45.

17 Backenstoss, G. et al. (1961) *Physical Review Letters*, **6**, 415; (a) Bardon, M. et al. (1961) *Physical Review Letters*, **7**, 23. Earlier experiments anticipated this result: (b) Goldhaber, M., Grodzins, L. and Snyder, A. W. (1958) *Physical Review*, **109**, 1015.

18 This comes from the angular part of the spatial wave function, $Y_l^m(\theta, \phi)$. See, for example, Tinkham, M. (2003) *Group Theory and Quantum Mechanics*, Dover, New York; (a) Lipkin, H. J. (2002) *Lie Groups for Pedestrians*, Dover, New York, p. 186, (or Problem 5.3 below).

19 Perkins, D. H. (1982) *Introduction to High-Energy Physics*, 2 d edn, Addison-Wesley, Reading, MA, p. 99.

20 Baltay C. et al. (1965) *Physical Review Letters*, **15**, 591.

21 Gell-Mann, M. and Pais, A. (1955) *Physical Review*, **97**, 1387. This paper was written before the overthrow of parity, but the essential idea remains unchanged if we substitute *CP* for their *C*. Of course, they didn't draw a quark diagram like Figure 4.12; they based their argument for Equation 4.64 on the fact that both K^0 and $\overline{K}^0$ can decay into $\pi^+ + \pi^-$, so $K^0 \Leftrightarrow \pi^+ + \pi^- \Leftrightarrow \overline{K}^0$.

22 Cronin J. W. and Greenwood, M. S. (July 1982) *Physics Today*, 38. Cronin uses an unorthodox sign convention, putting a -1 into our Equation 4.66, but the physics is still the same.

23 Lande, K. et al. (1956) *Physical Review*, **103**, 1901.

24 The detection of so minute a mass difference is itself a fascinating story. See, for example, Wu, C. S. et al. (1957) *Physical Review*, **105**, 1413. Sect. 16.13.1.

25 Christenson, J. H. et al. (1964) *Physical Review Letters*, **13**, 138.

26 Alavi-Harati, A. et al. (1999) *Physical Review Letters*, **83**, 22; (a) Fanti, V. et al. (1999) *Physics Letters B*, **465**, 335. Earlier claims (see (b) Schwarzschild, B. (October 1988) *Physics Today*, 17. appear to have been premature.)

27 Kobayashi, M. and Maskawa, T. (1973) *Progress in Theoretical Physics*, **49**, 652.

28 Gjesdal S. et al. (1974) *Physics Letters*, **52B**, 113; (a) Bennett, S. et al. (1967) *Physical Review Letters*, **19**, 993; (b) Dorfan, D. et al. (1967) *Physical Review Letters*, **19**, 987.

29 Wilczek, F. (December 1980) *Scientific American*, 82.

30 Carter, A. B. and Sanda, A. I. (1981) *Physical Review D*, **23**, 1567. For a comprehensive treatment of *CP* violation see (a) Bigi, I. I. and Sanda, A. I. (2000) *CP Violation*, Cambridge University Press, Cambridge, UK.

31 Quinn, H. R. and Witherell, M. S. (October 1998) *Scientific American*, 76.

32 Abashian, A. et al. (2001) *Physical Review Letters*, **86**, 2509; (a) Aubert, B. et al. (2001) *Physical Review Letters*, **86**, 2515.

33 Aubert, B. et al. (2004) *Physical Review Letters*, **93**, 131801; (a) Chao, Y. et al. (2005) *Physical Review D*, **71**, 031502.

34 Schneider, O. (2006) *Review of Particle Physics*, 836.

35 Aubert, B. et al. (2007) *Physical Review Letters*, **98**, 211802.

36 Excellent informal treatments of time reversal are Overseth, O. E. (October 1969) *Scientific American*, **88**; and (a) Sachs, R. G. (1972) *Science*, **176**, 587.

37 Harris, P. G. et al. (1999) *Physical Review Letters*, **82**, 904; (a) Ramsey, N. F. (1982) *Reports of Progress in Physics*, **45**, 95.

38 Regan, B. et al. (2002) *Physical Review Letters*, **88**, 071805.

39 The *TCP* theorem was discovered by Schwinger, J. and Lüders, G. and perfected by Pauli, W. (1955) *Niels Bohr and the Development of Physics*, (ed W. Pauli), McGraw-Hill, New

York. At first no one paid much attention to the *TCP* theorem, because at that time everyone thought *T*, *C*, and *P* were all perfect symmetries individually. It was only with the fall of parity, and especially with the failure of *CP*, that the importance of this theorem was fully appreciated.

Problems

4.1 Prove that I, R_+, R_-, R_a, R_b, and R_c are *all* the symmetries of the equilateral triangle. [*Hint*: One way to do this is to label the three corners, as in Figure 4.2 A given symmetry operation carries A into the position formerly occupied by A, B, or C. If $A \to A$, then either $B \to B$ and $C \to C$, or else $B \to C$ and $C \to B$. Take it from there.]

4.2 Construct a 'multiplication table' for the triangle group, filling in the blanks on the following diagram:

	I	R_+	R_-	R_a	R_b	R_c
I						
R_+						
R_-						
R_a						
R_b						
R_c						

In row i, column j, put the product R_iR_j. Is this an Abelian group? How can you tell, just by looking at the multiplication table?

4.3 (a) Construct a 3×3 representation of the triangle group as follows: let $D(R)$ be the matrix representing operation R. It acts on the column matrix $\begin{pmatrix} A \\ B \\ C \end{pmatrix}$ to produce a new column matrix $\begin{pmatrix} A' \\ B' \\ C' \end{pmatrix} = D(R)\begin{pmatrix} A \\ B \\ C \end{pmatrix}$, where A' is the vertex now occupying the location originally held by A. Thus, for example,

$$D(R_+) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Find the other five matrices. (You might want to check that multiplication of your matrices fits the table you constructed in Problem (4.2).)

(b) The triangle group, like any other group, has a trivial one-dimensional representation. It *also* has a *non*trivial, one-dimensional representation, in which the elements are *not* all represented by 1. Work out this second one-dimensional representation. That is, figure out what number (1×1 matrix) each group element is represented by. Is this representation faithful?

4.4 Work out the symmetry group of a square. How many elements does it have? Construct the multiplication table, and determine whether or not the group is Abelian.

4.5 (a) Show that the set of all unitary $n \times n$ matrices constitutes a group. (To prove closure, for instance, you must show that the product of two unitary matrices is itself unitary.)

(b) Show that the set of all $n \times n$ unitary matrices with determinant 1 constitutes a group.

(c) Show that $O(n)$ is a group.

(d) Show that $SO(n)$ is a group.

4.6 Consider a vector **A** in two dimensions. Suppose its components with respect to Cartesian axes x, y, are (a_x, a_y). What are its components (a'_x, a'_x) in a system x', y' which is rotated, counterclockwise, by an angle θ, with respect to x, y? Express your answer in the form of a 2×2 matrix $R(\theta)$:

$$\begin{pmatrix} a'_x \\ a'_y \end{pmatrix} = R \begin{pmatrix} a_x \\ a_y \end{pmatrix}$$

Show that R is an orthogonal matrix. What is its determinant? The set of *all* such rotations constitutes a group; what is the name of this group? By multiplying the matrices, show that $R(\theta_1)R(\theta_2) = R(\theta_1 + \theta_2)$; is this an Abelian group?

4.7 Consider the matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Is it in the group $O(2)$? How about $SO(2)$? What is its effect on the vector **A** of Problem (4.6)? Does it describe a possible rotation of the plane?

4.8 Suppose we interpret the electron literally as a classical solid sphere of radius r, mass m, spinning with angular momentum $\frac{1}{2}\hbar$. What is the speed, v, of a point on its 'equator'? Experimentally, it is known that r is less than 10^{-16} cm. What is the corresponding equatorial speed? What do you conclude from this?

4.9 When you are adding angular momenta, using Equation 4.12, it is useful to check your results by counting the number of states before and after the addition. For instance, in Example 4.1 we had two quarks to begin with, each could have $m_s = +\frac{1}{2}$ or $m_s = -\frac{1}{2}$, so there were four possibilities in all. *After* adding the spins, we had one combination with spin 1 (hence $m_s = 1, 0$, or -1) and one with spin 0 ($m_s = 0$) – again, four states in all.

(a) Apply this check to Example 4.2

(b) Add angular momenta 2, 1, and $\frac{1}{2}$. List the possible values of the total angular momentum, and check your answer by counting states.

4.10 Show that the 'original' beta-decay reaction $n \to p + e$ would violate conservation of angular momentum (all three particles have spin $\frac{1}{2}$). If you were Pauli, proposing that the reaction is really $n \to p + e + \overline{\nu}_e$, what spin would you assign to the neutrino?

4.11 In the decay $\Delta^{++} \to p + \pi^+$, what are the possible values of the (CM) orbital angular momentum quantum number, l, in the final state?

4.12 An electron in a hydrogen atom is in a state with orbital angular momentum quantum number $l = 1$. If the *total* angular momentum quantum number j is $\frac{3}{2}$, and the z component of total angular momentum is $\frac{1}{2}\hbar$, what is the probability of finding the electron with $m_s = +\frac{1}{2}$?

4.13 Suppose you had two particles of spin 2, each in a state with $S_z = 0$. If you measured the *total* angular momentum of this system, given that the *orbital* angular momentum is zero, what values might you get, and what is the probability of each? Check that they add up to 1.

4.14 Suppose you had a particle of spin $\frac{3}{2}$, and another of spin 2. If you knew that their orbital angular momentum was zero, and that the *total* spin of the composite system was $\frac{5}{2}$, and its z component was $-\frac{1}{2}$, what values might you get for a measurement of S_z on the spin-2 particle? What is the probability of each? Check that they add up to 1.

4.15 Check that $\chi_\pm$, Equation 4.22, are normalized eigenvectors of $\hat{S}_x$, Equation 4.21, and find the associated eigenvalues.

4.16 Show that $|a|^2 + |b|^2 = 1$ (Equation 4.24), provided the spinor in question is normalized (Equation 4.20).

4.17 (a) Find the eigenvalues and normalized eigenspinors of $\hat{S}_y$ (Equation 4.21).

(b) If you measured S_y on an electron in the state $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$, what values might you get, and what is the probability of each?

4.18 Suppose an electron is in the state $\begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix}$.

(a) If you measured S_x, what values might you get, and what is the probability of each?
(b) If you measured S_y, what values might you get, and what is the probability of each?
(c) If you measured S_z, what values might you get, and what is the probability of each?

4.19 (a) Show that $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1$. ('1' here really means the 2×2 unit matrix; if no matrix is specified, the unit matrix is understood.)
(b) Show that $\sigma_x\sigma_y = -\sigma_y\sigma_x = i\sigma_z, \sigma_y\sigma_z = -\sigma_z\sigma_y = i\sigma_x, \sigma_z\sigma_x = -\sigma_x\sigma_z = i\sigma_y$. These results are neatly summarized in the formula

$$\sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k$$

(summation over k implied), where δ_{ij} is the Kronecker delta:

$$\delta_{ij} = \left\{ \begin{array}{ll} 1, & \text{if} \quad i = j \\ 0, & \text{otherwise} \end{array} \right\}$$

and ϵ_{ijk} is the Levi–Civita symbol:

$$\epsilon_{ijk} = \left\{ \begin{array}{ll} 1, & \text{if } ijk = 123, 231, \text{or } 312 \\ -1, & \text{if } ijk = 132, 213, \text{or } 321 \\ 0, & \text{otherwise} \end{array} \right.$$

4.20 Use the results of Problem 4.19 to show that

(a) The commutator, $[A, B] \equiv AB - BA$, of two Pauli matrices is $[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k$.
(b) The *anticommutator*, $\{A, B\} \equiv AB + BA$, is $\{\sigma_i, \sigma_j\} = 2\delta_{ij}$.
(c) For any two vectors a and **b**, $(\boldsymbol{\sigma} \cdot \mathbf{a})(\boldsymbol{\sigma} \cdot \mathbf{b}) = \mathbf{a} \cdot \mathbf{b} + i\boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b})$.

4.21 (a) Show that $e^{i\pi\sigma_z/2} = i\sigma_z$.
(b) Find the matrix U representing a rotation by 180° about the y axis, and show that it converts 'spin up' into 'spin down', as we would expect.
(c) More generally, show that

$$U(\boldsymbol{\theta}) = \cos\frac{\theta}{2} - i(\hat{\theta}\cdot\boldsymbol{\sigma})\sin\frac{\theta}{2}$$

where $U(\boldsymbol{\theta})$ is given by Equation 4.28, θ is the magnitude of $\boldsymbol{\theta}$, and $\hat{\theta} \equiv \boldsymbol{\theta}/\theta$. [*Hint*: Use Problem 4.20, part (c).]

4.22 (a) Show that U, in Equation 4.28, is *unitary*.
(b) Show that $\det U = 1$. [*Hint*: You can either do this directly (however, see footnote after Equation 4.29), or else use the results of Problem 4.21.]

4.23 The extension of everything in Section 4.2.2 to higher spin is relatively straightforward. For spin 1 we have three states ($m_s = +1, 0, -1$), which we may represent by column vectors:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

respectively. The only problem is to construct the 3×3 matrices $\hat{S}_x$, $\hat{S}_y$ and $\hat{S}_z$. The latter is easy:

(a) Construct $\hat{S}_z$ for spin 1. To obtain $\hat{S}_x$ and $\hat{S}_y$ it is easiest to start with the 'raising' and 'lowering' operators, $S_\pm \equiv S_x \pm iS_y$, which have the property

$$S_\pm|sm\rangle = \hbar\sqrt{s(s+1) - m(m\pm 1)}|s(m\pm 1)\rangle \tag{4.77}$$

(b) Construct the matrices $\hat{S}_+$ and $\hat{S}_-$, for spin 1.
(c) Using (b), determine the spin-1 matrices $\hat{S}_x$ and $\hat{S}_y$.
(d) Carry out the same construction for spin $\frac{3}{2}$.

4.24 Determine the isospin assignments $|II_3\rangle$ for each of the following particles (refer to the Eightfold Way diagrams in Chapter 1): Ω^-, Σ^+, Ξ^0, ρ^+, η, $\overline{K}^0$.

4.25 (a) Check that the Gell-Mann–Nishijima formula works for the quarks u, d, and s.
(b) What are the appropriate isospin assignments, $|II_3\rangle$, for the *anti*quarks, $\overline{u}$, $\overline{d}$, and $\overline{s}$? Check that your assignment is consistent with the Gell-Mann–Nishijima formula. [Since Q, I_3, A, and S all *add*, when we combine quarks, it follows that the Gell-Mann–Nishijima formula holds for all hadrons made out of u, d, s, $\overline{u}$, $\overline{d}$, and $\overline{s}$.]

4.26 (a) The Gell-Mann–Nishijima formula, Equation 4.37, was proposed in the early fifties, which is to say long before the discovery of charm, beauty, or truth. Using the table of quark properties (in Section 1.11), and the quark isospin assignments, Equation 4.38, deduce the *general* formula expressing Q in terms of A, I_3, S, C, B, and T.
(b) Because u and d are the only quarks with nonzero isospin, it should be possible to express I_3 in terms of U ('upness') and D ('downness'). What's the formula? Likewise, express A in terms of the flavor numbers U, D, S, C, B, and T.
(c) Putting it all together, obtain the formula for Q in terms of the flavor numbers (that is, eliminate A and I_3 from your formula in part (a)). This final version represents the cleanest statement of the Gell-Mann–Nishijima formula, in the three-generation quark model.

4.27 For two isospin-$\frac{1}{2}$ particles, show that $\mathbf{I}^{(1)}\cdot\mathbf{I}^{(2)} = \frac{1}{4}$ in the triplet state and $-\frac{3}{4}$ in the singlet. [*Hint*: $\mathbf{I}_{tot} = \mathbf{I}^{(1)} + \mathbf{I}^{(2)}$; square both sides.]

4.28 (a) Referring to Equations 4.47 and 4.48, work out all the πN scattering amplitudes, $\mathscr{M}_a$ through $\mathscr{M}_j$, in terms of $\mathscr{M}_1$ and $\mathscr{M}_3$.
(b) Generalize Equation 4.49 to include all 10 cross sections.
(c) In the same way, generalize Equation 4.50

4.29 Find the ratio of the cross sections for the following reactions, assuming the CM energy is such that the $I = \frac{3}{2}$ channel dominates: (a) $\pi^- + p \to K^0 + \Sigma^0$; (b) $\pi^- + p \to K^+ + \Sigma^-$; (c) $\pi^+ + p \to K^+ + \Sigma^+$. What if the energy is such that the $I = \frac{1}{2}$ channel dominates?

4.30 What are the possible total isospins for the following reactions: (a) $K^- + p \to \Sigma^0 + \pi^0$; (b) $K^- + p \to \Sigma^+ + \pi^-$; (c) $\overline{K}^0 + p \to \Sigma^+ + \pi^0$; (d) $\overline{K}^0 + p \to \Sigma^0 + \pi^+$. Find the ratio of the cross sections, assuming one or the other isospin channel dominates.

4.31 On the graph in Figure 4.6, we see 'resonances' as 1525, 1688, 1920, and 2190 (as well as the one at 1232). By comparing the two curves, determine the isospin of each resonance. The nomenclature is N (followed by the mass) for any state with $I = \frac{1}{2}$, and Δ for any state with $I = \frac{3}{2}$. Thus the nucleon is $N(939)$, and the 'original' Δ is $\Delta(1232)$. Name the other resonances, and confirm your answers by looking in the *Particle Physics Booklet*.

4.32 The Σ^{*0} can decay into $\Sigma^+ + \pi^-$, $\Sigma^0 + \pi^0$, or $\Sigma^- + \pi^+$ (also $\Lambda + \pi^0$, but we're not concerned with that here). Suppose you observed 100 such disintegrations, how many would you expect to see of each type?

4.33 (a) The α particle is a bound state of two protons and two neutrons, that is, a ^{4}He nucleus. There is no isotope of hydrogen with an atomic weight of four (^{4}H), nor of lithium ^{4}Li. What do you conclude about the isospin of an α particle?
(b) The reaction $d + d \to \alpha + \pi^0$ has never been observed. Explain why.
(c) Would you expect ^{4}Be to exist? How about a bound state of four neutrons?

4.34 (a) Using Equation 4.52, prove that the eigenvalues of P are ± 1.
(b) Show that any scalar function $f(x, y, z)$ can be expressed as the sum of an eigenfunction $f_+(x, y, z)$ with eigenvalue +1 and an eigenfunction $f_-(x, y, z)$ with eigenvalue -1. Construct the functions f_+ and f_-, in terms of f. [*Hint:* $Pf(x, y, z) = f(-x, -y, -z)$.]

4.35 (a) Is the neutrino an eigenstate of P? If so, what is its intrinsic parity?
(b) Now that we know τ^+ and θ^+ are actually both the K^+, which of the decays in Equation 4.53 actually violates parity conservation?

4.36 (a) Using the information in Table 4.6, determine the G parity of the following mesons: $\pi, \rho, \omega, \eta, \eta', \phi, f_2$.
(b) Show that $R_2|I0\rangle = (-1)^I|I0\rangle$, and use this result to justify Equation 4.60

4.37 The dominant decays of the η meson are

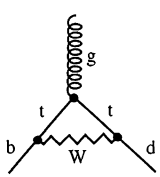
$$\eta \to 2\gamma(39\%), \quad \eta \to 3\pi(55\%), \quad \eta \to \pi\pi\gamma(5\%) \tag{4.78}$$

and it is classified as a 'stable' particle, so evidently none of these is a purely strong interaction. Offhand, this seems odd, since at 549 MeV/c^2. the η has plenty of mass to decay strongly into 2π or 3π.
(a) Explain why the 2π mode is forbidden, for both strong and electromagnetic interactions.
(b) Explain why the 3π mode is forbidden as a *strong* interaction, but allowed as an *electromagnetic* decay.

4.38 For two hadrons to interconvert, $A \rightleftharpoons B$, it is necessary that they have the same mass (which in practice means that they must be antiparticles of one another), the same charge, and the same baryon number. In the Standard Model, with the usual three generations, show that A and B would have to be neutral mesons, and identify their possible quark contents. What, then, are the candidate mesons? Why doesn't the neutron mix with the antineutron, in the same way as the K^0 and $\overline{K}^0$ mix to produce K_1 and K_2? Why don't we see mixing of the neutral strange *vector* mesons K^{0*} and $\overline{K}^{0*}$?

4.39 Suppose you wanted to inform someone in a distant galaxy that humans have their hearts on the left side. How could you communicate this unambiguously, without sending an actual 'handed' object (such as a corkscrew, a circularly polarized light beam, or a neutrino). For all you know their galaxy may be made of antimatter. You cannot afford to wait for any replies, but you are allowed to use English.

4.40 The charged weak interactions couple a d, s, or b to a u, c, or t, but a d (for example) cannot go directly to an s or a b. However, such a coupling can occur *in*directly, via a so-called "penguin" diagram, in which a quark emits a virtual W that it subsequently reabsorbs, having in the mean time interacted with a gluon:*



A 'tree' diagram is one with no closed loops. Construct a penguin diagram representing $\overline{B}^0 \to \pi^+ + \pi^-$, and a tree diagram for the same process (the latter should have no gluons). In both cases, let the $\overline{d}$ quark be a spectator. ['Direct' CP violation comes from the interference of these two diagrams.]

Don't look for anything resembling the bird here – the name is a joke. The story is told best by Woit, P. (2006) *Not Even Wrong*, Basic Books, New York, pp. 54–55.

5

Bound States

The first part of this chapter is devoted to the nonrelativistic theory of two-particle bound states – hydrogen (e^-p^+), positronium (e^-e^+), charmonium ($c\bar{c}$), and bottomonium ($b\bar{b}$). This material is not used in subsequent chapters and can be skimmed, saved for later, or skipped entirely. Some acquaintance with quantum mechanics is essential. The final two sections (5.5 and 5.6) concern relativistic light-quark systems – the familiar mesons and baryons – about which far less can be said with confidence. I concentrate on the spin/flavor/color structure of the wave functions and develop a model for estimating masses and magnetic moments.

5.1
The Schrödinger Equation

The analysis of a bound state is simplest when the constituents travel at speeds substantially less than c, for then the apparatus of nonrelativistic quantum mechanics can be brought to bear. Such is the case for hydrogen and for hadrons made out of heavy quarks (c and b). The more familiar light-quark states (made out of u, d, and s) are much more difficult to handle, because they are intrinsically relativistic, and quantum field theory (as currently practiced) is not well suited to the description of bound states. Most of the techniques available assume that the particles are initially free, and become free again after some brief interaction, whereas in a bound state the particles interact continuously over an extended period. Thus there exists a very rich theory of 'charmonium' ($c\bar{c}$, the ψ meson system), and 'bottomonium' ($b\bar{b}$, the Υ system), but comparatively little can be said about the excited states of $u\bar{u}$ (say) or $d\bar{d}$.

How can you tell whether a given bound state is relativistic or not? The simplest criterion is as follows: if the binding energy is small compared to the rest energies of the constituents, then the system is nonrelativistic.* For example,

* In general, the total energy of a composite system is the sum of three terms: (i) the rest energy of the constituents, (ii) the kinetic energy of the constituents, and (iii) the potential energy of the configuration. The latter two are typically comparable in size (the precise relation is given by the virial theorem). If the binding energy is much less than the constituent rest energies, so too is their kinetic energy, and hence the system is nonrelativistic. On the other hand, if the mass of the composite structure is substantially different from the sum of the rest masses of the constituents, then the kinetic energy is large and the system is relativistic.

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

the binding energy of hydrogen is 13.6 eV, whereas the rest energy of an electron is 511 000 eV – this is clearly a nonrelativistic system. On the other hand, quark–quark binding energies are on the order of a few hundred MeV, which is about the same as the effective rest energy of u, d, or s quarks, but substantially less than c, b, and t (see Table 4.4). So the light-quark hadrons are relativistic, but heavy-quark systems are not.

The foundation for nonrelativistic quantum theory is Schrödinger's equation [1].

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\Psi = i\hbar\frac{\partial}{\partial t}\Psi \tag{5.1}$$

It governs the time evolution of the *wave function* $\Psi(\mathbf{r}, t)$, describing a particle of mass m in the presence of a specified potential energy $V(\mathbf{r}, t)$. Specifically, $|\Psi(\mathbf{r}, t)|^2 d^3\mathbf{r}$ is the probability of finding the particle in the infinitesimal volume $d^3\mathbf{r}$, at time t. Since the particle must be *somewhere*, the integral of $|\Psi|^2$ over *all* space has to be 1:

$$\int |\Psi|^2 d^3\mathbf{r} = 1 \tag{5.2}$$

We say that the wave function is 'normalized'.*

If V does not depend explicitly on t, the Schrödinger equation can be solved by *separation of variables:*

$$\Psi(\mathbf{r}, t) = \psi(\mathbf{r})e^{-iEt/\hbar} \tag{5.3}$$

where ψ satisfies the *time-independent Schrödinger equation*

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\psi = E\psi \tag{5.4}$$

and the separation constant E is the energy of the particle. The operator on the left is the *Hamiltonian:*

$$H \equiv -\frac{\hbar^2}{2m}\nabla^2 + V \tag{5.5}$$

and the (time-independent) Schrödinger equation has the form of an eigenvalue equation:

$$H\psi = E\psi \tag{5.6}$$

ψ is an eigenfunction of H, and E is the eigenvalue.†

* A solution to the Schrödinger equation can be multiplied by any constant and remain a solution. In practice, we fix this constant by demanding that Equation 5.2 be satisfied; this process is called 'normalizing' the wave function.

† Notice that $|\Psi|^2 = |\psi|^2$. For most purposes it is only the absolute square of the wave function that matters, and we shall work almost exclusively with ψ. Casually, we often refer to ψ as 'the wave function', but remember that the *actual* wave function carries the exponential time dependence.

Table 5.1 Spherical harmonics for $l = 0, 1, 2$, and 3

$$Y_0^0 = \frac{1}{\sqrt{4\pi}}, \quad Y_1^0 = \sqrt{\frac{3}{4\pi}}\cos\theta, \quad Y_2^0 = \sqrt{\frac{5}{16\pi}}(3\cos^2\theta - 1),$$

$$Y_3^0 = \sqrt{\frac{7}{16\pi}}(5\cos^3\theta - 3\cos\theta),$$

$$Y_1^1 = -\sqrt{\frac{3}{8\pi}}\sin\theta e^{i\phi}, \quad Y_2^1 = -\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta e^{i\phi}$$

$$Y_3^1 = -\sqrt{\frac{21}{64\pi}}\sin\theta(5\cos^2\theta - 1)e^{i\phi}$$

$$Y_2^2 = \sqrt{\frac{15}{32\pi}}\sin^2\theta e^{2i\phi}, \quad Y_3^2 = \sqrt{\frac{105}{32\pi}}\sin^2\theta\cos\theta e^{i\phi}$$

$$Y_3^3 = -\sqrt{\frac{35}{64\pi}}\sin^3\theta^{3i\phi}$$

In the case of a spherically symmetrical (or 'central') potential, V is a function only of the distance from the origin, and the (time-independent) Schrödinger equation separates in spherical coordinates:

$$\psi(r,\theta,\phi) = \frac{u(r)}{r}\, Y_l^{m_l}(\theta,\phi) \tag{5.7}$$

Here Y is a *spherical harmonic*; these functions are tabulated in many places (including the *Particle Physics Booklet*); a few of the more useful ones are given in Table 5.1. The constants l and m_l correspond to the orbital angular momentum quantum numbers introduced in Chapter 4. Meanwhile $u(r)$ satisfies the *radial Schrödinger equation*,

$$-\frac{\hbar^2}{2m}\frac{d^2u}{dr^2} + \left[V(r) + \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}\right]u = Eu \tag{5.8}$$

Curiously, this has exactly the same form as Equation 5.4 for one dimension, except that the potential is augmented by the *centrifugal barrier*, $(\hbar^2/2m)l(l+1)/r^2$.

That is about as far as we can pursue the matter in general terms; at this point we have to put in the particular potential $V(r)$ for the problem at hand. The strategy is to solve the radial equation for $u(r)$, combine the result with the appropriate spherical harmonic, and multiply by the exponential factor $\exp(-iEt/\hbar)$, to get the full wave function Ψ. In the course of solving the radial equation, however, we discover that only certain special values of E lead to acceptable results. For *most* values of E the solution to Equation 5.8 blows up at large r, and yields a non-normalizable wave function. Such a solution does not represent a possible physical state. This rather

technical detail is the source of the most striking and important feature of quantum mechanics: a bound state cannot have just any old energy (as it could classically); instead, the energy can take on only certain specific values, the *allowed energies* of the system. Indeed, our real concern is not with the wave function itself, but with the *spectrum* of allowed energies.

5.2 Hydrogen

The hydrogen atom (electron plus proton) is not an elementary particle, of course, but it serves as the model for nonrelativistic bound systems. The proton is so heavy (relatively) that it just sits at the origin; the wave function in question is that of the *electron*. Its potential energy, due to the electrical attraction of the nucleus, is (in Gaussian units)

$$V(r) = -\frac{e^2}{r} \tag{5.9}$$

When this potential is put into the radial equation, it is found that normalizable solutions occur only when E assumes one of the special values

$$E_n = -\frac{me^4}{2\hbar^2 n^2} = -\alpha^2 mc^2 \left(\frac{1}{2n^2}\right) = -13.6\ \text{eV}/n^2 \tag{5.10}$$

where $n = 1, 2, 3, \ldots$, and

$$\alpha \equiv \frac{e^2}{\hbar c} = \frac{1}{137.036} \tag{5.11}$$

is the *fine structure constant*. The corresponding (normalized) wave function, $\Psi_{n,l,m_l}(r, \theta, \phi, t)$, is

$$\left\{\left(\frac{2}{na}\right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3}\right\}^{1/2} e^{-r/na} \left(\frac{2r}{na}\right)^l L_{n-l-1}^{2l+1}\left(\frac{2r}{na}\right) Y_l^{m_l}(\theta, \phi) e^{-iE_n t/\hbar} \tag{5.12}$$

where

$$a \equiv \frac{\hbar^2}{me^2} = 0.529 \times 10^{-8}\ \text{cm} \tag{5.13}$$

is the *Bohr radius* (roughly speaking, the size of the atom), and L is an *associated Laguerre polynomial*.

Obviously, the wave function itself is a mess, but that's not really what concerns us. The crucial thing is the formula of the allowed energies, Equation 5.10. It was first obtained by Bohr in 1913 (more than a decade before the Schrödinger equation was introduced) by a serendipitous amalgam of inapplicable classical

ideas and primitive quantum theory – an inspired blend, as Rabi put it, of 'artistry and effrontery'.

Notice that the wave function is labeled by three numbers: n (the *principal quantum number*), which can be any positive integer – it determines the energy of the state (Equation 5.10); l, an integer ranging from 0 to $n - 1$ that specifies the total orbital angular momentum (Equation 4.2); and m_l, an integer that can assume any value between $-l$ and $+l$, giving the z component of the angular momentum (Equation 4.4). Evidently, there are $2l + 1$ different m_l's, for each l, and n different l's, for each n. The total number of distinct states that share the same principal quantum number n (and hence the same energy) is, therefore

$$\sum_{l=0}^{n-1}(2l+1) = n^2 \tag{5.14}$$

This is called the *degeneracy* of the nth energy level. Hydrogen is a surprisingly degenerate system; spherical symmetry alone dictates that the $2l + 1$ states with a given value of the total angular momentum should be degenerate, since they differ only in the orientation of $\mathbf{L}$, but this suggests a sequence 1, 3, 5, 7, ..., whereas the energy levels of hydrogen have much higher degeneracies: 1, 4, 9, 16, ... This is because states with different l share the same n; it is an unusual feature of the Coulomb potential.

In practice, we do not measure the energies themselves, but rather the wavelength of the light emitted when the electron makes a transition from a higher level to a lower one (or the light *absorbed* when it goes the other way) [2]. The photon carries the *difference* in energy between the initial and final states. According to the Planck formula (Equation 1.1),

$$E_{\text{photon}} = h\nu = E_{\text{initial}} - E_{\text{final}} = -\frac{me^4}{2\hbar^2}\left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right) \tag{5.15}$$

The emitted wavelength is therefore given by

$$\frac{1}{\lambda} = R\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) \tag{5.16}$$

where

$$R \equiv \frac{me^4}{4\pi\hbar^3 c} = 1.09737 \times 10^5/\text{cm} \tag{5.17}$$

Equation 5.16 is the famous *Rydberg formula* for the spectrum of hydrogen. It was discovered experimentally by nineteenth-century spectroscopists, for whom R was simply an empirical constant. The greatest triumph of Bohr's theory was its derivation of the Rydberg formula, and the expression for R in terms of the fundamental constants m, e, c, and $\hbar$ (Figure 5.1).

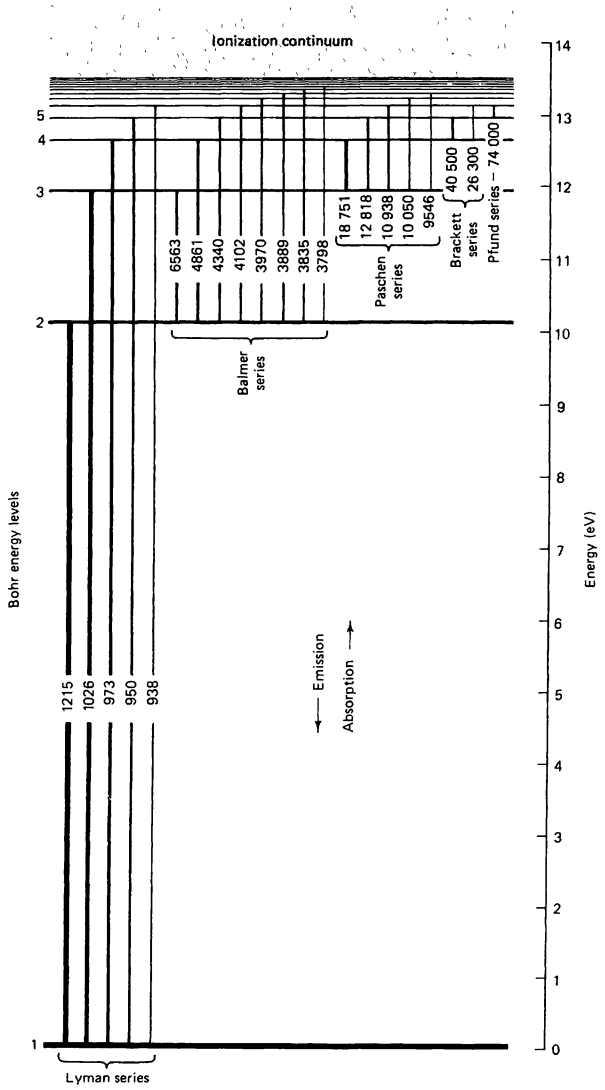


Fig. 5.1 The spectrum of hydrogen. When an atom changes from one state to another, the difference in energy appears as a quantum of radiation. The energy of the photon is directly proportional to the frequency of the radiation and inversely proportional to its wavelength. Absorption of radiation stimulates a transition to a state of higher energy; an atom falling to a state of lower energy emits radiation. The spectrum is organized into series of lines that share a common lower level. Wavelengths are given in angstroms; the relative intensity of the lines is indicated by thickness. (*Source*: Hansch, T. W., Schawlow, A. L. and Series, G. W. (March 1979) 'The Spectrum of Atomic Hydrogen', *Scientific American*, p. 94, reprinted by permission.)

5.2.1
Fine Structure

As the precision of experimental spectroscopy improved, small departures from the Rydberg formula were detected. Spectral lines were resolved into doublets, triplets, and even larger families of closely spaced peaks. This *fine structure* is actually attributable to two distinct mechanisms:

1. *Relativistic correction:* The first term in the Hamiltonian (Equation 5.5) comes from the classical expression for kinetic energy ($p^2/2m$), with the quantum replacement $p \to -i\hbar\nabla$. The lowest-order relativistic correction (Problem 5.4) is $-p^4/8m^3c^2$.
2. *Spin–orbit coupling:* The spinning electron constitutes a tiny magnet, with a dipole moment*

$$\boldsymbol{\mu}_e = -\frac{e}{mc}\mathbf{S} \tag{5.18}$$

From the electron's perspective the 'orbiting' proton sets up a magnetic field $\mathbf{B}$, and the spin–orbit term is the associated magnetic energy $-\boldsymbol{\mu}_e \cdot \mathbf{B}$.

The net result is a perturbation of the nth energy level by the amount [1]

$$\Delta E_{\text{fs}} = -\alpha^4 mc^2 \frac{1}{4n^4}\left(\frac{2n}{(j+\frac{1}{2})} - \frac{3}{2}\right) \tag{5.19}$$

where $j = l \pm \frac{1}{2}$ is the total angular momentum (spin plus orbital) of the electron (Equation 4.12). Recall that the Bohr energies go like $\alpha^2 mc^2$ (Equation 5.10); fine structure carries two more powers of α, and hence is smaller by a factor of about 10^{-4}. So we're talking about a tiny correction.† Since l can take on any integer value

* In the SI system the magnetic dipole moment is defined as current times area (Ia), but in Gaussian units it is Ia/c. The proportionality factor between the magnetic dipole moment and the angular momentum is known as the *gyromagnetic ratio*. Classically, it should have the (Gaussian) value [3] $-e/2mc$, and this is correct for *orbital* angular momentum. But it turns out that spin is 'twice as effective as it ought to be' in producing a magnetic dipole (one of the major successes of Dirac's original theory of the electron was its explanation of this extra 2). As it happens, however, even *this* is not quite right; there are minute corrections introduced by quantum electrodynamics (QED) that were first calculated by Schwinger in the late 1940s. By now, both experimental and theoretical determinations of the *anomalous magnetic moment* of the electron have been carried to fantastic precision, and stunning agreement [4].

† The *fine structure constant*, α, owes its name to the fact that it (or rather, α^2) sets the relative scale of the fine structure in hydrogen. However, one might equally well say that α^2 sets the scale of the Bohr levels themselves. Actually, the best way to characterize the fine structure constant is to say that it is the dimensionless measure (in units of $\hbar c$) of the (square of the) fundamental charge: $\alpha = e^2/\hbar c$.

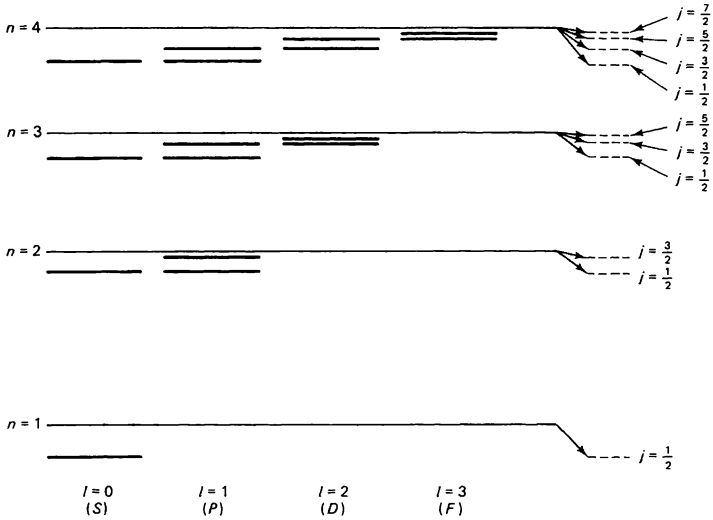


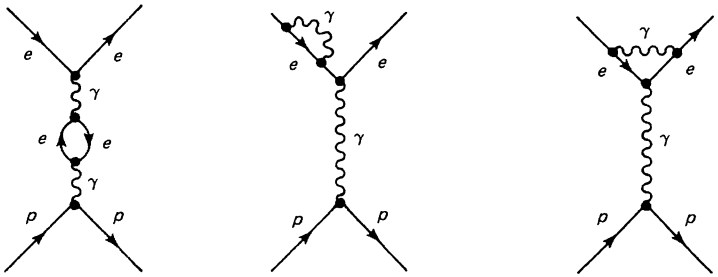
Fig. 5.2 Fine structure in hydrogen. The nth Bohr level (fine line) splits into n sublevels (dashed lines), characterized by $j = \frac{1}{2}, \frac{3}{2}, \ldots, (n - \frac{1}{2})$. Except for the last of these, two different values of l contribute to each level: $l = j - \frac{1}{2}$ and $l = j + \frac{1}{2}$. Spectroscopists' nomenclature – S for $l = 0$, P for $l = 1$, D for $l = 2$, F for $l = 3$ – is indicated. All levels are shifted downward, as shown (the diagram is not to scale, however).

from 0 to $n - 1$, j can be any half-integer from $\frac{1}{2}$ to $n - \frac{1}{2}$; thus the nth Bohr level, E_n, splits into n sublevels (see Figure 5.2).

5.2.2
The Lamb Shift

A striking feature of the fine structure formula (Equation 5.19) is that it depends only on j, not on l; in general, two different values of l share the same energy. For example, the $2S_{1/2}$ ($n = 2, l = 0, j = \frac{1}{2}$) and $2P_{1/2}$ ($n = 2, l = 1, j = \frac{1}{2}$) states remain perfectly degenerate. In 1947, Lamb and Retherford performed a classic experiment [5] which demonstrated that this is not, in fact, the case; the S state is slightly higher in energy than the P state. The explanation of the *Lamb shift* was provided by Bethe, Feynman, Schwinger, Tomonaga, and others; it is due to the *quantization of the electromagnetic field itself*. Everywhere else in the analysis – the Bohr levels, fine structure formula, and even hyperfine splitting (in the next section) – the electromagnetic field is treated entirely classically. The Lamb shift, by contrast, is an example of a *radiative correction* in QED, to which the semiclassical* theory is insensitive. In the Feynman formalism, it results from *loop diagrams*, such as those in Figure 5.3, which we shall discuss quantitatively later on.

* I call it *semi*classical because the *electron* is treated quantum mechanically, whereas the electromagnetic *field* is treated classically.



Vacuum polarization Electron mass renormalization Anomalous magnetic moment

Fig. 5.3 Some loop diagrams contributing to the Lamb shift.

Qualitatively, the first diagram in Figure 5.3 describes spontaneous production of electron–positron pairs in the neighborhood of the nucleus (misnamed *vacuum polarization*), leading to a partial screening of the proton's charge (Figure 2.1). The second diagram reflects the fact that the ground state of the electromagnetic field is not zero [6]; as the electron moves through the 'vacuum fluctuations' in the field, it jiggles slightly, and this alters its energy. The third diagram leads to a tiny modification of the electron's magnetic dipole moment (see footnote to Equation 5.18). We are not in a position to calculate these effects now, but here are the results [7]: For $l = 0$,

$$\Delta E_{\text{Lamb}} = \alpha^5 mc^2 \frac{1}{4n^3} k(n, 0) \tag{5.20}$$

where $k(n, 0)$ is a numerical factor that varies slightly with n, from 12.7 (for $n = 1$) to 13.2 (for $n \to \infty$). For $l \neq 0$,

$$\Delta E_{\text{Lamb}} = \alpha^5 mc^2 \frac{1}{4n^3} \left\{ k(n, l) \pm \frac{1}{\pi (j + \frac{1}{2})(l + \frac{1}{2})} \right\}, \quad \text{for } j = l \pm \tfrac{1}{2} \tag{5.21}$$

where $k(n, l)$ is a very small number (less than 0.05) that varies slightly with n and l. Evidently the Lamb shift is miniscule, except for states with $l = 0$, where it is about one-tenth the size of the fine structure. However, because it depends on l, it lifts the degeneracy of the pairs of states with common n and j, in Figure 5.2, and in particular it splits the $2S_{1/2}$ and $2P_{1/2}$ levels (Problem 5.6).

5.2.3
Hyperfine Splitting

Fine structure and the Lamb shift are minute corrections to the Bohr energy levels, but they are not the end of the story; there is a refinement that is smaller still (by a factor of 1000), due to the spin of the *nucleus*. The proton, like the electron,

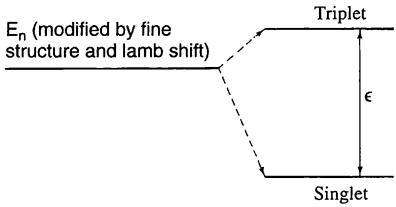


Fig. 5.4 Hyperfine splitting for $l = 0$.

constitutes a tiny magnet, but because it is so much heavier, its dipole moment is much smaller:

$$\boldsymbol{\mu}_p = \gamma_p \frac{e}{m_p c} \mathbf{S}_p \tag{5.22}$$

(The proton is a composite object, and its magnetic moment is not simply $e\hbar/2m_p c$, as it would be for a truly elementary particle of spin $\frac{1}{2}$. Hence the factor γ_p, whose experimental value is 2.7928. Later on we shall see how to calculate this quantity in the quark model.) The nuclear spin interacts with the electron's orbital motion by the same mechanism as the spin–orbit contribution to fine structure; in addition, it interacts directly with the electron spin. Together, the nuclear spin–orbit interaction and the proton–electron spin–spin coupling are responsible for the *hyperfine splitting* [8]:

$$\Delta E_{\text{hf}} = \left(\frac{m}{m_p}\right) \alpha^4 mc^2 \frac{\gamma_p}{2n^3} \frac{\pm 1}{(f + \frac{1}{2})(l + \frac{1}{2})}, \quad \text{for} \quad f = j \pm \tfrac{1}{2} \tag{5.23}$$

where f is the *total* angular momentum quantum number (orbital plus both spins).

Comparing the fine structure formula (Equation 5.19), we see that the difference in scale is due to the mass ratio (m/m_p) out front; it follows that hyperfine effects in hydrogen are about 1000 times smaller. If the orbital angular momentum is zero ($l = 0$), then f can take on two possible values: zero, in the singlet state (when the spins are oppositely aligned) and one, in the triplet state (when the spins are parallel). Thus, each $l = 0$ level splits into two, with the singlet pushed down and the triplet lifted up (Figure 5.4). In the ground state [9] $n = 1$ the energy gap is

$$\epsilon = E_{\text{triplet}} - E_{\text{singlet}} = \frac{32\gamma_p E_1^2}{3m_p c^2} \tag{5.24}$$

corresponding to a photon of wavelength

$$\lambda = \frac{2\pi\hbar c}{\epsilon} = 21.1 \text{ cm} \tag{5.25}$$

This is the transition that gives rise to the famous '21-centimeter line' in microwave astronomy [10].

5.3 Positronium

The theory of hydrogen carries over, with some modifications, to the so-called 'exotic' atoms, in which either the proton or the electron is replaced by some other particle. For instance, one can make muonic hydrogen ($p^+\mu^-$), pionic hydrogen ($p^+\pi^-$), positronium (e^+e^-), muonium (μ^+e^-), and so on. Of course, these exotic states are unstable, but many of them last long enough to exhibit a well-defined spectrum. In particular, positronium provides a rich testing ground for QED. It was analyzed theoretically by Pirenne in 1944, and first produced in the laboratory by Deutsch in 1951 [11]. In particle physics, positronium assumes special importance as the model for quarkonium.

The most conspicuous difference between positronium and hydrogen is that we are no longer dealing with a heavy, essentially stationary nucleus, around which the electron orbits, but rather with two particles of *equal* mass, both orbiting the common center. As in classical mechanics, this two-body problem can be converted into an equivalent one-body problem with the *reduced mass* [1]

$$m_{\text{red}} = \frac{m_1 m_2}{m_1 + m_2} \tag{5.26}$$

In the case of positronium $m_1 = m_2 = m$, so $m_{\text{red}} = m/2$, and we get the unperturbed energy levels by the simple substitution $m \to m/2$ in the Bohr formula (Equation 5.10):*

$$E_n^{\text{pos}} = \tfrac{1}{2} E_n = -\alpha^2 mc^2 \tfrac{1}{4n^2} \quad (n = 1, 2, 3, \ldots) \tag{5.27}$$

For example, the ground-state binding energy is $13.6\,\text{eV}/2 = 6.8\,\text{eV}$. The wave functions are the same as hydrogen's (Equation 5.12), except that the Bohr radius, which goes like $1/m$ (Equation 5.13), is doubled:

$$a^{\text{pos}} = 2a = 1.06 \times 10^{-8}\ \text{cm} \tag{5.28}$$

The perturbations run much as before, apart from pesky numerical factors, with one dramatic exception: in positronium, the hyperfine splitting is of the same order as the fine structure ($\alpha^4 mc^2$), since the mass ratio (m/m_p) that suppresses proton spin effects in hydrogen is *one* for positronium.† Meanwhile, since the 'nucleus' (e^+) is no longer stationary, there is a new correction due to the finite propagation

* In the case of hydrogen, the reduced mass differs from the electron mass by only a very small amount, about 0.05%. Nevertheless, technically the m in the Bohr formula is the reduced mass, and this leads to observable differences between the spectra of hydrogen and deuterium.

† This leads to some terminological confusion in the literature. I'll use the words 'fine structure' for *all* perturbations of order $\alpha^4 mc^2$, except the pair annihilation term (see below), including the spin–spin and positron spin–orbit couplings, whose analogs in hydrogen would be called 'hyperfine'.

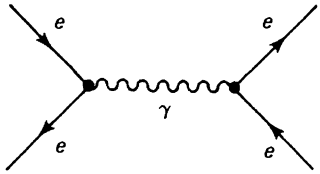


Fig. 5.5 Pair annihilation diagram, which affects the spectrum of positronium but does not occur in hydrogen.

time for the electromagnetic field; its contribution is also of order $(\alpha^4 mc^2)$. When all this is put together, the fine structure formula for positronium is found to be [11]

$$E_{\text{fs}}^{\text{pos}} = \alpha^4 mc^2 \frac{1}{2n^3}\left[\frac{11}{32n} - \frac{(1+\frac{1}{2}\epsilon)}{(2l+1)}\right] \tag{5.29}$$

where $\epsilon = 0$ for the singlet spin combination, whereas for the triplet*

$$\epsilon = \left\{\begin{array}{ll} \dfrac{-(3l+4)}{(l+1)(2l+3)}, & \text{for } j = l+1 \\ \dfrac{1}{l(l+1)}, & \text{for } j = l \\ \dfrac{(3l-1)}{l(2l-1)}, & \text{for } j = l-1 \end{array}\right\} \tag{5.30}$$

The Lamb shift, of order $\alpha^5 mc^2$, makes a smallish correction to this; however, since the 'accidental' degeneracy is already broken at the fine structure level in positronium, this contribution loses much of its interest. There is, however, an entirely new perturbation, with no analog in hydrogen, resulting from the fact that e^+ and e^- can annihilate temporarily to produce a virtual photon. In the Feynman picture, this process is represented by the diagram in Figure 5.5. Because it requires that the electron and positron coincide, this perturbation is proportional to $|\Psi\ (0)|^2$, and hence occurs only when $l = 0$ (Ψ goes like r^l near the origin – see Equation 5.12). Moreover, since the photon carries spin 1, it takes place only in the triplet configuration. This process raises the energy of the triplet S states by an amount

$$\Delta E_{\text{ann}} = \alpha^4 mc^2 \frac{1}{4n^3} \quad (l = 0, s = 1) \tag{5.31}$$

* In hydrogen, where the proton spin ($\mathbf{S}_p$) contributes only at the hyperfine level, we used $\mathbf{J}$ for the sum of the electron's spin and orbital angular momentum ($\mathbf{J} = \mathbf{L} + \mathbf{S}_e$); for the *total* angular momentum we needed a new letter: $\mathbf{F} = \mathbf{L} + \mathbf{S}_e + \mathbf{S}_p$. In positronium the two spins contribute on an equal footing, and it is customary to combine them first ($\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2$) and use $\mathbf{J}$ for the total: $\mathbf{J} = \mathbf{L} + \mathbf{S}_1 + \mathbf{S}_2$.

– the same order as fine structure. The complete splitting of the $n = 1$ and $n = 2$ Bohr levels in positronium is indicated on Figure (5.6)*

As in the case of hydrogen, positronium can make transitions from one state to another with the emission or absorption of a photon, whose wavelength is determined by the difference in energy between the two levels. *Unlike* hydrogen, positronium can also disintegrate completely, the positron annihilating the electron to produce two or more real photons. The charge conjugation number for positronium is $(-1)^{l+s}$, while for n photons $C = (-1)^n$ (see Section 4.4.2). Thus, charge conjugation invariance prescribes the selection rule

$$(-1)^{l+s} = (-1)^n \tag{5.32}$$

for the decay of positronium from the state l, s to n photons. Since the positron and electron overlap only when $l = 0$, such decays occur only from S states.† Evidently, the singlet ($s = 0$) must go to an even number of photons (typically two), whereas the triplet ($s = 1$) must go to an odd number (typically three). In Chapter 7 we will be in a position to calculate the lifetime of the ground state:

$$\tau = \frac{2\hbar}{\alpha^5 mc^2} = 1.25 \times 10^{-10} \text{ seconds} \tag{5.33}$$

5.4 Quarkonium

In the quark model all mesons are two-particle bound states, $q_1\bar{q}_2$, and it is natural to ask if the methods developed for hydrogen and positronium can be applied to mesons as well. Light-quark (u, d, s) states are intrinsically relativistic, so any analysis based on the Schrödinger equation is out of the question, but heavy-quark mesons ($c\bar{c}$, $b\bar{c}$, and $b\bar{b}$) should be suitable candidates. Even here, however, the interaction energy (E) is such a substantial fraction of the total that we are disposed to regard the various energy levels as representing *different particles*, with masses given by

$$M = m_1 + m_2 + E/c^2 \tag{5.34}$$

Unlike hydrogen and positronium, in which the forces at work are entirely electromagnetic, and the energy levels can be calculated to great precision, quarks are bound by the strong force; we don't know what potential to use, in place of Coulomb's law, or what the strong analog to magnetism might be, to obtain the spin couplings. In principle, these are derivable from chromodynamics, but no one

* Positronium states are conventionally labeled $n^{(2s+1)}l_j$, with l given in spectroscopist's notation (S for $l = 0$, P for $l = 1$, D for $l = 2$, etc.), and s the total spin (0 for the singlet, 1 for the triplet).

† Actually, positronium can in principle decay directly from a state with $l > 0$ by a higher-order process, but it is much more likely to cascade down to an S state first, and decay from there.

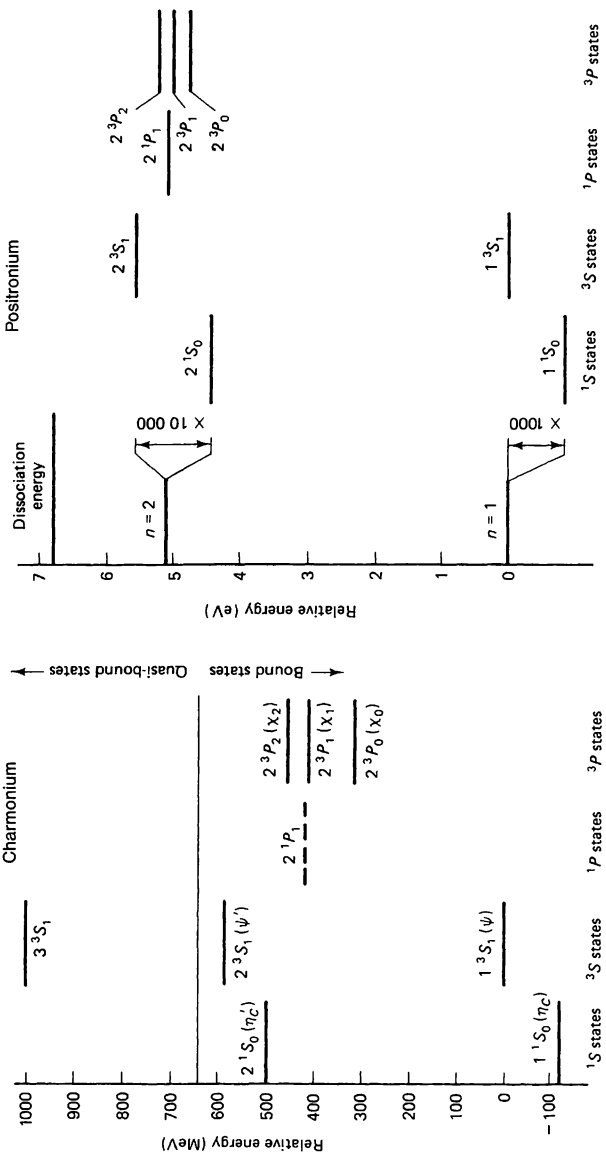


Fig. 5.6 Spectrum of energy levels in positronium and charmonium. Note that the scale is greater by a factor of 100 million for charmonium. In positronium, the various combinations of angular momentum cause only minuscule shifts in energy (shown by expanding the vertical scale), but in charmonium the shifts are much larger. All energies are given with reference to the 1^3S_1 state. At 6.8 eV positronium dissociates. At 633 MeV above the energy of the ψ charmonium becomes quasi-bound, because it can decay into D^0 and $\overline{D}^0$ mesons. (*Source*: Bloom, E. and Feldman, G. (May 1982) 'Quarkonium', *Scientific American*, p. 66, reprinted by permission.)

Table 5.2 'Bohr' Energy levels for linear-plus-Coulomb potential (Equation 5.83) with various values of F_0. They are for S-states ($l = 0$) and assume $\alpha_s = 0.2$, $m = 1500$ MeV/c^2 (reduced mass, 750 MeV/c^2).

F_0 (MeV fm^{-1})	E_1 (MeV)	E_2 (MeV)	E_3 (MeV)	E_4 (MeV)
500	307	677	961	1210
1000	533	1100	1550	1940
1500	727	1480	2040	2550

Numerical results from unpublished tables prepared by Richard E. Crandall.

yet knows how to do the calculation. Still, we can make some educated guesses, for chromodynamics is very similar in structure to electrodynamics, except for some nonlinear terms which, in the light of asymptotic freedom, probably don't contribute much at short distances.

In quantum chromodynamics (QCD), the short-distance behavior is dominated by one-gluon exchange, just as in QED it is dominated by one-photon exchange. Since the gluon and the photon are both massless spin-1 particles, the interactions are, in this approximation, identical, apart from the overall coupling strength and various so-called 'color factors', which result from counting the number of different gluons that contribute to a given process. At short range, therefore, we expect a Coulomb-like potential, $V \sim 1/r$, and a fine structure that is qualitatively similar to that of positronium [12]. On the other hand, at large distances we have to account for quark confinement: the potential must increase without limit. The precise functional form of $V(r)$ at large r is rather speculative; some authors favor a harmonic oscillator potential, $V \sim r^2$, others a logarithmic dependence, $V \sim \ln(r)$, still others a *linear* potential, $V \sim r$, corresponding to a constant force. The fact is, *any* of these can match the data reasonably well, because they do not differ substantially over the rather narrow range of distances for which we have sensitive probes.

For our purposes, we may as well choose

$$V(r) = -\frac{4}{3}\frac{\alpha_s \hbar c}{r} + F_0 r \tag{5.35}$$

where α_s is the chromodynamic analog to the fine structure constant, and $\frac{4}{3}$ is the appropriate color factor, which we'll calculate in Chapter 8. Unfortunately, exact solutions to the Schrödinger equation with linear-plus-Coulomb potential are not known, and I cannot give you a simple formula for the 'Bohr' energies. However, it can, of course, be done numerically (see Table 5.2), and F_0 can then be chosen so as to fit the data [13] (Problem 5.11). The result is about 16 tons(!), or, in more sensible units, 900 MeV fm^{-1}, which is to say that a quark and an antiquark attract one

another with a force of at least 16 tons, regardless of how far apart they are.* This perhaps makes it easier to understand why no one has ever managed to pull a free quark out of a meson.

5.4.1 Charmonium

Shortly before the discovery of the ψ, Appelquist and Politzer [14] suggested that if a heavy 'charm' quark existed (as Glashow and others had proposed) it should form a nonrelativistic bound state, $c\bar{c}$, with a spectrum of energy levels similar to positronium. They called the system 'charmonium' (which does more to emphasize the parallel than to beautify the language). When the ψ was found, in 1974, it was quickly identified as the 1^3S_1 state of charmonium.† (In the SLAC experiments the ψ was produced from e^+e^- annihilation through a virtual photon: $e^+e^- \to \gamma \to \psi$, so it has to carry the same quantum numbers as γ – in particular, spin 1. Thus it could not be the *ground* state of charmonium, but presumably it was the lowest-lying state with total angular momentum 1.) Consulting the positronium energy-level diagram (Figure 5.6), we immediately anticipate a spin-0 state at lower mass (the 1^1S_0) and six $n = 2$ configurations. Within two weeks the ψ' (2^3S_1) was found. This was easy, because it again carries the same spin – and parity – as the photon; it was produced in the same way as the ψ, simply by cranking up the beam energy.

In due course all the $n = 1$ and $n = 2$ states were discovered [15], save for the 2^1P_1 at a predicted mass of about 3500 MeV/c^2, which presents special experimental problems. The following nomenclature has been adopted: singlet S states (spin 0) are called η_c's, triplet S states (spin 1) are ψ's, and triplet P states (spin 0, 1, or 2) are designated χ_{c0}, χ_{c1}, χ_{c2}. For a while the value of n was indicated by primes, but this quickly got out of hand, and the current practice is simply to list the mass parenthetically; thus for $n = 1$ we have $\psi = \psi$ (3097); for $n = 2$, $\psi' = \psi$ (3686); for $n = 3$, $\psi'' = \psi$ (4040); for $n = 4$, $\psi''' = \psi$ (4160); and so on.‡ The correlation between states of charmonium and those of positronium is almost perfect (Figure 5.6). Bear in mind that the gap between the two $n = 1$ levels (which would be called hyperfine splitting in the case of hydrogen) is greater by a factor of 10^{11} in charmonium than in positronium. Yet even over so huge a change of scale, the ordering of the energy levels and, for a given value of n, their relative spacing, are strikingly similar.

All the charmonium states with $n = 1$ and $n = 2$ are relatively long-lived, because the OZI rule (Section 2.5) suppresses their strong decays. For $n \geq 3$ the charmonium masses lie above the threshold for (OZI-allowed) production of two

* At extremely short distances, F_0 and α_s themselves decrease, leading to asymptotic freedom, but for now we shall treat them as constants.

† The nomenclature is borrowed from that of positronium – see footnote after Equation 5.31.

‡ Some authors, including those of the *Particle Physics Booklet*, number states consecutively, starting with 1 for *each* combination of s, l, and j, so that what I call a $2P$ state (Figure 5.6) is listed as $1P$. Sorry about that. Incidentally, the $\psi(3770)$ is a displaced 3^3D_1 state, and does not really belong in this hierarchy.

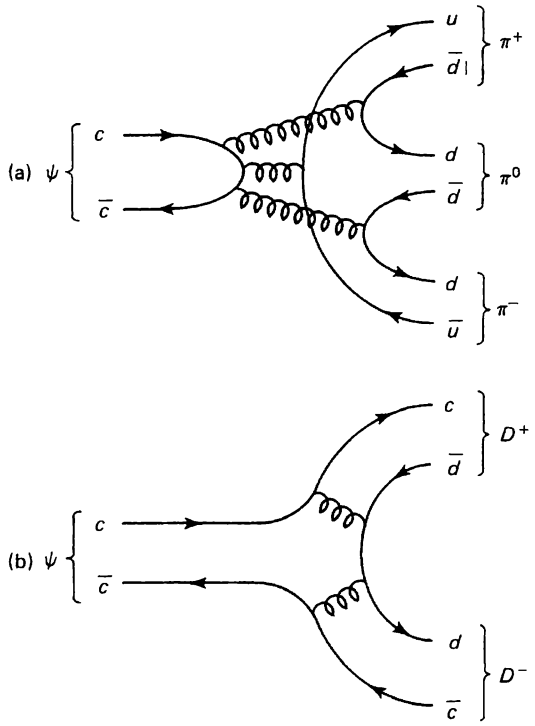


Fig. 5.7 (a) OZI-suppressed decay for charmonium below the $D\overline{D}$ threshold, (b) OZI-allowed decay for charmonium above the $D\overline{D}$ threshold.

charmed D mesons (D^0, $\overline{D}^0$ at a mass of 1865 MeV/c^2, or $D\pm$, at 1869 MeV/c^2). Their lifetimes are therefore much shorter, and we call them 'quasi-bound states' (see Figure 5.7). Quasi-bound states of charmonium have been observed going up at least as high as $n = 5$.

5.4.2
Bottomonium

In the aftermath of the November Revolution there was widespread speculation about the possible existence of a third-quark generation (b and t), and in 1976 Eichten and Gottfried [16] predicted that 'bottomonium' ($b\overline{b}$) would exhibit a hierarchy of bound states even richer than charmonium (Figure 5.8). The bottom analog to the D meson (to wit, the B) had an estimated mass large enough that not only the $n = 1$ and $n = 2$, but also the $n = 3$ levels should be bound. In 1977 the upsilon meson was discovered, and immediately interpreted as the 1^3S_1 state of bottomonium. At present, the $^3S_1$ states have been found for n up to 6, as well as

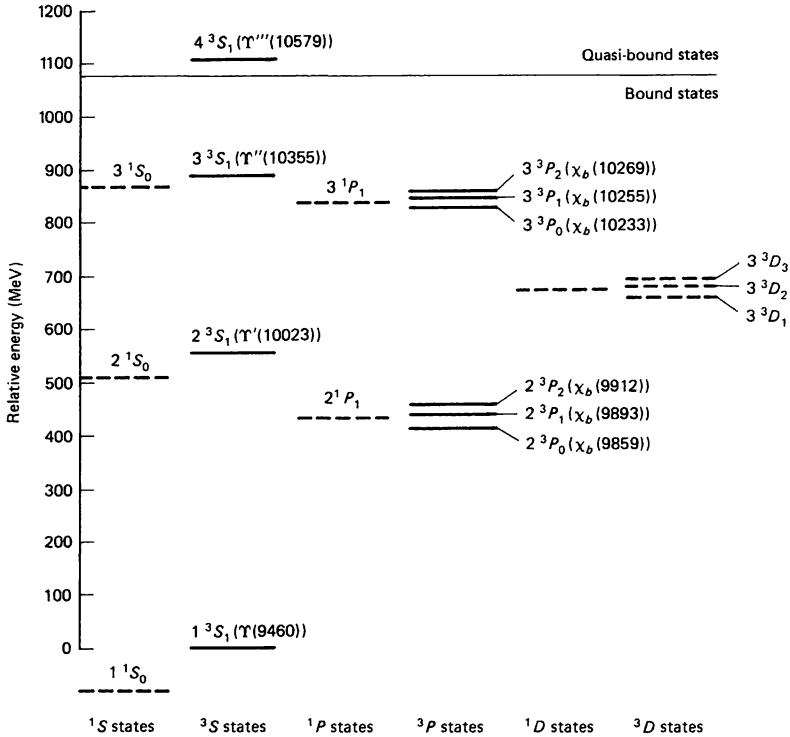


Fig. 5.8 Bottomonium. Note that there are far more bound states than for charmonium – compare Figure 5.6 (*Source*: Bloom, E. and Feldman, G. (May 1982) 'Quarkonium', *Scientific American*, p. 66, reprinted by permission. Corrected masses from *Particle Physics Booklet* (2006).)

the six ^3P states for $n = 2$ and $n = 3$. The level spacings in the ψ and Υ systems* are remarkably similar (Figure 5.9), in spite of the fact that the bottom quark is more than three times as heavy as the charm quark [17].

5.5 Light Quark Mesons

Consider now the mesons made entirely out of *light* quarks (u, d, s). These are relativistic systems, remember, so we cannot use the Schrödinger equation, and the theory is rather limited [18]. In particular, we shall not concern ourselves with the spectrum of excited states (Table 4.6), as we did in the case of the heavy

* In principle there should be a similar system for the $B_c^{\pm}$ mesons ($c\bar{b}$ and $b\bar{c}$), but so far only one of these, at 6286 MeV, has been produced in the laboratory.

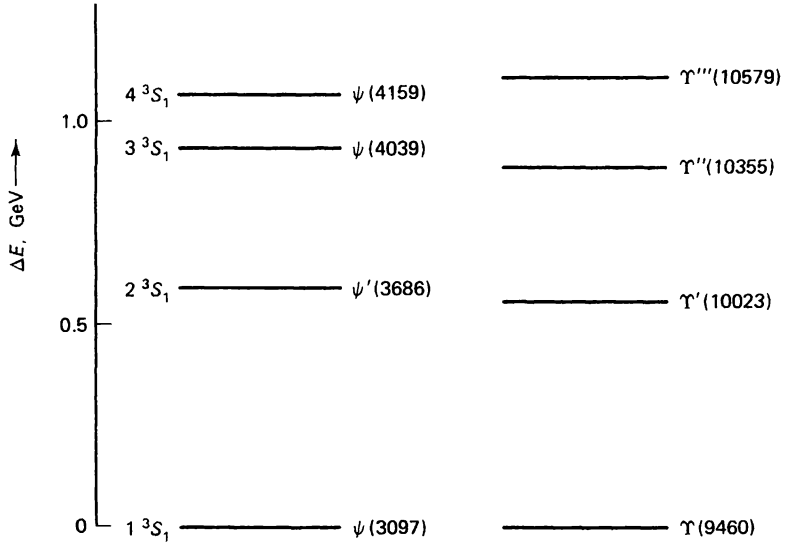


Fig. 5.9 Level spacings in the ψ and Υ systems. (*Source*: *Particle Physics Booklet* (2006).)

quarks, but will confine our attention to the ground state, with $l = 0$. The quark spins can be antiparallel (singlet state, $s = 0$) or parallel (triplet state, $s = 1$); the former configuration yields the pseudoscalar nonet, the latter gives the vector nonet (Figure 5.10).

To begin with, I want to clear up a problem that was not resolved in Chapter 1. We obtained nine mesons simply by combining a quark and an antiquark in all possible combinations (Section 1.8), but this left three neutral states with strangeness 0 ($u\bar{u}$, $d\bar{d}$, and $s\bar{s}$), and it was not clear which of these was the π^0, which the η, and which the η' (or, in the vector case, the ρ^0, ω, and ϕ). We are now in a position to resolve

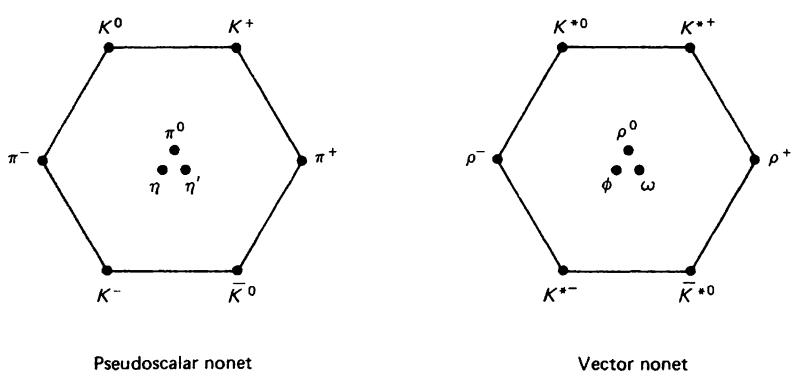


Fig. 5.10 Light-quark mesons with $l = 0$.

the ambiguity. The up and down quarks constitute an isospin doublet:

$$u = |\tfrac{1}{2}\tfrac{1}{2}), \quad d = |\tfrac{1}{2} - \tfrac{1}{2}) \tag{5.36}$$

So too do the antiquarks:

$$\bar{d} = -|\tfrac{1}{2}\tfrac{1}{2}), \quad \bar{u} = |\tfrac{1}{2} - \tfrac{1}{2}) \tag{5.37}$$

(Notice that $\bar{d}$ carries $I_3 = +\frac{1}{2}$, and $\bar{u}$ has $I_3 = -\frac{1}{2}$; within a multiplet, the particle with the higher charge is assigned the greater I_3. The minus sign is a technical detail [19], which does not affect the argument here in any essential way.) When we combine two particles with $I = \frac{1}{2}$, we obtain an isotriplet (Equation 4.15)

$$\begin{cases} |1 \quad 1\rangle = -u\bar{d} \\ |1 \quad 0) = (u\bar{u} - d\bar{d})/\sqrt{2} \\ |1 - 1\rangle = d\bar{u} \end{cases} \tag{5.38}$$

and an isosinglet (Equation 4.16)

$$|00) = (u\bar{u} + d\bar{d})/\sqrt{2} \tag{5.39}$$

In the case of the pseudoscalar mesons the triplet is the pion; for the vector mesons it is the ρ. Evidently, the π^0 (or the ρ^0) is neither $u\bar{u}$ nor $d\bar{d}$, but rather the linear combination

$$\pi^0, \rho^0 = (u\bar{u} - d\bar{d})/\sqrt{2} \tag{5.40}$$

If you could pull a π^0 apart, half the time you'd get a u in one hand and a $\bar{u}$ in the other, and half the time you'd get a d and a $\bar{d}$.

This leaves two $I = 0$ states (the isosinglet combination, Equation 5.39, and $s\bar{s}$) which must represent η and η' (or ω and ϕ). Here the situation is not so clean, for these particles carry identical quantum numbers, and they tend in practice to 'mix.' In the case of the pseudoscalars the physical states appear to be

$$\eta = (u\bar{u} + d\bar{d} - 2s\bar{s})/\sqrt{6} \tag{5.41}$$

$$\eta' = (u\bar{u} + d\bar{d} + s\bar{s})/\sqrt{3} \tag{5.42}$$

whereas for the vector mesons

$$\omega = (u\bar{u} + d\bar{d})/\sqrt{2} \tag{5.43}$$

$$\phi = s\bar{s} \tag{5.44}$$

To the extent that the Eightfold Way is a good symmetry, the pseudoscalar combinations are more 'natural', since the η', which treats u, d, and s symmetrically,

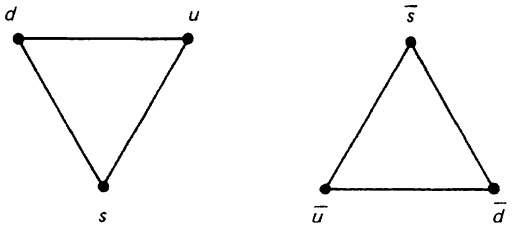


Fig. 5.11 Quarks and antiquarks.

is unaffected by $SU(3)$ transformations; it is a 'singlet' under $SU(3)$, in exactly the same sense that the π^0 is a singlet under $SU(2)$ (isospin). The η, meanwhile, transforms as part of an $SU(3)$ 'octet', whose other members are the three pions and the four K's. (This is, in fact, the original pseudoscalar octet.) By contrast, neither the ϕ nor the ω is an $SU(3)$ singlet. They are, you might say, 'maximally' mixed, since the strange quark is isolated from the other two. Incidentally, the other meson nonets seem to follow the $\phi - \omega$ mixing pattern [20].

Meanwhile, the strange mesons are constructed by combining an s quark with u or d

$$K^+ = u\bar{s}, \quad K^0 = d\bar{s}, \quad \overline{K}^0 = -s\bar{d}, \quad K^- = s\bar{u} \tag{5.45}$$

In the language of group theory, the three light quarks belong to the fundamental representation (denoted 3) of $SU(3)$, whereas the antiquarks belong to the conjugate representation ($\bar{3}$) (Figure 5.11). What we have done is combine these representations, obtaining an octet and a singlet:

$$3 \otimes \bar{3} = 8 \oplus 1 \tag{5.46}$$

just as in Chapter 4 we combined two two-dimensional (spin-$\frac{1}{2}$) representations of $SU(2)$ to obtain a triplet and a singlet:*

$$2 \otimes \bar{2} = 3 \oplus 1 \tag{5.47}$$

If $SU(3)$ were a perfect symmetry, all the particles in a given supermultiplet would have the same mass. But they obviously do *not*; the K weighs more than three times the π, for example. As I indicated in Chapter 4, the breaking of flavor symmetry is due to the fact that the quarks themselves have unequal masses; the u and d quarks weigh about the same, but the s quark is substantially heavier. Roughly speaking, the K's weigh more than the π's because they contain an s

* Unfortunately (from the point of view of notational consistency) representations of $SU(3)$ are customarily labeled by their dimension, whereas representations of $SU(2)$ are more often identified by their spin, so Equation 5.45 would usually be written as $\frac{1}{2} \otimes \frac{1}{2} = 1 \oplus 0$. By the way, it happens that the fundamental representation of $SU(2)$ is equivalent to its conjugate; there's only one kind of spin $\frac{1}{2}$. That's why we were able to represent $\bar{u}$ and $\bar{d}$ in Equation 5.79 in terms of ordinary isospin-$\frac{1}{2}$ states. For $SU(3)$ this is no longer the case.

Table 5.3 Pseudoscalar and vector meson masses. (MeV/c^2)

Meson	Calculated	Observed
π	139	138
K	487	496
η	561	548
ρ	775	776
ω	775	783
K^*	892	894
ϕ	1031	1020

in place of a u or d. But, that cannot be the whole story, for if it were, the ρ's would weigh the same as the π's; after all, they have the same quark content and are both in the spatial ground state ($n = 1, l = 0$). Since the pseudoscalar and vector mesons differ only in the relative orientation of the quark spins, the difference in their masses must be attributed to a spin–spin interaction, the QCD analog to hyperfine splitting in the ground state of hydrogen. This suggests the following meson mass formula:*

$$M(\text{meson}) = m_1 + m_2 + A\frac{(\mathbf{S}_1 \cdot \mathbf{S}_2)}{m_1 m_2} \tag{5.48}$$

where A is a constant [21]. By squaring $\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2$, we obtain

$$\mathbf{S}_1 \cdot \mathbf{S}_2 = \frac{1}{2}\left(S^2 - S_1^2 - S_2^2\right) = \left\{\begin{array}{ll} \frac{1}{4}\hbar^2, & \text{for } s = 1 \text{ (vector mesons)} \\ -\frac{3}{4}\hbar^2, & \text{for } s = 0 \text{ (pseudoscalars)} \end{array}\right\} \tag{5.49}$$

For constituent masses $m_u = m_d = 308$ MeV/c^2, $m_s = 483$ MeV/c^2, the best-fit value of A is $(2m_u/\hbar)^2 159$ MeV/c^2, and we obtain the results in Table 5.3.

5.6 Baryons

Some day, presumably, we shall be able to make nonrelativistic heavy-quark baryons – *ccc*, *ccb*, *cbb*, and *bbb*. These are the baryonic relatives of quarkonium – 'quarkelium', you might call it, since the nearest atomic analog would be helium. At present, though, it is hard enough to make a baryon with *one* heavy quark, never

* In $l = 0$ states the hyperfine correction is proportion to the dot product of the magnetic moments, $\mu_1 \cdot \mu_2$; dipole moments, in turn, are proportional to spin angular momentum and inversely proportional to mass. That's the inspiration behind Equation 5.46 Of course, this is for QED, not QCD. What's worse, it ignores the mass dependence of the wave function (contained in the 'constant' A), and it is based on nonrelativistic quantum mechanics. But nothing succeeds like success, and Equation 5.46 works surprisingly well. (Notice, however, that the η' is not included in the table – see Problem 5.12).

Table 5.4 Light-quark baryons (J = spin, P = parity, S = strangeness, I = isospin. This is not a complete list; baryons with spins as high as 11/2 have been observed)

			$S=-1$			
$SU(3)$ Representation	J^P	$S=0$	$I=0$	$I=1$	$S=-2$	$S=-3$
8	$\frac{1}{2}^+$	$N(939)$	Λ (1116)	Σ (1193)	Ξ (1318)	
10	$\frac{3}{2}^+$	Δ (1232)		Σ (1385)	Ξ (1530)	Ω (1672)
1	$\frac{1}{2}^-$		Λ (1405)			
	$\frac{3}{2}^-$		Λ (1520)			
8	$\frac{1}{2}^-$	$N(1535)$	Λ (1670)	Σ (1620)	?	
	$\frac{3}{2}^-$	$N(1520)$	Λ (1690)	Σ (1670)	Ξ (1820)	
	$\frac{5}{2}^-$	$N(1675)$	Λ (1830)	Σ (1775)	?	
10	$\frac{1}{2}^-$	Δ (1620)		?	?	?
	$\frac{3}{2}^-$	Δ (1700)		?	?	?
8	$\frac{3}{2}^+$	$N(1720)$	Λ (1890)	?	?	
	$\frac{5}{2}^+$	$N(1680)$	Λ (1820)	Σ (1915)	Ξ (2030)	
10	$\frac{5}{2}^+$	Δ (1905)		?	?	?
	$\frac{7}{2}^+$	Δ (1950)		Σ (2030)	?	?
8	$\frac{1}{2}^+$	$N(1440)$	Λ(1600)	Σ (1660)	?	

Source: Review of Particle Physics (2006), Section 14.4.

mind *three*, and I won't speculate here about the heavy-quark baryon spectrum. On the other hand, the array of observed *light* quark baryons is immense (see Table 5.4).

5.6.1
Baryon Wave Functions

Baryons are harder to analyze than mesons, for several reasons. In the first place, a baryon is a three-body system. There's not just *one* orbital angular momentum to consider, but *two* (see Figure 5.12). We'll concentrate on the ground state, for which $l = l' = 0$. In that case, the angular momentum of the baryon comes entirely from the combined spins of the three quarks. Now, the quarks carry spin $\frac{1}{2}$, so each can occupy either of two states: 'spin up' ($\uparrow$) or 'spin down' ($\downarrow$). Thus, we have *eight* possible states for the three quarks: ($\uparrow\uparrow\uparrow$), ($\uparrow\uparrow\downarrow$), ($\uparrow\downarrow\uparrow$), ($\uparrow\downarrow\downarrow$), ($\downarrow\uparrow\uparrow$), ($\downarrow\uparrow\downarrow$), ($\downarrow\downarrow\uparrow$), and ($\downarrow\downarrow\downarrow$). But these are not the most convenient configurations to work with, because they are not eigenstates of the total angular momentum. As we found in Example 4.2, the quark spins can combine to give a total of $\frac{3}{2}$ or $\frac{1}{2}$, and

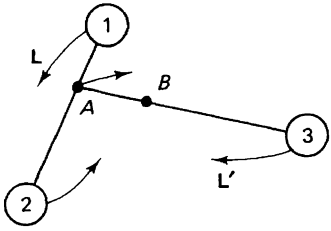


Fig. 5.12 Orbital angular momenta for a three-body system. **L** is the angular momentum of 1 and 2 about their center of mass (A); **L′** is the angular momentum of this combination and 3 about the center of mass of all three (B).

the latter can be achieved in two distinct ways. Specifically,

$$\left.\begin{aligned} |\tfrac{3}{2}\tfrac{3}{2}\rangle &= (\uparrow\uparrow\uparrow) \\ |\tfrac{3}{2}\tfrac{1}{2}\rangle &= (\uparrow\uparrow\downarrow + \uparrow\downarrow\uparrow + \downarrow\uparrow\uparrow)/\sqrt{3} \\ |\tfrac{3}{2}-\tfrac{1}{2}\rangle &= (\downarrow\downarrow\uparrow + \downarrow\uparrow\downarrow + \uparrow\downarrow\downarrow)/\sqrt{3} \\ |\tfrac{3}{2}-\tfrac{3}{2}\rangle &= (\downarrow\downarrow\downarrow) \end{aligned}\right\} \quad \text{spin } \tfrac{3}{2}\ (\psi_s) \tag{5.50}$$

$$\left.\begin{aligned} |\tfrac{1}{2}\tfrac{1}{2}\rangle_{12} &= (\uparrow\downarrow - \downarrow\uparrow)\uparrow/\sqrt{2} \\ |\tfrac{1}{2}-\tfrac{1}{2}\rangle_{12} &= (\uparrow\downarrow - \downarrow\uparrow)\downarrow/\sqrt{2} \end{aligned}\right\} \quad \text{spin } \tfrac{1}{2}\ (\psi_{12}) \tag{5.51}$$

$$\left.\begin{aligned} |\tfrac{1}{2}\tfrac{1}{2}\rangle_{23} &= \uparrow(\uparrow\downarrow - \downarrow\uparrow)/\sqrt{2} \\ |\tfrac{1}{2}-\tfrac{1}{2}\rangle_{23} &= \downarrow(\uparrow\downarrow - \downarrow\uparrow)/\sqrt{2} \end{aligned}\right\} \quad \text{spin } \tfrac{1}{2}\ (\psi_{23}) \tag{5.52}$$

The spin-$\frac{3}{2}$ combinations are completely *symmetric*, in the sense that interchanging any two particles leaves the state untouched. The spin-$\frac{1}{2}$ combinations are partially *anti*symmetric – interchange of two particles switches the sign. The first set is antisymmetric in particles 1 and 2 – hence the subscript; the second is antisymmetric in 2 and 3. We could also, of course, construct a pair of states antisymmetric in 1 and 3:

$$\left.\begin{aligned} |\tfrac{1}{2}\ \tfrac{1}{2}\rangle_{13} &= (\uparrow\uparrow\downarrow - \downarrow\uparrow\uparrow)/\sqrt{2} \\ |\tfrac{1}{2}-\tfrac{1}{2}\rangle_{13} &= (\uparrow\downarrow\downarrow - \downarrow\downarrow\uparrow)/\sqrt{2} \end{aligned}\right\} \quad \text{spin } \tfrac{1}{2}\ (\psi_{13}) \tag{5.53}$$

However, these are not independent of the other two; as you can check for yourself,

$$|\ \rangle_{13} = |\ \rangle_{12} + |\ \rangle_{23} \tag{5.54}$$

In the language of group theory, the direct product of three fundamental (two-dimensional) representations of $SU(2)$ decomposes into the direct sum of a four-dimensional representation and two two-dimensional representations:*

$$2 \otimes 2 \otimes 2 = 4 \oplus 2 \oplus 2 \tag{5.55}$$

A second respect in which baryons are more complicated than mesons has to do with the Pauli exclusion principle. In its original formulation the Pauli principle stated that no two electrons can occupy the same quantum state. It was designed to explain why all the electrons in an atom don't simply cascade down to the ground state, ψ_{100} (there wouldn't be much left of chemistry if they *did*): they *cannot*, because the ground state can only accommodate two of them – one spin up, one spin down. Once those positions are occupied, the next electrons are stuck in the first excited state, $n = 2, \ldots$, and so on. In this form, the Pauli principle seems a little *ad hoc*, but it is actually based on something far deeper: if two particles are *absolutely identical*, then the wave function should treat them on an equal footing. If someone secretly interchanges them, the physical state should not be altered. You might conclude from this that $\psi(1, 2) = \psi(2, 1)$, but that's a little *too* strong. Physical quantities are determined by the *square* of the wave function, so all we can say for sure is that $\psi(1, 2) = \pm\ \psi(2, 1)$: the wave function must either be *even* – symmetric – or *odd* – antisymmetric – under the interchange of two identical particles.† But which *is* it, even or odd? Nonrelativistic quantum mechanics offers no answer; there are simply two classes of particles – *bosons*, for which the wave function is even, and *fermions*, for which it is odd. It is an empirical fact that all particles of integer spin are bosons, whereas those of $\frac{1}{2}$-integer spin are fermions. One of the major achievements of quantum field theory was the rigorous *proof* of this connection between 'spin and statistics'.

$$\begin{aligned} \textbf{Boson}\ (\text{integer spin}) &\Rightarrow \text{symmetric wave function} : \psi(1,2) = \psi(2,1) \\ \textbf{Fermion}\ (\tfrac{1}{2}\text{-integer spin}) &\Rightarrow \text{antisymmetric wave function} : \psi(1,2) \\ &\qquad\qquad\qquad\quad = -\psi(2,1) \end{aligned}$$

Suppose we have two particles, one in state ψ_α and the other in state ψ_β. If the particles are *distinct* (one a muon and one an electron, say) then it makes sense to

* If the representations are labeled by spin, instead of dimensionality, Equation 5.55 reads $\frac{1}{2} \otimes \frac{1}{2} \otimes \frac{1}{2} = \frac{3}{2} \oplus \frac{1}{2} \oplus \frac{1}{2}$. Incidentally, it is *also* possible to construct a spin-$\frac{1}{2}$ combination that is *symmetric* in particles 1 and 2: $|\ \rangle = |\ \rangle_{13} + |\ \rangle_{23}$. Some authors prefer to use $|\ \rangle_{12}$ and $|\ \rangle$, instead of $|\ \rangle_{12}$ and $|\ \rangle_{23}$.

† From $|\psi(1, 2)|^2 = |\psi(2, 1)|^2$ it follows only that $\psi(1, 2) = e^{i\phi}\psi(2, 1)$. However, applying the interchange *twice* brings us back to where we started, so $e^{2i\phi} = 1$, and hence $e^{i\phi} = \pm 1$.

ask *which* is in state ψ_α and which is in state ψ_β. The wave function for the system is

$$\psi(1,2) = \psi_\alpha(1)\psi_\beta(2)$$

if particle 1 is in ψ_α and 2 is in ψ_β, or

$$\psi(1,2) = \psi_\beta(1)\psi_\alpha(2)$$

if it's the other way around. But if the two particles are *indistinguishable*, we cannot say which one is in which state. If the particles are identical bosons, the wave function is the symmetric combination

$$\psi(1,2) = (1/\sqrt{2})\left[\psi_\alpha(1)\psi_\beta(2) + \psi_\beta(1)\psi_\alpha(2)\right] \tag{5.56}$$

and if they are identical fermions, the wave function is the antisymmetric combination

$$\psi(1,2) = (1/\sqrt{2})\left[\psi_\alpha(1)\psi_\beta(2) - \psi_\beta(1)\psi_\alpha(2)\right] \tag{5.57}$$

In particular, if you try to put two fermions (electrons, say) into the *same* state ($\psi_\alpha = \psi_\beta$) you get *zero*; it can't be done. That's the original Pauli exclusion principle; but we see now that it is not an *ad hoc* assumption, but rather a consequence of a structural requirement on the wave functions of identical particles. Notice, by the way, that the Pauli principle does not apply to bosons; you can put as many pions into the same state as you like. Nor is there any symmetry requirement for *distinguishable* particles; that's why we didn't have to worry about it when we were constructing meson wave functions (since one constituent is a quark and the other an *antiquark*, they're *always* distinguishable). But in the case of the baryons we're putting three *quarks* together, and this time we must take the antisymmetrization requirement into account.

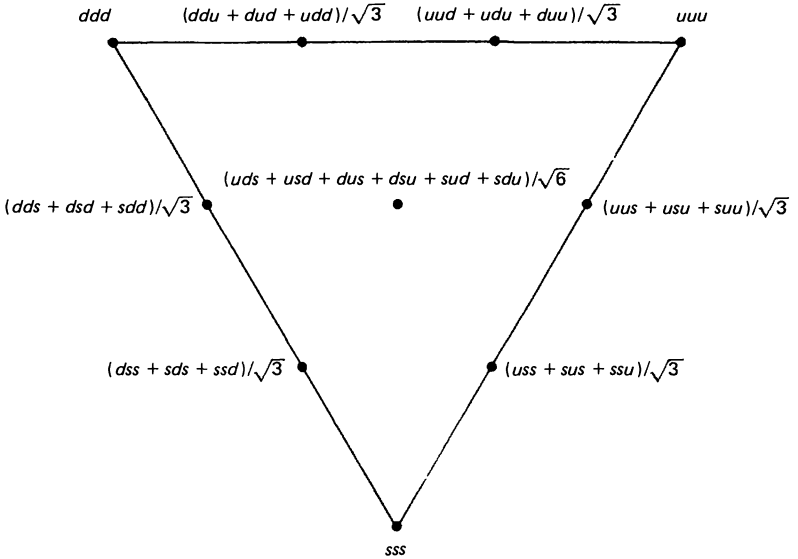
Now, the wave function of a baryon consists of several pieces; there is the *spatial* part, describing the locations of the three quarks; there is the *spin* part, representing their spins; there is a *flavor* component, indicating what combination of u, d, and s is involved; and there is a *color* term, specifying the colors of the quarks:

$$\psi = \psi(\text{space})\psi(\text{spin})\psi(\text{flavor})\psi(\text{color}) \tag{5.58}$$

It is the *whole works* that must be antisymmetric under the interchange of any two quarks.* We do not know the functional form of the spatial ground-state wave function, but it is surely symmetric; since $l = l' = 0$, there is no angular dependence at all. The spin state can either be completely symmetric ($j = \frac{3}{2}$) or of mixed symmetry ($j = \frac{1}{2}$). As for flavor, there are $3^3 = 27$ possibilities: *uuu, uud, udu, udd, ..., sss*, which we reshuffle into symmetric, antisymmetric, and mixed combinations; they form irreducible representations of $SU(3)$, just as the analogous

* Notice that a subtle extension of the notion of 'identical particle' has implicitly been made here, for we are treating all quarks, regardless of color or even flavor, as different *states* of a single particle [22].

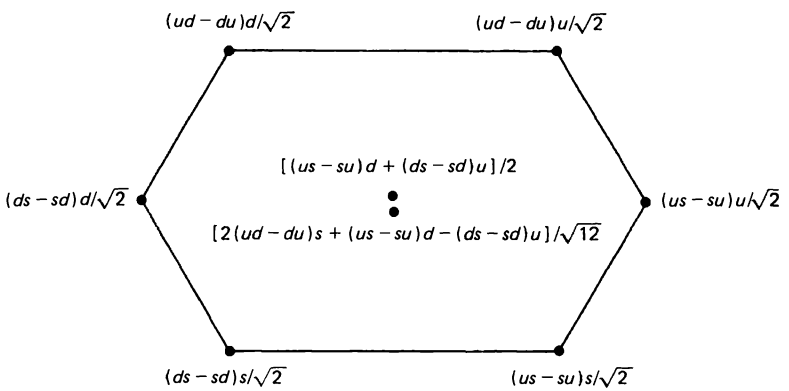
spin combinations form representations of $SU(2)$. These are conveniently displayed in Eightfold-Way patterns:



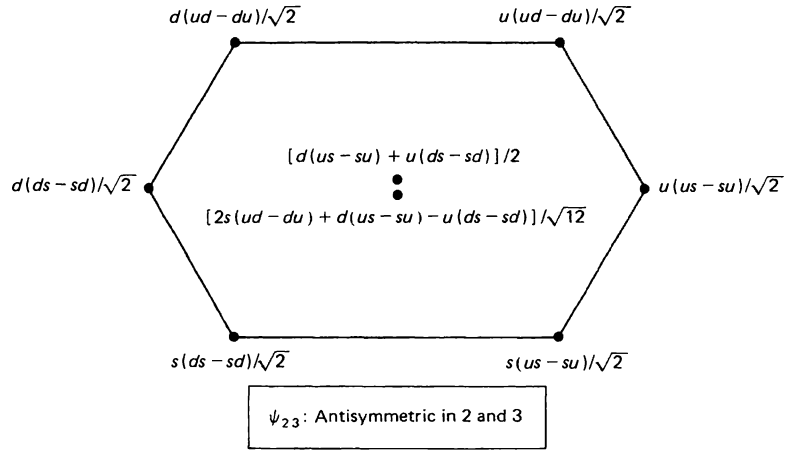
ψ_s: Completely symmetric states

$(uds - usd + dsu - dus + sud - sdu)/\sqrt{6}$

ψ_A: Completely antisymmetric state



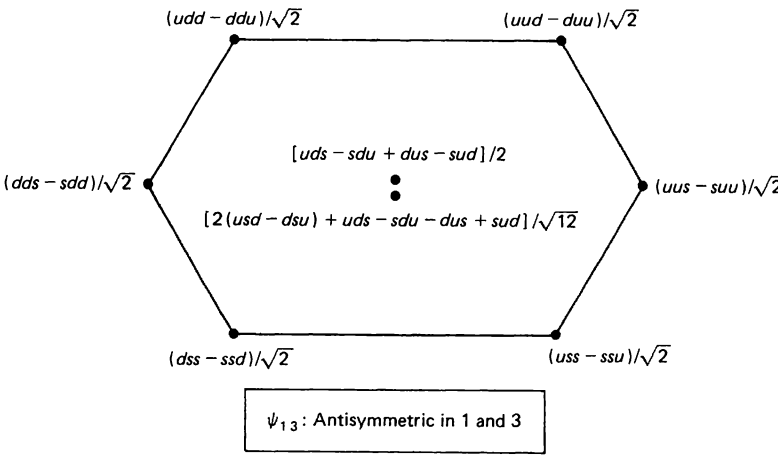
ψ_{12}: Antisymmetric in 1 and 2



Thus the combination of three light-quark flavors yields a decuplet, a singlet, and two octets;* in the language of group theory, the direct product of three fundamental representations of $SU(3)$ decomposes according to the rule

$$3 \otimes 3 \otimes 3 = 10 \oplus 8 \oplus 8 \oplus 1 \tag{5.59}$$

Incidentally, we can also construct an octet that is antisymmetric in 1 and 3, but this is not independent ($\psi_{13} = \psi_{12} + \psi_{23}$); we have already used up the 27 states available in making the four representations 10, 8, 8, and 1.



* As always in octet (and nonet) diagrams, I put the isotriplet ('Σ^0') above, and the isosinglet(s) ('Λ') beneath it, in the center.

Finally, there is the question of color. In Chapter 1, I stated a general rule that all naturally occurring particles are colorless; if a meson contains a red quark, it must also contain an *anti*red quark, and every baryon must harbor one quark of each color. Actually, this is a naive formulation of a deeper law:

Every naturally occurring particle is a color singlet.

The three *colors* generate a *color* $SU(3)$ symmetry, just as the three light-quark *flavors* generate *flavor* $SU(3)$. (The former is, however, an *exact* symmetry – quarks of different colors all weigh the same – whereas the latter is only *approximate*.) By putting together three colors, we obtain a color decuplet, two color octets, and a color singlet (simply make the flavor → color transcription, $u \to$ red, $d \to$ green, $s \to$ blue, in the diagrams above). But nature chooses the singlet, and so for baryons the color state is always

$$\psi(\text{color}) = (rgb - rbg + gbr - grb + brg - bgr)/\sqrt{6} \tag{5.60}$$

Because the color wave function is the same for *all* baryons, we generally do not bother to include it. However, it is absolutely crucial to remember that ψ(color) is *antisymmetric*, for this means the rest of the wave function must be *symmetric*. In particular, in the ground state, with ψ(space) symmetric, the product of ψ(spin) and ψ(flavor) has to be completely symmetric. Suppose we start with the symmetric spin configuration; this must go with the symmetric flavor state, and we obtain the spin-$\frac{3}{2}$ baryon decuplet:

$$\psi(\text{baryon decuplet}) = \psi_s(\text{spin})\psi_s(\text{flavor}) \tag{5.61}$$

Example 5.1 Write down the wave function for the Δ^+, in the spin state $m_j = -\frac{1}{2}$ (never mind the space and color parts).

Solution:

$$\begin{aligned}
|\Delta^+ : \tfrac{3}{2} - \tfrac{1}{2}\rangle &= \left\{(uud + udu + duu)/\sqrt{3}\right\}\left[(\downarrow\downarrow\uparrow + \downarrow\uparrow\downarrow + \uparrow\downarrow\downarrow)/\sqrt{3}\right] \\
&= \Big[u(\downarrow)u(\downarrow)d(\uparrow) + u(\downarrow)u(\uparrow)d(\downarrow) + u(\uparrow)u(\downarrow)d(\downarrow) \\
&\quad + u(\downarrow)d(\downarrow)u(\uparrow) + u(\downarrow)d(\uparrow)u(\downarrow) + u(\uparrow)d(\downarrow)u(\downarrow) \\
&\quad + d(\downarrow)u(\downarrow)u(\uparrow) + d(\downarrow)u(\uparrow)u(\downarrow) + d(\uparrow)u(\downarrow)u(\downarrow)\Big]/3
\end{aligned}$$

For instance, if you could pull such a particle apart, the probability is $\frac{1}{9}$ that the first quark would be a d with spin up, and $\frac{4}{9}$ that it would be a u with spin down. ▩

The baryon octet is a little trickier, for here we must put together states of mixed symmetry to make a completely symmetric combination. Notice first that

the product of two antisymmetric functions is itself *symmetric*. Thus $\psi_{12}(\text{spin}) \times \psi_{12}(\text{flavor})$ is symmetric in 1 and 2, for we pick up *two* minus signs when $1 \leftrightarrow 2$. Likewise, $\psi_{23}(\text{spin}) \times \psi_{23}(\text{flavor})$ is symmetric in 2 and 3, and $\psi_{13}(\text{spin}) \times \psi_{13}(\text{flavor})$ is symmetric in 1 and 3. If we now *add* these, the result will clearly be symmetric in all three (for the normalization factor, see Problem 5.16):

$$\psi(\text{baryon octet}) = (\sqrt{2}/3)[\psi_{12}(\text{spin})\psi_{12}(\text{flavor}) + \psi_{23}(\text{spin})\psi_{23}(\text{flavor}) + \psi_{13}(\text{spin})\psi_{13}(\text{flavor})] \tag{5.62}$$

Example 5.2 Write down the spin/flavor wave function for a proton with spin up.

Solution:

$$\begin{aligned}
|p : \tfrac{1}{2}\tfrac{1}{2}\rangle &= \left\{\tfrac{1}{2}(\uparrow\downarrow\uparrow - \downarrow\uparrow\uparrow)(udu - duu) + \tfrac{1}{2}(\uparrow\uparrow\downarrow - \uparrow\downarrow\uparrow)(uud - udu)\right. \\
&\quad \left. + \tfrac{1}{2}(\uparrow\uparrow\downarrow - \downarrow\uparrow\uparrow)(uud - duu)\right\}\frac{\sqrt{2}}{3} = \{uud(2\uparrow\uparrow\downarrow - \uparrow\downarrow\uparrow - \downarrow\uparrow\uparrow) \\
&\quad + udu(2\uparrow\downarrow\uparrow - \downarrow\uparrow\uparrow - \uparrow\uparrow\downarrow) + duu(2\downarrow\uparrow\uparrow - \uparrow\downarrow\uparrow - \uparrow\uparrow\downarrow)\}\frac{1}{3\sqrt{2}} \\
&= \frac{2}{3\sqrt{2}}(u(\uparrow)u(\uparrow)d(\downarrow)) - \frac{1}{3\sqrt{2}}(u(\uparrow)u(\downarrow)d(\uparrow)) \\
&\quad - \frac{1}{3\sqrt{2}}(u(\downarrow)u(\uparrow)d(\uparrow)) + \text{permutations.}
\end{aligned}$$

If nothing else, I hope you will have gathered from this exercise that the construction of baryon wave functions is a nontrivial business, in the quark model. Apart altogether from the spatial wave function, there are three spins to juggle, as well as three flavors and three colors, and it all has to be put together in a way that is consistent with the Pauli principle. Perhaps, also you will forgive me for deferring the explanation of how three quarks can generate the baryon octet (the decuplet, remember, we got by naive quark counting back in Chapter 1). The essential point is that the corners of the decuplet contain three identical quarks (*uuu*, *ddd*, and *sss*); they necessarily form a symmetric flavor state, and hence must go with the symmetric spin state ($j = \frac{3}{2}$). With *two* identical quarks (*uud*, say) there are three arrangements (*uud*, *udu*, *duu*); you can make a symmetric linear combination, which goes into the decuplet, and two of mixed symmetry, which belong to $SU(3)$ octets. Finally, with all three different, *uds*, there are six possibilities – the completely symmetric linear combination completes the decuplet, the completely antisymmetric combination makes an $SU(3)$ singlet, and the remaining four go into the two octets. Notice again the essential (if hidden) role of color in all this. Without it we would be looking for *anti*symmetric spin/flavor wave functions; spin $\frac{3}{2}$ (symmetric) would have to go with the flavor singlet (antisymmetric). It *is* possible to make a spin-$\frac{1}{2}$ octet without color (see Problem 5.18), but in place of the decuplet we would have just one spin-$\frac{3}{2}$ baryon. It was to avoid that disaster, without sacrificing the Pauli principle, that color was introduced in the first place.

5.6.2
Magnetic Moments

As an application of the baryon spin/flavor wave functions, we now calculate the magnetic dipole moments of the particles in the octet. In the absence of orbital motion, the net magnetic moment of a baryon is simply the vector sum of the moments of the three constituent quarks:

$$\boldsymbol{\mu} = \boldsymbol{\mu}_1 + \boldsymbol{\mu}_2 + \boldsymbol{\mu}_3 \tag{5.63}$$

It depends on the quark flavors (because the three flavors carry different magnetic moments) and on the spin configuration (because that determines the relative orientations of the three dipoles). Apart from minute radiative corrections, the magnetic dipole moment of a spin-$\frac{1}{2}$ point particle of charge q and mass m is (Equation 5.18):

$$\boldsymbol{\mu} = \frac{q}{mc}\mathbf{S} \tag{5.64}$$

Its magnitude is

$$\mu = \frac{q\hbar}{2mc} \tag{5.65}$$

More precisely, this is the value of μ_z in the spin-up state, for which $S_z = \hbar/2$. It is customary to refer to μ, rather than $\boldsymbol{\mu}$ itself, as 'the magnetic moment' of the particle. For the quarks,

$$\mu_u = \frac{2}{3}\frac{e\hbar}{2m_u c}, \quad \mu_d = -\frac{1}{3}\frac{e\hbar}{2m_d c}, \quad \mu_s = -\frac{1}{3}\frac{e\hbar}{2m_s c} \tag{5.66}$$

The magnetic moment of baryon B, then, is

$$\mu_B = \langle B\uparrow|(\mu_1 + \mu_2 + \mu_3)_z|B\uparrow\rangle = \frac{2}{\hbar}\sum_{i=1}^{3}\langle B\uparrow|(\mu_i S_{i_z})|B\uparrow\rangle \tag{5.67}$$

Example 5.3 Calculate the magnetic moment of the proton.

Solution: The wave function was found in Example 5.2. The first term is

$$\frac{2}{3\sqrt{2}}[u(\uparrow)u(\uparrow)d(\downarrow)]$$

Now $(\mu_1 S_{1_z} + \mu_2 S_{2_z} + \mu_3 S_{3_z})|u(\uparrow)u(\uparrow)d(\downarrow)\rangle$

$$= \left[\mu_u\frac{\hbar}{2} + \mu_u\frac{\hbar}{2} + \mu_d\left(-\frac{\hbar}{2}\right)\right]|u(\uparrow)u(\uparrow)d(\downarrow)\rangle$$

Table 5.5 Magnetic dipole moments of octet baryons

Baryon	Moment	Prediction	Experiment
p	$(\frac{4}{3})\mu_u - (\frac{1}{3})\mu_d$	2.79	2.793
n	$(\frac{4}{3})\mu_d - (\frac{1}{3})\mu_u$	−1.86	−1.913
Λ	μ_s	−0.58	−0.613
Σ^+	$(\frac{4}{3})\mu_u - (\frac{1}{3})\mu_s$	2.68	2.458
Σ^0	$(\frac{2}{3})(\mu_u + \mu_d) - (\frac{1}{3})\mu_s$	0.82	?
Σ^-	$(\frac{4}{3})\mu_d - (\frac{1}{3})\mu_s$	−1.05	−1.160
Ξ^0	$(\frac{4}{3})\mu_s - (\frac{1}{3})\mu_u$	−1.40	−1.250
Ξ^-	$(\frac{4}{3})\mu_s - (\frac{1}{3})\mu_d$	−0.47	−0.651

The numerical values are given as multiples of the nuclear magneton, $e\hbar/2m_pc$. *Source: Particle Physics Booklet* (2006).

so this term contributes an amount

$$\left(\frac{2}{3\sqrt{2}}\right)^2 \frac{\hbar}{2} \sum_{i=1}^{3} \langle u(\uparrow)u(\uparrow)d(\downarrow)|(\mu_i S_{i_z})|u(\uparrow)u(\uparrow)d(\downarrow)\rangle = \tfrac{2}{9}(2\mu_u - \mu_d)$$

Similarly, the second term $(u(\uparrow)u(\downarrow)d(\uparrow))$ gives $\frac{1}{18}\mu_d$, as does the third.* We could continue in this way to evaluate all nine terms, but the rest are simply permutations, in which d occupies position 2 or position 1. The result, then, is

$$\mu_p = 3[\tfrac{2}{9}(2\mu_u - \mu_d) + \tfrac{1}{18}\mu_d + \tfrac{1}{18}\mu_d] = \tfrac{1}{3}(4\mu_u - \mu_d)$$

In this way we can calculate all the octet magnetic moments in terms of μ_u, μ_d, and μ_s (Problem 5.19). The results are listed in the second column of Table 5.5. To get the actual *numbers*, we need to know the quark magnetic moments (Equation 5.66). Using the constituent quark masses $m_u = m_d = 336$ MeV/c^2, $m_s = 538$ MeV/c^2, we obtain the figures in the third column of Table 5.5. The comparison with experiment is reasonably good, considering the uncertainties in the quark masses. Somewhat better predictions are obtained if we take *ratios*. In particular, to the extent that $m_u = m_d$, we have

$$\frac{\mu_n}{\mu_p} = -\frac{2}{3} \tag{5.68}$$

which compares well with the experimental value, $-0.68497945 \pm 0.00000058$.

* Note that everything is normalized, so that for instance $\langle u(\uparrow)|u(\uparrow)\rangle = 1$, and the states are orthogonal $\langle u(\uparrow)|u(\downarrow)\rangle = 0$.

5.6.3 Masses

Finally, we turn to the problem of baryon masses. The situation is the same as for the mesons: if flavor $SU(3)$ were a perfect symmetry, all the octet baryons would weigh the same. But they don't. We attribute this in the first instance to the fact that the s quark is more massive than u and d. But that can't be the *whole* story, or the Λ would have the same mass as the Σ's, and the Δ's would match the proton. Evidently, there is a significant spin–spin ('hyperfine') contribution, which, as before, we take to be proportional to the dot product of the spins and inversely proportional to the product of the masses. The only difference is that this time there are *three* pairs of spins to contend with:

$$M(\text{baryon}) = m_1 + m_2 + m_3 + A'\left[\frac{\mathbf{S}_1 \cdot \mathbf{S}_2}{m_1 m_2} + \frac{\mathbf{S}_1 \cdot \mathbf{S}_3}{m_1 m_3} + \frac{\mathbf{S}_2 \cdot \mathbf{S}_3}{m_2 m_3}\right] \tag{5.69}$$

Here, A' (like A in Equation 5.46) is a constant, which we adjust to obtain the optimal fit to the data.

The spin products are easiest when the three quark masses are equal, for

$$J^2 = (\mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_3)^2 = S_1^2 + S_2^2 + S_3^2 + 2(\mathbf{S}_1 \cdot \mathbf{S}_2 + \mathbf{S}_1 \cdot \mathbf{S}_3 + \mathbf{S}_2 \cdot \mathbf{S}_3) \tag{5.70}$$

and hence

$$\begin{aligned} \mathbf{S}_1 \cdot \mathbf{S}_2 + \mathbf{S}_1 \cdot \mathbf{S}_3 + \mathbf{S}_2 \cdot \mathbf{S}_3 &= \frac{\hbar^2}{2}[j(j+1) - \tfrac{9}{4}] \\ &= \left\{ \begin{array}{ll} \frac{3}{4}\hbar^2, & \text{for } j = \frac{3}{2}\text{(decuplet)} \\ -\frac{3}{4}\hbar^2, & \text{for } j = \frac{1}{2}\text{(octet)} \end{array} \right\} \end{aligned} \tag{5.71}$$

Thus the nucleon (neutron or proton) mass is

$$M_N = 3m_u - \frac{3}{4}\frac{\hbar^2}{m_u^2}A' \tag{5.72}$$

the Δ is

$$M_\Delta = 3m_u + \frac{3}{4}\frac{\hbar^2}{m_u^2}A' \tag{5.73}$$

and the Ω^- is

$$M_\Omega = 3m_s + \frac{3}{4}\frac{\hbar^2}{m_s^2}A' \tag{5.74}$$

In the case of the decuplet the spins are all 'parallel' (every pair combines to make spin 1) so

$$(\mathbf{S}_1 + \mathbf{S}_2)^2 = S_1^2 + S_2^2 + 2\mathbf{S}_1 \cdot \mathbf{S}_2 = 2\hbar^2 \tag{5.75}$$

(and the same for 1 and 3, or 2 and 3). Hence for the decuplet

$$\mathbf{S}_1 \cdot \mathbf{S}_2 = \mathbf{S}_1 \cdot \mathbf{S}_3 = \mathbf{S}_2 \cdot \mathbf{S}_3 = \frac{\hbar^2}{4} \tag{5.76}$$

(which is consistent, notice, with Equation 5.71), and therefore

$$M_{\Sigma^*} = 2m_u + m_s + \frac{\hbar^2}{4} A' \left(\frac{1}{m_u^2} + \frac{2}{m_u m_s} \right) \tag{5.77}$$

while

$$M_{\Xi^*} = m_u + 2m_s + \frac{\hbar^2}{4} A' \left(\frac{2}{m_u m_s} + \frac{1}{m_s^2} \right) \tag{5.78}$$

The Σ and Λ can be done by noting that the up and down quarks combine to make isospin 1 and 0, respectively, and in order for the spin/flavor wave function to be symmetric, under the interchange of u and d, the *spins* must therefore combine to a total of 1 and 0, respectively. For the Σ's, then

$$(\mathbf{S}_u + \mathbf{S}_d)^2 = S_u^2 + S_d^2 + 2\mathbf{S}_u \cdot \mathbf{S}_d = 2\hbar^2, \quad \text{so} \quad \mathbf{S}_u \cdot \mathbf{S}_d = \frac{\hbar^2}{4} \tag{5.79}$$

whereas for the Λ

$$(\mathbf{S}_u + \mathbf{S}_d)^2 = 0, \quad \text{so} \quad \mathbf{S}_u \cdot \mathbf{S}_d = -\tfrac{3}{4}\hbar^2 \tag{5.80}$$

Using these results together with Equation 5.71, we find

$$\begin{aligned} M_\Sigma &= 2m_u + m_s + A' \left[\frac{\mathbf{S}_u \cdot \mathbf{S}_d}{m_u m_d} + \frac{(\mathbf{S}_1 \cdot \mathbf{S}_2 + \mathbf{S}_1 \cdot \mathbf{S}_3 + \mathbf{S}_2 \cdot \mathbf{S}_3 - \mathbf{S}_u \cdot \mathbf{S}_d)}{m_u m_s} \right] \\ &= 2m_u + m_s + \frac{\hbar^2}{4} A' \left(\frac{1}{m_u^2} - \frac{4}{m_u m_s} \right) \end{aligned} \tag{5.81}$$

Table 5.6 Baryon octet and decuplet masses. (MeV/c^2)

Baryon	Calculated	Observed
N	939	939
Λ	1114	1116
Σ	1179	1193
Ξ	1327	1318
Δ	1239	1232
Σ^*	1381	1385
Ξ^*	1529	1533
Ω	1682	1672

and

$$M_\Lambda = 2m_u + m_s - \frac{3}{4}\frac{\hbar^2}{m_u^2}A' \tag{5.82}$$

I'll let you figure out the mass of the Ξ's (Problem 5.22):

$$M_\Xi = 2m_s + m_u + \frac{\hbar^2}{4}A'\left(\frac{1}{m_s^2} - \frac{4}{m_u m_s}\right) \tag{5.83}$$

Using the constituent quark masses $m_u = m_d = 363\ \mathrm{MeV}/c^2$, $m_s = 538\ \mathrm{MeV}/c^2$, and picking $A' = (2m_u/\hbar)^2\ 50\ \mathrm{MeV}/c^2$, we obtain an excellent fit to the experimental data (Table 5.6).*

References

1 Readers unfamiliar with quantum mechanics will need to consult an appropriate textbook for background, and for details of the calculations presented here. See, for instance, Park, D. (**1992**) *Introduction to the Quantum Theory*, 3rd edn, McGraw-Hill, New York; (a) Townsend, J. S. (**2000**) *A Modern Approach to Quantum Mechanics*, University Science Books, Sausalito, CA; (b) Griffiths, D. J. (**2005**) *Introduction to Quantum Mechanics*, 2nd edn, Prentice Hall, Upper Saddle River, NJ.

2 For a fascinating account of the experimental study of the hydrogen spectrum, from its beginnings in the mid-nineteenth century up to the present day, see the article by Hänsch, T. W., Schawlow, A. L. and Series, G. W. (March **1979**) *Scientific American*, 94.

3 See, for example, Griffiths, D. J. (**1999**) *Introduction to Electrodynamics*, 3rd edn, Prentice Hall, Upper Saddle River, NJ; Problem 5.56.

4 For the calculation, see Kinoshita, T. and Nio, M. (**2006**) *Physical Review D*, **73**, 013003; for the measurement, see (a) Odom, B. *et al.* (**2006**) *Physical Review Letters*, **97**, 030801.

5 Lamb, W. E. Jr. and Retherford, R. C. (**1947**) *Physical Review*, **72**, 241.

6 Each normal mode of the electromagnetic field functions as an oscillator; in quantum mechanics the ground-state energy of a harmonic oscillator is not zero, but rather, $\frac{1}{2}\hbar\omega$. See Bjorken, J. D. and Drell, S. D. (**1964**) *Relativistic Quantum Mechanics*, McGraw-Hill, New York, p. 58.

7 Bethe, H. A. and Salpeter, E. E. (**1977**) *Quantum Mechanics of One- and Two-Electron Atoms*, Plenum, New York. Sect. 21, for details.

8 Bethe, H. A. and Salpeter, E. E. (**1977**) *Quantum Mechanics of One- and Two-Electron Atoms*, Plenum, New York, p. 110.

9 For an accessible account of hyperfine splitting in the ground state of hydrogen, see Griffiths, D. J. (**1982**) *American Journal of Physics*, **50**, 698.

10 Ewen, H. I. and Purcell, E. M. (**1951**) *Nature*, **168**, 356.

11 The spectrum of positronium is discussed in Bethe, H. A. and Salpeter, E. E. (**1977**) *Quantum Mechanics of One- and Two-Electron Atoms*, Plenum, New York; Sect. 21, for details. Sect. 23. See also (a) Berko, S. and Pendleton, H. N. (**1980**) *Annual*

* Notice, however, that we are obliged to use somewhat *different* constituent quark masses in Tables 5.3, 5.5, and 5.6, as I warned you in the footnote to Table 4.4.

Review of Nuclear and Particle Science, **30**, 543; and (b) Rich, A. (**1981**) *Reviews of Modern Physics*, **53**, 127.

12 The details of the fine structure are not quite the same in quarkonium as in positronium; see Eichten, E. and Feinberg, F. (**1979**) *Physical Review Letters*, **43**, 1205; and *Physical Review*, (**1981**) **D23**, 2724.

13 There are other ways to estimate F_0, which give roughly the same answer. See Perkins, D. H. (**2000**) *Introduction to High Energy Physics*, 4th edn, Cambridge University Press, Cambridge, UK; Section 6.3.

14 Appelquist, T. and Politzer, H. D. (**1975**) *Physical Review Letters*, **34**, 43; (**1975**) *Physical Review*, **D12**, 1404.

15 For an interesting account of these discoveries, see the article 'Quarkonium' by Bloom, E. D. and Feldman, G. J. (May **1982**) *Scientific American*, 66.

16 Eichten, E. and Gottfried, K. (**1977**) *Physics Letters B*, **66**, 286.

17 The effect of vacuum polarization on the spectrum of bottomonium is explored in an accessible article by Conway, J. *et al.* (**2001**) *European Journal of Physiology*, **22**, 533.

18 The 'MIT Bag Model' offers a possible approach to relativistic light-quark systems, but at the cost of vastly oversimplified dynamics. The quarks are treated as free particles confined within a spherical 'bag', which is stabilized by an *ad hoc* external pressure. Many interesting calculations have been carried out using the bag model, but no one would pretend that it is a realistic picture of hadron structure. See Close, F. E. (**1979**) *An Introduction to Quarks and Partons*, Academic, London; Chapter 18.

19 See, for example, Halzen, F. and Martin, A. D. (**1984**) *Quarks and Leptons*, John Wiley & Sons, New York, p. 42.

20 (**2006**) *Review of Particle Physics*; Section 14.2.

21 Gasiorowicz, S. and Rosner, J. L. (**1981**) *American Journal of Physics*, **49**, 954; an article that contains a wealth of useful and accessible information about the quark model.

22 Pais, A. (**1986**) *Inward Bound*, Oxford University Press, Oxford, UK, p. 425; attributes the extension to (a) Cassen, B. and Condon, E. U. (**1936**) *Physical Review*, **50**, 846.

Problems

5.1 (a) The deuteron's mass is 1875.6 MeV/c^2. What is its binding energy? Is this a relativistic system?

(b) If you take the up- and down-quark masses to be those given in Table 4.4, what is the binding energy of a pion? Is this a relativistic system?

5.2 Use Equation 5.12 to obtain the ground-state wave function ψ_{100}. Show that it satisfies the Schrödinger equation (Equation 5.1), with the appropriate energy, and check that it is properly normalized. [*Answer*: $\Psi_{100} = (1/\sqrt{\pi a^3})e^{-r/a}e^{-iE_1 t/\hbar}$]

5.3 Work out all of the hydrogen wave functions for $n = 2$, using Equation 5.12. (How many are there?)

5.4 Using Equation 3.43 to express the kinetic energy ($T = E - mc^2$) in terms of p (and m), show that the lowest-order relativistic correction to $T = p^2/2m$ is $-p^4/8m^3c^2$.

5.5 Find the energy splitting between the $j = \frac{3}{2}$ and $j = \frac{1}{2}$ levels for $n = 2$ (Figure 5.2), in electron volts. How does this compare with the spacing between the $n = 2$ and $n = 1$ Bohr energies?

5.6 Estimate the Lamb shift energy gap between the $2S_{1/2}$ and $2P_{1/2}$ levels in hydrogen, using Equations 5.20 and 5.21. What is the frequency of the photon emitted in such a transition? (The experimental value is 1057 MHz.)

5.7 If you include the fine structure, Lamb shift, and hyperfine splitting, how many different $n = 2$ energy levels are there altogether in hydrogen? Find the hyperfine splitting between the $2S_{1/2}$ and $2P_{1/2}$ levels, and compare the Lamb shift (Problem 5.6).

5.8 Analyze the splitting of the $n = 3$ Bohr level in positronium. How many different levels are there, and what are their relative energies? Construct the level diagram, analogous to Figure 5.6

5.9 Would you consider the $\phi(s\bar{s})$ meson *bound* or *quasi*-bound?

5.10 On dimensional grounds, show that the energy levels of a purely linear potential, $V(r) = F_0 r$, must be of the form

$$E_n = \left(\frac{(F_0\hbar)^2}{m}\right)^{1/3} a_n \tag{5.84}$$

where a_n is a dimensionless numerical factor.

5.11 Use the numerical results in Table 5.2 to 'predict' the masses of the four lightest ψ's and Υ's; compare the experimental results (Figure 5.9). What value of F_0 gives the best fit to the level spacings? Why aren't the calculated masses in better agreement with the experiments?

5.12 Using Equation 5.46, with the values of m_u, m_d, m_s, and A given in the text, calculate the meson masses in Table 5.3. [*Hint*: For the η, first find the mass for pure $u\bar{u}$, pure $d\bar{d}$, and pure $s\bar{s}$, and think of the η as being $\frac{1}{6}u\bar{u}$, $\frac{1}{6}d\bar{d}$, and $\frac{2}{3}s\bar{s}$.] Also apply the formula to the η', and note the disastrous result. [For commentary on the η' mass problem, see Quigg, C. (1983) *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions*, Benjamin, New York, p. 252.]

5.13 In the text, we used Equation 5.46 to calculate the masses of light-quark pseudoscalar and vector mesons. But the same formula can be applied to heavy-quark systems involving charm and beauty quarks.

(a) Calculate the masses of the pseudoscalar mesons $\eta_c(c\bar{c})$, $D^0(c\bar{u})$, $D_s^+(c\bar{s})$, and the corresponding vector mesons $\psi\ (c\bar{c})$, $D^{*0}\ (c\bar{u})$, and $D^{*+}\ (c\bar{s})$. Compare the experimental values, from the Particle Data Booklet.

(b) Do the same for the 'bottom' mesons $u\bar{b}$, $s\bar{b}$, $c\bar{b}$, and $b\bar{b}$. At present only the pseudoscalars $B^+(u\bar{b})$, $B_s^0(s\bar{b})$, $B_c^+(c\bar{b})$ and the vector $\Upsilon\ (b\bar{b})$ have been detected experimentally.

5.14 Construct the eight states ψ_{12} in Section 5.6.1. [*Hint*: The six outer ones are easy – the quark content is determined by Q and S, and all you have to do is antisymmetrize in 1 and 2. To get the two states in the center, remember that the one in the 'Σ^0' position forms an isotriplet with the 'Σ^+' and 'Σ^-'; the 'Λ' can then be constructed by orthogonalizing with respect to 'Σ^0' and ψ_Λ.]

5.15 Construct the (singlet) color wave function for mesons, analogous to Equation 5.60.

5.16 Check that the baryon octet spin/flavor wave function (Equation 5.60) is correctly normalized. Remember that ψ_{13} is not independent of ψ_{12} and ψ_{23}.

5.17 Construct the spin–flavor wave functions, as in Example 5.2, for Σ^+ with spin up and Λ with spin down.

5.18 Construct a totally *anti*symmetric spin/flavor baryon octet. (In this configuration we do not need color to antisymmetrize the wave function. However, an antisymmetric decuplet *cannot* be constructed. See Halzen and Martin, Reference [19], Exercise 2.18.)

5.19 (a) Derive the expressions in the second column of Table 5.5.

(b) From these results, calculate the numbers in the third column of Table 5.5, using the quark masses given in the text.

5.20 Calculate the ratio μ_n/μ_p in the configuration you found for Problem 5.18 Notice that μ_p is *negative* in this case! Is your result consistent with experiment? (Here, then, is a second strike against the quark model without color, the first strike being its failure to account for the decuplet.)

5.21 Show that $\mu_{\rho+} = -\mu_{\rho-} = \mu_\rho$. (See Halzen and Martin, Reference [19], Exercise 2.19). As far as I know, the magnetic dipole moments of vector mesons have not been measured.

5.22 Use Equation 5.69 to determine the mass of the Ξ.

5.23 Using Equations 5.12, 5.13, and 5.28, calculate the electron density at the location of the positron, in the ground state of positronium, $|\psi_{100}(0)|^2$.

6

The Feynman Calculus

In this chapter, we begin the quantitative formulation of elementary particle dynamics, which amounts, in practice, to the calculation of decay rates (Γ) *and scattering cross sections* (σ). *The procedure involves two distinct parts: (1) evaluation of the relevant Feynman diagrams to determine the 'amplitude'* ($\mathscr{M}$) *for the process in question and (2) insertion of* $\mathscr{M}$ *into Fermi's 'Golden Rule' to compute* Γ *or* σ, *as the case may be. To avoid distracting algebraic complications, I introduce here a simplified model. Realistic theories – QED, QCD, and GWS – are developed in succeeding chapters. If you like, Chapter 6 can be read immediately after Chapter 3. Study it with scrupulous care, or what follows will be unintelligible.*

6.1
Decays and Scattering

As I mentioned in the Introduction, we have three experimental probes of elementary particle interactions: bound states, decays, and scattering. Nonrelativistic quantum mechanics (in Schrödinger's formulation) is particularly well adapted to handle bound states, which is why we used it, as far as possible, in Chapter 5. By contrast, the relativistic theory (in Feynman's formulation) is especially well suited to describe decays and scattering. In this chapter I'll introduce the basic ideas and strategies of the Feynman 'calculus'; in subsequent chapters we will use it to develop the theories of strong, electromagnetic, and weak interactions.

6.1.1
Decay Rates

To begin with, we must decide what physical quantities we would like to calculate. In the case of decays, the item of greatest interest is the *lifetime* of the particle in question. What precisely do we mean by the lifetime of, say, the muon? We have in mind, of course, a muon at *rest*; a moving muon lasts longer (from our perspective) because of time dilation. But even stationary muons don't all last the same amount of time, for there is an intrinsically random element in the decay

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

process. We cannot hope to calculate the lifetime of any *particular* muon; rather, what we are after is the *average* (or 'mean') lifetime, τ, of the muons in any large sample.

Now, elementary particles have no memories, so the probability of a given muon decaying in the next microsecond is independent of how long ago that muon was created. (It's quite different in biological systems: an 80-year-old man is much more likely to die in the next year than is a 20-year-old, and his body shows the signs of eight decades of wear and tear. But all muons are identical, regardless of when they were produced; from an actuarial point of view they're all on an equal footing.) The critical parameter, then, is the *decay rate*, Γ, the *probability per unit time* that any given muon will disintegrate. If we had a large collection of muons, say, $N(t)$, at time t, then $N\Gamma \mathrm{d}t$ of them would decay in the next instant $\mathrm{d}t$. This would, of course, *decrease* the number remaining:

$$\mathrm{d}N = -\Gamma N \, \mathrm{d}t \tag{6.1}$$

It follows that

$$N(t) = N(0)e^{-\Gamma t} \tag{6.2}$$

Evidently, the number of particles left decreases exponentially with time. As you can check for yourself (Problem 6.1), the mean lifetime is simply the reciprocal of the decay rate:

$$\tau = \frac{1}{\Gamma} \tag{6.3}$$

Actually, most particles can decay by several different routes. The π^+, for instance, usually decays to $\mu^+ + \nu_\mu$, but sometimes one goes to $e^+ + \nu_e$; occasionally, a π^+ decays to $\mu^+ + \nu_\mu + \gamma$, and they have even been known to go to $e^+ + \nu_e + \pi^0$. In such circumstances, the *total* decay rate is the sum of the individual decay rates:

$$\Gamma_{\text{tot}} = \sum_{i=1}^{n} \Gamma_i \tag{6.4}$$

and the lifetime of the particle is the reciprocal of Γ_{tot}:

$$\tau = \frac{1}{\Gamma_{\text{tot}}} \tag{6.5}$$

In addition to τ, we want to calculate the various *branching ratios*, that is, the fraction of all particles of the given type that decay by each mode. Branching ratios are determined by the decay rates:

$$\text{Branching ratio for ith decay mode} = \Gamma_i / \Gamma_{\text{tot}} \tag{6.6}$$

For decays, then, the essential problem is to calculate the decay rate Γ_i for each mode; from there it is an easy matter to obtain the lifetime and branching ratios.

6.1.2
Cross Sections

How about scattering? What quantity should the experimentalist measure and the theorist calculate? If we were talking about an *archer* aiming at a 'bull's-eye', the parameter of interest would be the *size of the target* or, more precisely, the cross-sectional area it presents to a stream of incoming arrows. In a crude sense, the same goes for elementary particle scattering: if you fire a stream of electrons into a tank of hydrogen (which is essentially a collection of protons), the parameter of interest is the size of the proton – the cross-sectional area σ it presents to the incident beam. The situation is more complicated than in archery, however, for several reasons. First of all the target is 'soft'; it's not a simple case of 'hit-or-miss', but rather 'the closer you come the greater the deflection'. Nevertheless, it is still possible to define an 'effective' cross section; I'll show you how in a moment. Secondly, the cross section depends on the nature of the 'arrow' as well as the structure of the 'target'. Electrons scatter off hydrogen more sharply than neutrinos and less so than pions, because different interactions are involved. It depends, too, on the *outgoing* particles; if the energy is high enough we can have not only *elastic* scattering ($e + p \to e + p$), but also a variety of *inelastic* processes, such as $e + p \to e + p + \gamma$, or $e + p + \pi^0$, or even, in principle, $\nu_e + \Lambda$. Each one of these has its own ('exclusive') scattering cross section, σ_i (for process i). In some experiments, however, the final products are not examined, and we are interested only in the *total* ('inclusive') cross section:

$$\sigma_{\text{tot}} = \sum_{i=1}^{n} \sigma_i \tag{6.7}$$

Finally, each cross section typically depends on the *velocity* of the incident particle. At the most naive level we might expect the cross section to be proportional to the amount of time the incident particle spends in the vicinity of the target, which is to say that σ should be inversely proportional to v. But this behavior is dramatically altered in the neighborhood of a 'resonance' – a special energy at which the particles involved 'like' to interact, forming a short-lived semibound state before breaking apart. Such 'bumps' in the graph of σ versus v (or, as it is more commonly plotted, σ versus E) are in fact the principal means by which short-lived particles are discovered (see Figure 4.6). So, unlike the archer's target, there's a lot of physics in an elementary particle cross section.

Let's go back, now, to the question of what we mean by a 'cross section' when the target is soft. Suppose a particle (maybe an electron) comes along, encounters some kind of potential (perhaps the Coulomb potential of a stationary proton), and scatters off at an angle θ. This *scattering angle* is a function of the *impact parameter* b, the distance by which the incident particle would have missed the scattering center, had it continued on its original trajectory (Figure 6.1). Ordinarily, the smaller the impact parameter, the larger the deflection, but the actual functional form of $\theta(b)$ depends on the particular potential involved.

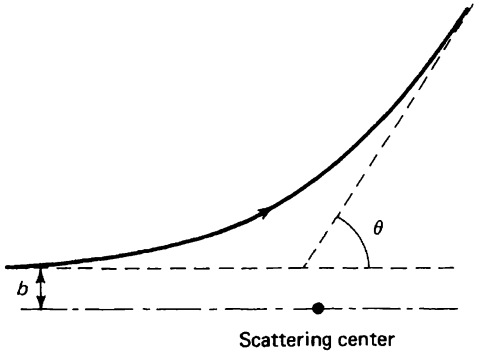


Fig. 6.1 Scattering from a fixed potential: θ is the scattering angle and b is the impact parameter.

Example 6.1 Hard-sphere Scattering Suppose the particle bounces elastically off a sphere of radius R. From Figure 6.2, we have

$$b = R\sin\alpha, \qquad 2\alpha + \theta = \pi$$

Thus,

$$\sin\alpha = \sin(\pi/2 - \theta/2) = \cos(\theta/2)$$

and hence

$$b = R\cos(\theta/2) \quad \text{or} \quad \theta = 2\cos^{-1}(b/R)$$

This is the relation between θ and b for classical hard-sphere scattering.

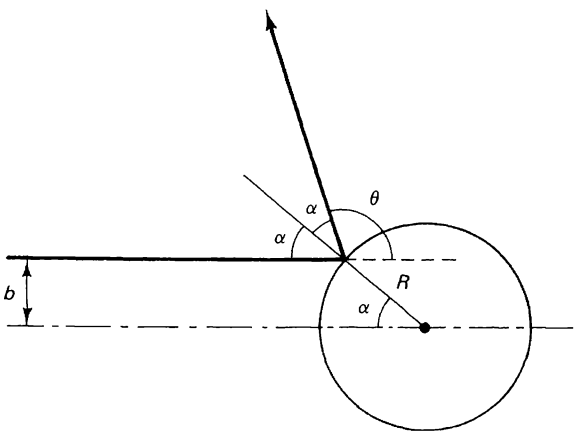


Fig. 6.2 Hard-sphere scattering.

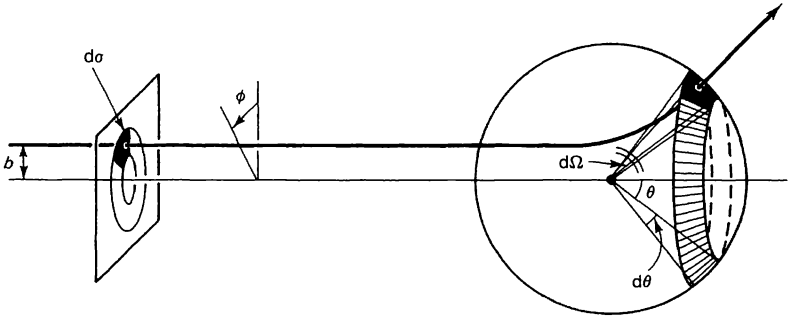


Fig. 6.3 Particle incident in area $d\sigma$ scatters into solid angle $d\Omega$.

If the particle comes in with an impact parameter between b and $b + db$, it will emerge with a scattering angle between θ and $\theta + d\theta$. More generally, if it passes through an infinitesimal *area* $d\sigma$, it will scatter into a corresponding *solid angle* $d\Omega$ (Figure 6.3). Naturally, the larger we make $d\sigma$, the larger $d\Omega$ will be. The proportionality factor is called the *differential (scattering) cross section, D*:*

$$d\sigma = D(\theta)\, d\Omega \tag{6.8}$$

The name is poorly chosen; it's not a differential, or even a derivative, in the mathematical sense. The words would apply more naturally to $d\sigma$ than to $d\sigma/d\Omega$ . . . but I'm afraid we're stuck with it.

Now, from Figure 6.3 we see that

$$d\sigma = |b\, db\, d\phi|, \quad d\Omega = |\sin\theta\, d\theta\, d\phi| \tag{6.9}$$

(Areas and solid angles are intrinsically positive, hence the absolute value signs.) Accordingly,

$$D(\theta) = \frac{d\sigma}{d\Omega} = \left|\frac{b}{\sin\theta}\left(\frac{db}{d\theta}\right)\right| \tag{6.10}$$

Example 6.2 In the case of hard-sphere scattering, Example 6.1, we find

$$\frac{db}{d\theta} = -\frac{R}{2}\sin\left(\frac{\theta}{2}\right)$$

* In principle D can depend on the azimuthal angle ϕ; however, most potentials of interest are spherically symmetrical, in which case the differential cross section depends only on θ (or, if you prefer, on b). By the way, the notation (D) is my own; most people just write $d\sigma/d\Omega$, and in the rest of the book I'll do the same.

and hence

$$D(\theta) = \frac{Rb\sin(\theta/2)}{2\sin\theta} = \frac{R^2}{2}\frac{\cos(\theta/2)\sin(\theta/2)}{\sin\theta} = \frac{R^2}{4}$$

Finally, the *total* cross section is the integral of $d\sigma$ over all solid angles:

$$\sigma = \int d\sigma = \int D(\theta)\, d\Omega \tag{6.11}$$

Example 6.3 For hard-sphere scattering,

$$\sigma = \int \frac{R^2}{4}\, d\Omega = \pi R^2$$

which is, of course, the total cross section the sphere presents to an incoming beam: any particles *within* this area will scatter, and any *outside* will pass by unaffected.

As Example 6.3 indicates, the formalism developed here is consistent with our naive sense of the term 'cross section', in the case of a 'hard' target; its virtue is that it applies as well to 'soft' targets, which do not have sharp edges.

Example 6.4 Rutherford Scattering A particle of charge q_1 scatters off a stationary particle of charge q_2. In classical mechanics, the formula relating the impact parameter to the scattering angle is [1]

$$b = \frac{q_1 q_2}{2E}\cot(\theta/2)$$

where E is the initial kinetic energy of the incident charge. The differential cross section is therefore

$$D(\theta) = \left(\frac{q_1 q_2}{4E\sin^2(\theta/2)}\right)^2$$

In this case, the total cross section is actually *infinite*:*

$$\sigma = 2\pi\left(\frac{q_1 q_2}{4E}\right)^2 \int_0^\pi \frac{1}{\sin^4(\theta/2)}\sin\theta\, d\theta = \infty$$

Suppose we have a *beam* of incoming particles, with uniform *luminosity* $\mathcal{L}$ ($\mathcal{L}$ is the number of particles passing down the line per unit time, per unit area). Then

* This is related to the fact that the Coulomb potential has infinite range (see footnote in Section 1.3).

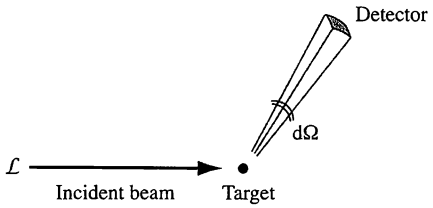


Fig. 6.4 Scattering of a beam with luminosity $\mathcal{L}$.

$dN = \mathcal{L}\, d\sigma$ is the number of particles per unit time passing through area $d\sigma$, and hence also the number per unit time scattered into solid angle $d\Omega$:

$$dN = \mathcal{L}\, d\sigma = \mathcal{L} D(\theta)\, d\Omega \tag{6.12}$$

Suppose I set up a detector that subtends a solid angle $d\Omega$ with respect to the collision point (Figure 6.4). I count the number of particles per unit time (dN) reaching my detector – what an experimentalist would call the *event rate*. Equation 6.12 says that the event rate is equal to the luminosity times the differential cross section times the solid angle. Whoever is operating the accelerator controls the luminosity; whoever set up the detector determined the solid angle. With these parameters established, the differential cross section can be measured by simply counting the number of particles entering the detector:

$$\frac{d\sigma}{d\Omega} = \frac{dN}{\mathcal{L}\, d\Omega} \tag{6.13}$$

If the detector completely surrounds the target, then $N = \sigma\mathcal{L}$; as accelerator physicists like to say, 'the event rate is the cross section times the luminosity'.*

6.2 The Golden Rule

In Section 6.1 I introduced the physical quantities we need to calculate: decay rates and cross sections. In both cases there are two ingredients in the recipe: (i) the *amplitude* ($\mathcal{M}$) for the process and (ii) the *phase space* available.† The amplitude contains all the *dynamical* information; we calculate it by evaluating the relevant Feynman diagrams, using the *Feynman rules* appropriate to the interaction in

* In this discussion, I have assumed that the target itself is *stationary* and that the incident particle is simply *deflected* as it passes through the scattering potential. My purpose was to introduce the essential ideas in the simplest possible context. But in Section 6.2 the formalism is completely general; it includes the recoil of the target, and allows for a change in the identity of the participants during the scattering process (in the reaction $\pi^- + p^+ \to K^+ + \Sigma^-$, for example, $d\Omega$ might represent the solid angle into which the K^+ scatters).

† The *amplitude* is also called the *matrix element*; the *phase space* is sometimes called the *density of final states*.

question. The phase space factor is purely *kinematic*; it depends on the masses, energies, and momenta of the participants, and reflects the fact that a given process is more likely to occur the more 'room to maneuver' there is in the final state. For example, the decay of a heavy particle into light secondaries involves a large phase space factor, for there are many different ways to apportion the available energy. By contrast, the decay of the neutron ($n \to p + e + \overline{\nu}_e$), in which there is almost no extra mass to spare, is tightly constrained and the phase space factor is very small.*

The ritual for calculating reaction rates was dubbed the *Golden Rule* by Enrico Fermi. In essence, Fermi's Golden Rule says that a transition rate is given by the *product* of the phase space and the (absolute) square of the amplitude. You may have encountered the nonrelativistic version, in the context of time-dependent perturbation theory [2]. We need the relativistic version, which comes from quantum field theory [3]. I can't *derive* it here; what I will do is *state* the Golden Rule and try to make it plausible. Actually, I'll do it twice: once in a form appropriate to decays and again in a form suitable for scattering.

6.2.1
Golden Rule for Decays

Suppose particle 1 (at rest)† decays into several other particles 2, 3, 4, . . . , n:

$$1 \to 2 + 3 + 4 + \cdots + n \tag{6.14}$$

The decay rate is given by the formula

$$\Gamma = \frac{S}{2\hbar m_1} \int |\mathscr{M}|^2 (2\pi)^4 \delta^4 \left(p_1 - p_2 - p_3 \cdots - p_n\right) \times \prod_{j=2}^{n} 2\pi\delta\left(p_j^2 - m_j^2 c^2\right) \theta\left(p_j^0\right) \frac{\mathrm{d}^4 p_j}{(2\pi)^4} \tag{6.15}$$

where m_i is the mass of the ith particle and p_i is its four-momentum. S is a statistical factor that corrects for double-counting when there are identical particles in the final state: for each such group of s particles, S gets a factor of $(1/s!)$. For instance, if $a \to b + b + c + c + c$, then $S = (1/2!)(1/3!) = 1/12$. If there are *no* identical particles in the final state (the most common circumstance), then $S = 1$.

Remember: The *dynamics* of the process is contained in the *amplitude*, $\mathscr{M}(p_1, p_2, \ldots, p_n)$, which is a function of the various momenta; we'll calculate it (later)

* For a more extreme case, consider the (kinematically forbidden) decay $\Omega^- \to \Xi^- + \overline{K}^0$. Since the final products weigh more than the Ω, there is no phase space available at all and the decay rate is zero.

† There is no loss of generality in assuming particle 1 is at rest; this is simply an astute choice of reference frame.

by evaluating the appropriate Feynman diagrams. The rest is *phase space*; it tells us to *integrate over all outgoing four-momenta*, subject to three kinematical constraints:

1. Each outgoing particle lies on its mass shell: $p_j^2 = m_j^2 c^2$ (which is to say, $E_j^2 - \mathbf{p}_j^2 c^2 = m_j^2 c^4$). This is enforced by the delta function $\delta\left(p_j^2 - m_j^2 c^2\right)$, which is zero unless its argument vanishes.*
2. Each outgoing energy is positive: $p_j^0 = E_j/c > 0$. Hence the θ function.†
3. Energy and momentum must be conserved: $p_1 = p_2 + p_3 \cdots + p_n$. This is ensured by the factor $\delta^4(p_1 - p_2 - p_3 \cdots - p_n)$.

The Golden Rule (Equation 6.15) may *look* forbidding, but what it actually *says* could hardly be simpler: all outcomes consistent with the three natural kinematic constraints are a priori equally likely. To be sure, the *dynamics* (contained in $\mathscr{M}$) may favor some combinations of momenta over others, but with that modulation you just *add up all the possibilities*. How about all those factors of 2π? These are easy to keep track of if you adhere scrupulously to the following rule:‡

$$\text{Every } \delta \text{ gets } (2\pi); \text{ every d gets } 1/(2\pi). \tag{6.16}$$

Four-dimensional 'volume' elements can be split into spatial and temporal parts:

$$\mathrm{d}^4 p = \mathrm{d}p^0\, \mathrm{d}^3\mathbf{p} \tag{6.17}$$

(I'll drop the subscript j, for simplicity – this argument applies to *each* of the outgoing momenta). The p^0 integrals§ can be performed immediately, by exploiting the delta function

$$\delta\left(p^2 - m^2c^2\right) = \delta\left[(p^0)^2 - \mathbf{p}^2 - m^2c^2\right] \tag{6.18}$$

Now

$$\delta(x^2 - a^2) = \frac{1}{2a}\left[\delta(x-a) + \delta(x+a)\right] \quad (a > 0) \tag{6.19}$$

* If you are unfamiliar with the Dirac delta function, you *must* study Appendix A carefully before proceeding.

† $\theta(x)$ is the (Heaviside) step function: 0 if $x < 0$ and 1 if $x > 0$ (see Appendix A).

‡ Some of these factors eventually cancel out, and you might wonder if there is a more efficient way to manage them. I don't think so. Feynman is supposed to have shouted in exasperation (at a graduate student who 'couldn't be bothered with such trivial matters') 'If you can't get the 2π's right, you don't know *nothing!*'

§ The integral sign in Equation 6.15 actually stands for $4(n-1)$ integrations – one for each component of the $n-1$ outgoing momenta.

(see Problem A.7), so

$$\theta\left(p^0\right)\delta\left[\left(p^0\right)^2-\mathbf{p}^2-m^2c^2\right]=\frac{1}{2\sqrt{\mathbf{p}^2+m^2c^2}}\delta\left(p^0-\sqrt{\mathbf{p}^2+m^2c^2}\right) \tag{6.20}$$

(the theta function kills the spike at $p^0=-\sqrt{\mathbf{p}^2+m^2c^2}$, and it's 1 at $p^0=\sqrt{\mathbf{p}^2+m^2c^2}$). Thus Equation 6.15 reduces to

$$\Gamma=\frac{S}{2\hbar m_1}\int|\mathscr{M}|^2(2\pi)^4\delta^4\left(p_1-p_2-p_3\cdots-p_n\right)$$
$$\times\prod_{j=2}^{n}\frac{1}{2\sqrt{\mathbf{p}_j^2+m_j^2c^2}}\frac{\mathrm{d}^3\mathbf{p}_j}{(2\pi)^3} \tag{6.21}$$

with

$$p_j^0\rightarrow\sqrt{\mathbf{p}_j^2+m_j^2c^2} \tag{6.22}$$

wherever it appears (in $\mathscr{M}$ and in the remaining delta function). This is a more useful way to express the Golden Rule, though it obscures the physical content.*

6.2.1.1 Two-particle Decays

In particular, if there are only two particles in the final state

$$\Gamma=\frac{S}{32\pi^2\hbar m_1}\int|\mathscr{M}|^2\frac{\delta^4(p_1-p_2-p_3)}{\sqrt{\mathbf{p}_2^2+m_2^2c^2}\sqrt{\mathbf{p}_3^2+m_3^2c^2}}\,\mathrm{d}^3\mathbf{p}_2x\,\mathrm{d}^3\mathbf{p}_3 \tag{6.23}$$

The four-dimensional delta function is a product of temporal and spatial parts:

$$\delta^4(p_1-p_2-p_3)=\delta\left(p_1^0-p_2^0-p_3^0\right)\delta^3\left(\mathbf{p}_1-\mathbf{p}_2-\mathbf{p}_3\right) \tag{6.24}$$

But particle 1 is at rest, so $\mathbf{p}_1=\mathbf{0}$ and $p_1^0=m_1c$. Meanwhile, p_2^0 and p_3^0 have been replaced (Equation 6.22), so†

$$\Gamma=\frac{S}{32\pi^2\hbar m_1}\int|\mathscr{M}|^2\frac{\delta\left(m_1c-\sqrt{\mathbf{p}_2^2+m_2^2c^2}-\sqrt{\mathbf{p}_3^2+m_3^2c^2}\right)}{\sqrt{\mathbf{p}_2^2+m_2^2c^2}\sqrt{\mathbf{p}_3^2+m_3^2c^2}}$$
$$\times\delta^3\left(\mathbf{p}_2+\mathbf{p}_3\right)\,\mathrm{d}^3\mathbf{p}_2\,\mathrm{d}^3\mathbf{p}_3 \tag{6.25}$$

* You might recognize the quantity $\sqrt{\mathbf{p}_j^2+m_j^2c^2}$ as E_j/c, and many books write it this way. It's dangerous notation: $\mathbf{p}_j$ is an integration variable, so E_j is not some constant you can take outside the integral. Use it as shorthand, if you like, but remember that E_j is a function of $\mathbf{p}_j$, not an independent variable.

† We can drop the minus sign in the final delta function, since $\delta(-x)=\delta(x)$.

The $\mathbf{p}_3$ integral is now trivial: in view of the final delta function it simply makes the replacement

$$\mathbf{p}_3 \to -\mathbf{p}_2 \tag{6.26}$$

leaving

$$\Gamma = \frac{S}{32\pi^2\hbar m_1}\int |\mathscr{M}|^2 \frac{\delta\left(m_1 c - \sqrt{\mathbf{p}_2^2 + m_2^2c^2} - \sqrt{\mathbf{p}_2^2 + m_3^2c^2}\right)}{\sqrt{\mathbf{p}_2^2 + m_2^2c^2}\sqrt{\mathbf{p}_2^2 + m_3^2c^2}}\, d^3\mathbf{p}_2 \tag{6.27}$$

For the remaining integral we adopt spherical coordinates, $\mathbf{p}_2 \to (r,\ \theta,\ \phi)$, $d^3\mathbf{p}_2 \to r^2 \sin\theta dr\, d\theta\, d\phi$ (this is *momentum* space, of course: $r = |\mathbf{p}_2|$).

$$\Gamma = \frac{S}{32\pi^2\hbar m_1}\int |\mathscr{M}|^2 \frac{\delta\left(m_1 c - \sqrt{r^2 + m_2^2c^2} - \sqrt{r^2 + m_3^2c^2}\right)}{\sqrt{r^2 + m_2^2c^2}\sqrt{r^2 + m_3^2c^2}}$$
$$\times\, r^2 \sin\theta\, dr\, d\theta\, d\phi \tag{6.28}$$

Now, $\mathscr{M}$ was originally a function of the four-momenta p_1, p_2, and p_3, but $p_1 = (m_1 c, \mathbf{0})$ is a constant (as far as the integration is concerned), and the integrals already performed have made the replacements $p_2^0 \to \sqrt{\mathbf{p}_2^2 + m_2^2c^2}$, $p_3^0 \to \sqrt{\mathbf{p}_3^2 + m_3^2c^2}$, and $\mathbf{p}_3 \to -\mathbf{p}_2$, so by now $\mathscr{M}$ depends only on $\mathbf{p}_2$. As we shall see, however, amplitudes must be *scalars*, and the only scalar you can make out of a vector is the dot product with itself:* $\mathbf{p}_2 \cdot \mathbf{p}_2 = r^2$. At this stage, then, $\mathscr{M}$ is a function only of r (not of θ or ϕ). That being the case we can do the angular integrals

$$\int_0^\pi \sin\theta\, d\theta = 2, \quad \int_0^{2\pi} d\phi = 2\pi \tag{6.29}$$

and there remains only the r integral:

$$\Gamma = \frac{S}{8\pi\hbar m_1}\int_0^\infty |\mathscr{M}(r)|^2 \frac{\delta\left(m_1 c - \sqrt{r^2 + m_2^2c^2} - \sqrt{r^2 + m_3^2c^2}\right)}{\sqrt{r^2 + m_2^2c^2}\sqrt{r^2 + m_3^2c^2}}\, r^2\, dr \tag{6.30}$$

To simplify the argument of the delta function, let

$$u \equiv \sqrt{r^2 + m_2^2c^2} + \sqrt{r^2 + m_3^2c^2} \tag{6.31}$$

* If the particles carry spin, then $\mathscr{M}$ might depend also on $(\mathbf{p}_i \cdot \mathbf{S}_j)$ and $(\mathbf{S}_i \cdot \mathbf{S}_j)$. However, since experiments rarely measure the spin orientation, we almost always work with the spin-averaged amplitude. In that case, and of course in the case of spin 0, the only vector in sight is $\mathbf{p}_2$ and the only scalar variable is $(\mathbf{p}_2)^2$.

so

$$\frac{du}{dr} = \frac{ur}{\sqrt{r^2 + m_2^2c^2}\sqrt{r^2 + m_3^2c^2}} \tag{6.32}$$

Then

$$\Gamma = \frac{S}{8\pi\hbar m_1}\int_{(m_2+m_3)c}^{\infty} |\mathscr{M}(r)|^2\delta\left(m_1c - u\right)\frac{r}{u}\,du \tag{6.33}$$

The last integral sends* u to m_1c, and hence r to

$$r_0 = \frac{c}{2m_1}\sqrt{m_1^4 + m_2^4 + m_3^4 - 2m_1^2m_2^2 - 2m_1^2m_3^2 - 2m_2^2m_3^2} \tag{6.34}$$

(Problem 6.5). Remember that r was short for the variable $|\mathbf{p}_2|$; r_0 is the *particular* value of $|\mathbf{p}_2|$ that is consistent with conservation of energy, and Equation 6.25 simply reproduces the result we obtained back in Chapter 3 (Problem 3.19). In more comprehensible notation, then,

$$\Gamma = \frac{S|\mathbf{p}|}{8\pi\hbar m_1^2c}|\mathscr{M}|^2 \tag{6.35}$$

where $|\mathbf{p}|$ is the magnitude of either outgoing momentum, given in terms of the three masses by Equation 6.34, and $\mathscr{M}$ is evaluated at the momenta dictated by the conservation laws. The various substitutions (Equations 6.22, 6.26, and 6.34) have systematically enforced these conservation laws – hardly a surprise, since they were built into the Golden Rule.

The two-body decay formula (Equation 6.35) is surprisingly simple; we were able to carry out all the integrals *without ever knowing the functional form of* $\mathscr{M}$! Mathematically, there were just enough delta functions to cover all the variables; physically, two-body decays are *kinematically determined*: the particles have to come out back-to-back with opposite three-momenta – the *direction* of this axis is not fixed, but since the initial state was symmetric, it doesn't matter. We will use Equation 6.35 frequently. Unfortunately, when there are three or more particles in the final state, the integrals cannot be done until we know the specific functional form of $\mathscr{M}$. In such cases (of which we shall encounter mercifully few), you have to go back to the Golden Rule and work it out from scratch.

6.2.2
Golden Rule for Scattering

Suppose particles 1 and 2 collide, producing particles 3, 4, . . . , n:

$$1 + 2 \rightarrow 3 + 4 + \cdots + n \tag{6.36}$$

* This assumes $m_1 > (m_2 + m_3)$; otherwise the delta function spike is outside the domain of integration and we get $\Gamma = 0$, recording the fact that a particle cannot decay into heavier secondaries.

The scattering cross section is given by the formula

$$\sigma = \frac{S\hbar^2}{4\sqrt{(p_1 \cdot p_2)^2 - (m_1 m_2 c^2)^2}} \int |\mathscr{M}|^2 (2\pi)^4 \delta^4(p_1 + p_2 - p_3 \cdots - p_n) \times \prod_{j=3}^{n} 2\pi\delta\left(p_j^2 - m_j^2 c^2\right) \theta\left(p_j^0\right) \frac{d^4 p_j}{(2\pi)^4} \tag{6.37}$$

where p_i is the four-momentum of particle i (mass m_i) and the statistical factor (S) is the same as before (Equation 6.15). The phase space is essentially the same as before: integrate over all outgoing momenta, subject to the three kinematical constraints (every outgoing particle is on its mass shell, every outgoing energy is positive, and energy and momentum are conserved), which are enforced by the delta and theta functions. Once again, we can simplify matters by performing the p_j^0 integrals:

$$\sigma = \frac{S\hbar^2}{4\sqrt{(p_1 \cdot p_2)^2 - (m_1 m_2 c^2)^2}} \int |\mathscr{M}|^2 (2\pi)^4 \delta^4(p_1 + p_2 - p_3 \cdots - p_n) \times \prod_{j=3}^{n} \frac{1}{2\sqrt{\mathbf{p}_j^2 + m_j^2 c^2}} \frac{d^3\mathbf{p}_j}{(2\pi)^3} \tag{6.38}$$

with

$$p_j^0 = \sqrt{\mathbf{p}_j^2 + m_j^2 c^2} \tag{6.39}$$

wherever it occurs in $\mathscr{M}$ and the delta function.

6.2.2.1 Two-body Scattering in the CM Frame

Consider the process

$$1 + 2 \rightarrow 3 + 4 \tag{6.40}$$

in the CM frame, $\mathbf{p}_2 = -\mathbf{p}_1$ (Figure 6.5), where

$$\sqrt{(p_1 \cdot p_2)^2 - (m_1 m_2 c^2)^2} = (E_1 + E_2)|\mathbf{p}_1|/c \tag{6.41}$$

(Problem 6.7). In this case, Equation 6.38 reduces to

$$\sigma = \frac{S\hbar^2 c}{64\pi^2 (E_1 + E_2)|\mathbf{p}_1|} \int |\mathscr{M}|^2 \frac{\delta^4(p_1 + p_2 - p_3 - p_4)}{\sqrt{\mathbf{p}_3^2 + m_3^2 c^2}\sqrt{\mathbf{p}_4^2 + m_4^2 c^2}} d^3\mathbf{p}_3\, d^3\mathbf{p}_4 \tag{6.42}$$

Fig. 6.5 Two-body scattering in the *CM* frame.

As before, we begin by rewriting the delta function:*

$$\delta^4(p_1 + p_2 - p_3 - p_4) = \delta\left(\frac{E_1 + E_2}{c} - p_3^0 - p_4^0\right)\delta^3(\mathbf{p}_3 + \mathbf{p}_4) \tag{6.43}$$

Next we insert Equation 6.39 and carry out the $\mathbf{p}_4$ integral (which sends $\mathbf{p}_4 \to -\mathbf{p}_3$):

$$\begin{aligned}\sigma &= \left(\frac{\hbar}{8\pi}\right)^2 \frac{Sc}{(E_1 + E_2)|\mathbf{p}_1|}\int |\mathscr{M}|^2 \\ &\quad \times \frac{\delta\left[(E_1 + E_2)/c - \sqrt{\mathbf{p}_3^2 + m_3^2c^2} - \sqrt{\mathbf{p}_3^2 + m_4^2c^2}\right]}{\sqrt{\mathbf{p}_3^2 + m_3^2c^2}\sqrt{\mathbf{p}_3^2 + m_4^2c^2}}\, \mathrm{d}^3\mathbf{p}_3\end{aligned} \tag{6.44}$$

This time, however, $|\mathscr{M}|^2$ depends on the *direction* of $\mathbf{p}_3$ as well as its magnitude,† so we cannot carry out the angular integration. But that's all right – we didn't really want σ in the first place; what we're after is $d\sigma/d\Omega$. Adopting spherical coordinates, as before,

$$\mathrm{d}^3\mathbf{p}_3 = r^2\, \mathrm{d}r\, \mathrm{d}\Omega \tag{6.45}$$

(where r is shorthand for $|\mathbf{p}_3|$ and $\mathrm{d}\Omega = \sin\theta\, \mathrm{d}\theta\, \mathrm{d}\phi$), we obtain

$$\begin{aligned}\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} &= \left(\frac{\hbar}{8\pi}\right)^2 \frac{Sc}{(E_1 + E_2)|\mathbf{p}_1|}\int_0^\infty |\mathscr{M}|^2 \\ &\quad \times \frac{\delta\left[(E_1 + E_2)/c - \sqrt{r^2 + m_3^2c^2} - \sqrt{r^2 + m_4^2c^2}\right]}{\sqrt{r^2 + m_3^2c^2}\sqrt{r^2 + m_4^2c^2}}\, r^2\, \mathrm{d}r\end{aligned} \tag{6.46}$$

* Observe that $\mathbf{p}_1$ and $\mathbf{p}_2$ are *fixed* vectors (related by our choice of reference frame: $\mathbf{p}_2 = -\mathbf{p}_1$), but at this stage $\mathbf{p}_3$ and $\mathbf{p}_4$ are integration variables. It is only *after* the $\mathbf{p}_4$ integration that they are restricted ($\mathbf{p}_4 = -\mathbf{p}_3$), and after the $|\mathbf{p}_3|$ integration that they are determined by the scattering angle θ.

† In general, $|\mathscr{M}|^2$ depends on all four-momenta. However, in this case $\mathbf{p}_2 = -\mathbf{p}_1$ and $\mathbf{p}_4 = -\mathbf{p}_3$, so it remains a function only of $\mathbf{p}_1$ and $\mathbf{p}_3$ (assuming again that spin does not come into it). From these vectors we can construct three scalars: $\mathbf{p}_1\cdot\mathbf{p}_1 = |\mathbf{p}_1|^2$, $\mathbf{p}_3\cdot\mathbf{p}_3 = |\mathbf{p}_3|^2$, and $\mathbf{p}_1\cdot\mathbf{p}_3 = |\mathbf{p}_1||\mathbf{p}_3|\cos\theta$. But $\mathbf{p}_1$ is fixed, so the only *integration* variables on which $|\mathscr{M}|^2$ can depend are $|\mathbf{p}_3|$ and θ.

The integral over r is the same as in Equation 6.30, with $m_2 \to m_4$ and $m_1 \to (E_1 + E_2)/c^2$. Quoting our previous result (Equation 6.35), I conclude that

$$\frac{d\sigma}{d\Omega} = \left(\frac{\hbar c}{8\pi}\right)^2 \frac{S|\mathscr{M}|^2}{(E_1 + E_2)^2} \frac{|\mathbf{p}_f|}{|\mathbf{p}_i|} \tag{6.47}$$

where $|\mathbf{p}_f|$ is the magnitude of either outgoing momentum and $|\mathbf{p}_i|$ is the magnitude of either incoming momentum.

As in the case of decays, the two-body final state is peculiarly simple, in the sense that we are able to carry the calculation through to the end *without knowing the explicit functional form of* $\mathscr{M}$. We will be using Equation 6.47 frequently in later chapters.

By the way, lifetimes obviously carry the dimensions of *time* (seconds); decay rates ($\Gamma = 1/\tau$), therefore, are measured in inverse seconds. Cross sections have dimensions of *area* – cm^2, or, more conveniently, 'barns':

$$1\ \text{b} = 10^{-24}\ \text{cm}^2 \tag{6.48}$$

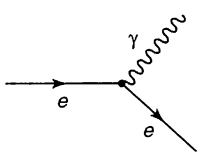
Differential cross sections, $d\sigma/d\Omega$, are given in barns per steradian or simply barns (steradians, like radians, being dimensionless). The amplitude, $\mathscr{M}$, has units that depend on the number of particles involved: if there are n external lines (incoming plus outgoing), the dimensions of $\mathscr{M}$ are those of momentum raised to the power $4 - n$:

$$\text{Dimensions of } \mathscr{M} = (mc)^{4-n} \tag{6.49}$$

For example, in a three-body process ($A \to B + C$), $\mathscr{M}$ has dimensions of momentum; in a four-body process ($A \to B + C + D$ or $A + B \to C + D$), $\mathscr{M}$ is dimensionless. You can check for yourself that the two Golden Rules then yield the correct units for Γ and σ.

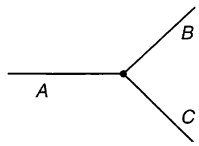
6.3 Feynman Rules for a Toy Theory

In Section 6.2, we learned how to calculate decay rates and scattering cross sections, in terms of the amplitude $\mathscr{M}$ for the process in question. Now I'll show you how to determine $\mathscr{M}$ itself, using the 'Feynman rules' to evaluate the relevant diagrams. We could go straight to a 'real-life' system, such as quantum electrodynamics, with electrons and photons interacting via the primitive vertex:

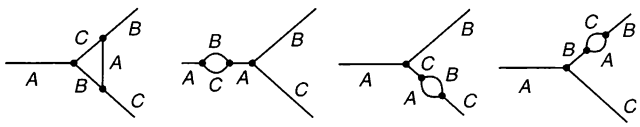


This is the original, the most important, and the best understood application of Feynman's technique. Unfortunately, it involves diverting complications (the electron has spin $\frac{1}{2}$, the photon is massless and carries spin 1), which have nothing to do with the Feynman calculus as such. In Chapter 7, I'll show you how to handle particles with spin, but for the moment I don't want to confuse the issue, so I'm going to introduce a 'toy' theory, which does not pretend to represent the real world, but will serve to illustrate the *method*, with a minimum of extraneous baggage [4].

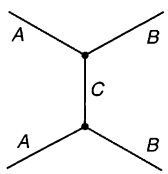
Imagine a world in which there are just three kinds of particles – call them A, B, and C – with masses m_A, m_B, and m_C. They all have spin 0 and each is its own antiparticle (so we don't need arrows on the lines). There is one primitive vertex, by which the three particles interact:



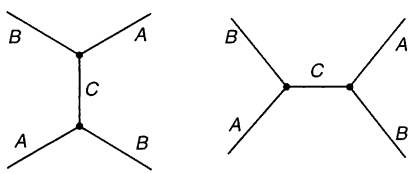
I shall assume that A is the heaviest of the three and in fact weighs more than B and C combined, so that it can decay into $B + C$. The lowest-order diagram describing this disintegration is the primitive vertex itself; to this there are (small) third-order corrections:



and even smaller ones of higher order. Our first project will be to calculate the lifetime of the A, to lowest order. After that, we'll look at various scattering processes, such as $A + A \to B + B$:



$A + B \to A + B$:



and so on.

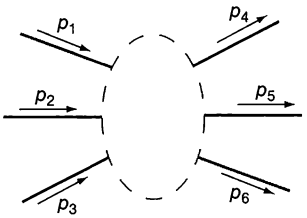


Fig. 6.6 A generic Feynman diagram, with external lines labeled (internal lines not shown).

Our problem is to find the amplitude $\mathscr{M}$ associated with a given Feynman diagram. The ritual is as follows [5]:

1. *Notation*: Label the incoming and outgoing four-momenta $p_1, p_2, \ldots, p_n$ (Figure 6.6). Label the internal momenta q_1, $q_2, \ldots$ Put an arrow beside each line, to keep track of the 'positive' direction (forward in time for external lines, arbitrary for internal lines).
2. *Vertex factors*: For each vertex, write down a factor

$$-ig$$

   g is called the *coupling constant*; it specifies the strength of the interaction between A, B, and C. In this toy theory, g has the dimensions of momentum; in the 'real-world' theories, we shall encounter later on, the coupling constant is always dimensionless.
3. *Propagators*: For each internal line, write a factor

$$\frac{i}{q_j^2 - m_j^2 c^2}$$

   where q_j is the four-momentum of the line and m_j is the mass of the particle the line describes. (Note that $q_j^2 \neq m_j^2 c^2$, because a virtual particle does not lie on its mass shell.)
4. *Conservation of energy and momentum*: For each vertex, write a delta function of the form

$$(2\pi)^4 \delta^4(k_1 + k_2 + k_3)$$

   where the k's are the three four-momenta coming *into* the vertex (if the arrow leads outward, then k is *minus* the four-momentum of that line). This factor imposes conservation of energy and momentum at each vertex, since the delta function is zero unless the sum of the incoming momenta equals the sum of the outgoing momenta.

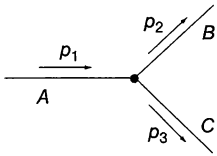


Fig. 6.7 Lowest-order contribution to $A \to B + C$.

5. *Integration over internal momenta*: For each internal line, write down a factor*

$$\frac{1}{(2\pi)^4}\, d^4 q_j$$

and integrate over all internal momenta.

6. *Cancel the delta function*: The result will include a delta function

$$(2\pi)^4 \delta^4(p_1 + p_2 + \cdots - p_n)$$

reflecting overall conservation of energy and momentum. Erase this factor† and multiply by i. The result is $\mathscr{M}$.

6.3.1
Lifetime of the *A*

The simplest possible diagram, representing the lowest-order contribution to $A \to B + C$, has no internal lines at all (Figure 6.7). There is one vertex, at which we pick up a factor of $-ig$ (Rule 2) and a delta function

$$(2\pi)^4 \delta^4(p_1 - p_2 - p_3)$$

(Rule 4), which we promptly discard (Rule 6). Multiplying by i, we get

$$\mathscr{M} = g \tag{6.50}$$

This is the *amplitude* (to lowest order); the decay rate is found by plugging $\mathscr{M}$ into Equation 6.35:

$$\Gamma = \frac{g^2 |\mathbf{p}|}{8\pi \hbar m_A^2 c} \tag{6.51}$$

* Notice (again) that every δ gets a factor of (2π) and every d gets a factor of $1/(2\pi)$.

† Of course, the Golden Rule immediately puts this factor back in Equations 6.15 and 6.37, and you might wonder why we don't just keep it in $\mathscr{M}$. The problem is that $|\mathscr{M}|^2$, not $\mathscr{M}$, comes into the Golden Rule and the *square* of a delta function is undefined. So you have to remove it here, even though you'll be putting it back at the next stage.

where $|\mathbf{p}|$ (the magnitude of either outgoing momentum) is

$$|\mathbf{p}| = \frac{c}{2m_A}\sqrt{m_A^4 + m_B^4 + m_C^4 - 2m_A^2 m_B^2 - 2m_A^2 m_C^2 - 2m_B^2 m_C^2} \tag{6.52}$$

The lifetime of the A, then, is

$$\tau = \frac{1}{\Gamma} = \frac{8\pi\hbar m_A^2 c}{g^2|\mathbf{p}|} \tag{6.53}$$

You should check for yourself that τ comes out with the correct units.

6.3.2
$A + A \to B + B$ Scattering

The lowest-order contribution to the process $A + A \to B + B$ is shown in Figure 6.8 In this case, there are two vertices (hence two factors of $-ig$), one internal line, with the propagator

$$\frac{i}{q^2 - m_C^2 c^2}$$

two delta functions:

$$(2\pi)^4\delta^4(p_1 - p_3 - q) \quad \textit{and} \quad (2\pi)^4\delta^4(p_2 + q - p_4)$$

and one integration:

$$\frac{1}{(2\pi)^4}\mathrm{d}^4 q$$

Rules 1–5, then, yield

$$-i(2\pi)^4 g^2 \int \frac{1}{q^2 - m_C^2 c^2}\,\delta^4(p_1 - p_3 - q)\,\delta^4(p_2 + q - p_4)\,\mathrm{d}^4 q$$

Doing the integral, the second delta function sends $q \to p_4 - p_2$, and we have

$$-ig^2 \frac{1}{(p_4 - p_2)^2 - m_C^2 c^2}(2\pi)^4\delta^4(p_1 + p_2 - p_3 - p_4)$$

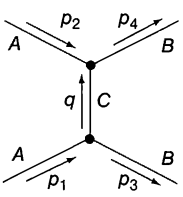


Fig. 6.8 Lowest-order contribution to $A + A \to B + B$.

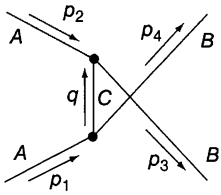


Fig. 6.9 Second diagram contributing in lowest order to $A + A \to B + B$.

As promised, there is one remaining delta function, reflecting overall conservation of energy and momentum. Erasing it and multiplying by i (Rule 6), we are left with

$$\mathscr{M} = \frac{g^2}{(p_4 - p_2)^2 - m_C^2 c^2} \tag{6.54}$$

But that's not the whole story, for there is another diagram of order g^2, obtained by 'twisting' the B lines (Figure 6.9).* Since this differs from Figure 6.8 only by the interchange $p_3 \leftrightarrow p_4$, there is no need to compute it from scratch; quoting Equation 6.54, we can write down immediately the total amplitude (to order g^2) for the process $A + A \to B + B$:

$$\mathscr{M} = \frac{g^2}{(p_4 - p_2)^2 - m_C^2 c^2} + \frac{g^2}{(p_3 - p_2)^2 - m_C^2 c^2} \tag{6.55}$$

Notice, incidentally, that $\mathscr{M}$ is a Lorentz-invariant (scalar) quantity. This is *always* the case; it is built into the Feynman rules.

Suppose we are interested in the differential cross section $(\mathrm{d}\sigma/\mathrm{d}\Omega)$ for this process, in the CM system (Figure 6.10). Say, for simplicity, that $m_A = m_B = m$ and $m_C = 0$. Then

$$(p_4 - p_2)^2 - m_C^2 c^2 = p_4^2 + p_2^2 - 2p_2 \cdot p_4 = -2\mathbf{p}^2(1 - \cos\theta) \tag{6.56}$$

$$(p_3 - p_2)^2 - m_C^2 c^2 = p_3^2 + p_2^2 - 2p_3 \cdot p_2 = -2\mathbf{p}^2(1 + \cos\theta) \tag{6.57}$$

(where $\mathbf{p}$ is the incident momentum of particle 1), and hence

$$\mathscr{M} = -\frac{g^2}{\mathbf{p}^2 \sin^2\theta} \tag{6.58}$$

According to Equation 6.47, then,

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{2}\left(\frac{\hbar c g^2}{16\pi E \mathbf{p}^2 \sin^2\theta}\right)^2 \tag{6.59}$$

(there are two identical particles in the final state, so $S = 1/2$). As in the case of Rutherford scattering (Example 6.4), the *total* cross section is infinite.

* You don't get yet another new diagram by twisting the A lines; the only choice here is whether p_3 connects to p_1 or to p_2.

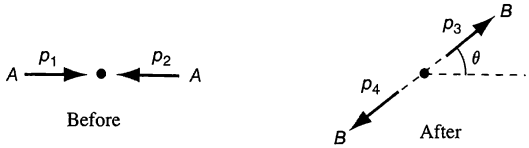
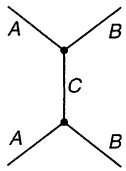


Fig. 6.10 $A + A \to B + B$ in the *CM* frame.

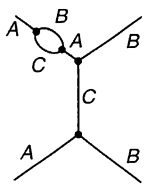
6.3.3
Higher-order Diagrams

So far we have looked only at lowest-order ('tree level') Feynman diagrams; in the case of $A + A \to B + B$, for instance, we considered the graph:

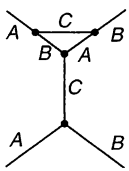


This diagram has two vertices, so $\mathscr{M}$ is proportional to g^2. But there are eight diagrams with *four* vertices (and eight more with the external B lines 'twisted'):

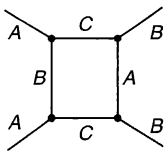
- five 'self-energy' diagrams, in which one of the lines sprouts a loop:



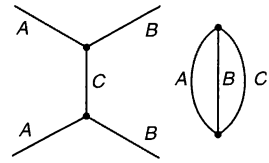
- two 'vertex corrections', in which a vertex becomes a triangle:



- and one 'box' diagram:

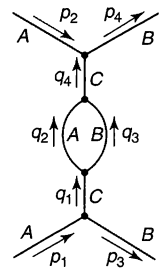


(Disconnected diagrams, such as



don't count.)

I am certainly not going to evaluate all these 'one-loop' diagrams (or even *think* about *two*-loop diagrams), but I would like to take a closer look at *one* of them – the one with a bubble on the virtual C line:



Applying Feynman rules 1–5, we obtain

$$g^4 \int \frac{\delta^4(p_1 - q_1 - p_3)\,\delta^4(q_1 - q_2 - q_3)\,\delta^4(q_2 + q_3 - q_4)\,\delta^4(q_4 + p_2 - p_4)}{(q_1^2 - m_C^2c^2)(q_2^2 - m_A^2c^2)(q_3^2 - m_B^2c^2)(q_4^2 - m_C^2c^2)} \times \mathrm{d}^4q_1\,\mathrm{d}^4q_2\,\mathrm{d}^4q_3\,\mathrm{d}^4q_4 \tag{6.60}$$

Integration over q_1, using the first delta function, replaces q_1 by $(p_1 - p_3)$; integration over q_4, using the last delta function, replaces q_4 by $(p_4 - p_2)$:

$$\frac{g^4}{[(p_1 - p_3)^2 - m_C^2c^2][(p_4 - p_2)^2 - m_C^2c^2]} \times \int \frac{\delta^4(p_1 - p_3 - q_2 - q_3)\,\delta^4(q_2 + q_3 - p_4 + p_2)}{(q_2^2 - m_A^2c^2)(q_3^2 - m_B^2c^2)}\,\mathrm{d}^4q_2\,\mathrm{d}^4q_3 \tag{6.61}$$

Here, the first delta function sends $q_2 \to p_1 - p_3 - q_3$, and the second delta function becomes

$$\delta^4(p_1 + p_2 - p_3 - p_4)$$

which, by Rule 6, we erase, leaving

$$\mathscr{M} = i\left(\frac{g}{2\pi}\right)^4 \frac{1}{[(p_1 - p_3)^2 - m_C^2c^2]^2} \int \frac{1}{[(p_1 - p_3 - q)^2 - m_A^2c^2](q^2 - m_B^2c^2)}\,\mathrm{d}^4q \tag{6.62}$$

(I drop the subscript on q_3 at this point.)

You can try calculating this integral, if you've got the energy, but I'll tell you right now you're going to hit a snag. The four-dimensional volume element could be written as $d^4q = q^3\, dq\, d\Omega'$ (where $d\Omega'$ stands for the angular part), just as in two-dimensional polar coordinates the element of area is $r\, dr\, d\theta$ and in three-dimensional spherical coordinates the volume element is $r^2\, dr \sin\theta\, d\theta\, d\phi$. At large q the integrand is essentially just $1/q^4$, so the q integral has the form

$$\int^{\infty} \frac{1}{q^4} q^3\, dq = \ln q|^{\infty} = \infty \tag{6.63}$$

The integral is logarithmically divergent at large q. This disaster, in one form or another, held up the development of quantum electrodynamics for nearly two decades, until, through the combined efforts of many great physicists – from Dirac, Pauli, Kramers, Weisskopf, and Bethe through Tomonaga, Schwinger, and Feynman – systematic methods were developed for 'sweeping the infinities under the rug'. The first step is to *regularize* the integral, using a suitable cutoff procedure that renders it finite without spoiling other desirable features (such as Lorentz invariance). In the case of Equation 6.62, this can be accomplished by introducing a factor

$$\frac{-M^2c^2}{(q^2 - M^2c^2)} \tag{6.64}$$

under the integral sign. The *cutoff* mass M is assumed to be very large, and will be taken to infinity at the end of the calculation (note that the 'fudge factor', Equation 6.64, goes to 1 as $M \to \infty$).* The integral can now be calculated [6] and separated into two parts: a finite term, independent of M, and a term involving (in this case) the logarithm of M, which blows up as $M \to \infty$.

At this point, a miraculous thing happens: all the divergent, M-dependent terms appear in the final answer in the form of *additions to the masses and the coupling constant.* If we take this seriously, it means that the *physical* masses and couplings

* No one would deny that this procedure is artificial. Still, it can be argued that the inclusion of Equation 6.64 merely confesses our ignorance of the high-energy (short distance) behavior of quantum field theory. Perhaps the Feynman propagators are not quite right in this regime, and M is simply a crude way of accounting for the unknown modification. (This would be the case, for example, if the 'particles' have substructure that becomes relevant at extremely close range.) Dirac said, of renormalization,

> *It's just a stop-gap procedure. There must be some fundamental change in our ideas, probably a change just as fundamental as the passage from Bohr's orbit theory to quantum mechanics. When you get a number turning out to be infinite which ought to be finite, you should admit that there is something wrong with your equations, and not hope that you can get a good theory just by doctoring up that number.*

P. Buckley and F. D. Peat, *A Question of Physics* (Toronto: University of Toronto Press, 1979), page 39.

are not the m's and g's that appeared in the original Feynman rules, but rather the 'renormalized' ones, containing these extra factors:

$$m_{\text{physical}} = m + \delta m; \quad g_{\text{physical}} = g + \delta g \tag{6.65}$$

The fact that δm and δg are infinite (in the limit $M \to \infty$) is disturbing, but not catastrophic, for we never measure them anyway; all we ever see in the laboratory are the *physical* values, and these are (obviously) finite (evidently the unmeasurable 'bare' masses and couplings, m and g, contain compensating infinities).* As a practical matter, we take account of the infinities by using the *physical* values of m and g in the Feynman rules, and then systematically ignoring the divergent contributions from higher-order diagrams.

Meanwhile, there remain the *finite* (M-*in*dependent) contributions from the loop diagrams. They, too, lead to modifications in m and g (perfectly calculable ones, in this case) – which, however, are functions of the four-momentum of the line in which the loop is inserted ($p_1 - p_3$ in the example). This means that the *effective* masses and coupling constants actually depend on the *energies* of the particles involved; we call them 'running' masses and 'running' coupling constants. The dependence is typically rather slight, at low energies, and can ordinarily be ignored, but it does have observable consequences, in the form of the Lamb shift (in QED) and asymptotic freedom (in QCD).†

* In case it is some comfort, essentially the same thing occurs in *classical* electrodynamics: the electrostatic energy of a point charge is infinite, and makes an infinite contribution (via $E = mc^2$) to the particle's mass. Perhaps this means that there are no true point charges, in classical electrodynamics; perhaps that's what it means in quantum field theory, too. In neither case, however, do we know how to avoid the point particle as a theoretical construct.

† A *physical* interpretation of the running coupling constant in *QED* and *QCD* was suggested in Chapter 2, Section 2.3. A nice explanation of mass renormalization is given by P. Nelson in *American Scientist*, 73, 66 (1985):

> *According to renormalization theory, not only the strengths of the various interactions but the masses of the participating particles appear to vary on differing length scales. To get a feel for this seemingly paradoxical statement, imagine firing a cannon underwater. Even neglecting friction, the trajectory will be very different from the corresponding one on land, since the cannonball must now drag with it a considerable amount of water, modifying its apparent, or "effective," mass. We can experimentally measure the cannonball's effective mass by shaking it to and fro at a rate ω, computing the mass from $F = ma$. (This is how astronauts "weigh" themselves in space.) Having found the effective mass, we can now replace the difficult problem of underwater ballistics by a simplified approximation: we ignore the water altogether, but in Newton's equations we simply replace the true cannonball mass by the effective mass. The complicated details of the interaction with the medium are*

The procedure I have sketched in the last three paragraphs is called *renormalization* [7]. If *all* the infinities arising from higher-order diagrams can be accommodated in this way, we say that the theory is *renormalizable*. *ABC* theory and quantum electrodynamics are renormalizable. In the early 1970s, 't Hooft showed that all *gauge* theories, including chromodynamics and the electroweak theory of Glashow, Weinberg, and Salam, are renormalizable. This was a profoundly important discovery, because, beyond lowest-order calculations, a nonrenormalizable theory yields answers that are cutoff-dependent and, therefore, really, quite meaningless.

References

1 See, for example, Taylor, J. R.(**2005**) *Classical Mechanics*, University Science Books, Sausalito, C.A. Sect.14.6, or; (a) Goldstein, H., Poole, C. and Safko, J. (**2002**) *Classical Mechanics*, 3rd edn, Addison-Wesley, San Francisco, C.A. Sect. 3-10.

2 See, for example, Park, D. (**1992**) *Introduction to the Quantum Theory*, 3rd edn, McGraw-Hill, New York. Sect. 7.9; (a) Townsend, J. S. (**2000**) *A Modern Approach to Quantum Mechanics*, University Science Books, Sausalito, C.A. Sect. 14.7; (b) Griffiths, D. J. (**2005**) *Introduction to Quantum Mechanics*, 2nd edn, Prentice Hall, Upper Saddle River, N.J. Sect. 9.2.3.

3 Convincing and accessible accounts are in fact difficult to find. Feynman, R. P. (**1961**) *Theory of Fundamental Processes*, Benjamin, Reading, M.A. Chapters 15 and 16, is a good place to start. See also; (a) Jauch, J. M. and Rohrlich, F. (**1976**) *Theory of Photons and Electrons*, 2nd edn, Springer-Verlag, Berlin. Sect. 8.6; (b) Hagedorn, R. (**1963**) *Relativistic Kinematics*, Benjamin, New York. Chap. 7; (c) Ryder, L. H. (**1985**) *Quantum Field Theory*, Cambridge University Press, Cambridge. Sect. 6.10; (d) Peskin, M. E. and Schroeder, D. V. (**1995**) *An Introduction to Quantum Field Theory*, Perseus, Cambridge. Chap. 4; (e) Weinberg, S. (**1995**) *The Quantum Theory of Fields*, vol. I, Cambridge University Press, Cambridge. Sect. 3.4.

4 This model was shown to me by Max Dresden. For a more sophisticated treatment see Aitchison, I. J. R. and Hey, A. J. G. (**2003**) *Gauge Theories in Particle Physics*, 3rd edn, vol. I, Institute of Physics, Bristol, UK. Section 6.3.

5 You may well ask where these rules *come* from, and I am not sure Feynman himself could have given you

thus reduced to determining one effective parameter. A key feature of this approach is that the effective mass so computed depends on ω, since as ω approaches zero, for example, the water has no effect whatever. In other words, the presence of a medium can introduce a scale-dependent effective mass. We say that the effective mass is "renormalized" by the medium. In quantum physics, every particle moves through a "medium" consisting of the quantum fluctuations of all particles present in the theory. We again take into account this medium by ignoring it but changing the values of our parameters to scale-dependent "effective" values.

a completely satisfactory answer in 1949, when he first published the rules for QED Feynman, R. P. (1949) *Physical Review*, **76**, 749-769. It remained for Freeman Dyson to show how the Feynman rules could be obtained from quantum field theory; (a) Dyson, F. J. (1949) *Physical Review* **75**, 486 and 1376; and (1951) **82**, 428. For a fascinating personal history of these events, see Chapters 5 and 6 of Dyson's book (1979) *Disturbing The Universe*, Harper & Row, New York. We shall return to the question of how Feynman's rules are derived in Chapter 11; for now I will simply treat them as axioms.

6 The method is explained in Sakurai, J. J. (1967) *Advanced Quantum Mechanics*, Addison-Wesley, Reading, M.A.; see particularly the useful collection of formulas in Appendix E.

7 Renormalization of *ABC* theory is studied in Griffiths, D. J. and Kraus, P. (1992) *American Journal of Physical*, **60**, 1013.

Problems

6.1 Derive Equation 6.3. [*Hint*: What fraction of the original sample decays between t and $t + \mathrm{d}t$? What, then, is the (initial) probability, $p(t)\,\mathrm{d}t$, of any given particle decaying between t and $t + \mathrm{d}t$? The *average* lifetime is $\int_0^\infty tp(t)\,\mathrm{d}t$.]

6.2 Nuclear physicists traditionally work with 'half-life' ($t_{1/2}$) instead of mean life (τ); $t_{1/2}$ is the time it takes for half the members of a large sample to decay. For exponential decay (Equation 6.2), derive the formula for $t_{1/2}$ (as a multiple of τ).

6.3 (a) Suppose you started out with a million muons (at rest); how many would still be around 2.2×10^{-5} seconds later?

(b) What is the probability of a π^- lasting more than 1 second (express your answer as a power of 10)?

6.4 A nonrelativistic particle of mass m and (kinetic) energy E scatters from a fixed repulsive potential, $V(r) = k/r^2$, where k is a constant.

(a) Find the scattering angle, θ, as a function of the impact parameter, b.

(b) Determine the differential cross section $\mathrm{d}\sigma/\mathrm{d}\Omega$, as a function of θ (not b).

(c) Find the total cross section.

[References: Goldstein, H., Poole, C. and Safko, J. (2002) *Classical Mechanics*, 3rd edn, Addison-Wesley, San Francisco, CA. Sect. 3-10., Equation 3-97; Becker, R.A. (1954) *Introduction to Theoretical Mechanics*, McGraw-Hill, New York, Example 10-3.]

6.5 Derive Equation 6.34, starting from Equation 6.31 with $u = m_1 c$.

6.6 As an application of the Golden Rule, consider the decay of $\pi^0 \to \gamma + \gamma$. Of course, the π^0 is a composite object, so Equation 6.35 does not really apply, but let's pretend that it's a true elementary particle, and see how close we come. Unfortunately, we don't know the amplitude $\mathscr{M}$; however, it must have the dimensions of mass times velocity (Equation 6.49), and there is only one mass and one velocity available. Moreover, the emission of each photon introduces a factor of $\sqrt{\alpha}$ (the fine structure constant) into $\mathscr{M}$, as we shall see in Chapter 7, so the amplitude must be proportional to α. On this basis, estimate the lifetime of the π^0. Compare the experimental value. [Evidently, the decay of the π^0 is a much more complicated process than this crude model suggests. See Quigg, C. (1997) *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions*, Addison-Wesley, Reading, M.A, Equation 1.2.25 – but beware of the misprint: f_π should be squared.]

6.7 (a) Derive Equation 6.41 for scattering of particles 1 and 2 in the CM.

(b) Obtain the corresponding formula for the lab frame (particle 2 at rest).

[*Answer* : $m_2|\mathbf{p}_1|c$]

6.8 Consider elastic scattering, $a + b \to a + b$, in the lab frame (b initially at rest), assuming the target is so heavy ($m_b c^2 \gg E_a$) that its recoil is negligible. Determine the differential scattering cross section. [*Hint:* In this limit the lab frame and the CM frame are the same.]

$$[\textit{Answer}: \ (\mathrm{d}\sigma/\mathrm{d}\Omega) = (\hbar/8\pi m_b c)^2 |\mathscr{M}|^2]$$

6.9 Consider the collision $1 + 2 \to 3 + 4$ in the lab frame (2 at rest), with particles 3 and 4 massless. Obtain the formula for the differential cross section.

$$\left[\textit{Answer}: \ \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left(\frac{\hbar}{8\pi}\right)^2 \frac{S|\mathscr{M}|^2|\mathbf{p}_3|}{m_2|\mathbf{p}_1|(E_1 + m_2c^2 - |\mathbf{p}_1|c\cos\theta)}\right]$$

6.10 (a) Analyze the problem of elastic scattering ($m_3 = m_1, m_4 = m_2$) in the lab frame (particle 2 at rest). Derive the formula for the differential cross section.

$$\left[\textit{Answer}: \ \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left(\frac{\hbar}{8\pi}\right)^2 \frac{\mathbf{p}_3^2 S|\mathscr{M}|^2}{m_2|\mathbf{p}_1|\,|(E_1 + m_2c^2)|\mathbf{p}_3| - |\mathbf{p}_1|E_3\cos\theta|}\right]$$

(b) If the incident particle is massless ($m_1 = 0$), show that the result in part (a) simplifies to

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right) = S\left(\frac{\hbar E_3}{8\pi m_2 c E_1}\right)^2 |\mathscr{M}|^2$$

6.11 (a) Is $A \to B + B$ a possible process in the ABC theory?
(b) Suppose a diagram has n_A external A lines, n_B external B lines, and n_C external C lines. Develop a simple criterion for determining whether it is an allowed reaction.
(c) Assuming A is heavy enough, what are the next most likely decay modes, after $A \to B + C$? Draw a Feynman diagram for each decay.

6.12 (a) Draw all the lowest-order diagrams for $A + A \to A + A$. (There are six of them.)
(b) Find the amplitude for this process, in lowest order, assuming $m_B = m_C = 0$. Leave your answer in the form of an integral over one remaining four-momentum, q.

6.13 Calculate $\mathrm{d}\sigma/\mathrm{d}\Omega$ for $A + A \to B + B$, in the CM frame, assuming $m_B = m_C = 0$. Find the total cross section, σ.

6.14 Find $\mathrm{d}\sigma/\mathrm{d}\Omega$ and σ for $A + A \to B + B$ in the lab frame. (Let E be the energy, and $\mathbf{p}$ the momentum, of the incident A. Assume $m_B = m_C = 0$.) Determine the nonrelativistic and ultrarelativistic limits of your formula.

6.15 (a) Determine the lowest-order amplitude for $A + B \to A + B$. (There are two diagrams.)
(b) Find the differential cross section for this process in the CM frame, assuming $m_A = m_B = m, m_C = 0$. Express your answer in terms of the incident energy (of A), E, and the scattering angle (for particle A), θ.
(c) Find $\mathrm{d}\sigma/\mathrm{d}\Omega$ for this process in the lab frame, assuming B is much heavier than A and remains stationary. A is incident with energy E. [*Hint:* See Problem (6.8). Assume $m_B \gg m_A, m_C$, and E/c^2.]
(d) In case (c), find the total cross section, σ.

7

Quantum Electrodynamics

In this chapter I introduce the Dirac equation, state the Feynman rules for quantum electrodynamics, develop useful calculational tools, and derive some of the classic QED results. The treatment leans heavily on material from Chapters 2, 3, and 6, as well as the spin-$\frac{1}{2}$ formalism in Chapter 4. In turn, Chapter 7 is the indispensable foundation for everything that follows (however, you may want to skip Example 7.8 and Section 7.9, together with the related passages in Chapters 8 and 9).

7.1
The Dirac Equation

Although the 'ABC' model in Chapter 6 is a perfectly legitimate quantum field theory, it does not describe the real world, because the particles A, B, and C have spin 0, whereas quarks and leptons carry spin $\frac{1}{2}$ and mediators carry spin 1. The inclusion of spin can be algebraically cumbersome; that's why I introduced the Feynman calculus in the context of a 'toy' theory free of such distractions.

In *nonrelativistic* quantum mechanics, particles are described by *Schrödinger's* equation; in *relativistic* quantum mechanics, particles of spin 0 are described by the *Klein–Gordon* equation, particles of spin $\frac{1}{2}$ by the *Dirac* equation, and particles of spin 1 by the *Proca* equation. Once the Feynman rules have been established, however, the underlying field equation fades into the background – that's how we got through Chapter 6 without ever mentioning the Klein–Gordon equation. But for spin $\frac{1}{2}$ the very *notation* of the Feynman rules presupposes some familiarity with the Dirac equation. So for the next three sections we'll study the Dirac theory in its own right.

One way to 'derive' the Schrödinger equation is to start with the *classical* energy–momentum relation:

$$\frac{\mathbf{p}^2}{2m} + V = E \tag{7.1}$$

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

apply the quantum prescription

$$\mathbf{p} \to -i\hbar\nabla, \qquad E \to i\hbar\frac{\partial}{\partial t} \tag{7.2}$$

and let the resulting operators act on the 'wave function', Ψ:

$$-\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi = i\hbar\frac{\partial\Psi}{\partial t} \qquad \textbf{(Schrödinger equation)} \tag{7.3}$$

The Klein–Gordon equation can be obtained in exactly the same way, beginning with the *relativistic* energy–momentum relation, $E^2 - \mathbf{p}^2c^2 = m^2c^4$, or (better)

$$p^\mu p_\mu - m^2c^2 = 0 \tag{7.4}$$

(I'll leave out the potential energy, from now on; we'll stick to *free* particles). Surprisingly, the quantum substitution (Equation 7.2) requires no relativistic modification; in four-vector notation, it reads

$$p_\mu \to i\hbar\,\partial_\mu \tag{7.5}$$

Here*

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} \tag{7.6}$$

which is to say

$$\partial_0 = \frac{1}{c}\frac{\partial}{\partial t}, \qquad \partial_1 = \frac{\partial}{\partial x}, \qquad \partial_2 = \frac{\partial}{\partial y}, \qquad \partial_3 = \frac{\partial}{\partial z} \tag{7.7}$$

Putting Equation 7.5 into Equation 7.4, and letting the derivatives act on a wave function ψ,† we obtain

$$-\hbar^2\partial^\mu\partial_\mu\psi - m^2c^2\psi = 0 \tag{7.8}$$

or

$$-\frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} + \nabla^2\psi = \left(\frac{mc}{\hbar}\right)^2\psi \qquad \textbf{(Klein–Gordon equation)} \tag{7.9}$$

Schrödinger actually discovered this equation even before the nonrelativistic one that bears his name; he abandoned it when (with the Coulomb potential included)

* The gradient with respect to a *contravariant* position-time four-vector x^μ is itself a *covariant* four-vector; hence the placement of the index. Written out in full, Equation (7.5) says $(E/c, -\mathbf{p}) \to i\hbar\left(\frac{1}{c}\frac{\partial}{\partial t}, \nabla\right)$. Of course, $\partial^\mu \equiv \partial/\partial x_\mu$. See Problem 7.1.

† In nonrelativistic quantum mechanics we customarily use the capital letter (Ψ) for the wave function, and reserve ψ for its spatial part (Equation 5.3). In the relativistic theory it is more common to use ψ for the wave function itself.

it failed to reproduce the Bohr energy levels for hydrogen. The problem is that the electron has spin $\frac{1}{2}$, and the Klein–Gordon equation applies to particles with spin 0. Moreover, the Klein–Gordon equation is incompatible with Born's statistical interpretation, which says that $|\psi(\mathbf{r})|^2$ gives the probability of finding the particle at the point $\mathbf{r}$. The source of this difficulty was traced to the fact that the Klein–Gordon equation is second order in t.* So Dirac set out to find an equation consistent with the relativistic energy–momentum formula, and yet first order in time. Ironically, in 1934, Pauli and Weisskopf showed that the statistical interpretation itself must be reformulated in relativistic quantum theory,† and restored the Klein–Gordon equation to its rightful place, while keeping the Dirac equation for particles of spin $\frac{1}{2}$.

Dirac's strategy was to 'factor' the energy–momentum relation (Equation 7.4). This would be easy if we had only p^0 (that is, if $\mathbf{p}$ were zero):

$$(p^0)^2 - m^2c^2 = (p^0 + mc)(p^0 - mc) = 0 \tag{7.10}$$

We obtain two first-order equations:

$$(p^0 - mc) = 0 \quad \text{or} \quad (p^0 + mc) = 0 \tag{7.11}$$

either one of which guarantees that $p^\mu p_\mu - m^2c^2 = 0$. But it's a different matter when the spatial components are included; in that case we are looking for something of the form

$$(p^\mu p_\mu - m^2c^2) = (\beta^\kappa p_\kappa + mc)(\gamma^\lambda p_\lambda - mc) \tag{7.12}$$

where β^κ and γ^λ are eight coefficients yet to be determined.‡ Multiplying out the right-hand side, we have

$$\beta^\kappa \gamma^\lambda p_\kappa p_\lambda - mc(\beta^\kappa - \gamma^\kappa)p_\kappa - m^2c^2$$

We don't want any terms linear in p_κ, so we must choose $\beta^\kappa = \gamma^\kappa$; to finish the job, we need to find coefficients γ^κ such that

$$p^\mu p_\mu = \gamma^\kappa \gamma^\lambda p_\kappa p^\lambda$$

* Notice that the Schrödinger equation is *first* order in t.

† A relativistic theory has to account for pair production and annihilation, and hence the number of particles is not a conserved quantity.

‡ In case the notation confuses you, let me write Equation (7.12) 'long-hand':

$$(p^0)^2 - (p^1)^2 - (p^2)^2 - (p^3)^2 - m^2c^2 = (\beta^0 p^0 - \beta^1 p^1 - \beta^2 p^2 - \beta^3 p^3 + mc) \times (\gamma^0 p^0 - \gamma^1 p^1 - \gamma^2 p^2 - \gamma^3 p^3 - mc)$$

which is to say

$$
\begin{aligned}
(p^0)^2 - (p^1)^2 - (p^2)^2 - (p^3)^2 = {} & (\gamma^0)^2(p^0)^2 + (\gamma^1)^2(p^1)^2 + (\gamma^2)^2(p^2)^2 \\
& + (\gamma^3)^2(p^3)^2 + (\gamma^0\gamma^1 + \gamma^1\gamma^0)p_0p_1 \\
& + (\gamma^0\gamma^2 + \gamma^2\gamma^0)p_0p_2 + (\gamma^0\gamma^3 + \gamma^3\gamma^0)p_0p_3 \\
& + (\gamma^1\gamma^2 + \gamma^2\gamma^1)p_1p_2 + (\gamma^1\gamma^3 + \gamma^3\gamma^1)p_1p_3 \\
& + (\gamma^2\gamma^3 + \gamma^3\gamma^2)p_2p_3
\end{aligned} \tag{7.13}
$$

You see the problem: we could pick $\gamma^0 = 1$ and $\gamma^1 = \gamma^2 = \gamma^3 = i$, but there doesn't seem to be any way to get rid of the cross terms.

At this point Dirac had a brilliant inspiration: what if the γ's are *matrices*, instead of *numbers*? Since matrices don't commute, we might be able to find a set such that

$$
\begin{aligned}
&(\gamma^0)^2 = 1, \qquad (\gamma^1)^2 = (\gamma^2)^2 = (\gamma^3)^2 = -1, \\
&\qquad \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 0, \qquad \text{for } \mu \neq \nu
\end{aligned} \tag{7.14}
$$

Or, more succinctly,

$$
\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} \tag{7.15}
$$

where $g^{\mu\nu}$ is the Minkowski metric (Equation 3.13), and curly brackets denote the *anticommutator*:

$$
\{A, B\} \equiv AB + BA \tag{7.16}
$$

You might try fiddling with this problem for yourself. It turns out that it *can* be done, although the smallest matrices that work are 4×4. There are a number of essentially equivalent sets of 'gamma matrices'; we'll use the standard 'Bjorken and Drell' convention [1]:

$$
\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \tag{7.17}
$$

where $\sigma^i (i = 1, 2, 3)$ is the indicated Pauli matrix (Equation 4.26), 1 denotes the 2×2 unit matrix, and 0 is the 2×2 matrix of zeroes.*

* When the context allows no room for ambiguity, I'll use 1 and 0 this way for 2×2 or 4×4 matrices; also, a unit matrix of the appropriate dimension is *implied*, when necessary, as on the right-hand side of Equation 7.15 Incidentally, since σ is not the spatial part of a four-vector, we do not distinguish upper and lower indices: $\sigma^i \equiv \sigma_i$.

As a 4×4 matrix equation, then, the relativistic energy–momentum relation *does* factor:

$$(p^\mu p_\mu - m^2c^2) = (\gamma^\kappa p_\kappa + mc)(\gamma^\lambda p_\lambda - mc) = 0 \tag{7.18}$$

We obtain the Dirac equation, now, by peeling off one term (it doesn't really matter *which* one, but this is the conventional choice – see Problem 7.10):

$$\gamma^\mu p_\mu - mc = 0 \tag{7.19}$$

Finally, we make the quantum substitution $p_\mu \to i\hbar\,\partial_\mu$ (Equation 7.5), and let the result act on the wave function ψ:

$$i\hbar\gamma^\mu \partial_\mu \psi - mc\,\psi = 0 \qquad \textbf{(Dirac equation)} \tag{7.20}$$

Note that ψ is now a four-element column matrix:

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} \tag{7.21}$$

We call it a 'bi-spinor', or 'Dirac spinor'. (Although it has four components, this object is *not* a four-vector. In Section 7.3 I'll show you how it *does* transform when you change inertial systems; it's *not* going to be an ordinary Lorentz transformation.)

7.2 Solutions to the Dirac Equation

Suppose that ψ is *independent of position*:

$$\frac{\partial\psi}{\partial x} = \frac{\partial\psi}{\partial y} = \frac{\partial\psi}{\partial z} = 0 \tag{7.22}$$

In view of Equation 7.5, this describes a state with zero momentum ($\mathbf{p} = 0$), which is to say, a particle at rest. The Dirac Equation (7.20) reduces to

$$\frac{i\hbar}{c}\gamma^0 \frac{\partial\psi}{\partial t} - mc\,\psi = 0 \tag{7.23}$$

or

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \partial\psi_A/\partial t \\ \partial\psi_B/\partial t \end{pmatrix} = -i\frac{mc^2}{\hbar} \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix} \tag{7.24}$$

where

$$\psi_A = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \tag{7.25}$$

consists of the upper two components, and

$$\psi_B = \begin{pmatrix} \psi_3 \\ \psi_4 \end{pmatrix} \tag{7.26}$$

comprises the lower two. Thus

$$\frac{\partial \psi_A}{\partial t} = -i\left(\frac{mc^2}{\hbar}\right)\psi_A, \qquad -\frac{\partial \psi_B}{\partial t} = -i\left(\frac{mc^2}{\hbar}\right)\psi_B \tag{7.27}$$

and the solutions are

$$\psi_A(t) = e^{-i(mc^2/\hbar)t}\psi_A(0), \qquad \psi_B(t) = e^{+i(mc^2/\hbar)t}\psi_B(0) \tag{7.28}$$

Referring to Equation 5.10, we recognize the factor

$$e^{-iEt/\hbar} \tag{7.29}$$

as the characteristic time dependence of a quantum state with energy E. For a particle at rest, $E = mc^2$, so ψ_A is exactly what we should have expected, in the case $\mathbf{p} = 0$. But what about ψ_B? It ostensibly represents a state with *negative energy* ($E = -mc^2$). This is the famous disaster I mentioned back in Chapter 1, which Dirac at first tried to avoid by postulating an unseen infinite 'sea' of negative-energy particles, which fill up all those unwanted states.* Instead, we now take the solutions with 'anomalous' time dependence to represent *anti*particles with *positive* energy.† Thus ψ_A describes electrons (for example), whereas ψ_B describes *positrons*; each is a two-component spinor, just right for a system of spin $\frac{1}{2}$. *Conclusion:* The Dirac equation with $\mathbf{p} = 0$ admits four independent solutions (ignoring normalization factors, for the moment):

$$\psi^{(1)} = e^{-i(mc^2/\hbar)t}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad \psi^{(2)} = e^{-i(mc^2/\hbar)t}\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\psi^{(3)} = e^{+i(mc^2/\hbar)t}\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \psi^{(4)} = e^{+i(mc^2/\hbar)t}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{7.30}$$

* You might ask why we don't simply stipulate that $\psi_B(0) = 0$ – call the 'negative-energy' solutions 'physically unacceptable', and forget about them. Unfortunately, this can't be done. In a quantum system we need a *complete* set of states, and the positive energy states by themselves are *not* complete.

† In the Schrödinger equation the sign of i is purely conventional. Had Schrödinger made the opposite choice, $e^{iEt/\hbar}$ would be the 'normal' time dependence for a stationary state of energy E. In the relativistic theory *both* signs arise, and this, when properly interpreted, implies the existence of antiparticles.

They describe, respectively, an electron with spin up, an electron with spin down, a positron with spin down,* and a positron with spin up.

We look next for *plane-wave* solutions:†

$$\psi(x) = ae^{-ik\cdot x}u(k) \tag{7.31}$$

We're hoping to find a four-vector k^μ and an associated bispinor $u(k)$ such that $\psi(x)$ satisfies the Dirac equation (a is a normalization factor, irrelevant to our present purpose but necessary later to keep the units consistent). Because the x dependence is confined to the exponent‡

$$\partial_\mu \psi = -ik_\mu \psi \tag{7.32}$$

Putting this into the Dirac Equation (7.20), we get

$$\hbar\gamma^\mu k_\mu e^{-ik\cdot x}u - mce^{-ik\cdot x}u = 0$$

or

$$(\hbar\gamma^\mu k_\mu - mc)u = 0 \tag{7.33}$$

Notice that this equation is purely algebraic – no derivatives. If u satisfies Equation 7.33, then ψ (Equation 7.31) satisfies the Dirac equation.

Now

$$\gamma^\mu k_\mu = \gamma^0 k^0 - \boldsymbol{\gamma}\cdot\mathbf{k} = k^0\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} - \mathbf{k}\cdot\begin{pmatrix}0 & \boldsymbol{\sigma}\\ -\boldsymbol{\sigma} & 0\end{pmatrix} = \begin{pmatrix}k^0 & -\mathbf{k}\cdot\boldsymbol{\sigma}\\ \mathbf{k}\cdot\boldsymbol{\sigma} & -k^0\end{pmatrix} \tag{7.34}$$

so

$$(\hbar\gamma^\mu k_\mu - mc)u = \begin{pmatrix}(\hbar k^0 - mc) & -\hbar\mathbf{k}\cdot\boldsymbol{\sigma}\\ \hbar\mathbf{k}\cdot\boldsymbol{\sigma} & (-\hbar k^0 - mc)\end{pmatrix}\begin{pmatrix}u_A\\ u_B\end{pmatrix}$$

$$= \begin{pmatrix}(\hbar k^0 - mc)u_A - \hbar\mathbf{k}\cdot\boldsymbol{\sigma}u_B\\ \hbar\mathbf{k}\cdot\boldsymbol{\sigma}u_A - (\hbar k^0 + mc)u_B\end{pmatrix}$$

* Notice the 'backward' antiparticle spin orientations. In Dirac's interpretation (which remains a handy mnemonic device) $\psi^{(3)}$ is a *negative*-energy *electron* state with spin *up*, whose *absence* (a 'hole' in the 'sea') behaves as a *positive*-energy *positron* with spin *down* [2].

† Here $k\cdot x = k_\mu x^\mu = k^0ct - \mathbf{k}\cdot\mathbf{r}$, so the real part of the exponential is $\cos(k^0ct - \mathbf{k}\cdot\mathbf{r})$, which represents a sinusoidal plane wave of (angular) frequency $\omega = ck^0$ and wavelength $\lambda = 2\pi/|\mathbf{k}|$, propagating in the direction $\mathbf{k}$.

‡ This *looks* right, but if it makes you nervous you can easily check it:

$$\partial_0 e^{-ik\cdot x} = (1/c)\frac{\partial}{\partial t}e^{-ik^0ct+i\mathbf{k}\cdot\mathbf{r}} = -ik^0e^{-ik\cdot x}$$

(and $k^0 = k_0$). Similarly

$$\partial_1 e^{-ik\cdot x} = ik^1e^{-ik\cdot x}$$

(but $k^1 = -k_1$).

where, as before, the subscript A denotes the upper two components, and B stands for the lower two. In order to satisfy Equation 7.33, then, we must have

$$u_A = \frac{1}{k^0 - mc/\hbar}(\mathbf{k}\cdot\boldsymbol{\sigma})u_B \quad \text{and} \quad u_B = \frac{1}{k^0 + mc/\hbar}(\mathbf{k}\cdot\boldsymbol{\sigma})u_A \tag{7.35}$$

Substituting the second of these into the first gives

$$u_A = \frac{1}{(k^0)^2 - (mc/\hbar)^2}(\mathbf{k}\cdot\boldsymbol{\sigma})^2 u_A \tag{7.36}$$

But

$$\mathbf{k}\cdot\boldsymbol{\sigma} = k_x\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} + k_y\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} + k_z\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} = \begin{pmatrix}k_z & (k_x - ik_y)\\ (k_x + ik_y) & -k_z\end{pmatrix} \tag{7.37}$$

so

$$(\mathbf{k}\cdot\boldsymbol{\sigma})^2 = \begin{pmatrix}k_z^2 + (k_x - ik_y)(k_x + ik_y) & k_z(k_x - ik_y) - k_z(k_x - ik_y)\\ k_z(k_x + ik_y) - k_z(k_x + ik_y) & (k_x + ik_y)(k_x - ik_y) + k_z^2\end{pmatrix} = \mathbf{k}^2 1 \tag{7.38}$$

where 1 is the 2×2 unit matrix (written in explicitly, just this once). Thus

$$u_A = \frac{\mathbf{k}^2}{(k^0)^2 - (mc/\hbar)^2}u_A \tag{7.39}$$

and hence*

$$(k^0)^2 - (mc/\hbar)^2 = \mathbf{k}^2, \quad \text{or} \quad k^2 = k^\mu k_\mu = (mc/\hbar)^2 \tag{7.40}$$

In order for $\psi = \exp(-ik\cdot x)u(k)$ to satisfy the Dirac equation, then, $\hbar k^\mu$ must be a four-vector, associated with the particle, whose 'square' is m^2c^2. Of course, we *know* such a quantity: the energy–momentum four-vector. Evidently

$$k^\mu = \pm p^\mu/\hbar \tag{7.41}$$

The positive sign (time dependence $e^{-iEt/\hbar}$) is associated with particle states, and the negative sign (time dependence $e^{+iEt/\hbar}$) with antiparticle states.

Returning to Equation 7.35 (and using Equation 7.37), it is a simple matter to construct four independent solutions to the Dirac equation:

$$\text{(1) Pick } u_A = \begin{pmatrix}1\\0\end{pmatrix} : u_B = \frac{\mathbf{p}\cdot\boldsymbol{\sigma}}{p^0 + mc}\begin{pmatrix}1\\0\end{pmatrix} = \frac{c}{E + mc^2}\begin{pmatrix}p_z\\ p_x + ip_y\end{pmatrix}$$

* Equation 7.39 would also allow $u_A = 0$ as a solution. However, the same argument, starting with Equation 7.35 but inserting the first into the second, yields Equation 7.39 with u_B in place of u_A. So, unless u_A and u_B are *both* zero (in which case we have no solution at all) Equation 7.40 must hold.

$$(2)\ \text{Pick } u_A = \begin{pmatrix} 0 \\ 1 \end{pmatrix} : u_B = \frac{\mathbf{p}\cdot\boldsymbol{\sigma}}{p^0 + mc}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{c}{E + mc^2}\begin{pmatrix} p_x - ip_y \\ -p_z \end{pmatrix}$$

$$(3)\ \text{Pick } u_B = \begin{pmatrix} 1 \\ 0 \end{pmatrix} : u_A = \frac{\mathbf{p}\cdot\boldsymbol{\sigma}}{p^0 + mc}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{c}{E + mc^2}\begin{pmatrix} p_z \\ p_x + ip_y \end{pmatrix}$$

$$(4)\ \text{Pick } u_B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} : u_A = \frac{\mathbf{p}\cdot\boldsymbol{\sigma}}{p^0 + mc}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{c}{E + mc^2}\begin{pmatrix} p_x - ip_y \\ -p_z \end{pmatrix} \tag{7.42}$$

For (1) and (2) we were obliged to use the plus sign in Equation 7.41 – otherwise u_B blows up as $\mathbf{p} \to 0$; these are *particle* solutions. For (3) and (4) we must use the minus sign; these are *anti*particle states.

A convenient normalization for these spinors is*

$$u^\dagger u = 2E/c \tag{7.43}$$

Here the dagger signifies the transpose conjugate (or *Hermitian conjugate*):

$$u = \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} \Rightarrow u^\dagger = (\alpha^* \beta^* \gamma^* \delta^*)$$

so

$$u^\dagger u = |\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 \tag{7.44}$$

With the resulting normalization factor (Problem 7.3)

$$N \equiv \sqrt{(E + mc^2)/c} \tag{7.45}$$

the four canonical solutions become:

$$u^{(1)} = N\begin{pmatrix} 1 \\ 0 \\ \dfrac{c(p_z)}{E + mc^2} \\ \dfrac{c(p_x + ip_y)}{E + mc^2} \end{pmatrix}, \quad u^{(2)} = N\begin{pmatrix} 0 \\ 1 \\ \dfrac{c(p_x - ip_y)}{E + mc^2} \\ \dfrac{c(-p_z)}{E + mc^2} \end{pmatrix} \tag{7.46}$$

* Notice that any multiple of u is still a solution to Equation 7.33; normalization merely fixes the overall constant. Actually, there are at least three different conventions in the literature: $u^\dagger u = 2E/c$ (Halzen and Martin), $u^\dagger u = E/mc^2$ (Bjorken and Drell), $u^\dagger u = 1$ (Bogoliubov and Shirkov). In this one instance I depart from Bjorken and Drell, whose choice introduces spurious difficulties when $m \to 0$.

$$v^{(1)} = N \begin{pmatrix} \dfrac{c(p_x - ip_y)}{E + mc^2} \\ \dfrac{c(-p_z)}{E + mc^2} \\ 0 \\ 1 \end{pmatrix}, \quad v^{(2)} = -N \begin{pmatrix} \dfrac{c(p_z)}{E + mc^2} \\ \dfrac{c(p_x + ip_y)}{E + mc^2} \\ 1 \\ 0 \end{pmatrix} \tag{7.47}$$

$$\psi = ae^{-ip\cdot x/\hbar}u \quad \text{(particles)}, \qquad \psi = ae^{ip\cdot x/\hbar}v \quad \text{(antiparticles)} \tag{7.48}$$

It is customary, from here on, to use the letter v for antiparticles (and to include a minus sign in $v^{(2)}$), as indicated. Notice that whereas particle states satisfy the *momentum space Dirac equation* (see Equation 7.33) in the form

$$(\gamma^\mu p_\mu - mc)u = 0 \tag{7.49}$$

antiparticles (v's) satisfy:

$$(\gamma^\mu p_\mu + mc)v = 0 \tag{7.50}$$

You might guess that $u^{(1)}$ describes an electron with spin up, $u^{(2)}$ an electron with spin down, $v^{(1)}$ a positron with spin up, and $v^{(2)}$ a positron with spin down,* but this is not quite the case. For Dirac particles the spin matrices (generalizing Equation 4.21) are

$$\mathbf{S} = \frac{\hbar}{2}\mathbf{\Sigma}, \quad \text{with } \mathbf{\Sigma} \equiv \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix} \tag{7.51}$$

and it's easy to check that $u^{(1)}$, for instance, is *not* an eigenstate of Σ_z. However, if we orient the z axis so that it points along the direction of motion (in which case $p_x = p_y = 0$) then $u^{(1)}$, $u^{(2)}$, $v^{(1)}$, and $v^{(2)}$ *are* eigenspinors of S_z; $u^{(1)}$ and $v^{(1)}$ are spin up, $u^{(2)}$ and $v^{(2)}$ are spin down† (Problem 7.6).

Incidentally, plane waves are, of course, rather special solutions to the Dirac equation. They are the ones of interest to us, however, because they describe particles with specified energies and momenta, and in a typical experiment these are the parameters we control and measure.

* See the footnote to Equation 7.30 for positron spin orientations.

† As a matter of fact, it is *impossible* to construct plane-wave solutions to the Dirac equation and are, at the same time, eigenstates of S_z (except in the special case $\mathbf{p} = p_z\hat{\mathbf{z}}$). The reason is that $\mathbf{S}$ by itself is not a conserved quantity; only the *total* angular momentum, $\mathbf{L} + \mathbf{S}$, is conserved here (see Problem 7.8). It is possible to construct eigenstates of *helicity*, $\mathbf{\Sigma} \cdot \hat{\mathbf{p}}$ (there's no *orbital* angular momentum about the direction of motion), but these are rather cumbersome (see Problem 7.7), and in practice it is easier to work with the spinors in Equations 7.46 and 7.47, even though their physical interpretation is not so clean. All that *really* matters is that we have a *complete* set of solutions.

7.3 Bilinear Covariants

I mentioned in Section 7.1 that the components of a Dirac spinor do not transform as a four-vector, when you go from one inertial system to another. How, then, *do* they transform? I shall not work it out here (you get to do it, in Problem 7.11), but merely quote the result: if you go to a system moving with speed v in the x direction

$$\psi \to \psi' = S\psi \tag{7.52}$$

where S is the following 4×4 matrix:

$$S = a_+ + a_-\gamma^0\gamma^1 = \begin{pmatrix} a_+ & a_-\sigma_1 \\ a_-\sigma_1 & a_+ \end{pmatrix} = \begin{pmatrix} a_+ & 0 & 0 & a_- \\ 0 & a_+ & a_- & 0 \\ 0 & a_- & a_+ & 0 \\ a_- & 0 & 0 & a_+ \end{pmatrix} \tag{7.53}$$

with

$$a_\pm = \pm\sqrt{\tfrac{1}{2}(\gamma \pm 1)} \tag{7.54}$$

and $\gamma = 1/\sqrt{1 - v^2/c^2}$, as usual.

Suppose we want to construct a scalar quantity out of a spinor ψ. It would be reasonable to try the expression

$$\psi^\dagger\psi = (\psi_1^*\psi_2^*\psi_3^*\psi_4^*)\begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = |\psi_1|^2 + |\psi_2|^2 + |\psi_3|^2 + |\psi_4|^2 \tag{7.55}$$

Unfortunately, this is *not* invariant, as you can check by applying the transformation rule:*

$$(\psi^\dagger\psi)' = (\psi')^\dagger\psi' = \psi^\dagger S^\dagger S\psi \neq (\psi^\dagger\psi) \tag{7.56}$$

* Note that the transpose of a product is the product of the transposes *in reverse order*:

$$(\widetilde{AB})_{ij} = (AB)_{ji} = \sum_k A_{jk}B_{ki}$$

$$= \sum_k \tilde{B}_{ik}\tilde{A}_{kj} = (\tilde{B}\tilde{A})_{ij}$$

The same goes for the Hermitian conjugate:

$$(AB)^\dagger = B^\dagger A^\dagger$$

In fact (Problem 7.13):

$$S^\dagger S = S^2 = \gamma \begin{pmatrix} 1 & -(v/c)\sigma_1 \\ -(v/c)\sigma_1 & 1 \end{pmatrix} \neq 1 \tag{7.57}$$

Of course, the sum of the squares of the elements of a four-vector is not invariant either; we need minus signs for the spatial components (Equation 3.12). With a little trial and error you will discover that in the case of spinors we need minus signs for the third and fourth components. Just as we introduced *covariant* four-vectors to keep track of the signs in Chapter 3, we now introduce the *adjoint* spinor:

$$\overline{\psi} \equiv \psi^\dagger \gamma^0 = (\psi_1^* \ \psi_2^* - \psi_3^* - \psi_4^*) \tag{7.58}$$

I claim that the quantity

$$\overline{\psi}\psi = \psi^\dagger \gamma^0 \psi = |\psi_1|^2 + |\psi_2|^2 - |\psi_3|^2 - |\psi_4|^2 \tag{7.59}$$

is a relativistic invariant. For $S^\dagger \gamma^0 S = \gamma^0$ (Problem 7.13), and hence

$$(\overline{\psi}\psi)' = (\psi')^\dagger \gamma^0 \psi' = \psi^\dagger S^\dagger \gamma^0 S \psi = \psi^\dagger \gamma^0 \psi = \overline{\psi}\psi \tag{7.60}$$

In Chapter 4 we learned to distinguish scalars and *pseudo*scalars, according to their behavior under the parity transformation, P: $(x, y, z) \to (-x, -y, -z)$. Pseudoscalars change sign; scalars do not. It is natural to ask whether $\overline{\psi}\psi$ is the former type, or the latter. First, we need to know how Dirac spinors transform under P. Again, I won't derive it, but simply quote the result (Problem 7.12):*

$$\psi \to \psi' = \gamma^0 \psi \tag{7.61}$$

It follows that

$$(\overline{\psi}\psi)' = (\psi')^\dagger \gamma^0 \psi' = \psi^\dagger (\gamma^0)^\dagger \gamma^0 \gamma^0 \psi = \psi^\dagger \gamma^0 \psi = \overline{\psi}\psi \tag{7.62}$$

so $(\overline{\psi}\psi)$ is *invariant* under P – it's a 'true' scalar. But we can *also* make a *pseudo*scalar out of ψ:

$$\overline{\psi}\gamma^5 \psi \tag{7.63}$$

where

$$\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{7.64}$$

* The *sign* in Equation 7.61 is pure convention; $-\gamma^0\psi$ would do just as well.

I'll let you check that it is Lorentz invariant (Problem 7.14). As for its behavior under parity

$$(\overline{\psi}\gamma^5\psi)' = (\psi')^\dagger\gamma^0\gamma^5\psi' = \psi^\dagger\gamma^0\gamma^0\gamma^5\gamma^0\psi = \psi^\dagger\gamma^5\gamma^0\psi \tag{7.65}$$

(I used the fact that $(\gamma^0)^2 = 1$ in the last step.) Now, the γ^0 is on the 'wrong side' of the γ^5, but we can 'pull it through' by noting that it *anti*commutes with γ^1, γ^2, and γ^3 (Equation 7.15) and *commutes* (of course) with itself ($\gamma^3\gamma^0 = -\gamma^0\gamma^3, \gamma^2\gamma^0 = -\gamma^0\gamma^2, \gamma^1\gamma^0 = -\gamma^0\gamma^1, \gamma^0\gamma^0 = \gamma^0\gamma^0$), so

$$\gamma^5\gamma^0 = i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^0 = (-1)^3\gamma^0(i\gamma^0\gamma^1\gamma^2\gamma^3) = -\gamma^0\gamma^5$$

By the same token, γ^5 anticommutes with all the other γ matrices:

$$\{\gamma^\mu, \gamma^5\} = 0 \tag{7.66}$$

At any rate

$$(\overline{\psi}\gamma^5\psi)' = -\psi^\dagger\gamma^0\gamma^5\psi = -(\overline{\psi}\gamma^5\psi) \tag{7.67}$$

so it's a *pseudo*scalar.

All told, there are 16 products of the form $\psi_i^*\psi_j$ (taking one component from ψ^* and one from ψ), since i and j run from 1 to 4. These 16 products can be assembled in various linear combinations to construct quantities with distinct transformation behavior, as follows:

$$\left\{\begin{array}{lll} (1) & \overline{\psi}\psi = \text{scalar} & \text{(one component)} \\ (2) & \overline{\psi}\gamma^5\psi = \text{pseudoscalar} & \text{(one component)} \\ (3) & \overline{\psi}\gamma^\mu\psi = \text{vector} & \text{(four components)} \\ (4) & \overline{\psi}\gamma^\mu\gamma^5\psi = \text{pseudovector} & \text{(four components)} \\ (5) & \overline{\psi}\sigma^{\mu\nu}\psi = \text{antisymmetric tensor} & \text{(six components)} \end{array}\right\} \tag{7.68}$$

where

$$\sigma^{\mu\nu} \equiv \frac{i}{2}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu) \tag{7.69}$$

This gives 16 terms, so it's all we can hope to make. You cannot, for example, construct a *symmetric* tensor bilinear in ψ^* and ψ, and if you're looking for a vector, $\overline{\psi}\gamma^\mu\psi$ is the only candidate.* (Another way to think of it is this: $1, \gamma^5, \gamma^\mu, \gamma^\mu\gamma^5$, and $\sigma^{\mu\nu}$ constitute a 'basis' for the space of all 4×4 matrices; *any* 4×4 matrix can be written as a linear combination of these 16. In particular, if you ever encounter a

* Notice that $\overline{\psi}\gamma^0\psi = \psi^\dagger\gamma^0\gamma^0\psi = \psi^\dagger\psi$, so $\psi^\dagger\psi$ is actually the zeroth component of a four-vector. That's why the normalization convention (Equation 7.43), which no doubt looked peculiar at the time, is actually very sensible. By normalizing $u^\dagger u$ to the zeroth component of the four-vector p^μ, we obtain a relativistically 'natural' convention (see Problem 7.16).

product of five γ matrices, say, you may be sure that it can be reduced to a product of no more than two.)

Pause a moment to admire the ingenious notation in Equation 7.68. The tensorial character of the *bilinear covariants*, and even their behavior under parity, is indicated at a glance: $\overline{\psi}\gamma^\mu\psi$ *looks* like a four-vector, and it *is* a four-vector. But γ^μ by itself is certainly *not* a four-vector; it's a collection of four fixed matrices (Equation 7.17); *they* don't change when you go to a different inertial system – it's ψ that changes, and in just such a way as to give the whole 'sandwich' the tensorial taste of the jam inside.

7.4 The Photon

In classical electrodynamics, the electric and magnetic fields (**E** and **B**) produced by a charge density ρ and a current density **J** are determined by Maxwell's equations:*

$$\left\{\begin{array}{llll} \text{(i)} & \nabla\cdot\mathbf{E}=4\pi\rho & \text{(iii)} & \nabla\cdot\mathbf{B}=0 \\ \text{(ii)} & \nabla\times\mathbf{E}+\dfrac{1}{c}\dfrac{\partial\mathbf{B}}{\partial t}=0 & \text{(iv)} & \nabla\times\mathbf{B}-\dfrac{1}{c}\dfrac{\partial\mathbf{E}}{\partial t}=\dfrac{4\pi}{c}\mathbf{J} \end{array}\right\} \tag{7.70}$$

In relativistic notation, **E** and **B** together form an antisymmetric second-rank tensor, the 'field strength tensor', $F^{\mu\nu}$:

$$F^{\mu\nu}=\begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \tag{7.71}$$

(that is, $F^{01}=-E_x$, $F^{12}=-B_z$, etc.), while ρ and **J** constitute a four-vector:

$$J^\mu=(c\rho,\mathbf{J}) \tag{7.72}$$

The inhomogeneous Maxwell equations, (i) and (iv) in Equation 7.70, can be written more neatly in tensor notation (Problem 7.20)

$$\partial_\mu F^{\mu\nu}=\frac{4\pi}{c}J^\nu \tag{7.73}$$

From the antisymmetry of $F^{\mu\nu}$ ($F^{\nu\mu}=-F^{\mu\nu}$) it follows (Problem 7.20) that J^μ is divergenceless:

$$\partial_\mu J^\mu=0 \tag{7.74}$$

Or, in three-vector notation, $\nabla\cdot\mathbf{J}=-\partial\rho/\partial t$; this is the 'continuity equation', expressing local conservation of charge (Problem 7.21).

* This section presupposes some familiarity with classical electrodynamics; it is designed to make the description of photons in quantum electrodynamics more plausible. As always, I use Gaussian cgs units.

As for the homogeneous Maxwell equations, (iii) in Equation 7.70 is equivalent to the statement that **B** can be written as the curl of a vector potential, **A**:

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{7.75}$$

With this, (ii) becomes

$$\nabla \times \left(\mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \right) = 0 \tag{7.76}$$

which is equivalent to the statement that $\mathbf{E} + (1/c)(\partial \mathbf{A}/\partial t)$ can be written as the gradient of a scalar potential, V:

$$\mathbf{E} = -\nabla V - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \tag{7.77}$$

In relativistic notation, Equations 7.75 and 7.77 become

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{7.78}$$

where

$$A^\mu = (V, \mathbf{A}) \tag{7.79}$$

In terms of this four-vector potential, the inhomogeneous Maxwell Equations (7.73) read:

$$\partial_\mu \partial^\mu A^\nu - \partial^\nu (\partial_\mu A^\mu) = \frac{4\pi}{c} J^\nu \tag{7.80}$$

In classical electrodynamics the *fields* are the physical entities; the potentials are simply useful mathematical constructs. The *virtue* of the potential formulation is that it *automatically* takes care of the homogeneous Maxwell equations: given Equations 7.75 and 7.77, (ii) and (iii) in Equation 7.70 follow immediately, no matter what V and **A** might be. This leaves us only the inhomogeneous Equation (7.80) to worry about. The *defect* of the potential formulation is that V and **A** are not uniquely determined. Indeed, it is clear from Equation 7.78 that new potentials

$$A'_\mu = A_\mu + \partial_\mu \lambda \tag{7.81}$$

(where λ is any function of position and time) would do just as well, since $\partial^\mu A^{\nu'} - \partial^\nu A^{\mu'} = \partial^\mu A^\nu - \partial^\nu A^\mu$. Such a change of potentials, which has no effect on the fields, is called a *gauge transformation*. We can exploit this gauge freedom to impose an extra constraint on the potential [3]:

$$\partial_\mu A^\mu = 0 \tag{7.82}$$

This is called the *Lorentz condition*; with it Maxwell's equations (7.80) simplify still further:

$$\Box A^\mu = \frac{4\pi}{c} J^\mu \tag{7.83}$$

Here

$$\Box \equiv \partial^\mu \partial_\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 \tag{7.84}$$

is the relativistic extension of the Laplacian (∇^2); it is called the *d'Alembertian*.

Even the Lorentz condition, however, does not uniquely specify A^μ. Further gauge transformations are possible, without disturbing Equation 7.82, provided that the gauge function λ satisfies the wave equation:

$$\Box \lambda = 0 \tag{7.85}$$

Unfortunately, there is no clean way to eliminate the residual ambiguity in A^μ, and one must choose either to live with the indeterminacy, which means carrying along spurious degrees of freedom, or to impose an additional constraint, which spoils the manifest Lorentz covariance of the theory. Both approaches have been used in formulating QED; we shall follow the latter course. In empty space, where $J^\mu = 0$, we pick (see Problem 7.22)

$$A^0 = 0 \tag{7.86}$$

The Lorentz condition then reads

$$\nabla \cdot \mathbf{A} = 0 \tag{7.87}$$

This choice (the *Coulomb gauge*) is attractively simple, but by selecting one component (A^0) for special treatment, it ties us down to a particular inertial system (or rather, it obliges us to perform a gauge transformation in conjunction with every Lorentz transformation, in order to restore the Coulomb gauge condition). In practice, this is very seldom a problem, but it is aesthetically displeasing.

In QED, A^μ becomes the wave function of the photon. The *free* photon satisfies Equation 7.83 with $J^\mu = 0$,

$$\Box A^\mu = 0 \tag{7.88}$$

which we recognize in this context as the Klein–Gordon Equation (7.9) for a massless particle. As in the case of the Dirac equation, we look for plane-wave solutions with four-momentum $p = (E/c, \mathbf{p})$:

$$A^\mu(x) = a e^{-(i/\hbar)p\cdot x} \epsilon^\mu(p) \tag{7.89}$$

Here ϵ^μ is the *polarization vector* – it characterizes the spin of the photon – and a is a normalization factor. Substituting Equation 7.89 into Equation 7.88, we obtain

a constraint on p^μ:

$$p^\mu p_\mu = 0, \quad \text{or} \quad E = |\mathbf{p}|c \tag{7.90}$$

which is as it should be for a massless particle.

Meanwhile, ϵ^μ has four components, but they are not all independent. The Lorentz condition (Equation 7.82) requires that

$$p^\mu \epsilon_\mu = 0 \tag{7.91}$$

In the Coulomb gauge, moreover,

$$\epsilon^0 = 0, \quad \text{so} \quad \boldsymbol{\epsilon} \cdot \mathbf{p} = 0 \tag{7.92}$$

which is to say that the polarization three-vector ($\boldsymbol{\epsilon}$) is perpendicular to the direction of propagation; we say that a free photon is *transversely polarized.** Now, there are *two* linearly independent three-vectors perpendicular to $\mathbf{p}$; for example, if $\mathbf{p}$ points in the z direction, we might choose

$$\boldsymbol{\epsilon}^{(1)} = (1, 0, 0), \quad \boldsymbol{\epsilon}^{(2)} = (0, 1, 0) \tag{7.93}$$

Instead of *four* independent solutions for a given momentum (too many, for a particle of spin 1), we are left with only *two*. That sounds like too *few* – shouldn't the photon have *three* spin states? The answer is *no*: a *massive* particle of spin s admits $2s + 1$ different spin orientations, but a *massless* particle has only two, regardless of its spin (except for $s = 0$, which has only *one*). Along its direction of motion it can only have $m_s = +s$ or $m_s = -s$; its helicity, in other words, can only be $+1$ or -1.†

7.5 The Feynman Rules for QED

In Section 7.2 we found that free electrons and positrons of momentum $p = (E/c, \mathbf{p})$, with $E = \sqrt{m^2c^4 + \mathbf{p}^2c^2}$, are represented by the wave functions‡

* This corresponds to the fact that electromagnetic waves are transverse.

† Photon states with $m_s = \pm 1$ correspond to right- and left-circular polarization; the respective polarization vectors are $\boldsymbol{\epsilon}_\pm = \mp(\boldsymbol{\epsilon}^{(1)} \pm i\boldsymbol{\epsilon}^{(2)})/\sqrt{2}$. Notice that it was by specifying a particular gauge that we eliminated the nonphysical ($m_s = 0$) solution. If we were to follow a 'covariant' approach, in which we avoid imposing the Coulomb gauge condition, longitudinal free photons would be present in the theory. But these 'ghosts' decouple from everything else, and they do not affect the final results.

‡ For reference, I begin with a summary of the essential results from earlier sections. I speak of 'electrons' and 'positrons', but they could as well be μ^- and μ^+, or τ^- and τ^+, or (with the appropriate electric charges) quarks and antiquarks – in short, any point charges of spin $\frac{1}{2}$.

Electrons **Positrons**

$$\psi(x) = ae^{-(i/\hbar)p\cdot x}u^{(s)}(p) \qquad \psi(x) = ae^{(i/\hbar)p\cdot x}v^{(s)}(p) \tag{7.94}$$

where $s = 1, 2$ for the two spin states. The spinors $u^{(s)}$ and $v^{(s)}$ satisfy the momentum space Dirac equation(s):

$$(\gamma^\mu p_\mu - mc)u = 0 \qquad (\gamma^\mu p_\mu + mc)v = 0 \tag{7.95}$$

their adjoints, $\overline{u} = u^\dagger\gamma^0, \overline{v} = v^\dagger\gamma^0$, satisfy

$$\overline{u}(\gamma^\mu p_\mu - mc) = 0 \qquad \overline{v}(\gamma^\mu p_\mu + mc) = 0 \tag{7.96}$$

They are orthogonal,

$$\overline{u}^{(1)}u^{(2)} = 0 \qquad \overline{v}^{(1)}v^{(2)} = 0 \tag{7.97}$$

normalized,

$$\overline{u}u = 2mc \qquad \overline{v}v = -2mc \tag{7.98}$$

and *complete*, in the sense that

$$\sum_{s=1,2} u^{(s)}\overline{u}^{(s)} = (\gamma^\mu p_\mu + mc) \qquad \sum_{s=1,2} v^{(s)}\overline{v}^{(s)} = (\gamma^\mu p_\mu - mc) \tag{7.99}$$

(see Problem 7.24). A convenient explicit set $\{u^{(1)}, u^{(2)}, v^{(1)}, v^{(2)}\}$ is given in Equations 7.46 and 7.47. Ordinarily, we'll be averaging over electron and positron spins, and in that case it doesn't matter that these are not pure spin up and spin down – all we really need is completeness. For the occasional problem in which the spins are specified, we must, of course, use the spinors appropriate to the case at hand.

Meanwhile, a free photon of momentum $p = (E/c, \mathbf{p})$, with $E = |\mathbf{p}|c$, is represented by the wave function

Photons

$$A_\mu(x) = ae^{-(i/\hbar)p\cdot x}\epsilon_\mu^{(s)} \tag{7.100}$$

where $s = 1, 2$ for the two spin states (polarizations). The polarization vectors $\epsilon_\mu^{(s)}$ satisfy the momentum space Lorentz condition:

$$p^\mu\epsilon_\mu = 0 \tag{7.101}$$

They are orthogonal, in the sense that

$$\epsilon_\mu^{(1)*}\epsilon^{(2)\mu} = 0 \tag{7.102}$$

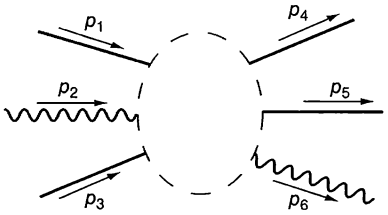


Fig. 7.1 A generic QED diagram, with external lines labeled. (Internal lines not shown.)

and normalized

$$\epsilon^{\mu *}\epsilon_\mu = -1 \tag{7.103}$$

In the Coulomb gauge

$$\epsilon^0 = 0, \quad \boldsymbol{\epsilon} \cdot \mathbf{p} = 0 \tag{7.104}$$

and the polarization three-vectors obey the *completeness relation* (Problem 7.25)

$$\sum_{s=1,2} \epsilon_i^{(s)} \epsilon_j^{(s)*} = \delta_{ij} - \hat{p}_i \hat{p}_j \tag{7.105}$$

A convenient explicit pair $(\epsilon^{(1)}, \epsilon^{(2)})$ is given in Equation 7.93.

To calculate the amplitude, $\mathscr{M}$, associated with a particular Feynman diagram, proceed as follows:

Feynman Rules

1. *Notation:* To each *external* line associate a momentum $p_1, p_2, \ldots, p_n$, and draw an arrow next to the line, indicating the positive direction (forward in time).* To each *internal* line associate a momentum $q_1, q_2, \ldots$; again draw an arrow next to the line indicating the positive direction (arbitrarily assigned). See Figure 7.1.
2. *External lines:* External lines contribute factors as follows:

$$\begin{array}{ll} \textit{Electrons}: & \begin{cases} \text{Incoming}(\rightarrow\!\bullet) : u \\ \text{Outgoing}(\bullet\!\rightarrow) : \bar{u} \end{cases} \\ \textit{Positrons}: & \begin{cases} \text{Incoming}(\leftarrow\!\bullet) : \bar{v} \\ \text{Outgoing}(\bullet\!\leftarrow) : v \end{cases} \\ \textit{Photons}: & \begin{cases} \text{Incoming}(\sim\!\bullet) : \epsilon_\mu \\ \text{Outgoing}(\bullet\!\sim) : \epsilon_\mu{}^* \end{cases} \end{array}$$

* For a fermion, of course, there will *already* be an arrow *on* the line, telling us whether it is an electron or a positron. The two arrows have nothing to do with one another; they may or may not point in the same direction.

3. *Vertex factors:* Each vertex contributes a factor

$$ig_e\gamma^\mu$$

The dimensionless coupling constant g_e is related to the charge of the electron: $g_e = e\sqrt{4\pi/\hbar c} = \sqrt{4\pi\alpha}$.*

4. *Propagators:* Each internal line contributes a factor as follows:

Electrons and positrons: $$\frac{i(\gamma^\mu q_\mu + mc)}{q^2 - m^2c^2}$$

Photons: $$\frac{-ig_{\mu\nu}}{q^2}$$

5. *Conservation of energy and momentum:* For each vertex, write a delta function of the form

$$(2\pi)^4\delta^4(k_1 + k_2 + k_3)$$

where the k's are the three four-momenta coming *into* the vertex (if an arrow leads *outward*, then k is *minus* the four-momentum of that line).

6. *Integrate over internal momenta:* For each internal momentum q, write a factor

$$\frac{d^4q}{(2\pi)^4}$$

and integrate.

7. *Cancel the delta function:* The result will include a factor

$$(2\pi)^4\delta^4(p_1 + p_2 + \cdots - p_n)$$

corresponding to overall energy–momentum conservation. Cancel this factor, and multiply by i; what remains is $\mathscr{M}$.

It is critically important that the pieces be assembled in the correct order – otherwise the matrix multiplications will be gibberish. The safest procedure is

* In Heaviside–Lorentz units, with $\hbar$ and c set equal to 1, g_e is the charge of the positron, and hence is written 'e' in most texts. In this book I use Gaussian units, and keep all factors of $\hbar$ and c. The easiest way to avoid trouble over units is to express all results in terms of the dimensionless number α. In writing the Feynman rules for QED I assume we are dealing with electrons and positrons. In *general*, the QED coupling constant is $-q\sqrt{4\pi/\hbar c}$, where q is the charge of the *particle* (as opposed to the *anti*particle). For electrons, $q = -e$, but for 'up' quarks, say, $q = \frac{2}{3}e$.

to *track each fermion line backward through the diagram.* Start (for example) with an outgoing electron line, and follow the arrow (the one *on* the line) back until it emerges, either as an incoming electron or as an outgoing positron, writing down the various line factors, vertex factors, and propagators, left to right, as you encounter them. Each fermion line produces a 'sandwich', of the form adjoint spinor, 4×4 matrix, spinor (row · matrix · column = number). Meanwhile, each vertex carries a contravariant vector index (μ, ν, λ, ...), which contracts with the covariant index of the associated photon line or propagator. (Don't worry: all of this will make much better sense when we work some examples, but I wanted to prescribe the ritual, for future reference.)

As before, the idea is to draw all the diagrams contributing to the process in question (up to the desired order), calculate the amplitude ($\mathscr{M}$) for each one, and add them up to get the *total* amplitude, which is then inserted into the appropriate Golden Rule for the decay rate or the scattering cross section, as the case may be. There's one new twist that occasionally arises: the antisymmetrization of fermion wave functions requires that we insert a minus sign in combining amplitudes that differ only in the interchange of two identical external fermions. It doesn't matter *which* diagram you associate the minus sign with, since the total will be squared eventually anyway; but there must be a *relative* minus sign between them:

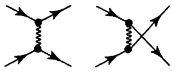
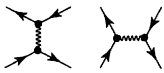
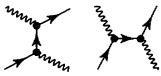
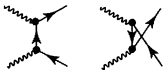
8. *Antisymmetrization:* Include a minus sign between diagrams that differ only in the interchange of two incoming (or outgoing) electrons (or positrons), or of an incoming electron with an outgoing positron (or vice versa).

7.6 Examples

We are now in a position to reproduce many of the classic calculations in quantum electrodynamics. Just so you don't get lost in the details, let me begin by giving you a catalog of the most important processes (Table 7.1). The simplest case is electron–muon scattering, for here only one diagram contributes in second order.*

* It doesn't have to be an e and a μ, of course. Any spin-$\frac{1}{2}$ point charges would do (e and τ, for instance, or μ and τ, or electron and quark, etc.), as long as you put in the correct masses and charges. As a matter of fact, most books use electron–*proton* scattering as the canonical example, but that is actually a rather *in*appropriate choice, since the proton is a composite structure, not a point particle. Still, to the extent that the internal structure of the proton can be ignored, it is not a bad approximation (rather like treating the sun as a point mass in the theory of the solar system). If the 'muon' is much heavier than the 'electron', we have *Mott scattering*; if, moreover, the 'electron' is nonrelativistic, we get *Rutherford scattering*, for which QED reproduces precisely the classical formula (Example 6.4).

Table 7.1 Catalog of basic quantum electrodynamic processes.

Second-order processes	
Elastic	
	Electron–muon scattering($e + \mu \to e + \mu$) (Mott scattering($M \gg m$) $\Rightarrow$ Rutherford scattering($v \ll c$))
	Electron–electron scattering($e^- + e^- \to e^- + e^-$) (Møller scattering)
	Electron–positron scattering($e^- + e^+ \to e^- + e^+$) (Bhabha scattering)
	Compton scattering($\gamma + e^- \to \gamma + e^-$)
Inelastic	
	Pair annihilation($e^- + e^+ \to \gamma + \gamma$)
	Pair production($\gamma + \gamma \to e^- + e^+$)
Most important third-order process	
	$\Rightarrow$ Anomalous magnetic moment of electron

Example 7.1 Electron–Muon Scattering Walking 'backward' along each fermion line (Figure 7.2), and applying the Feynman rules as we go:

$$(2\pi)^4 \int [\overline{u}^{(s_3)}(p_3)(ig_e\gamma^\mu)u^{(s_1)}(p_1)]\frac{-ig_{\mu\nu}}{q^2}[\overline{u}^{(s_4)}(p_4)(ig_e\gamma^\nu)u^{(s_2)}(p_2)]$$
$$\times\, \delta^4(p_1 - p_3 - q)\delta^4(p_2 + q - p_4)\, \mathrm{d}^4 q$$

Notice how the space-time indices on the photon propagator contract with those of the vertex factors at either end of the photon line. Carrying out the (trivial) q integration, and dropping the overall delta function, we find

$$\mathscr{M} = -\frac{g_e^2}{(p_1 - p_3)^2}[\overline{u}^{(s_3)}(p_3)\gamma^\mu u^{(s_1)}(p_1)][\overline{u}^{(s_4)}(p_4)\gamma_\mu u^{(s_2)}(p_2)] \tag{7.106}$$

In spite of its complicated appearance, with four spinors and eight γ matrices, this is just a *number*, which you can work out once the spins are specified (see Problem 7.26).

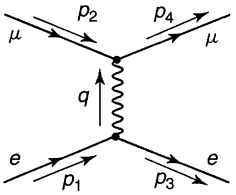


Fig. 7.2 Electron–muon scattering.

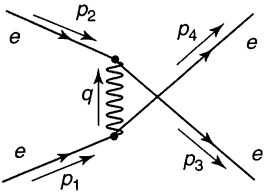


Fig. 7.3 'Twisted' diagram for electron–electron scattering.

Example 7.2 Electron–Electron Scattering In this case there is a second diagram, in which the electron that emerges with momentum p_3 and spin s_3 comes from the p_2, s_2 electron, instead of the p_1, s_1 electron (Figure 7.3). We can obtain this amplitude from Equation 7.106 simply by the replacement $p_3, s_3 \leftrightarrow p_4, s_4$. According to Rule 8, the two diagrams are to be *subtracted*, so the total amplitude is

$$\begin{aligned}\mathscr{M} = &-\frac{g_e^2}{(p_1 - p_3)^2}[\overline{u}(3)\gamma^\mu u(1)][\overline{u}(4)\gamma_\mu u(2)] \\ &+ \frac{g_e^2}{(p_1 - p_4)^2}[\overline{u}(4)\gamma^\mu u(1)][\overline{u}(3)\gamma_\mu u(2)] \qquad (7.107)\end{aligned}$$

(Note the transparent shorthand I have adopted to label the spinors.)

Example 7.3 Electron–Positron Scattering Again, there are two diagrams.* The first is similar to the electron–muon diagram (Figure 7.4):

$$\begin{aligned}(2\pi)^4 \int [\overline{u}(3)(ig_e\gamma^\mu)u(1)]\frac{-ig_{\mu\nu}}{q^2}[\overline{v}(2)(ig_e\gamma^\nu)v(4)] \\ \times\, \delta^4(p_1 - p_3 - q)\delta^4(p_2 + q - p_4)\, \mathrm{d}^4 q\end{aligned}$$

Notice that 'proceeding backwards' along an *anti*particle line means working *forward* in time; the order is always *adjoint/matrix/spinor*. The amplitude for this diagram

* The fact that there are *two* diagrams for electron–electron and electron–positron scattering, but only *one* for electron–muon scattering, would appear offhand to be inconsistent with the classical limit. After all, Coulomb's law says that the force of attraction or repulsion between two particles depends only on their charges, not on whether they happen to be identical (or antiparticles of one another); in the nonrelativistic limit, then, we should get the same answer whether we use the electron–muon formula or the electron–electron formula. The *amplitudes*, it is true, are *not* the same, but the cross section formula (Equation 6.34) carries a factor of S, which is $\frac{1}{2}$ for electron–electron scattering and 1 for electron–muon scattering. For electron–positron scattering, $S = 1$, but the second amplitude (Equation 7.109) is smaller than the first (Equation 7.108) by a factor $(v/c)^2$, so only $\mathscr{M}_1$ contributes, in the nonrelativistic limit.

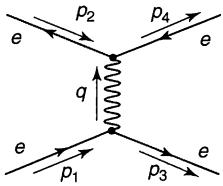


Fig. 7.4 Electron–positron scattering.

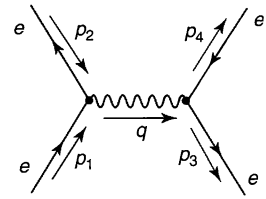


Fig. 7.5 Second diagram contributing to electron–positron scattering.

is thus

$$\mathscr{M}_1 = -\frac{g_e^2}{(p_1 - p_3)^2}[\overline{u}(3)\gamma^\mu u(1)][\overline{v}(2)\gamma_\mu v(4)] \tag{7.108}$$

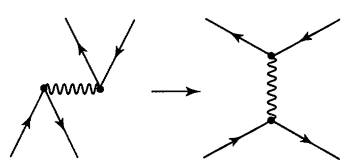
The other diagram represents virtual annihilation of the electron and positron, followed by pair production (Figure 7.5):

$$(2\pi)^4 \int [\overline{u}(3)(ig_e\gamma^\mu)v(4)]\frac{-ig_{\mu\nu}}{q^2}[\overline{v}(2)(ig_e\gamma^\nu)u(1)] \\ \times \delta^4(q - p_3 - p_4)\delta^4(p_1 + p_2 - q)\,\mathrm{d}^4q$$

The amplitude for this diagram is therefore

$$\mathscr{M}_2 = -\frac{g_e^2}{(p_1 + p_2)^2}[\overline{u}(3)\gamma^\mu v(4)][\overline{v}(2)\gamma_\mu u(1)] \tag{7.109}$$

Now, should we *add* these amplitudes, or *subtract* them? Interchanging the incoming positron and the outgoing electron in the second diagram (Figure 7.5), and then redrawing it in a more customary configuration



we recover the first diagram (Figure 7.4). According to Rule 8, then, we need a minus sign:

$$\mathscr{M} = -\frac{g_e^2}{(p_1 - p_3)^2}[\overline{u}(3)\gamma^\mu u(1)][\overline{v}(2)\gamma_\mu v(4)] \\ + \frac{g_e^2}{(p_1 + p_2)^2}[\overline{u}(3)\gamma^\mu v(4)][\overline{v}(2)\gamma_\mu u(1)] \tag{7.110}$$

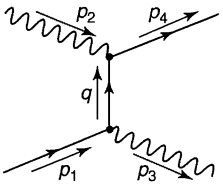


Fig. 7.6 Compton scattering.

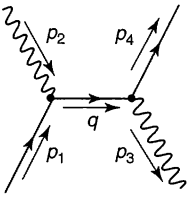


Fig. 7.7 Second diagram for Compton scattering.

Example 7.4 Compton Scattering For an example involving the electron propagator and photon polarization, consider Compton scattering, $\gamma + e \to \gamma + e$. Again there are two diagrams, but they do *not* differ by the interchange of fermions, and the amplitudes *add*. The first diagram (Figure 7.5) yields

$$(2\pi)^4 \int \epsilon_\mu(2) \left[\overline{u}(4)(ig_e\gamma^\mu) \frac{i(\not{q} + mc)}{(q^2 - m^2c^2)} (ig_e\gamma^\nu) u(1) \right] \epsilon_\nu(3)^* \\ \times \delta^4(p_1 - p_3 - q)\delta^4(p_2 + q - p_4)\, \mathrm{d}^4 q$$

Notice that the space-time index on each photon polarization vector is contracted with the index of the γ matrix at the vertex where the photon was created or absorbed. Notice also how the electron propagator fits in as we work our way backward along the electron line. I have introduced here the very convenient 'slash' notation:

$$\not{a} \equiv a^\mu \gamma_\mu \tag{7.111}$$

Evidently, the amplitude associated with Figure 7.6 is*

$$\mathscr{M}_1 = \frac{g_e^2}{(p_1 - p_3)^2 - m^2c^2} [\overline{u}(4)\not{\epsilon}(2)(\not{p}_1 - \not{p}_3 + mc)\not{\epsilon}(3)^* u(1)] \tag{7.112}$$

Meanwhile, the second diagram (Figure 7.7) yields

$$\mathscr{M}_2 = \frac{g_e^2}{(p_1 + p_2)^2 - m^2c^2} [\overline{u}(4)\not{\epsilon}(3)^*(\not{p}_1 + \not{p}_2 + mc)\not{\epsilon}(2) u(1)] \tag{7.113}$$

and the total amplitude is $\mathscr{M} = \mathscr{M}_1 + \mathscr{M}_2$.

7.7 Casimir's Trick

In some experiments the incoming and outgoing electron (or positron) spins are specified, and the photon polarizations are given. If so, the next thing to do is to

* Here and below, $\not{\epsilon}^*$ means $\gamma^\mu(\epsilon_\mu^*)$; the γ-matrix is *not* conjugated.

insert the appropriate spinors and polarization vectors into the expression for $\mathscr{M}$, and compute $|\mathscr{M}|^2$ – the quantity we actually need, to determine cross sections and lifetimes. More often, however, we are not interested in the spins. A typical experiment starts out with a beam of particles whose orientations are random, and simply counts the *number* of particles scattered in a given direction. In this case the relevant cross section is the *average* over all initial spin configurations, s_i, and the *sum* over all final spin configurations, s_f. In principle, we could compute $|\mathscr{M}(s_i \to s_f)|^2$ for every possible combination, and *then* do the summing and averaging:

$$\langle|\mathscr{M}|^2\rangle \equiv \textit{average} \text{ over initial spins, } \textit{sum} \text{ over final spins, of } |\mathscr{M}(s_i \to s_f)|^2 \tag{7.114}$$

In practice, it is much easier to compute $\langle|\mathscr{M}|^2\rangle$ directly, without ever evaluating the individual amplitudes.

Consider, for instance, the electron–muon scattering amplitude (Equation 7.106). Squaring, we have

$$|\mathscr{M}|^2 = \frac{g_e^4}{(p_1 - p_3)^4}[\overline{u}(3)\gamma^\mu u(1)][\overline{u}(4)\gamma_\mu u(2)][\overline{u}(3)\gamma^\nu u(1)]^*[\overline{u}(4)\gamma_\nu u(2)]^* \tag{7.115}$$

(I use ν for the second contraction, since μ has been preempted.) A glance at the first and third 'sandwiches' (or the second and fourth) reveals that we must handle quantities of the generic form

$$G \equiv [\overline{u}(a)\Gamma_1 u(b)][\overline{u}(a)\Gamma_2 u(b)]^* \tag{7.116}$$

where a and b stand for the appropriate spins and momenta, and Γ_1, and Γ_2 are 4×4 matrices. All the other processes described in Section 7.6 – Møller, Bhabha, and Compton scattering, as well as pair production and annihilation – lead to expressions with similar structure. To begin with, we evaluate the complex conjugate (which is the same as the Hermitian conjugate, since the quantity in square brackets is a 1×1 'matrix'):

$$[\overline{u}(a)\Gamma_2 u(b)]^* = [u(a)^\dagger \gamma^0 \Gamma_2 u(b)]^\dagger = u(b)^\dagger \Gamma_2^\dagger \gamma^{0\dagger} u(a) \tag{7.117}$$

Now, $\gamma^{0\dagger} = \gamma^0$, and $(\gamma^0)^2 = 1$, so

$$u(b)^\dagger \Gamma_2^\dagger \gamma^{0\dagger} u(a) = u(b)^\dagger \gamma^0 \gamma^0 \Gamma_2^\dagger \gamma^0 u(a) = \overline{u}(b)\overline{\Gamma}_2 u(a) \tag{7.118}$$

where*

$$\overline{\Gamma}_2 \equiv \gamma^0 \Gamma_2^\dagger \gamma^0 \tag{7.119}$$

* Observe that the overbar now serves two different functions. On a *spinor* it denotes the *adjoint*: $\overline{\psi} \equiv \psi^\dagger \gamma^0$ (Equation 7.58); on a 4×4 *matrix* it defines a *new* matrix: $\overline{\Gamma} \equiv \gamma^0 \Gamma^\dagger \gamma^0$.

Thus

$$G = [\overline{u}(a)\Gamma_1 u(b)][\overline{u}(b)\overline{\Gamma}_2 u(a)] \tag{7.120}$$

We are now ready to sum over the spin orientations of particle b. Using the completeness relation (Equation 7.99), we have

$$\sum_{b \text{ spins}} G = \overline{u}(a)\Gamma_1 \left\{ \sum_{s_b=1,2} u^{(s_b)}(p_b)\overline{u}^{(s_b)}(p_b) \right\} \overline{\Gamma}_2 u(a)$$

$$= \overline{u}(a)\Gamma_1(\not{p}_b + m_b c)\overline{\Gamma}_2 u(a) = \overline{u}(a) Q u(a) \tag{7.121}$$

where Q is a temporary shorthand for the 4×4 matrix

$$Q \equiv \Gamma_1(\not{p}_b + m_b c)\overline{\Gamma}_2 \tag{7.122}$$

Next, we do the same for particle a:

$$\sum_{a \text{ spins}} \sum_{b \text{ spins}} G = \sum_{s_a=1,2} \overline{u}^{(s_a)}(p_a) Q u^{(s_a)}(p_a)$$

Or, writing out the matrix multiplication explicitly:*

$$\sum_{s_a=1,2} \sum_{i,j=1}^{4} \overline{u}^{(s_a)}(p_a)_i Q_{ij} u^{(s_a)}(p_a)_j = \sum_{i,j=1}^{4} Q_{ij} \left\{ \sum_{s_a=1,2} u^{(s_a)}(p_a)\overline{u}^{(s_a)}(p_a) \right\}_{ji}$$

$$= \sum_{i,j=1}^{4} Q_{ij}(\not{p}_a + m_a c)_{ji} = \sum_{i=1}^{4} [Q(\not{p}_a + m_a c)]_{ii} = \mathrm{Tr}[Q(\not{p}_a + m_a c)] \tag{7.123}$$

where 'Tr' denotes the *trace* of the matrix (the sum of its diagonal elements):

$$\mathrm{Tr}(A) \equiv \sum_i A_{ii} \tag{7.124}$$

Conclusion:

$$\sum_{\text{all spins}} [\overline{u}(a)\Gamma_1 u(b)][\overline{u}(a)\Gamma_2 u(b)]^* = \mathrm{Tr}[\Gamma_1(\not{p}_b + m_b c)\overline{\Gamma}_2(\not{p}_a + m_a c)] \tag{7.125}$$

This may not look like much of a simplification, but it is actually *huge*. Notice that there are *no spinors left* – once we do the summation over spins, all that remains is to multiply matrices and take the trace. This is sometimes called 'Casimir's trick', since Casimir was apparently the first one to use it [4]. Incidentally, if either u (in

* This is fancy footwork, so watch closely. You can't mess with the ordering of two spinors, but their *components* are just *numbers* ... *they* can be written either way: $\overline{u}_i u_j = u_j \overline{u}_i$. In the second step we recognize this product as the ji element of the *matrix* $u\overline{u}$ (note the unusual matrix multiplication here, column times row: 4×1 times 1×4 produces 4×4).

Equation 7.125) is replaced by a v, the corresponding mass on the right-hand side switches sign (see Problem 7.28).

Example 7.5 In the case of electron–muon scattering (Equation 7.115), $\Gamma_2 = \gamma^\nu$, and hence $\overline{\Gamma}_2 = \gamma^0\gamma^{\nu\dagger}\gamma^0 = \gamma^\nu$ (Problem 7.29). Applying Casimir's trick twice, we find

$$\langle|\mathscr{M}|^2\rangle = \frac{g_e^4}{4(p_1 - p_3)^4}\mathrm{Tr}[\gamma^\mu(\not{p}_1 + mc)\gamma^\nu(\not{p}_3 + mc)] \times \mathrm{Tr}[\gamma_\mu(\not{p}_2 + Mc)\gamma_\nu(\not{p}_4 + Mc)] \tag{7.126}$$

where m is the mass of the electron and M is the mass of the muon. The factor of $\frac{1}{4}$ is included because we want the *average* over the initial spins; since there are two particles, each with two possible orientations, the average is a quarter of the sum. ▒

Casimir's trick reduces everything to an exercise in calculating the trace of some complicated product of γ matrices. This algebra is facilitated by a number of theorems, which I shall now list (I'll leave the proofs to you – see Problems 7.31–7.34). First of all, I should mention three facts about traces in general: if A and B are any two matrices, and α is any number

1. $\mathrm{Tr}(A + B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$
2. $\mathrm{Tr}(\alpha A) = \alpha\,\mathrm{Tr}(A)$
3. $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$

It follows from number 3 that $\mathrm{Tr}(ABC) = \mathrm{Tr}(CAB) = \mathrm{Tr}(BCA)$, but this is *not* equal, in general, to the trace of the matrices taken in the other order: $\mathrm{Tr}(ACB) = \mathrm{Tr}(BAC) = \mathrm{Tr}(CBA)$. You can 'peel' matrices off the back end of a product and move them around to the front, but you must preserve the ordering. It is useful to note that

4. $g_{\mu\nu}g^{\mu\nu} = 4$

and to recall the fundamental anticommutation relation for the γ matrices (together with an associated rule for 'slash' products):

5. $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}$ 5′. $\not{a}\not{b} + \not{b}\not{a} = 2a \cdot b$

From these there follows a sequence of 'contraction theorems':

6. $\gamma_\mu\gamma^\mu = 4$
7. $\gamma_\mu\gamma^\nu\gamma^\mu = -2\gamma^\nu$ 7′. $\gamma_\mu\not{a}\gamma^\mu = -2\not{a}$

8. $\gamma_\mu \gamma^\nu \gamma^\lambda \gamma^\mu = 4g^{\nu\lambda}$ 8′. $\gamma_\mu \not{a}\not{b}\gamma^\mu = 4(a \cdot b)$

9. $\gamma_\mu \gamma^\nu \gamma^\lambda \gamma^\sigma \gamma^\mu = -2\gamma^\sigma \gamma^\lambda \gamma^\nu$ 9′. $\gamma_\mu \not{a}\not{b}\not{c}\gamma^\mu = -2\not{c}\not{b}\not{a}$

and a collection of 'trace theorems':

10. The trace of the product of an *odd* number of gamma matrices is *zero*.

11. $\mathrm{Tr}(1) = 4$

12. $\mathrm{Tr}(\gamma^\mu \gamma^\nu) = 4g^{\mu\nu}$ 12′. $\mathrm{Tr}(\not{a}\not{b}) = 4(a \cdot b)$

13. $\mathrm{Tr}(\gamma^\mu \gamma^\nu \gamma^\lambda \gamma^\sigma)$ 13′. $\mathrm{Tr}(\not{a}\not{b}\not{c}\not{d})$

$= 4(g^{\mu\nu}g^{\lambda\sigma} - g^{\mu\lambda}g^{\nu\sigma} + g^{\mu\sigma}g^{\nu\lambda})$ $= 4(a \cdot b\, c \cdot d - a \cdot c\, b \cdot d + a \cdot d\, b \cdot c)$

Finally, since $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ is the product of an *even* number of γ matrices, it follows from Rule 10 that $\mathrm{Tr}\,(\gamma^5\gamma^\mu) = \mathrm{Tr}\,(\gamma^5\gamma^\mu\gamma^\nu\gamma^\lambda) = 0$. When γ^5 is multiplied by an *even* number of γ's, we find

14. $\mathrm{Tr}(\gamma^5) = 0$

15. $\mathrm{Tr}(\gamma^5\gamma^\mu\gamma^\nu) = 0$ 15′. $\mathrm{Tr}(\gamma^5\not{a}\not{b}) = 0$

16. $\mathrm{Tr}(\gamma^5\gamma^\mu\gamma^\nu\gamma^\lambda\gamma^\sigma) = 4i\epsilon^{\mu\nu\lambda\sigma}$ 16′. $\mathrm{Tr}(\gamma^5\not{a}\not{b}\not{c}\not{d}) = 4i\epsilon^{\mu\nu\lambda\sigma}a_\mu b_\nu c_\lambda d_\sigma$

where*

$$\epsilon^{\mu\nu\lambda\sigma} \equiv \begin{cases} -1, & \text{if } \mu\nu\lambda\sigma \text{ is an } \textit{even} \text{ permutation of } 0123, \\ +1, & \text{if } \mu\nu\lambda\sigma \text{ is an } \textit{odd} \text{ permutation}, \\ 0, & \text{if any two indices are the same.} \end{cases} \tag{7.127}$$

Example 7.6 Evaluate the traces in electron–muon scattering (Equation 7.126):

$$\mathrm{Tr}[\gamma^\mu(\not{p}_1 + mc)\gamma^\nu(\not{p}_3 + mc)]$$
$$= \mathrm{Tr}(\gamma^\mu\not{p}_1\gamma^\nu\not{p}_3) + mc\left[\mathrm{Tr}(\gamma^\mu\not{p}_1\gamma^\nu) + \mathrm{Tr}(\gamma^\mu\gamma^\nu\not{p}_3)\right] + (mc)^2\mathrm{Tr}(\gamma^\mu\gamma^\nu)$$

* By 'even permutation' I mean an even number of interchanges of two indices. Thus $\epsilon^{\mu\nu\lambda\sigma} = -\epsilon^{\nu\mu\lambda\sigma} = \epsilon^{\nu\lambda\mu\sigma} = -\epsilon^{\nu\lambda\sigma\mu}$, and so on – in other words, $\epsilon^{\mu\nu\lambda\sigma}$ is antisymmetric in every pair of superscripts. It might seem strange that ϵ^{0123} is *minus* 1; why not make it *plus* 1? It's purely conventional, of course – evidently, whoever established the definition wanted ϵ_{0123} to be plus 1, and from that it follows that $\epsilon^{0123} = -1$, since three spatial indices are raised. By the way, if you are used to working with the three-dimensional Levi–Civita symbol ϵ_{ijk} (Problem 4.19), be warned that although an even permutation on *three* indices corresponds to preservation of cyclic order ($\epsilon_{ijk} = \epsilon_{jki} = \epsilon_{kij}$), this is *not* the case for *four* indices: $\epsilon^{\mu\nu\lambda\sigma} = -\epsilon^{\nu\lambda\sigma\mu} = \epsilon^{\lambda\sigma\mu\nu} = -\epsilon^{\sigma\mu\nu\lambda}$.

Solution: According to Rule 10, the terms in square brackets are zero. The last term can be evaluated using Rule 12, and the first by Rule 13:

$$\begin{aligned}\mathrm{Tr}(\gamma^\mu \not{p}_1 \gamma^\nu \not{p}_3) &= (p_1)_\lambda (p_3)_\sigma \mathrm{Tr}(\gamma^\mu \gamma^\lambda \gamma^\nu \gamma^\sigma) \\ &= (p_1)_\lambda (p_3)_\sigma 4(g^{\mu\lambda} g^{\nu\sigma} - g^{\mu\nu} g^{\lambda\sigma} + g^{\mu\sigma} g^{\lambda\nu}) \\ &= 4\left[p_1^\mu p_3^\nu - g^{\mu\nu}(p_1 \cdot p_3) + p_3^\mu p_1^\nu\right]\end{aligned}$$

Thus

$$\begin{aligned}&\mathrm{Tr}\left[\gamma^\mu(\not{p}_1 + mc)\gamma^\nu(\not{p}_3 + mc)\right] \\ &\qquad = 4\left\{p_1^\mu p_3^\nu + p_3^\mu p_1^\nu + g^{\mu\nu}\left[(mc)^2 - (p_1 \cdot p_3)\right]\right\}\end{aligned} \tag{7.128}$$

The second trace in Equation 7.126 is the same, with $m \to M$, $1 \to 2$, $3 \to 4$, and the Greek indices lowered. So

$$\begin{aligned}\langle|\mathscr{M}|^2\rangle &= \frac{4g_e^4}{(p_1 - p_3)^4}\left\{p_1^\mu p_3^\nu + p_3^\mu p_1^\nu + g^{\mu\nu}\left[(mc)^2 - (p_1 \cdot p_3)\right]\right\} \\ &\quad \times \left\{p_{2\mu} p_{4\nu} + p_{4\mu} p_{2\nu} + g_{\mu\nu}\left[(Mc)^2 - (p_2 \cdot p_4)\right]\right\} \\ &= \frac{8g_e^4}{(p_1 - p_3)^4}\left[(p_1 \cdot p_2)(p_3 \cdot p_4) + (p_1 \cdot p_4)(p_2 \cdot p_3)\right. \\ &\quad \left. - (p_1 \cdot p_3)(Mc)^2 - (p_2 \cdot p_4)(mc)^2 + 2(mMc^2)^2\right]\end{aligned} \tag{7.129}$$

7.8 Cross Sections and Lifetimes

We are now back on familiar turf. Having calculated $|\mathscr{M}|^2$ (or, where appropriate, $\langle|\mathscr{M}|^2\rangle$), we simply plug it into the relevant cross section formula from Chapter 6: Equation 6.38 in the general case, Equation 6.47 for two-body scattering in the CM, or one of the equations from Problems 6.8, 6.9, or 6.10 in the lab frame.

Example 7.7 Mott and Rutherford Scattering An electron (mass m) scatters off a much heavier 'muon' (mass $M \gg m$). Assuming that the recoil of M can be neglected, find the differential scattering cross section in the lab frame (M at rest).

Solution: According to Problem 6.8, the cross section is given by

$$\frac{d\sigma}{d\Omega} = \left(\frac{\hbar}{8\pi Mc}\right)^2 \langle|\mathscr{M}|^2\rangle$$

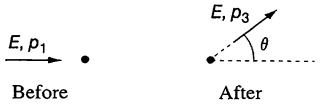


Fig. 7.8 Electron scattering from a heavy target.

Because the target is stationary, we have (Figure 7.8):

$$p_1 = (E/c, \mathbf{p}_1), \quad p_2 = (Mc, \mathbf{0}), \quad p_3 = (E/c, \mathbf{p}_3), \quad p_4 = (Mc, \mathbf{0})$$

where E is the incident (and scattered) electron energy, $\mathbf{p}_1$ is the incident momentum, and $\mathbf{p}_3$ is the scattered momentum (their magnitudes are equal, $|\mathbf{p}_1| = |\mathbf{p}_3| \equiv |\mathbf{p}|$, and the angle between them is θ so $\mathbf{p}_1 \cdot \mathbf{p}_3 = \mathbf{p}^2 \cos\theta$). Thus

$$\begin{aligned}
(p_1 - p_3)^2 &= -(\mathbf{p}_1 - \mathbf{p}_3)^2 = -\mathbf{p}_1^2 - \mathbf{p}_3^2 + 2\mathbf{p}_1 \cdot \mathbf{p}_3 \\
&= -2\mathbf{p}^2(1 - \cos\theta) = -4\mathbf{p}^2 \sin^2(\theta/2) \\
(p_1 \cdot p_3) &= (E/c)^2 - \mathbf{p}_1 \cdot \mathbf{p}_3 = \mathbf{p}^2 + m^2c^2 - \mathbf{p}^2 \cos\theta = m^2c^2 + 2\mathbf{p}^2 \sin^2(\theta/2) \\
(p_1 \cdot p_2)(p_3 \cdot p_4) &= (p_1 \cdot p_4)(p_2 \cdot p_3) = (ME)^2 \\
(p_2 \cdot p_4) &= (Mc)^2
\end{aligned}$$

Putting this into Equation 7.129, we have

$$\langle |\mathscr{M}|^2 \rangle = \left(\frac{g_e^2 Mc}{\mathbf{p}^2 \sin^2(\theta/2)} \right)^2 \left[(mc)^2 + \mathbf{p}^2 \cos^2(\theta/2) \right] \tag{7.130}$$

and therefore (recalling that $g_e = \sqrt{4\pi\alpha}$)

$$\frac{d\sigma}{d\Omega} = \left(\frac{\alpha\hbar}{2\mathbf{p}^2 \sin^2(\theta/2)} \right)^2 \left[(mc)^2 + \mathbf{p}^2 \cos^2(\theta/2) \right] \tag{7.131}$$

This is the Mott formula. It gives, to good approximation, the differential cross section for electron–proton scattering. If the incident electron is nonrelativistic, so $\mathbf{p}^2 \ll (mc)^2$, Equation 7.131 reduces to the Rutherford formula (compare Example 6.4):

$$\frac{d\sigma}{d\Omega} = \left(\frac{e^2}{2mv^2 \sin^2(\theta/2)} \right)^2 \tag{7.132}$$

What about decays? Actually, there is no such thing, in pure QED, for if a single fermion goes *in*, that same fermion must eventually come *out*; a fermion line

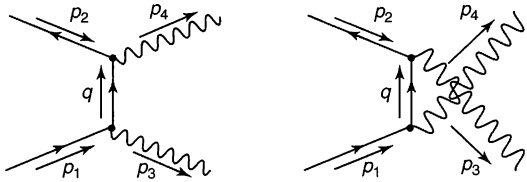


Fig. 7.9 Two contributions to pair annihilation.

cannot simply terminate within a diagram, nor is there any mechanism in QED for converting one fermion (say, a muon) into another (such as an electron). To be sure, there exist electromagnetic decays of *composite* particles, for example, $\pi^0 \to \gamma + \gamma$; but the *electromagnetic* component in this process is nothing but quark–antiquark pair annihilation, $q + \overline{q} \to \gamma + \gamma$. It is really a *scattering* event, in which the two colliding particles happen to be in a bound state.

The cleanest example of such a process is the decay of positronium: $e^+ + e^- \to \gamma + \gamma$, which we consider in the following example. We'll do the analysis in the positronium rest frame (which is to say, in the CM frame of the electron–positron pair). The electron and positron are moving rather slowly – indeed, for purposes of calculating the amplitude we shall assume they are at rest. On the other hand, this is one of those cases in which we *cannot* average over initial spins, because the composite system is either in the singlet configuration – spins antiparallel – or in the triplet configuration – spins parallel – and the formula for the cross section (and hence the lifetime) is quite different in the two cases.*

Example 7.8 Pair Annihilation† Compute the amplitude, $\mathscr{M}$, for $e^+ + e^- \to \gamma + \gamma$, assuming that the electron and positron are at rest, and in the singlet spin configuration.

Solution: Two diagrams contribute (Figure 7.9). The amplitudes are (for simplicity I'll suppress the complex conjugate signs on the ϵ's):

$$\mathscr{M}_1 = \frac{g_e^2}{(p_1 - p_3)^2 - m^2c^2}\,\overline{v}(2)\not{\epsilon}_4(\not{p}_1 - \not{p}_3 + mc)\not{\epsilon}_3 u(1) \tag{7.133}$$

$$\mathscr{M}_2 = \frac{g_e^2}{(p_1 - p_4)^2 - m^2c^2}\,\overline{v}(2)\not{\epsilon}_3(\not{p}_1 - \not{p}_4 + mc)\not{\epsilon}_4 u(1) \tag{7.134}$$

* As a matter of fact, you *can* do this particular problem by Casimir's trick, because of a rather special circumstance: the singlet state can only decay to an *even* number of photons (predominantly *two*) and the triplet to an *odd* number (usually *three*). So in calculating the matrix element for $e^+ + e^- \to \gamma + \gamma$, we are *automatically* selecting out the singlet configuration even if the triplet was included in the sum over spins. See Problem 7.40.

† *Warning:* This is not an easy calculation, though every step is reasonably straightforward. You may prefer to skim it (or skip it altogether). The final result will be used once or twice later on, but it is not necessary to master the details at this stage. (However, I do think it is an illuminating application of the Feynman rules.)

and they *add*

$$\mathscr{M} = \mathscr{M}_1 + \mathscr{M}_2 \tag{7.135}$$

With the initial particles at rest, the photons come out back-to-back, and we may as well choose the z axis to coincide with the direction of the first photon; then

$$\begin{aligned} p_1 &= mc(1,0,0,0), \qquad & p_2 &= mc(1,0,0,0), \\ p_3 &= mc(1,0,0,1), \qquad & p_4 &= mc(1,0,0,-1) \end{aligned} \tag{7.136}$$

and hence

$$(p_1 - p_3)^2 - m^2c^2 = (p_1 - p_4)^2 - m^2c^2 = -2(mc)^2 \tag{7.137}$$

The amplitudes simplify somewhat if we exploit Rule 5′ from Section 7.7:

$$\not{p}_1 \not{\epsilon}_3 = -\not{\epsilon}_3 \not{p}_1 + 2(p_1 \cdot \epsilon_3)$$

But ϵ_3 has only spatial components (in the Coulomb gauge), whereas p_1 is purely temporal, so $p_1 \cdot \epsilon_3 = 0$, and hence

$$\not{p}_1 \not{\epsilon}_3 = -\not{\epsilon}_3 \not{p}_1 \tag{7.138}$$

Similarly

$$\not{p}_3 \not{\epsilon}_3 = -\not{\epsilon}_3 \not{p}_3 + 2(p_3 \cdot \epsilon_3)$$

but $(p_3 \cdot \epsilon_3) = 0$ by virtue of the Lorentz condition (Equation 7.91), so

$$\not{p}_3 \not{\epsilon}_3 = -\not{\epsilon}_3 \not{p}_3 \tag{7.139}$$

Therefore

$$(\not{p}_1 - \not{p}_3 + mc)\not{\epsilon}_3 = \not{\epsilon}_3(-\not{p}_1 + \not{p}_3 + mc)$$

But $(\not{p}_1 - mc)u(1) = 0$ (Equation 7.33), so

$$(\not{p}_1 - \not{p}_3 + mc)\not{\epsilon}_3 u(1) = \not{\epsilon}_3 \not{p}_3 u(1) \tag{7.140}$$

By the same token

$$(\not{p}_1 - \not{p}_4 + mc)\not{\epsilon}_4 u(1) = \not{\epsilon}_4 \not{p}_4 u(1) \tag{7.141}$$

Putting all this together, we find

$$\mathscr{M} = -\frac{g_e^2}{2(mc)^2}\,\bar{v}(2)[\not{\epsilon}_4\not{\epsilon}_3\not{p}_3 + \not{\epsilon}_3\not{\epsilon}_4\not{p}_4]u(1) \tag{7.142}$$

Now

$$\not{p}_3 = mc(\gamma^0 - \gamma^3), \qquad \not{p}_4 = mc(\gamma^0 + \gamma^3)$$

so the expression in square brackets (Equation 7.142) can be written as

$$mc\left[(\not{\epsilon}_4\not{\epsilon}_3 + \not{\epsilon}_3\not{\epsilon}_4)\gamma^0 - (\not{\epsilon}_4\not{\epsilon}_3 - \not{\epsilon}_3\not{\epsilon}_4)\gamma^3\right] \tag{7.143}$$

But

$$\not{\epsilon} = -\epsilon\cdot\gamma = -\begin{pmatrix} 0 & \sigma\cdot\epsilon \\ -\sigma\cdot\epsilon & 0 \end{pmatrix} \tag{7.144}$$

and therefore

$$\begin{aligned}\not{\epsilon}_3\not{\epsilon}_4 &= \begin{pmatrix} 0 & \sigma\cdot\epsilon_3 \\ -\sigma\cdot\epsilon_3 & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma\cdot\epsilon_4 \\ -\sigma\cdot\epsilon_4 & 0 \end{pmatrix} \\ &= -\begin{pmatrix} (\sigma\cdot\epsilon_3)(\sigma\cdot\epsilon_4) & 0 \\ 0 & (\sigma\cdot\epsilon_3)(\sigma\cdot\epsilon_4) \end{pmatrix}\end{aligned} \tag{7.145}$$

In Chapter 4 (Problem 4.20) we encountered the useful theorem

$$(\sigma\cdot\mathbf{a})(\sigma\cdot\mathbf{b}) = \mathbf{a}\cdot\mathbf{b} + i\sigma\cdot(\mathbf{a}\times\mathbf{b}) \tag{7.146}$$

It follows that

$$(\not{\epsilon}_4\not{\epsilon}_3 + \not{\epsilon}_3\not{\epsilon}_4) = -2\epsilon_3\cdot\epsilon_4 \tag{7.147}$$

(which we could also have obtained from Rule 5′), and

$$(\not{\epsilon}_4\not{\epsilon}_3 - \not{\epsilon}_3\not{\epsilon}_4) = 2i(\epsilon_3\times\epsilon_4)\cdot\mathbf{\Sigma} \tag{7.148}$$

where $\mathbf{\Sigma} = \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix}$, as before. Accordingly

$$\mathscr{M} = \frac{g_e^2}{mc}\,\bar{v}(2)\left[(\epsilon_3\cdot\epsilon_4)\gamma^0 + i(\epsilon_3\times\epsilon_4)\cdot\mathbf{\Sigma}\gamma^3\right]u(1) \tag{7.149}$$

So far I have said nothing about the spins of the electron and positron. Remember that we are interested in the *singlet* state:

$$(\uparrow\downarrow - \downarrow\uparrow)/\sqrt{2}$$

Symbolically

$$\mathscr{M}_{\text{singlet}} = (\mathscr{M}_{\uparrow\downarrow} - \mathscr{M}_{\uparrow\downarrow})/\sqrt{2} \tag{7.150}$$

$\mathscr{M}_{\uparrow\downarrow}$ is obtained from Equation 7.149 with 'spin up' for the electron ($u^{(1)}$ in Equation 7.46)

$$u(1) = \sqrt{2mc}\begin{pmatrix}1\\0\\0\\0\end{pmatrix} \tag{7.151}$$

and 'spin down' for the positron ($v^{(2)}$ in Equation 7.47)

$$\bar{v}(2) = \sqrt{2mc}(0\ 0\ 1\ 0) \tag{7.152}$$

Using these spinors, we find

$$\bar{v}(2)\gamma^0 u(1) = 0 \tag{7.153}$$

$$\bar{v}(2)\Sigma\gamma^3 u(1) = -2mc\hat{z} \tag{7.154}$$

So

$$\mathscr{M}_{\uparrow\downarrow} = -2ig_e^2(\epsilon_3 \times \epsilon_4)_z \tag{7.155}$$

Meanwhile, for $\mathscr{M}_{\downarrow\uparrow}$ we have

$$u(1) = \sqrt{2mc}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \quad \bar{v}(2) = -\sqrt{2mc}(0\ 0\ 0\ 1) \tag{7.156}$$

from which it follows that

$$\mathscr{M}_{\downarrow\uparrow} = 2ig_e^2(\epsilon_3 \times \epsilon_4)_z = -\mathscr{M}_{\uparrow\downarrow} \tag{7.157}$$

Thus the amplitude for annihilation of a stationary e^+e^- pair into two photons, which emerge in the directions $\pm\hat{z}$, is

$$\mathscr{M}_{\text{singlet}} = -2\sqrt{2}\, ig_e^2(\epsilon_3 \times \epsilon_4)_z \tag{7.158}$$

(I note in passing that since $\mathscr{M}_{\uparrow\downarrow} = -\mathscr{M}_{\downarrow\uparrow}$, the *triplet* configuration $(\uparrow\downarrow + \downarrow\uparrow)/\sqrt{2}$ gives *zero*, confirming our earlier observation that the two-photon decay is forbidden in that case.)

Finally, we must put in the appropriate photon polarization vectors. Recall that for 'spin up' ($m_s = +1$) we have (see footnote to Equation 7.94)

$$\epsilon_+ = -(1/\sqrt{2})(1, i, 0) \tag{7.159}$$

whereas for 'spin down' ($m_s = -1$)

$$\epsilon_- = (1/\sqrt{2})(1, -i, 0) \tag{7.160}$$

If the photon is traveling in the $+z$ direction, these correspond to right and left-circular polarization, respectively. Since the z component of the total angular momentum must be zero, the photon spins must be oppositely aligned: $\uparrow\downarrow$ or $\downarrow\uparrow$. In the first case we have

$$\epsilon_3 = -(1/\sqrt{2})(1, i, 0), \qquad \epsilon_4 = (1/\sqrt{2})(1, -i, 0),$$

so

$$\epsilon_3 \times \epsilon_4 = i\hat{k} \tag{7.161}$$

In the second case 3 and 4 are interchanged:

$$\epsilon_3 \times \epsilon_4 = -i\hat{k} \tag{7.162}$$

Evidently we need the antisymmetric combination, $(\uparrow\downarrow - \downarrow\uparrow)/\sqrt{2}$, which should come as no surprise: this corresponds to a total spin of zero, just as it did when we combined two particles of spin $\frac{1}{2}$. Again, the amplitude is $(\mathscr{M}_{\uparrow\downarrow} - \mathscr{M}_{\downarrow\uparrow})/\sqrt{2}$, only this time the arrows refer to *photon* polarization. Finally, then:

$$\mathscr{M}_{\text{singlet}} = -4g_e^2 \tag{7.163}$$

(I have restored the complex conjugation of the polarization vectors, suppressed until now; this simply reverses the signs in Equation 7.161 and 7.162.)

That was a lot of work, for a modest-looking answer.* What can we *do* with it? In the first place, we can calculate the total cross section for electron–positron annihilation. In the CM frame, the differential cross section is (Equation 6.47)

$$\frac{d\sigma}{d\Omega} = \left(\frac{\hbar c}{8\pi(E_1 + E_2)}\right)^2 \frac{|\mathbf{p}_f|}{|\mathbf{p}_i|} |\mathscr{M}|^2 \tag{7.164}$$

Here

$$E_1 = E_2 = mc^2, \quad |\mathbf{p}_f| = mc \tag{7.165}$$

and, since the collision is nonrelativistic

$$|\mathbf{p}_i| = mv \tag{7.166}$$

where v is the incident electron (or positron) speed.† Putting all this together, we find

$$\frac{d\sigma}{d\Omega} = \frac{1}{cv}\left(\frac{\hbar\alpha}{m}\right)^2 \tag{7.167}$$

* Once you get used to it the evaluation of Feynman diagrams becomes a tedious and mechanical process, and there exist a number of computer programs that will do the hard work for you. In particular, Mathematica and Maple both support useful packages [5].

† We used $v = 0$ in calculating $\mathscr{M}$, but obviously we cannot do so here. Is there an inconsistency in this? Not really. Think of it this way: $\mathscr{M}$ (and also $E_1, E_2, |\mathbf{p}_f|$, and $|\mathbf{p}_i|$) could be expanded in powers of v/c. What we have done is to calculate the leading term in each expansion.

Since there is no angular dependence, the total cross section is 4π times this [6]:

$$\sigma = \frac{4\pi}{cv}\left(\frac{\hbar\alpha}{m}\right)^2 \tag{7.168}$$

Does it make sense that the cross section is inversely proportional to the incoming velocity? *Yes:* the more slowly the electron and positron approach one another, the more time there is for them to interact, and the greater is the likelihood of annihilation.

Finally, we can calculate the lifetime of positronium, in the singlet state. This is clearly *related* to the cross section for pair annihilation (Equation 7.168), but what is the precise connection? Well, going back to Equation 6.13,

$$\frac{d\sigma}{d\Omega} = \frac{1}{\mathscr{L}}\frac{dN}{d\Omega}$$

we see that the total number of scattering events per unit time is equal to the luminosity times the total cross section:

$$N = \mathscr{L}\sigma \tag{7.169}$$

If ρ is the number of incident particles per unit volume, and if they are traveling at speed v, then the luminosity (Figure 7.10) is

$$\mathscr{L} = \rho v \tag{7.170}$$

For a single 'atom', the electron density is $|\psi(0)|^2$, and N represents the probability of a disintegration, per unit time – which is to say, the decay rate. Thus

$$\Gamma = v\sigma|\psi(0)|^2 = \frac{4\pi}{c}\left(\frac{\hbar\alpha}{m}\right)^2|\psi(0)|^2 \tag{7.171}$$

Now, in the ground state

$$|\psi(0)|^2 = \frac{1}{\pi}\left(\frac{\alpha mc}{2\hbar}\right)^3 \tag{7.172}$$

(Problem 5.23), so the lifetime of positronium is

$$\tau = \frac{1}{\Gamma} = \frac{2\hbar}{\alpha^5 mc^2} = 1.25 \times 10^{-10}\,\text{s} \tag{7.173}$$

which is the result I quoted back in Chapter 5 (Equation 5.33).

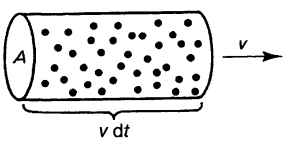
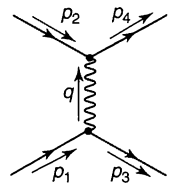


Fig. 7.10 The number of particles in the cylinder is $\rho A v\,dt$, so the *luminosity* (number per unit area per unit time) is ρv.

7.9 Renormalization

In Section 7.6 we considered 'electron–muon' scattering, described in lowest order by the diagram



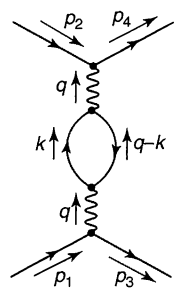
and by the corresponding amplitude

$$\mathscr{M} = -g_e^2[\overline{u}(p_3)\gamma^\mu u(p_1)]\frac{g_{\mu\nu}}{q^2}[\overline{u}(p_4)\gamma^\nu u(p_2)] \tag{7.174}$$

with

$$q = p_1 - p_3 \tag{7.175}$$

There are a number of fourth-order corrections, of which perhaps the most interesting is 'vacuum polarization':



Here the virtual photon momentarily splits into an electron–positron pair, leading (as we saw qualitatively in Chapter 2) to a modification in the effective charge of the electron. My purpose now is to indicate how this works out quantitatively.

The amplitude for this diagram is (Problem 7.42)

$$\begin{aligned}\mathscr{M} =& \frac{-ig_e^4}{q^4}[\overline{u}(p_3)\gamma^\mu u(p_1)] \\ &\times \left\{ \int \frac{d^4k}{(2\pi)^4} \frac{\mathrm{Tr}[\gamma_\mu(\not{k} + mc)\gamma_\nu(\not{k} - \not{q} + mc)]}{(k^2 - m^2c^2)[(k-q)^2 - m^2c^2]} \right\} [\overline{u}(p_4)\gamma^\nu u(p_2)]\end{aligned} \tag{7.176}$$

Its inclusion amounts to a modification of the photon propagator:

$$\frac{g_{\mu\nu}}{q^2} \to \frac{g_{\mu\nu}}{q^2} - \frac{i}{q^4} I_{\mu\nu} \tag{7.177}$$

where (comparing Equations 7.174 and 7.176):

$$I_{\mu\nu} = -g_e^2 \int \frac{\mathrm{d}^4 k}{(2\pi)^4} \frac{\mathrm{Tr}[\gamma_\mu(\not{k} + mc)\gamma_\nu(\not{k} - \not{q} + mc)]}{(k^2 - m^2c^2)[(k-q)^2 - m^2c^2]} \tag{7.178}$$

Unfortunately, this integral is divergent. Naively, it should go like

$$\int |k|^3 \, \mathrm{d}|k| \frac{|k|^2}{|k|^4} = \int |k| \, \mathrm{d}k = |k|^2, \quad \text{as } |k| \to \infty \tag{7.179}$$

(that is, it should be 'quadratically divergent'). In actual fact, because of cancellations in the algebra, it only goes like $\ln|k|$ (it is 'logarithmically divergent'). But never mind – either way, it blows up. We encountered a similar problem in Chapter 6; it seems to be characteristic of closed-loop diagrams in the Feynman calculus. Once again, the strategy will be to absorb the infinities into 'renormalized' masses and coupling constants.

The integral in Equation 7.178 carries two space-time indices; once we have integrated over k, the only four-vector left is q^μ, so $I_{\mu\nu}$ must have the generic form $g_{\mu\nu}(\ ) + q_\mu q_\nu(\ )$, where the parentheses contain some functions of q^2. We write it thus [7]:

$$I_{\mu\nu} = -ig_{\mu\nu} q^2 I(q^2) + q_\mu q_\nu J(q^2) \tag{7.180}$$

The second term contributes nothing to $\mathscr{M}$, since the q_μ contracts with γ^μ in Equation 7.176, giving

$$[\overline{u}(p_3) \not{q} u(p_1)] = \overline{u}(p_3)(\not{p}_1 - \not{p}_3)u(p_1)$$

while, from Equations 7.95 and 7.96,

$$\not{p}_1 u(p_1) = mc\, u(p_1), \qquad \overline{u}(p_3)\not{p}_3 = \overline{u}(p_3)mc$$

and hence

$$[\overline{u}(p_3)\not{q} u(p_1)] = 0 \tag{7.181}$$

So we can forget about the second term in Equation 7.180 As for the first term, appropriate massaging of the integral (7.174) reduces it to the form (Problem 7.43)

$$I(q^2) = \frac{g_e^2}{12\pi^2} \left\{ \int_{m^2}^{\infty} \frac{\mathrm{d}z}{z} - 6 \int_0^1 z(1-z) \ln\left[1 - \frac{q^2}{m^2c^2} z(1-z)\right] \mathrm{d}z \right\} \tag{7.182}$$

The first integral cleanly isolates the logarithmic divergence. To handle it, we temporarily impose a cutoff M (not to be confused with the mass of the muon), which we will send to infinity at the end of the calculation:

$$\int_{m^2}^{\infty} \frac{dz}{z} \to \int_{m^2}^{M^2} \frac{dz}{z} = \ln \frac{M^2}{m^2} \tag{7.183}$$

The second integral

$$\begin{aligned} f(x) &\equiv 6 \int_0^1 z(1-z)\ \ln[1 + xz(1-z)]\, dz \\ &= -\frac{5}{3} + \frac{4}{x} + \frac{2(x-2)}{x}\sqrt{\frac{x+4}{x}}\tanh^{-1}\sqrt{\frac{x}{x+4}} \end{aligned} \tag{7.184}$$

is cumbersome but perfectly finite (Figure 7.11); the limiting expressions for large and small x are

$$f(x) \cong \begin{Bmatrix} x/5 & (x \ll 1) \\ \ln x & (x \gg 1) \end{Bmatrix} \tag{7.185}$$

Thus

$$I(q^2) = \frac{g_e^2}{12\pi^2}\left\{\ln\left(\frac{M^2}{m^2}\right) - f\left(\frac{-q^2}{m^2c^2}\right)\right\} \tag{7.186}$$

Notice that q^2 is *negative*, here: if the incident electron's three-momentum in the CM is $\mathbf{p}$, and the scattering angle is θ, then (Problem 7.44)

$$q^2 = -4\mathbf{p}^2 \sin^2\frac{\theta}{2} \tag{7.187}$$

Thus $-q^2/m^2c^2 \sim v^2/c^2$, and the limiting cases in Equation 7.185 correspond to nonrelativistic and ultrarelativistic scattering, respectively.

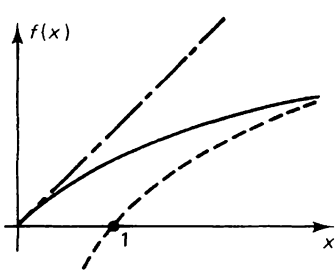


Fig. 7.11 Graph of $f(x)$ (Equation 7.184). The solid line is the numerical result; the dashed line below it is $\ln x$ (which approximates $f(x)$ at large x); the straight line above it is $x/5$ (which approximates $f(x)$ at small x).

The amplitude for electron–muon scattering, including vacuum polarization, is therefore

$$\mathscr{M} = -g_e^2[\overline{u}(p_3)\gamma^\mu u(p_1)]\frac{g_{\mu\nu}}{q^2}\left\{1 - \frac{g_e^2}{12\pi^2}\left[\ln\left(\frac{M^2}{m^2}\right) - f\left(\frac{-q^2}{m^2c^2}\right)\right]\right\}$$
$$\times [\overline{u}(p_4)\gamma^\nu u(p_2)] \tag{7.188}$$

Now comes the critical step, in which we 'sop up' the infinity (contained for the moment in the cutoff M) by introducing the 'renormalized' coupling constant

$$g_R \equiv g_e\sqrt{1 - \frac{g_e^2}{12\pi^2}\ln\left(\frac{M^2}{m^2}\right)} \tag{7.189}$$

Rewriting Equation 7.188 in terms of g_R, we have

$$\mathscr{M} = -g_R^2[\overline{u}(p_3)\gamma^\mu u(p_1)]\frac{g_{\mu\nu}}{q^2}\left\{1 + \frac{g_R^2}{12\pi^2}f\left(\frac{-q^2}{m^2c^2}\right)\right\}[\overline{u}(p_4)\gamma^\nu u(p_2)] \tag{7.190}$$

(Equation 7.188 is only valid to order g_e^4 anyway, so it doesn't matter whether we use g_e or g_R inside the curly brackets.)

There are two important things to notice about this result:

1. The infinities are gone. There is no M in Equation 7.190. All reference to the cutoff has been absorbed into the coupling constant. To be sure, everything is now written in terms of g_R, instead of g_e. But that's all to the good: g_R, not g_e, is what we actually measure in the laboratory (in Heaviside–Lorentz units it is the charge of the electron – or muon – and we determine it experimentally as the coefficient of attraction or repulsion between two such particles). If, in our theoretical analysis, we look only at 'tree level' (lowest-order) diagrams, we are led to suppose that the physical charge is the same as the 'bare' coupling constant, g_e. But as soon as we include higher-order effects we find that it is really g_R, not g_e, that corresponds to the measured electric charge. Does this mean that our earlier results are all wrong? No. What it means is that by naively interpreting g_e as the physical electric charge we were unwittingly taking into account the divergent part of the higher-order diagrams.
2. There remains the *finite* correction term, and here the important thing to notice is that it *depends on* q^2. We can absorb this, too, into the coupling constant, but the 'constant' is now a function of q^2; we call it a 'running' coupling constant:

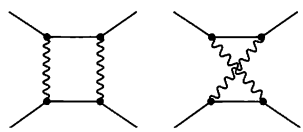
$$g_R(q^2) = g_R(0)\sqrt{1 + \frac{g_R(0)^2}{12\pi^2}f\left(\frac{-q^2}{m^2c^2}\right)} \tag{7.191}$$

or, in terms of the fine structure 'constant' ($g_e = \sqrt{4\pi\alpha}$):

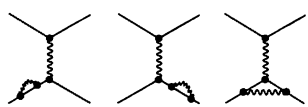
$$\alpha(q^2) = \alpha(0)\left\{1 + \frac{\alpha(0)}{3\pi} f\left(\frac{-q^2}{m^2c^2}\right)\right\} \tag{7.192}$$

The effective charge of the electron (and the muon), then, depends on the momentum transferred in the collision. Higher momentum transfer means closer approach, so another way of saying it is that the effective charge of each particle depends on how far apart they are. This is a consequence of vacuum polarization, which 'screens' each charge. We now have an explicit formula for what was, in Chapter 2, a purely qualitative description. How come Millikan and Rutherford, or even Coulomb, never noticed this effect? If the electron's charge is not a constant, why doesn't this foul up everything from electronics to chemistry? The answer is that the variation is extremely slight, in nonrelativistic situations. Even in a head-on collision at $\frac{1}{10}c$, the correction term in Equation 7.192 is only about 6×10^{-6} (Problem 7.45). For most purposes, therefore, $\alpha(0) = \frac{1}{137}$ will do just fine. Nevertheless, the second term in Equation 7.192 makes a detectable contribution to the Lamb shift [8], and it has been measured directly in inelastic e^+e^- scattering [9]. Moreover, we shall encounter the same problem in quantum chromodynamics, where (because of quark confinement) the short-distance, relativistic regime is the case of interest.

I have concentrated on one particular fourth-order process (vacuum polarization), but there are, of course, several others. There are the 'ladder-diagrams':

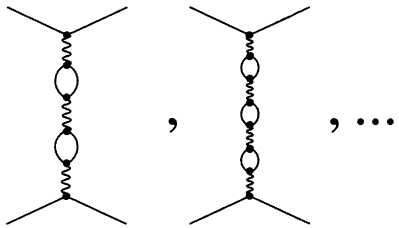


These are finite and present no particular problems. But there are also three divergent graphs:



(and of course three more in which the extra virtual photon couples to the muon). The first two renormalize the electron's mass; the third modifies its magnetic moment. In addition, all three, considered separately, contribute to the

renormalization of the electron's charge. Luckily, the latter contributions cancel one another, so Equation 7.189 remains valid. (I say 'luckily', for these corrections depend on the mass of the particle to which the virtual photon line attaches, and if they did *not* cancel we would have a different renormalization for the muon than for the electron. The Ward identity (the official name for this cancellation) guarantees that renormalization preserves the equality of electric charges, irrespective of the mass of the carrier).* And then, there are even higher-order diagrams, such as



These introduce further terms in Equation 7.192, of order α^2, α^3, and so on, but I won't pursue the matter here; the essential *ideas* are all on the table.

References

1 Bjorken, J. D. and Drell, S. D. (1964) *Relativistic Quantum Mechanics and Relativistic Quantum Fields*, McGraw-Hill, New York.

2 For further discussion see Bjorken, J. D. and Drell, S. D. (1964) *Relativistic Quantum Mechanics and Relativistic Quantum Fields*, McGraw-Hill, New York. Chapter 5, or Seiden, A. (2005) *Particle Physics: A Comprehensive Introduction*, Addison Wesley, San Francisco, CA, Section 2.14.2.

3 Jackson, J. D. (1999) *Classical Electrodynamics*, 3rd edn, John Wiley & Sons, New York. Sect. 6.3.

4 Pais, A. (1986) *Inward Bound*, Oxford, New York, p. 375.

5 For a review article on automatic computation of Feynman diagrams see Harlander, R. and Steinhauser, M. (1999) *Progress in Particle and Nuclear Physics*, 43, 167.

6 For more elegant derivations of Equation 7.168, see Jauch, J. M. and Rohrlich, F. (1975) *The Theory of Photons and Electrons*, 2nd edn, Springer-Verlag, New York. Sect. 12-6, or; (a) Sakurai, J. J. (1967) *Advanced Quantum Mechanics*, Addison-Wesley, Reading, M.A., pp. 216 ff.

7 My notation follows that of Halzen, F. and Martin, A. D. (1984) *Quarks and Leptons*, John Wiley & Sons, New York. Chap. 7, and Bjorken, J. D. and Drell, S. D. (1964) *Relativistic Quantum Mechanics and Relativistic Quantum Fields*, McGraw-Hill, New York. Chap. 8. I refer the reader to these texts, or to Sakurai, J. J. (1967) *Advanced Quantum Mechanics*, Addison-Wesley, Reading, M.A., pp. 216 ff, for further discussion.

* Of course, we can put a *muon* bubble onto a photon line, just as we did with the *electron* in Equation 7.176 But this will modify the electron and muon charge (and for that matter the tau and the quarks) by the same amount. The electron insertion is the dominant correction, simply because it is the lightest charged particle.

8 See, for example, Halzen, F. and Martin, A. D. (1984) *Quarks and Leptons*, John Wiley & Sons, New York. Sect. 7.3.

9 Levine, I. *et al* (1997) *Physical Review Letters*, 78, 424.

Problems

7.1 Show that $\partial\phi/\partial x^\mu$ is a *covariant* four-vector (ϕ is a scalar function of x, y, z, and t). [*Hint*: First determine (from Equation 3.8) how covariant four-vectors transform; then use $\partial\phi/\partial x^{\mu'} = (\partial\phi/\partial x^\nu)(\partial x^\nu/\partial x^{\mu'})$ to find out how $\partial\phi/\partial x^\mu$ transforms.]

7.2 Show that Equation 7.17 satisfies Equation 7.15

7.3 Derive Equation 7.45, using Equations 7.43, 7.46, and 7.47

7.4 Show that $u^{(1)}$ and $u^{(2)}$ (Equation 7.46) are *orthogonal*, in the sense that $u^{(1)\dagger}u^{(2)} = 0$. Likewise, show that $v^{(1)}$ and $v^{(2)}$ are orthogonal. Are $u^{(1)}$ and $v^{(1)}$ orthogonal?

7.5 Show that for $u^{(1)}$ and $u^{(2)}$ (Equation 7.46) the lower components (u_B) are smaller than the upper ones (u_A), in the nonrelativistic limit, by a factor v/c. [This simplifies matters, when we are doing nonrelativistic approximations; we think of u_A as the 'big' components and u_B as the 'little' components. (For $v^{(1)}$ and $v^{(2)}$ the roles are reversed.) In the relativistic limit, by contrast, u_A and u_B are comparable in size.]

7.6 If the z axis points along the direction of motion, show that $u^{(1)}$ (Equation 7.46) reduces to

$$u^{(1)} = \begin{pmatrix} \sqrt{(E+mc^2)/c} \\ 0 \\ \sqrt{(E-mc^2)/c} \\ 0 \end{pmatrix}$$

and construct $u^{(2)}$, $v^{(1)}$, and $v^{(2)}$. Confirm that they are all eigenspinors of S_z, and find the eigenvalues.

7.7 Construct the normalized spinors $u^{(+)}$ and $u^{(-)}$ representing an electron of momentum $\mathbf{p}$ with helicity ± 1. That is, find the u's that satisfy Equation 7.49 and are eigenspinors of the helicity operator ($\hat{p}\cdot\boldsymbol{\Sigma}$ with eigenvalues ± 1.

$$\left[\textit{Solution}:\ u^{(\pm)} = A\begin{pmatrix} u \\ \frac{\pm c|\mathbf{p}|}{(E+mc^2)}u \end{pmatrix},\right.$$

$$\left.\text{where } u = \begin{pmatrix} p_z \pm |\mathbf{p}| \\ p_x + ip_y \end{pmatrix} \text{ and } A^2 = \frac{(E+mc^2)}{2|\mathbf{p}|c(|\mathbf{p}| \pm p_z)}\right]$$

7.8 *The purpose of this problem is to demonstrate that particles described by the Dirac equation carry 'intrinsic' angular momentum (S) in addition to their orbital angular momentum (L), neither of which is separately conserved, although their sum is. It should be attempted only if you are reasonably familiar with quantum mechanics.*

(a) Construct the Hamiltonian, H, for the Dirac equation. [*Hint*: Solve Equation 7.19 for p^0c. *Solution*: $H = c\gamma^0(\boldsymbol{\gamma}\cdot\mathbf{p} + mc)$, where $\mathbf{p} \equiv (\hbar/i)\nabla$ is the momentum operator.]

(b) Find the commutator of H with the orbital angular momentum $\mathbf{L} \equiv \mathbf{r}\times\mathbf{p}$. [*Solution*: $[H, \mathbf{L}] = -i\hbar c\gamma^0(\boldsymbol{\gamma}\times\mathbf{p})$] Since $[H, \mathbf{L}]$ is not zero, $\mathbf{L}$ by itself is not conserved. Evidently

there is some *other* form of angular momentum lurking here. Introduce the 'spin angular momentum', $\mathbf{S}$, defined by Equation 7.51

(c) Find the commutator of H with the spin angular momentum, $\mathbf{S} \equiv (\hbar/2)\boldsymbol{\Sigma}$. [*Solution:* $[H, \mathbf{S}] = i\hbar c\gamma^0(\gamma \times \mathbf{p})$] It follows that the *total* angular momentum, $\mathbf{J} = \mathbf{L} + \mathbf{S}$, *is* conserved.

(d) Show that every bispinor is an eigenstate of $\mathbf{S}^2$, with eigenvalue $\hbar^2 s(s+1)$, and find s. What, then, is the spin of a particle described by the Dirac equation?

7.9 The charge conjugation operator (C) takes a Dirac spinor ψ into the 'charge-conjugate' spinor ψ_c, given by

$$\psi_c = i\gamma^2\psi^*$$

where γ^2 is the third Dirac gamma matrix. [See Halzen and Martin [7], Sect. 5.4.] Find the charge-conjugates of $u^{(1)}$ and $u^{(2)}$, and compare them with $v^{(1)}$ and $v^{(2)}$.

7.10 In going from Equation 7.18 to Equation 7.19, we (arbitrarily) chose to work with the factor containing the minus sign. How would Section 7.2 be changed if we were to replace Equation 7.19 by $\gamma^\mu p_\mu + mc = 0$?

7.11 Confirm the transformation rule (Equation 7.52, with 7.53 and 7.54) for spinors. [*Hint:* we want it to carry solutions to the Dirac equation in the original frame to solutions in the primed frame:

$$i\hbar\gamma^\mu\partial_\mu\psi - mc\,\psi = 0 \quad \leftrightarrow \quad i\hbar\gamma^\mu\partial'_\mu\psi' - mc\,\psi' = 0$$

where $\psi' = S\psi$ and

$$\partial'_\mu = \frac{\partial}{\partial x^{\mu\prime}} = \frac{\partial x^\nu}{\partial x^{\mu\prime}}\frac{\partial}{\partial x^\nu} = \frac{\partial x^\nu}{\partial x^{\mu\prime}}\partial_\nu$$

It follows that

$$(S^{-1}\gamma^\mu S)\frac{\partial x^\nu}{\partial x^{\mu\prime}} = \gamma^\nu$$

The (inverse) Lorentz transformations tell us $\partial x^\nu/\partial x^{\mu\prime}$. Take it from there.]

7.12 Derive the transformation rule for parity, Equation 7.61, using the method in Problem 7.11.

7.13 (a) Starting with Equation 7.53, calculate $S^\dagger S$, and confirm Equation 7.57.

(b) Show that $S^\dagger\gamma^0 S = \gamma^0$.

7.14 Show that $\overline{\psi}\gamma^5\psi$ is invariant under the transformation 7.52.

7.15 Show that the adjoint spinors $\overline{u}^{(1,2)}$ and $\overline{v}^{(1,2)}$ satisfy the equations

$$\overline{u}(\gamma^\mu p_\mu - mc) = 0, \quad \overline{v}(\gamma^\mu p_\mu + mc) = 0$$

[*Hint:* Take the transpose conjugate of Equations 7.49 and 7.50; multiply from the right by γ^0, and show that $(\gamma^\mu)^\dagger\gamma^0 = \gamma^0\gamma^\mu$.]

7.16 Show that the normalization condition (Equation 7.43), expressed in terms of the adjoint spinors, becomes

$$\overline{u}u = -\overline{v}v = 2mc$$

7.17 Show that $\overline{\psi}\gamma^\mu\psi$ is a four-vector, by confirming that its components obey the Lorentz transformations (Equation 3.8). Check that it transforms as a (polar) vector under parity

(that is, the 'time' component is invariant, whereas the 'spatial' components switch sign).

7.18 Show that the spinor representing an electron at rest (Equation 7.30) is an eigenstate of the parity operator, P. What is its intrinsic parity? How about the positron? What if you changed the sign convention in Equation 7.61 ? Notice that whereas the *absolute* parity of a spin-$\frac{1}{2}$ particle is in a sense arbitrary, the fact that particles and antiparticles carry *opposite* parity is *not* arbitrary.

7.19 (a) Express $\gamma^\mu\gamma^\nu$ as a linear combination of $1, \gamma^5, \gamma^\mu, \gamma^\mu\gamma^5$, and $\sigma^{\mu\nu}$.
(b) Construct the matrices σ^{12}, σ^{13}, and σ^{23} (Equation 7.69), and relate them to Σ^1, Σ^2, and Σ^3 (Equation 7.51).

7.20 (a) Derive Equations 7.70 (i and iv) from Equation 7.73.
(b) Prove Equation 7.74, from Equation 7.73.

7.21 Show that the continuity equation (Equation 7.74) enforces conservation of charge. [If you don't see how to do this, look in any electrodynamics textbook.]

7.22 Show that we are always free to pick $A^0 = 0$, in free space. That is, given a potential A^μ which does *not* satisfy this constraint, find a gauge function λ, consistent with Equation 7.85, such that A'_0 (in Equation 7.81) is zero.

7.23 Suppose we apply a gauge transformation (Equation 7.81) to the plane-wave potential (Equation 7.89), using as the gauge function

$$\lambda = i\hbar\kappa a e^{-(i/\hbar)p\cdot x}$$

where κ is an arbitrary constant and p is the photon four-momentum.
(a) Show that this λ satisfies Equation 7.85.
(b) Show that this gauge transformation has the effect of modifying $\epsilon^\mu : \epsilon^\mu \to \epsilon^\mu + \kappa p^\mu$. (In particular, if we choose $\kappa = -\epsilon^0/p^0$ we obtain the Coulomb gauge polarization vector, Equation 7.92) This observation leads to a beautifully simple test for the gauge invariance of QED results: the answer must be unchanged if you replace ϵ^μ by $\epsilon^\mu + \kappa p^\mu$.

7.24 Using $u^{(1)}, u^{(2)}$ (Equation 7.46) and $v^{(1)}, v^{(2)}$ (Equation 7.47), prove the completeness relations for spinors (7.99). [*Note:* $u\overline{u}$ is the 4×4 matrix defined by $(u\overline{u})_{ij} \equiv u_i\overline{u}_j$.]

7.25 Using $\epsilon^{(1)}$ and $\epsilon^{(2)}$ (Equation 7.93), confirm the completeness relation for photons (Equation 7.105).

7.26 Evaluate the amplitude for electron–muon scattering (Equation 7.106) in the CM system, assuming the e and μ approach one another along the z axis, repel, and return back along the z axis. Assume the initial and final particles all have helicity $+1$. [*Answer:* $\mathscr{M} = -2g_e^2$]

7.27 Derive the amplitudes (Equations 7.133 and 7.134) for pair annihilation, $e^+ + e^- \to \gamma + \gamma$.

7.28 Work out the analog to Casimir's trick (Equation 7.125) for *antiparticles*

$$\sum_{\text{all spins}} [\overline{v}(a)\Gamma_1 v(b)][\overline{v}(a)\Gamma_2 v(b)]^*$$

and for the 'mixed' cases

$$\sum_{\text{all spins}} [\overline{u}(a)\Gamma_1 v(b)][\overline{u}(a)\Gamma_2 v(b)]^* \quad \text{and} \quad \sum_{\text{all spins}} [\overline{v}(a)\Gamma_1 u(b)][\overline{v}(a)\Gamma_2 u(b)]^*$$

7.29 (a) Show that $\gamma^0\gamma^{\nu\dagger}\gamma^0 = \gamma^\nu$, for $\nu = 0, 1, 2$, and 3.
(b) If Γ is any product of γ matrices ($\Gamma = \gamma_a\gamma_b\cdots\gamma_c$) show that $\overline{\Gamma}$ (Equation 7.119) is the same product in reverse order, $\overline{\Gamma} = \gamma_c\cdots\gamma_b\gamma_a$.

7.30 Use Casimir's trick to obtain an expression analogous to Equation 7.126 for Compton scattering. Note that there are four terms here:

$$|\mathscr{M}|^2 = |\mathscr{M}_1|^2 + |\mathscr{M}_2|^2 + \mathscr{M}_1\mathscr{M}_2^* + \mathscr{M}_1^*\mathscr{M}_2$$

7.31 (a) Prove trace theorems 1, 2, and 3, in Section 7.7.
(b) Prove Equation 4.
(c) Using the anticommutation relation 5, prove 5′.

7.32 (a) Use the anticommutation relation 5 to prove the contraction theorems 6, 7, 8, and 9.
(b) From 7, prove 7′; from 8, prove 8′; from 9, prove 9′.

7.33 (a) Confirm the trace theorems 10, 11, 12, and 13.
(b) From 12, prove 12′; from 13, prove 13′.

7.34 (a) Prove theorems 14, 15, and 16.
(b) From 15, prove 15′; from 16, prove 16′.

7.35 (a) Show that $\epsilon^{\mu\nu\lambda\sigma}\epsilon_{\mu\nu\lambda\tau} = -6\,\delta^\sigma_\tau$ (summation over μ, ν, λ implied).
(b) Show that $\epsilon^{\mu\nu\lambda\sigma}\epsilon_{\mu\nu\theta\tau} = -2(\delta^\lambda_\theta\,\delta^\sigma_\tau - \delta^\lambda_\tau\,\delta^\sigma_\theta)$.
(c) Find the analogous formula for $\epsilon^{\mu\nu\lambda\sigma}\epsilon_{\mu\phi\theta\tau}$.
(d) Find the analogous formula for $\epsilon^{\mu\nu\lambda\sigma}\epsilon_{\omega\phi\theta\tau}$.
[Here δ^μ_ν is the Kronecker delta: 1 if $\mu = \nu$, 0 otherwise. It can also be written in terms of the mixed (co/contravariant) metric tensor: $\delta^\mu_\nu = g^\mu{}_\nu = g_\nu{}^\mu$.]

7.36 Evaluate the following traces:
(a) $\mathrm{Tr}\left[\gamma^\mu\gamma^\nu\left(1-\gamma^5\right)\gamma^\lambda\left(1+\gamma^5\right)\gamma_\lambda\right]$
(b) $\mathrm{Tr}\left[\left(\not{p}+mc\right)\left(\not{q}+Mc\right)\left(\not{p}+mc\right)\left(\not{q}+Mc\right)\right]$, where p is the four-momentum of a (real) particle of mass m and q is the four-momentum of a (real) particle of mass M. Express your answer in terms of m, M, c, and $(p \cdot q)$.

7.37 Starting with Equation 7.107, determine the spin-averaged amplitude, (analogous to Equation 7.129) for elastic electron–electron scattering. Assume we're working at high energies, so that the mass of the electron can be ignored (i.e., set $m = 0$). [*Hint*: You can read $\langle|\mathscr{M}_1|^2\rangle$ and $\langle|\mathscr{M}_2|^2\rangle$ from Equation 7.129 For $\langle\mathscr{M}_1\mathscr{M}_2^*\rangle$ use the same strategy as Casimir's trick to get

$$\langle\mathscr{M}_1\mathscr{M}_2^*\rangle = \frac{-g_e^4}{4(p_1-p_3)^2(p_1-p_4)^2}\mathrm{Tr}(\gamma^\mu\not{p}_1\gamma^\nu\not{p}_4\gamma_\mu\not{p}_2\gamma_\nu\not{p}_3)$$

Then exploit the contraction theorems to evaluate the trace. Notice that for *massless* particles the conservation of momentum $(p_1 + p_2 = p_3 + p_4)$ implies that $p_1 \cdot p_2 = p_3 \cdot p_4$, $p_1 \cdot p_3 = p_2 \cdot p_4$, and $p_1 \cdot p_4 = p_2 \cdot p_3$.]

$$\left[\textit{Answer}: \langle|\mathscr{M}|^2\rangle = \frac{2g_e^4}{(p_1\cdot p_3)^2(p_1\cdot p_4)^2}\left[(p_1\cdot p_2)^4 + (p_1\cdot p_3)^4 + (p_1\cdot p_4)^4\right]\right]$$

7.38 (a) Starting with Equation 7.129, find the spin-averaged amplitude for electron–muon scattering in the CM frame, in the high-energy regime $(m, M \to 0)$.
(b) Find the CM differential cross section for electron–muon scattering at high energy. Let E be the electron energy and θ the scattering angle.

$$\left[\textit{Answer}: \frac{d\sigma}{d\Omega} = \left(\frac{\hbar c}{8\pi}\right)^2\frac{g_e^4}{2E^2}\left(\frac{1+\cos^4\theta/2}{\sin^4\theta/2}\right)\right]$$

7.39 (a) Using the result of Problem 7.37, determine the spin-averaged amplitude for electron–electron scattering in the CM in the high-energy regime $(m \to 0)$.

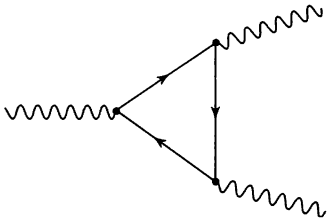


Fig. 7.12 Decay of the photon, $\gamma \to 2\gamma$ – a process forbidden by Furry's theorem (Problem 7.46).

(b) Find the CM differential cross section for electron–electron scattering at high energy.

$$\left[\textit{Answer}: \frac{d\sigma}{d\Omega} = \left(\frac{\hbar c}{8\pi}\right)^2 \frac{g_e^4}{2E^2}\left(1 - \frac{4}{\sin^2\theta}\right)^2\right]$$

Compare your answer to Problem 7.38 (see footnote to Example 7.3).

7.40 Starting with Equation 7.158, calculate $|\mathscr{M}|^2$, and use Equation 7.105 to sum over photon polarizations. Check that the answer is consistent with Equation 7.163, and explain why this method gives the correct answer (note that we are now summing over *all* photon polarizations, whereas in *fact* the photons must be in the singlet configuration).

7.41 Starting with Equation 7.149, calculate $\langle|\mathscr{M}|^2\rangle$ for $e^+ + e^- \to \gamma + \gamma$, and use it to get the differential cross section for pair annihilation. Compare Equation 7.167 (see footnote before Example 7.8).

7.42 Derive Equation 7.176 You'll need one last Feynman rule: for a closed fermion loop include a factor -1 and take the trace.

7.43 Derive Equation 7.182 [*Hint*: use the integral theorems in Appendix E of Sakurai [6].]

7.44 Derive Equation 7.187.

7.45 Evaluate the correction term in Equation 7.192 for the case of a head-on collision in the CM; assume the electron is traveling at $\frac{1}{10}c$. In the experiment [9], the beam energies were 57.8 GeV; what should the measured fine structure 'constant' have been? Look up the actual result, and compare it with your prediction.

7.46 Why can't the photon 'decay', by the process $\gamma \to \gamma + \gamma$ (Figure 7.12)? Calculate the amplitude for this diagram. [This is an example of *Furry's theorem*, which says that any diagram containing a closed electron loop with an odd number of vertices has an amplitude of zero.]

7.47 Starting with your answer to Problem 7.30, derive the *Klein–Nishina formula* for Compton scattering (in the rest frame of the target):

$$\frac{d\sigma}{d\Omega} = \frac{\pi\alpha^2}{m^2}\left(\frac{\omega'}{\omega}\right)^2\left[\frac{\omega'}{\omega} + \frac{\omega}{\omega'} - \sin^2\theta\right]$$

where ω and ω' are the frequencies of the incident and scattered photons (Problem 3.27).

Problems 48–50 pertain to the following model: Imagine that the photon, instead of being a massless vector (spin 1) particle, were a massive scalar (spin 0). Specifically, suppose the QED vertex factor were

$$ig_e 1$$

(where 1 is the 4 × 4 unit matrix), and the 'photon' propagator were

$$\frac{-i}{q^2 - (m_\gamma c)^2}.$$

There is no photon polarization vector now, and hence no factor for external photon lines. Apart from that, the Feynman rules for QED are unchanged.

7.48 Assuming it is heavy enough, this 'photon' can decay.

(a) Calculate the decay rate for $\gamma \to e^+ + e^-$.

(b) If $m_\gamma = 300\ \mathrm{MeV}/c^2$, find the lifetime of the 'photon', in seconds.

7.49 (a) Find the amplitude, $\mathscr{M}$, for electron–muon scattering, in this theory.

(b) Calculate the spin-averaged quantity, $\langle|\mathscr{M}|^2\rangle$.

(c) Determine the differential cross section for electron–muon scattering in the CM frame. Assume the energy is high enough so that the electron and muon masses can be neglected: $m_e, m_\mu \to 0$. Express your answer in terms of the incident electron energy, E and the scattering angle, θ.

(d) From your result in (c), calculate the total cross section, assuming the 'photon' is extremely heavy, $m_\gamma c^2 \gg E$.

(e) Going back to (b), consider now the case of *low*-energy scattering from an extremely heavy 'muon': $|\mathbf{p}_e|/c \ll m_e \ll m_\gamma \ll m_\mu$. Find the differential cross section in the *lab* frame (muon at rest), assuming the muon does not recoil appreciably. Compare the Rutherford formula (Example 7.7), and calculate the total cross section. [Actually, if you set $m_\gamma \to 0$ and then take $|\mathbf{p}| \ll mc$, you get *precisely* the Rutherford formula.]

7.50 (a) Find the amplitude, $\mathscr{M}$, for pair annihilation ($e^+ + e^- \to \gamma + \gamma$), in this theory.

(b) Determine $\langle|\mathscr{M}|^2\rangle$, assuming the energy is high enough that we can ignore both the electron and the 'photon' mass ($m_e, m_\gamma \to 0$).

(c) Evaluate your result, in (b), in the CM system. Express your answer in terms of the incident electron energy, E, and the scattering angle, θ.

(d) Find the differential cross section for pair annihilation, in the CM system, still assuming $m_e = m_\gamma = 0$. Is the total cross section finite?

7.51 Spin-$\frac{1}{2}$ particles that are electrically neutral could conceivably be their own antiparticles (if so, they are called "Majorana" fermions – in the Standard Model the only possible candidates are the neutrinos)

(a) According to Problem 7.9, the charge conjugate spinor is $\psi_c = i\gamma^2\psi^*$. Evidently, if a particle is the same as its antiparticle, then $\psi = \psi_c$. Show that this condition is Lorentz invariant (if true in one inertial frame, it is true in any inertial frame). [*Hint:* Use Equations 7.52 and 7.53.]

(b) Show that if $\psi = \psi_c$, the "lower" two elements of ψ are related to the "upper" two by $\psi_B = -i\sigma_y\psi_A^*$. For Majorana particles, then, we only need a two-component spinor, $\chi \equiv \psi_A$. This makes sense: A Dirac spinor takes four elements to represent the two spin states (each) of the particle and the antiparticle, but in this case the latter two are redundant. Show that the Dirac equation for a Majorana particle can be written in 2-component form as

$$i\hbar\left[\partial_0\chi + i(\sigma\cdot\nabla)\sigma_y\chi^*\right] - mc\chi = 0$$

Check that the equation you get for the "lower" elements is consistent with this.

(c) Construct spinors χ representing plane wave Majorana states. [*Hint:* Form the general linear combination $\psi = a_1\psi^{(1)} + a_2\psi^{(2)} + a_3\psi^{(3)} + a_4\psi^{(4)}$ (Equations 7.46 and 7.47), impose the constraint in part **(b)**, and solve for a_3 and a_4 (in terms of a_1 and a_2); then pick (say) $a_1 = 1$, $a_2 = 0$ for $\chi^{(1)}$, and $a_1 = 0$, $a_2 = 1$ for $\chi^{(2)}$.]

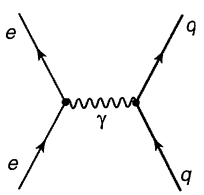
8

Electrodynamics and Chromodynamics of Quarks

Because the electromagnetic interactions of electrons are well understood, they serve as useful probes of the structure of hadrons. Everything I said in Chapter 7 about leptons applies just as well to quarks (using, of course, the appropriate charge: $\frac{2}{3}e$ or $-\frac{1}{3}e$). However, the experimental situation is complicated by the fact that the quarks themselves never see the light of day, and we are obliged to infer from the observed behavior of mesons and baryons what their constituents are up to. In this chapter we shall consider two important examples: the production of hadrons in electron–positron collisions (Section 8.1), and elastic electron–proton scattering (Section 8.2). We then turn to chromodynamics: the Feynman rules (Section 8.3), color factors (Section 8.4), pair annihilation in QCD (Section 8.5), and asymptotic freedom (Section 8.6).

8.1
Hadron Production in e^+e^- Collisions

When electrons and positrons collide, they can (of course) scatter elastically, $e^+ + e^- \to e^+ + e^-$ (Bhabha scattering), or they could produce two photons, $e^+ + e^- \to \gamma + \gamma$ (pair annihilation), or – if the energy is sufficiently high – they could make a pair of muons (or taus), $e^+ + e^- \to \mu^+ + \mu^-$. But they can also produce a pair of quarks: $e^+ + e^- \to q + \bar{q}$, and it is this process that I want to consider next. The lowest-order QED diagram is



For a brief moment the quarks fly apart as free particles, but when they reach a separation distance of around 10^{-15} m (the diameter of a hadron), their (strong) interaction is so great that new quark–antiquark pairs are produced – this time mainly from gluons (Figure 8.1). These quarks and antiquarks (literally *dozens* of

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

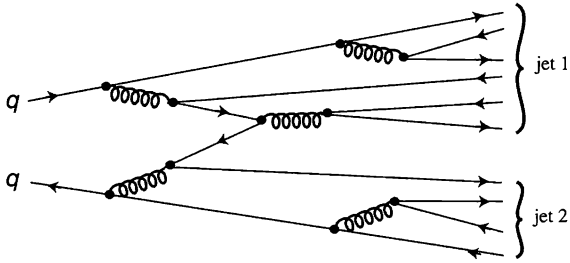
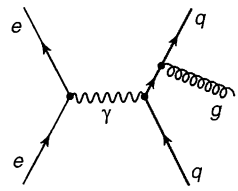


Fig. 8.1 Hadronization and jet formation.

them, in a typical modern experiment) join together in myriad combinations to make the mesons and baryons that are actually recorded in the detector – a process known as 'hadronization'. What we observe in the laboratory, then, is $e^+ + e^- \to$ hadrons.

In all the debris there is often an unmistakable footprint left behind by the original quark–antiquark pair: the hadrons emerge in two back-to-back 'jets', one along the direction of the primordial quark,* the other along the direction of the antiquark (Figure 8.2). Sometimes one sees a three-jet event (Figure 8.3), indicating that a gluon carrying a substantial fraction of the total energy was emitted in conjunction with the original $q\bar{q}$ production:



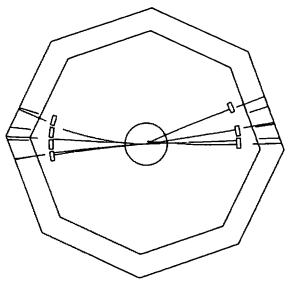


Fig. 8.2 A typical two-jet event. (*Source*: J. Dorfan, SLAC.)

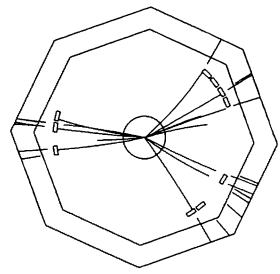
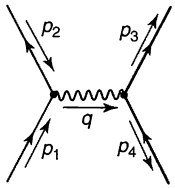


Fig. 8.3 A three-jet event. (*Source*: J. Dorfan, SLAC.)

* Notice that the quark (say) has to 'reach back' and pick up an antiquark from the other branch, to make each jet colorless, but as long as the energy transferred is relatively small this does not disrupt the jet structure.

Indeed, the observation of three-jet events is generally regarded as our most direct evidence for the existence of gluons.

Now, the *first* stage in all this ($e^+ + e^- \to \gamma \to q + \bar{q}$) is ordinary QED; the calculation is exactly the same as for $e^+ + e^- \to \mu^+ + \mu^-$:



The amplitude is

$$\mathscr{M} = \frac{Qg_e^2}{(p_1+p_2)^2}[\bar{v}(p_2)\gamma^\mu u(p_1)][\bar{u}(p_3)\gamma_\mu v(p_4)] \tag{8.1}$$

where Q is the quark charge, in units of e ($\frac{2}{3}$, for u, c, and t; $-\frac{1}{3}$ for d, s, and b). Exploiting Casimir's trick, we obtain

$$\begin{aligned}\langle|\mathscr{M}|^2\rangle = \frac{1}{4}\left[\frac{Qg_e^2}{(p_1+p_2)^2}\right]^2 &\mathrm{Tr}[\gamma^\mu(\not{p}_1+mc)\gamma^\nu(\not{p}_2-mc)]\\ &\times\mathrm{Tr}[\gamma_\mu(\not{p}_4-Mc)\gamma_\nu(\not{p}_3+Mc)]\end{aligned} \tag{8.2}$$

where m is the mass of the electron and M is that of the quark (Problem 8.1). Invoking the trace theorems of Chapter 7, we can reduce this to

$$\begin{aligned}\langle|\mathscr{M}|^2\rangle = 8\left[\frac{Qg_e^2}{(p_1+p_2)^2}\right]^2 &[(p_1\cdot p_3)(p_2\cdot p_4)+(p_1\cdot p_4)(p_2\cdot p_3)\\ &+(mc)^2(p_3\cdot p_4)+(Mc)^2(p_1\cdot p_2)+2(mc)^2(Mc)^2]\end{aligned} \tag{8.3}$$

or, in terms of the incident (CM) electron energy E and the angle θ between the incoming electron and the outgoing quark:

$$\begin{aligned}\langle|\mathscr{M}|^2\rangle = Q^2g_e^4\Bigg\{&1+\left(\frac{mc^2}{E}\right)^2+\left(\frac{Mc^2}{E}\right)^2\\ &+\left[1-\left(\frac{mc^2}{E}\right)^2\right]\left[1-\left(\frac{Mc^2}{E}\right)^2\right]\cos^2\theta\Bigg\}\end{aligned} \tag{8.4}$$

The differential scattering cross section is given by Equation 6.47; integrating over θ and ϕ, we obtain the total cross section (Problem 8.2):

$$\sigma = \frac{\pi Q^2}{3}\left(\frac{\hbar c\alpha}{E}\right)^2\sqrt{\frac{1-(Mc^2/E)^2}{1-(mc^2/E)^2}}\left[1+\frac{1}{2}\left(\frac{Mc^2}{E}\right)^2\right]\left[1+\frac{1}{2}\left(\frac{mc^2}{E}\right)^2\right] \tag{8.5}$$

Notice the threshold at $E = Mc^2$; for energies less than this the square root is imaginary, reflecting the fact that the process is kinematically forbidden when there is not enough energy to create the $q\bar{q}$ pair. If we are substantially above threshold ($E > Mc^2 \gg mc^2$), Equation 8.5 simplifies considerably:*

$$\sigma = \frac{\pi}{3}\left(\frac{\hbar Qc\alpha}{E}\right)^2 \tag{8.6}$$

As we crank up the beam energy, we encounter a succession of such thresholds – first the muon and the light quarks, later (at about 1300 MeV) the charm quark, the tau (at 1777 MeV), the bottom quark (4500 MeV), and eventually the top quark. There is a beautiful way to display this structure: suppose we examine the *ratio* of the rate of hadron production to that for muon pairs:

$$R \equiv \frac{\sigma(e^+e^- \to \text{ hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)} \tag{8.7}$$

Since the numerator includes all the quark–antiquark events,† Equation 8.6 gives

$$R(E) = 3\sum Q_i^2 \tag{8.8}$$

in which the sum is over all quark flavors with thresholds below E. Notice the 3 in front – it records the fact that there are three *colors* for each flavor. We anticipate a 'staircase' graph for $R(E)$, then, ascending one step at each new quark threshold, with the height of the rise determined by the quark's charge. At low energy where only the u, d, and s quarks contribute, we expect

$$R = 3\left[\left(\frac{2}{3}\right)^2 + \left(-\frac{1}{3}\right)^2 + \left(-\frac{1}{3}\right)^2\right] = 2 \tag{8.9}$$

Between the c threshold and the b threshold we should have

$$R = 2 + 3\left(\frac{2}{3}\right)^2 = \frac{10}{3} = 3.33 \tag{8.10}$$

at the b threshold it goes up slightly,

$$R = \frac{10}{3} + 3\left(-\frac{1}{3}\right)^2 = \frac{11}{3} = 3.67 \tag{8.11}$$

and the top quark should produce a jump to $R = 5$.

* This approximation is actually better than it looks, because of a lucky algebraic cancellation: expanding the radical, $\sqrt{1-(Mc^2/E)^2}[1+\frac{1}{2}(Mc^2/E)^2] = 1 - \frac{3}{8}(Mc^2/E)^4 \ldots$, so the error is of order $(Mc^2/E)^4$, not $(Mc^2/E)^2$. As for the *electron* mass terms, these are smaller to begin with, though there is a second-order correction; however, these terms cancel exactly in the calculation of R (Equation 8.7).

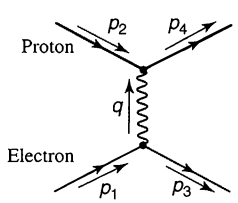
† The τ lepton decays predominantly into hadrons, and this adds a bit to R, above 1777 MeV; that's why the experimental numbers are somewhat above the '$u + d + s + c$' line in Figure 8.4

The experimental results are shown in Figure 8.4. The agreement between theory and experiment is pretty good, especially at high energy. But you may well ask why it is not *perfect*. Apart from the approximation in going from Equation 8.5 to Equation 8.6 (which artificially sharpens the corners at each threshold), and the neglect of the tau, we have made a fundamental oversimplification in assuming that we could treat the process as a sequence of two independent operations: $e^+e^- \to q\bar{q}$ (QED) followed by $q\bar{q} \to$ hadrons (QCD). In point of fact, the quarks produced in the first step are *not* free particles, obeying the Dirac equation; rather, they are *virtual* particles, on their way to a second interaction. This is particularly critical when the energy is right for formation of a bound state ($\phi = s\bar{s}$, $\psi = c\bar{c}$, $\Upsilon = b\bar{b}$); in the vicinity of such a 'resonance', the interaction of the two quarks can scarcely be ignored. Hence the sharp spikes in the graph, which typically occur just below each threshold. Finally, above about 50 GeV, the graph starts to rise toward the Z^0 peak, at 91 GeV.

But, really, all this is quibbling anyway, for the importance of Figure (8.4) lies not in what the small discrepancies whisper, but in what the overall agreement shouts: the factor of 3 in Equation 8.8 clearly belongs there. Without it the theory would be wildly off (look at the dashed line in Figure 8.4) – and not just at isolated resonances, but across-the-board. That 3, remember, counts the number of colors. Here, then, is compelling *experimental* evidence for the color hypothesis – a hypothesis that was introduced originally for esoteric theoretical reasons but is now an indispensable ingredient in the recipe for strong interactions.

8.2 Elastic Electron–Proton Scattering

We now turn to electron–proton scattering, our best probe of the internal structure of the proton. If the proton were a simple point charge, obeying the Dirac equation, we could just copy our analysis of electron–muon scattering, with M now the mass of the proton. The lowest-order Feynman diagram would be



and the (spin-averaged) amplitude would be (Equation 7.126)

$$\langle |\mathscr{M}|^2 \rangle = \frac{g_e^4}{q^4} L^{\mu\nu}_{\text{electron}} L_{\mu\nu\ \text{proton}} \tag{8.12}$$

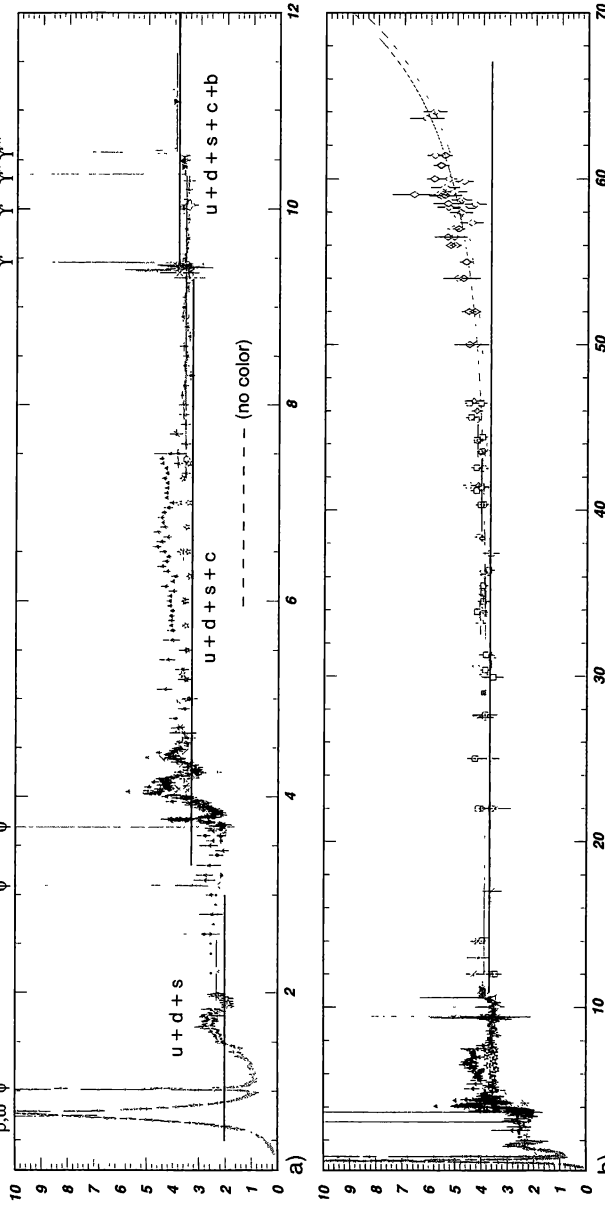


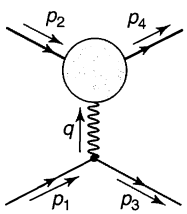
Fig. 8.4 Graph of R, based on experimental data, plotted against total energy ($2E$), in GeV. (Courtesy of the COMPAS (IHEP, Russia) and HEPDATA (Durham, UK) groups, with corrections by P. Janot (CERN) and M. Schmitt (Northwestern).)

where $q = p_1 - p_3$ and (Equation 7.128)

$$L^{\mu\nu}_{\text{electron}} = 2\left\{p_1^\mu p_3^\nu + p_1^\nu p_3^\mu + g^{\mu\nu}[(mc)^2 - (p_1 \cdot p_3)]\right\} \tag{8.13}$$

(and a similar expression for $L^{\mu\nu}_{\text{proton}}$, only with $m \to M$ and $1, 3 \to 2, 4$). We used these results in Example 7.7 to derive the Mott and Rutherford scattering formulas.

But the proton is *not* a simple point charge, and so, long before the advent of the quark model, a more flexible formalism was introduced for describing electron–proton scattering. We might represent the process, in lowest-order QED, by a diagram like this:



where the blob serves to remind us that we don't really know how the (virtual) photon interacts with the proton. (However, we *do* assume, that the scattering is *elastic*: $e + p \to e + p$; *inelastic* electron–proton scattering, $e + p \to e + X$, is much more complicated, and we will not consider it in this book.) Now, the essential point is that the *electron* vertex and the photon propagator are unchanged, and therefore, since $\langle|\mathscr{M}|^2\rangle$ neatly factors (Equation 8.12),

$$\langle|\mathscr{M}|^2\rangle = \frac{g_e^4}{q^4} L^{\mu\nu}_{\text{electron}} K_{\mu\nu\ \text{proton}} \tag{8.14}$$

where $K_{\mu\nu}$ is an unknown quantity describing the photon–proton vertex.

Well ... not *completely* unknown, for this much we can say: it is certainly a second-rank tensor, and the only variables that it can possibly depend on are p_2, p_4, and q. Since $q = p_4 - p_2$, these three are not independent, and we are free to use any two of them; the customary choice is q and p_2 (I'll drop the subscript from here on: $p \equiv p_2$ is the initial proton momentum). Now, there aren't many tensors that can be constructed out of just two four-vectors; the most general possible form is

$$K^{\mu\nu} = -K_1 g^{\mu\nu} + \frac{K_2}{(Mc)^2} p^\mu p^\nu + \frac{K_4}{(Mc)^2} q^\mu q^\nu + \frac{K_5}{(Mc)^2}(p^\mu q^\nu + p^\nu q^\mu) \tag{8.15}$$

where the K_i are (unknown) functions of the only scalar variable in the problem: q^2.* The factors $(Mc)^{-2}$ have been pulled out, in defining K_2, K_4, and K_5, just so all the K's will have the same dimensions.† In principle, we could add an

* Notice that $p^2 = (Mc)^2$ is a constant, and $p_4^2 = (q + p)^2 = q^2 + 2q \cdot p + p^2 = (Mc)^2 \Rightarrow q \cdot p = -q^2/2$.

† The subscript 3 is traditionally reserved for a term that enters in the corresponding analysis of *neutrino*–proton scattering, but does not occur here.

antisymmetric combination $(p^\mu q^\nu - p^\nu q^\mu)$, but since $L^{\mu\nu}$ is symmetric (Equation 8.13), such a term would contribute nothing to $\langle|\mathscr{M}|^2\rangle$. Now, these four functions are not independent; it can be shown (Problem 8.4) that

$$q_\mu K^{\mu\nu} = 0 \tag{8.16}$$

from which it follows (Problem 8.5) that

$$K_4 = \frac{(Mc)^2}{q^2} K_1 + \frac{1}{4} K_2 \quad \text{and} \quad K_5 = \frac{1}{2} K_2 \tag{8.17}$$

Thus $K^{\mu\nu}$ can be expressed in terms of just *two* (unknown) functions, $K_1(q^2)$ and $K_2(q^2)$:

$$K^{\mu\nu} = K_1 \left(-g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2} \right) + \frac{K_2}{(Mc)^2} \left(p^\mu + \frac{1}{2} q^\mu \right) \left(p^\nu + \frac{1}{2} q^\nu \right) \tag{8.18}$$

The 'form factors' K_1 and K_2 are directly related to the electron–proton elastic scattering cross section. According to Equations 8.13 and (8.18) (Problem 8.7)

$$\langle|\mathscr{M}|^2\rangle = \left(\frac{2g_e^2}{q^2} \right)^2 \left\{ K_1[(p_1 \cdot p_3) - 2(mc)^2] + K_2 \left[\frac{(p_1 \cdot p)(p_3 \cdot p)}{(Mc)^2} + \frac{q^2}{4} \right] \right\} \tag{8.19}$$

We shall work in the laboratory frame, with the target proton at rest, $p = (Mc, 0, 0, 0)$. An electron with incident energy E scatters at an angle θ, emerging with energy E'. Let us assume it's a moderately energetic collision $(E, E') \gg mc^2$, so that we can safely ignore the mass of the electron (set $m = 0$);* then $p_1 = (E/c)(1, \hat{\mathbf{p}}_i)$ and $p_3 = (E'/c)(1, \hat{\mathbf{p}}_f)$, with $\hat{\mathbf{p}}_i \cdot \hat{\mathbf{p}}_f = \cos\theta$, and we find (Problem 8.8)

$$\langle|\mathscr{M}|^2\rangle = \frac{g_e^4 c^2}{4EE' \sin^4(\theta/2)} \left(2K_1 \sin^2 \frac{\theta}{2} + K_2 \cos^2 \frac{\theta}{2} \right) \tag{8.20}$$

The outgoing electron energy, E', is not an independent variable; it is kinematically determined by E and θ (Problem 8.9):

$$E' = \frac{E}{1 + (2E/Mc^2) \sin^2(\theta/2)} \tag{8.21}$$

For a massless incident particle we have (Problem 6.10)

$$\frac{d\sigma}{d\Omega} = \left(\frac{\hbar E'}{8\pi McE} \right)^2 \langle|\mathscr{M}|^2\rangle \tag{8.22}$$

* The Mott formula (Equation 7.131) neglects proton structure and proton recoil; it applies to the regime $E \ll Mc^2$, but it does not assume $E \gg mc^2$. We now work in the regime $E \gg mc^2$, but do *not* ignore proton structure and recoil (i.e. we do *not* assume $E \ll Mc^2$). In the intermediate range, $mc^2 \ll E \ll Mc^2$, the two results agree (Problem 8.10).

and so, for elastic electron–proton scattering

$$\frac{d\sigma}{d\Omega} = \left(\frac{\alpha\hbar}{4ME\sin^2(\theta/2)}\right)^2 \frac{E'}{E}[2K_1\sin^2(\theta/2) + K_2\cos^2(\theta/2)] \tag{8.23}$$

where E' is given by Equation 8.21 This is known as the *Rosenbluth formula*; it was first derived in 1950 [1]. By counting the number of electrons scattered in a given direction, for a range of incident energies, we can determine $K_1(q^2)$ and $K_2(q^2)$ experimentally. Actually, it is traditional to work instead with the 'electric' and 'magnetic' form factors, $G_E(q^2)$ and $G_M(q^2)$:

$$K_1 = -q^2 G_M^2, \quad K_2 = (2Mc)^2 \frac{G_E^2 - [q^2/(2Mc)^2]G_M^2}{1 - [q^2/(2Mc)^2]} \tag{8.24}$$

G_E and G_M are related to the charge and magnetic moment distributions of the proton, respectively [2].

There is precious little *physics* in all of this; what we have done is to *set the agenda* for a model of the proton. A successful theory must enable us to *calculate* the form factors, which at this stage are completely arbitrary. The most naive model treats the proton as a simple point charge; in this case (Problem 8.6)

$$K_1 = -q^2, \quad K_2 = (2Mc)^2 \quad \Rightarrow G_E = G_M = 1 \tag{8.25}$$

It's not a bad approximation at low energies, where the electron never gets close enough to 'see' inside the proton. But it is grossly inadequate at high energies (Figure 8.5). Evidently the proton has a rich internal structure. That's no surprise in light of the quark model, but it would shock anyone who still thinks the proton is a truly elementary particle.

8.3 Feynman Rules For Chromodynamics

Quantum *electro*dynamics (QED) describes the interactions of *charged* particles; quantum *chromo*dynamics (QCD) describes the interactions of *colored* particles. Electromagnetic interactions are mediated by *photons*, chromodynamic interactions by *gluons*. The strength of the electromagnetic force is set by the coupling constant

$$g_e = \sqrt{4\pi\alpha} \tag{8.26}$$

In appropriate units g_e is the fundamental charge (the charge of the positron). The strength of the chromodynamic force is set by the 'strong' coupling constant

$$g_s = \sqrt{4\pi\alpha_s} \tag{8.27}$$

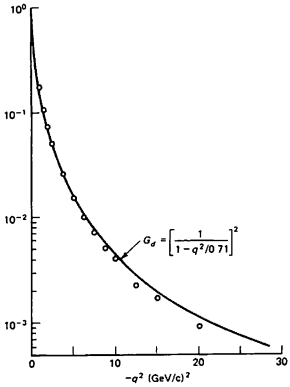


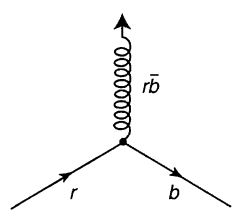
Fig. 8.5 Proton elastic form factors. Apart from an overall constant, the electric and magnetic form factors G_E and G_M are practically identical, and – at least, up to about 10 $(\text{GeV}/c)^2$ – are well fit by the phenomenological 'dipole' function G_d (solid line). Circles are experimental values of $G_M/(1+K)(\approx G_E)$. [*Source*: Frauenfelder, H. and Henley, E. M. (1991) *Subatomic Physics*, 2nd edn, Prentice-Hall, Englewood Cliffs, NJ, p. 141. Based on data of Kirk, P. N. *et al.*, (1973) *Physical Review*, **D8**, 63.]

which may be thought of as the fundamental unit of color. Quarks come in three colors,* 'red' (r), 'blue' (b), and 'green' (g). Thus the specification of a quark state in QCD requires not only the Dirac spinor $u^{(s)}(p)$, giving its momentum and spin, but also a three-element column vector c, giving its color:

$$c = \begin{pmatrix}1\\0\\0\end{pmatrix} \text{ for red,} \quad \begin{pmatrix}0\\1\\0\end{pmatrix} \text{ for blue,} \quad \begin{pmatrix}0\\0\\1\end{pmatrix} \text{ for green} \tag{8.28}$$

I'll label the elements of c by a Roman subscript near the middle of the alphabet – c_i, for example – so that $i, j, k, \ldots$ run from 1 to 3 over quark colors.†

Typically, quark color changes at a quark–gluon vertex, and the difference is carried off by the gluon. For example:



* Quarks also come in different *flavors*, of course, but this is irrelevant in QCD, except insofar as the different quark flavors carry different *masses*. Just as QED only looks at the *charge* of a particle, QCD cares only about its *color*.

† I should perhaps warn you that most books do not specify quark color states explicitly; they are 'implied', or 'understood to be contained in $u(p)$'. I think it is wiser at this stage to write them out explicitly, even at the cost of some extra notational baggage.

In this diagram, a red quark turned into a blue quark emitting a red-antiblue gluon. Each gluon carries one unit of color and one unit of anticolor. It would appear, then, that there should be nine species of gluons – $r\bar{r}$, $r\bar{b}$, $r\bar{g}$, $b\bar{r}$, $b\bar{b}$, $b\bar{g}$, $g\bar{r}$, $g\bar{b}$, $g\bar{g}$. Such a nine-gluon theory is perfectly possible in *principle*, but it would describe a world very different from our own. In terms of color $SU(3)$ symmetry (on which, as we shall see, QCD is based), these nine states constitute a 'color octet':

$$\left\{\begin{array}{ll} |1\rangle = (r\bar{b} + b\bar{r})/\sqrt{2} & |5\rangle = -i(r\bar{g} - g\bar{r})/\sqrt{2} \\ |2\rangle = -i(r\bar{b} - b\bar{r})/\sqrt{2} & |6\rangle = (b\bar{g} + g\bar{b})/\sqrt{2} \\ |3\rangle = (r\bar{r} - b\bar{b})/\sqrt{2} & |7\rangle = -i(b\bar{g} - g\bar{b})/\sqrt{2} \\ |4\rangle = (r\bar{g} + g\bar{r})/\sqrt{2} & |8\rangle = (r\bar{r} + b\bar{b} - 2g\bar{g})/\sqrt{6} \end{array}\right\} \tag{8.29}$$

and a 'color singlet':

$$|9\rangle = (r\bar{r} + b\bar{b} + g\bar{g})/\sqrt{3} \tag{8.30}$$

(See Section 5.5; there we were concerned with *flavor*, not color, but the mathematics is identical – just let u, d, $s \to r$, b, g. We're not dealing with isotopic spin, here, and I have used different linear combinations of states within the octet. This simplifies the notation later on.) If the singlet gluon existed, it would be as common and conspicuous as the photon.* Confinement requires that all naturally occurring particles be color singlets, and this 'explains' why the octet gluons never appear as free particles.† But $|9\rangle$ *is* a color singlet, and if it exists as a mediator it should also occur as a free particle. Moreover, it could be exchanged between two color singlets (a proton and a neutron, say), giving rise to a long-range force with strong coupling,‡ whereas in fact we know that the strong force is

* Maybe the 'ninth gluon' *is* the photon! That would make for a beautiful unification of the strong and electromagnetic interactions. Of course, the coupling strength isn't quite right, but that's a problem with all unification schemes, and could presumably be managed. There's a much more serious difficulty with this idea, which I'll let you figure out (see Problem 8.10).

† Notice the distinction between 'colorless' and 'color singlet'. Gluons $|3\rangle$ and $|8\rangle$ are colorless, in the sense that the net amount of each color is zero, but they are not color singlets. This situation has an analog in the theory of spin: we can have a state with $S_z = 0$, but this does not prove it has spin 0 (although spin 0 certainly implies $S_z = 0$, and by the same token a color singlet is necessarily colorless). Many authors use the word 'colorless' to *mean* 'color singlet', but this can lead to misunderstanding. (I was sloppy myself, back in Chapters 1 and 2, because at that stage it was not possible to explain the idea of a color singlet.) You might prefer the word 'color-invariant' (instead of 'color singlet') or even 'color scalar'; the essential point is that such a state is unaffected by the transformations of color $SU(3)$ (see Problem 8.12).

‡ Because gluons are massless, they mediate a force of infinite range (the same as electrodynamics). In this sense the force between two quarks is actually *long* range. However, confinement, and the absence of a singlet gluon, conceals this from us. A singlet state (such as the proton) can only emit and absorb a singlet (such as the pion), so individual gluons cannot be exchanged between a proton and a neutron. That's why the force we *observe* is of *short* range. If the singlet gluon existed, it *could* be exchanged between singlets, and the strong force would have a component of infinite range.

of very *short* range. In our world, then, there are evidently only eight kinds of gluons.*

Like the photon, gluons are massless particles of spin 1; they are represented by a polarization vector, ϵ^μ, which is orthogonal to the gluon momentum, p:

$$\epsilon^\mu p_\mu = 0 \quad \text{(Lorentz condition)} \tag{8.31}$$

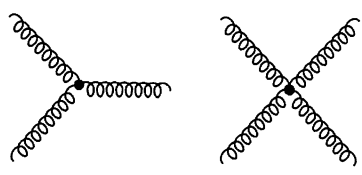
As before, we adopt the Coulomb gauge:†

$$\epsilon^0 = 0, \quad \text{so that} \quad \boldsymbol{\epsilon} \cdot \mathbf{p} = 0 \tag{8.32}$$

This spoils manifest Lorentz covariance, but it cannot be helped (see Section 7.4). In order to describe the color state of the gluon, we need in addition an eight-element column vector, a:

$$a = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \text{ for } |1\rangle, \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \text{ for } |7\rangle, \quad \text{and so on} \tag{8.33}$$

Elements of a will be labeled by a Greek superscript near the front of the alphabet (a^α): $\alpha, \beta, \gamma, \ldots$ run from 1 to 8 over gluon color states. Because the gluons themselves carry color (in contrast to the photon, which is electrically neutral), they couple directly to one another. In fact, there is a three-gluon vertex and a four-gluon vertex:



Before I can state the Feynman rules for QCD, I need to introduce two pieces of notation. First, the Gell-Mann 'λ-matrices', which are to $SU(3)$ what the Pauli spin

* In group-theoretical terms, the issue here is whether the symmetry of QCD is $U(3)$ (which would require all nine gluons) or $SU(3)$ (which calls for only eight). The experimental situation resolves the question decisively in favor of the latter.

† There is a subtle problem here, because gauge transformations in chromodynamics are more complicated than Equation 7.81, and in fact the Coulomb gauge *cannot* be consistently imposed. However, the *correction* to Equation 7.81 contains a factor of g_s, and hence, in the Feynman calculus, the 'error' introduced by using the Coulomb gauge can be compensated for by appropriate modification of the rules for computing higher-order (loop) diagrams.

matrices are to $SU(2)$:

$$\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \lambda^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad \lambda^5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \qquad \lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{8.34}$$

$$\lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \qquad \lambda^8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Second, the *commutators* of the λ matrices define the 'structure constants' ($f^{\alpha\beta\gamma}$) of the group $SU(3)$:

$$[\lambda^\alpha, \lambda^\beta] = 2if^{\alpha\beta\gamma}\lambda^\gamma \tag{8.35}$$

(summation over γ – from 1 to 8 – implied by the repeated index). The structure constants are completely antisymmetric, $f^{\beta\alpha\gamma} = f^{\alpha\gamma\beta} = -f^{\alpha\beta\gamma}$. You can work them out for yourself (Problem 8.15). Since each index runs from 1 to 8, there are $8 \times 8 \times 8 = 512$ structure constants in all, but most of them are zero, and the rest can be obtained by antisymmetry from the following set:

$$f^{123} = 1, \quad f^{147} = f^{246} = f^{257} = f^{345} = f^{516} = f^{637} = \frac{1}{2},$$
$$f^{458} = f^{678} = \sqrt{3}/2 \tag{8.36}$$

I can now state the Feynman rules for evaluating tree-level diagrams in QCD:

1. *External Lines*. For an external quark with momentum p, spin s, and color c:

$$\textit{Quark}: \begin{Bmatrix} \text{Incoming } (\rightarrow\!\bullet) : u^{(s)}(p)c \\ \text{Outgoing } (\bullet\!\rightarrow) : \overline{u}^{(s)}(p)c^\dagger \end{Bmatrix} \tag{8.37}$$

(note that $c^\dagger = \tilde{c}^*$ will be a *row* matrix). For an external *anti*quark:

$$\textit{Antiquark}: \begin{Bmatrix} \text{Incoming } (\leftarrow\!\bullet) : \overline{v}^{(s)}(p)c^\dagger \\ \text{Outgoing } (\bullet\!\leftarrow) : v^{(s)}(p)c \end{Bmatrix} \tag{8.38}$$

where c represents the color of the corresponding *quark*. For an external gluon of momentum p, polarization ϵ, and color

a, include a factor

$$\text{Gluon}: \left\{ \begin{array}{l} \text{Incoming}(\xrightarrow{\alpha,\,\mu} \bullet) : \epsilon_\mu(p)a^\alpha \\ \text{Outgoing}(\bullet \xleftarrow{\alpha,\,\mu}) : \epsilon^*_\mu(p)a^{\alpha *} \end{array} \right\} \tag{8.39}$$

(To avoid confusion it is helpful to indicate on the diagram the indices – space-time and color – you are using for each gluon.)

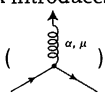
2. *Propagators.* Each internal line contributes a factor

$$\text{Quarks and antiquarks: } (\bullet \xrightarrow{q} \bullet) : \frac{i(\not{q} + mc)}{q^2 - m^2c^2} \tag{8.40}$$

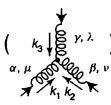
$$\text{Gluons: } (\underset{\alpha,\,\mu}{\bullet} \overset{q}{\sim\sim} \underset{\beta,\,\nu}{\bullet}) : \frac{-ig_{\mu\nu}\delta^{\alpha\beta}}{q^2} \tag{8.41}$$

3. *Vertices.* Each vertex introduces a factor

$$\text{Quark-gluon: } (\quad) : \quad \frac{-ig_s}{2}\lambda^\alpha\gamma^\mu \tag{8.42}$$



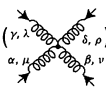
$$\text{Three gluon: } (\quad) :$$



$$-g_s f^{\alpha\beta\gamma}[g_{\mu\nu}(k_1 - k_2)_\lambda + g_{\nu\lambda}(k_2 - k_3)_\mu + g_{\lambda\mu}(k_3 - k_1)_\nu] \tag{8.43}$$

Here the gluon momenta (k_1, k_2, k_3) are assumed to point *into* the vertex; if any point *outward* in your diagram, change their signs.

$$\text{Four gluon: } (\quad) :$$



$$-ig_s^2[f^{\alpha\beta\eta}f^{\gamma\delta\eta}(g_{\mu\lambda}g_{\nu\rho} - g_{\mu\rho}g_{\nu\lambda}) + f^{\alpha\delta\eta}f^{\beta\gamma\eta}(g_{\mu\nu}g_{\lambda\rho} - g_{\mu\lambda}g_{\nu\rho})$$
$$+f^{\alpha\gamma\eta}f^{\delta\beta\eta}(g_{\mu\rho}g_{\nu\lambda} - g_{\mu\nu}g_{\lambda\rho})] \tag{8.44}$$

(summation over η implied).

Everything else is the same as for QED*: impose conservation of energy and momentum at each vertex to determine the internal four momenta; follow each fermion line 'backward' along the arrow, erase the overall delta function, and multiply by i to get $\mathscr{M}$. In the next two sections I'll work out some examples to show you how it goes.

* Loop diagrams in QCD require special rules, including the introduction of so-called 'Faddeev–Popov ghosts'. These are deep waters, into which we shall not venture [3].

8.4 Color Factors

In this section, we consider the interaction between two quarks (also a quark and an antiquark) in lowest-order QCD. Of course, we cannot observe quark–quark scattering directly in the laboratory (although hadron–hadron scattering is an indirect manifestation), so we won't be looking for cross sections here. Instead, we concentrate on the effective *potentials* between quarks – the QCD analog of the Coulomb potential in electrodynamics. We used such potentials, with a promise to derive them later, back in Chapter 5, in the analysis of quarkonium. Bear in mind that this is a perturbation theory calculation, valid only insofar as the coupling α_s is small. We cannot hope to get the confining term in the potential by this route – we are implicitly relying on asymptotic freedom, and all we're going to find is the short-range behavior. Nevertheless, we will obtain a very suggestive result: *Quarks attract one another most strongly when they are in the color singlet configuration* (indeed, in other arrangements they generally repel). At very short range, then, the color singlet is the 'maximally attractive channel' – an indication that binding is more likely, at least, for singlet states.*

8.4.1 Quark and Antiquark

Consider first the interaction of a quark and an antiquark, in QCD. We shall assume that they have different flavors, so the only diagram (in lowest order) is the one in Figure (8.6),† representing, for instance, $u + \bar{d} \to u + \bar{d}$. The amplitude is given by

$$\begin{aligned}\mathscr{M} &= i[\bar{u}(3)c_3^\dagger]\left[-i\frac{g_s}{2}\lambda^\alpha\gamma^\mu\right][u(1)c_1]\left[\frac{-ig_{\mu\nu}\delta^{\alpha\beta}}{q^2}\right] \\ &\quad\times[\bar{v}(2)c_2^\dagger]\left[-i\frac{g_s}{2}\lambda^\beta\gamma^\nu\right][v(4)c_4] \end{aligned} \tag{8.45}$$

Thus

$$\mathscr{M} = -\frac{g_s^2}{4q^2}[\bar{u}(3)\gamma^\mu u(1)][\bar{v}(2)\gamma_\mu v(4)](c_3^\dagger\lambda^\alpha c_1)(c_2^\dagger\lambda^\alpha c_4) \tag{8.46}$$

(summation over α implied). This is exactly what we had for electron–positron scattering (Equation 7.108), except that g_e is replaced by g_s (of course), and we have

* This is a very pleasing conclusion, but it does not prove that binding *must* occur in the color singlet, or that it *cannot* occur in other configurations. For this we would have to know the *long*-range behavior of the potential, about which, at present, we can only speculate.

† In principle, for the *same* flavor (e.g. $u + \bar{u} \to u + \bar{u}$) we should include a second diagram, as in electron–positron scattering (Figure 7.5). However, in the nonrelativistic limit of interest here this second diagram does not contribute anyway (see footnote to Example 7.3), so in *practice* what we're doing applies just as well *whatever* the quark flavors. (See also Problem 8.17)

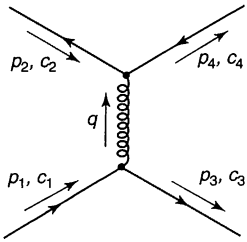


Fig. 8.6 The quark–antiquark interaction.

in addition the 'color factor'

$$f = \frac{1}{4}(c_3^\dagger \lambda^\alpha c_1)(c_2^\dagger \lambda^\alpha c_4) \tag{8.47}$$

Therefore, the *potential* describing the $q\bar{q}$ interaction is the same as that acting in electrodynamics between two opposite charges (to wit: the Coulomb potential), only with α replaced by $f\alpha_s$:

$$V_{q\bar{q}}(r) = -f\frac{\alpha_s \hbar c}{r} \tag{8.48}$$

Now, the color factor depends on the color state of the interacting quarks. From a quark and an antiquark we can make a color singlet, Equation 8.30, and a color octet, Equation 8.29 (all members of which yield the same f). I'll calculate the octet color factor first, because it's a little easier [4].

Example 8.1 Color Factor for the Octet Configuration A typical octet state (Equation 8.29) is $r\bar{b}$ (any of the others would do just as well; see Problem 8.16). Here the incoming quark is red, and the incoming antiquark is antiblue. Because color is conserved,* the outgoing quark must also be red and the antiquark antiblue. Thus

$$c_1 = c_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad c_2 = c_4 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

and hence

$$f = \frac{1}{4}\left[(1\ 0\ 0)\lambda^\alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right]\left[(0\ 1\ 0)\lambda^\alpha \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right] = \frac{1}{4}\lambda_{11}^\alpha \lambda_{22}^\alpha$$

* Yes, quark color *can* change at a QCD vertex, but in this case the outgoing antiquark cannot carry off the positive unit of redness, so the outgoing quark is forced to do that job.

A glance at the λ matrices reveals that the only ones with entries in the 11 and 22 positions are λ^3 and λ^8. So

$$f = \frac{1}{4}(\lambda^3_{11}\lambda^3_{22} + \lambda^8_{11}\lambda^8_{22}) = \frac{1}{4}[(1)(-1) + (1/\sqrt{3})(1/\sqrt{3})] = -\frac{1}{6} \tag{8.49}$$

Example 8.2 Color Factor for the Singlet Configuration The color singlet state (Equation 8.30) is

$$(1/\sqrt{3})(r\bar{r} + b\bar{b} + g\bar{g})$$

If the incoming quarks are in the singlet state (as they would be for a meson, say), the color factor is a sum of three terms:

$$f = \frac{1}{4}\frac{1}{\sqrt{3}}\left\{\left[c_3^\dagger\lambda^\alpha\begin{pmatrix}1\\0\\0\end{pmatrix}\right][(1\ 0\ 0)\lambda^\alpha c_4] + \left[c_3^\dagger\lambda^\alpha\begin{pmatrix}0\\1\\0\end{pmatrix}\right][(0\ 1\ 0)\lambda^\alpha c_4]\right.$$
$$\left. + \left[c_3^\dagger\lambda^\alpha\begin{pmatrix}0\\0\\1\end{pmatrix}\right][(0\ 0\ 1)\lambda^\alpha c_4]\right\}$$

The outgoing quarks are necessarily also in the singlet state, and we get *nine* terms in all, which can be written compactly as follows:

$$f = \frac{1}{4}\frac{1}{\sqrt{3}}\frac{1}{\sqrt{3}}(\lambda^\alpha_{ij}\lambda^\alpha_{ji}) = \frac{1}{12}\mathrm{Tr}(\lambda^\alpha\lambda^\alpha) \tag{8.50}$$

(summation over i and j, from 1 to 3, implied in the second expression). Now

$$\mathrm{Tr}(\lambda^\alpha\lambda^\beta) = 2\delta^{\alpha\beta} \tag{8.51}$$

(Problem 8.23), so, with the summation over α (from 1 to 8),

$$\mathrm{Tr}(\lambda^\alpha\lambda^\alpha) = 16 \tag{8.52}$$

Evidently, then, for the color singlet

$$f = \frac{4}{3} \tag{8.53}$$

Putting Equations 8.49 and 8.53 into Equation 8.48, we conclude that the quark–antiquark potentials are

$$V_{q\bar{q}}(r) = -\frac{4}{3}\frac{\alpha_s\hbar c}{r} \quad \text{(color singlet)} \tag{8.54}$$

$$V_{q\bar{q}}(r) = \frac{1}{6}\frac{\alpha_s\hbar c}{r} \quad \text{(color octet)} \tag{8.55}$$

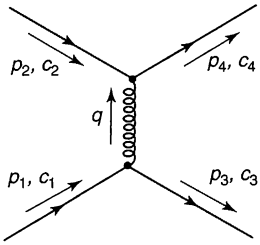


Fig. 8.7 The quark–quark interaction.

From the signs it appears that the force is *attractive* in the color singlet but *repulsive* for the octet. This helps to explain why quark–antiquark binding (to form mesons) occurs in the singlet configuration but not in the (color) octet (which would have produced colored mesons).

8.4.2
Quark and Quark

We turn now to the interaction of two quarks. Again, we shall assume that they have different flavors, so the only diagram (in lowest order) is the one indicated in Figure (8.7),* representing, say, $u + d \to u + d$. The amplitude is

$$\mathscr{M} = \frac{-g_s^2}{4}\frac{1}{q^2}[\overline{u}(3)\gamma^\mu u(1)][\overline{u}(4)\gamma_\mu u(2)](c_3^\dagger \lambda^\alpha c_1)(c_4^\dagger \lambda^\alpha c_2) \tag{8.56}$$

This is the same as for electron–muon scattering (Equation 7.106), except that g_e is replaced by g_s, and there is a color factor

$$f = \frac{1}{4}(c_3^\dagger \lambda^\alpha c_1)(c_4^\dagger \lambda^\alpha c_2) \tag{8.57}$$

The potential, therefore, takes the same form as that for like charges in electrodynamics:

$$V_{qq}(r) = f\frac{\alpha_s \hbar c}{r} \tag{8.58}$$

Again, the color factor depends on the configuration of the quarks. However, from two quarks you can't make a singlet and an octet (as for $q\overline{q}$) – rather, we obtain a triplet (the antisymmetric combinations):

$$\left\{\begin{array}{l} (rb - br)/\sqrt{2} \\ (bg - gb)/\sqrt{2} \\ (gr - rg)/\sqrt{2} \end{array}\right\} \quad \text{(triplet)} \tag{8.59}$$

* For identical quarks there is also a 'crossed' diagram. However, inclusion of this diagram, together with the statistical factor S in the cross section formula, leads to the same nonrelativistic limit (see footnote to Example 7.3), so in fact our potentials are correct even for same-flavor quarks.

and a sextet (the symmetric combinations):*

$$\left\{ \begin{array}{l} rr, bb, gg, \\ (rb+br)/\sqrt{2}, (bg+gb)/\sqrt{2}, (gr+rg)/\sqrt{2} \end{array} \right\} \quad \text{(sextet)} \tag{8.60}$$

Example 8.3 Color Factor for the Sextet Configuration A typical sextet state is rr (use any of the others if you prefer – you'll get the same result for f). In this case

$$c_1 = c_2 = c_3 = c_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

and hence

$$\begin{aligned} f &= \frac{1}{4}\left[(1 \quad 0 \quad 0)\lambda^\alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right]\left[(1 \quad 0 \quad 0)\lambda^\alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right] = \frac{1}{4}(\lambda^\alpha_{11}\lambda^\alpha_{11}) \\ &= \frac{1}{4}\left(\lambda^3_{11}\lambda^3_{11} + \lambda^8_{11}\lambda^8_{11}\right) = \frac{1}{4}\left[(1)(1) + (1/\sqrt{3})(1/\sqrt{3})\right] = \frac{1}{3} \end{aligned} \tag{8.61}$$

Example 8.4 Color Factor for the Triplet Configuration A typical triplet state is $(rb - br)/\sqrt{2}$, so†

$$\begin{aligned} f = \frac{1}{4}\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} \Bigg\{ & \left[(1 \quad 0 \quad 0)\lambda^\alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right]\left[(0 \quad 1 \quad 0)\lambda^\alpha \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right] \\ & - \left[(0 \quad 1 \quad 0)\lambda^\alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right]\left[(1 \quad 0 \quad 0)\lambda^\alpha \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right] \\ & - \left[(1 \quad 0 \quad 0)\lambda^\alpha \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right]\left[(0 \quad 1 \quad 0)\lambda^\alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right] \\ & + \left[(0 \quad 1 \quad 0)\lambda^\alpha \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right]\left[(1 \quad 0 \quad 0)\lambda^\alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right] \Bigg\} \\ = \frac{1}{8}\left(\lambda^\alpha_{11}\lambda^\alpha_{22} - \lambda^\alpha_{21}\lambda^\alpha_{12} - \lambda^\alpha_{12}\lambda^\alpha_{21} + \lambda^\alpha_{22}\lambda^\alpha_{11}\right) \end{aligned}$$

* In group-theoretical language, $3 \otimes \bar{3} = 1 \oplus 8$, but $3 \otimes 3 = \bar{3} \oplus 6$.

† Here $(rb - br) \to (rb - br)$, so there are four terms. Schematically, $rb \to rb$, $rb \to -br$, $-br \to rb$, and $-br \to -br$ (in the last term the factors of -1 cancel).

$$= \frac{1}{4}(\lambda^{\alpha}_{11}\lambda^{\alpha}_{22} - \lambda^{\alpha}_{12}\lambda^{\alpha}_{21})$$

$$= \frac{1}{4}(\lambda^{3}_{11}\lambda^{3}_{22} + \lambda^{8}_{11}\lambda^{8}_{22} - \lambda^{1}_{12}\lambda^{1}_{21} - \lambda^{2}_{12}\lambda^{2}_{21})$$

$$= \frac{1}{4}\left(-1 + \frac{1}{3} - 1 - 1\right) = -\frac{2}{3} \tag{8.62}$$

Putting Equations 8.61 and 8.62 into Equation 8.58, we conclude that the quark–quark potentials are

$$V_{qq}(r) = -\frac{2}{3}\frac{\alpha_s \hbar c}{r} \quad \text{(color triplet)} \tag{8.63}$$

$$V_{qq}(r) = \frac{1}{3}\frac{\alpha_s \hbar c}{r} \quad \text{(color sextet)} \tag{8.64}$$

In particular, the signs indicate that the force is *attractive* for the triplet and *repulsive* for the sextet. Of course, that's not too helpful as it stands, because *neither* combination occurs in nature.* However, it does have interesting implications for the binding of *three* quarks. This time we can make a singlet (completely antisymmetric), a decuplet (completely symmetric), and two octets (of mixed symmetry), as we found in Section 5.6.1.† Since the singlet is completely antisymmetric, every *pair* of quarks is in the (antisymmetric) triplet state – the *attractive* channel. In the decuplet, every pair is in the (symmetric) sextet state – they *repel*. As for the two octets, some pairs are triplet and some are sextet; we expect some attraction, then, and some repulsion. *Only in the singlet configuration*, though, *do we get complete mutual attraction* of the three quarks. Again, this is a comforting result: as in the case of mesons, the potential is most favorable for binding when the quarks are in the color singlet configuration.

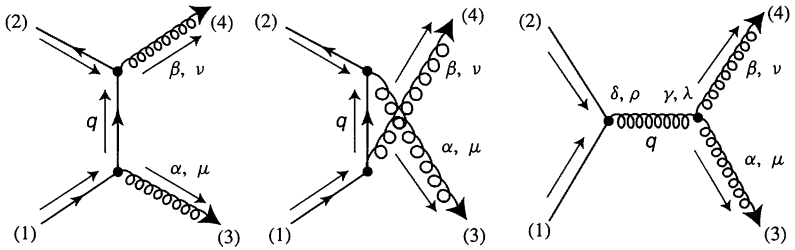
8.5 Pair Annihilation in QCD

In this section we consider the process quark plus antiquark → two gluons – the QCD analog of pair annihilation. The calculation is quite similar to Example 7.8;

* If you don't heed the warning in footnote (*) to the first paragraph of Section 8.4, you may be alarmed to find that two quarks in the triplet state attract one another. There is some comfort in the observation that the singlet $q\bar{q}$ coupling is *twice as strong*; but still, if this were the whole story we might very well expect triplet qq binding to occur, leading to free 'diquark' states. There has, in fact, been some speculation about the possible existence of diquarks within nuclei [5].

† In Chapter 5 we were dealing with *flavor*, not *color*, but the *mathematics* is the same. Group theoretically, $3 \otimes 3 \otimes 3 = 1 \oplus 8 \oplus 8 \oplus 10$.

however, in QCD there are *three* contributing diagrams, in lowest order:



The amplitude for the first diagram is

$$\mathscr{M}_1 = i\bar{v}(2)c_2^\dagger\left[-i\frac{g_s}{2}\lambda^\beta\gamma^\nu\right][\epsilon_{4\nu}^* a_4^{\beta *}]\left[\frac{i(\not{q}+mc)}{q^2-m^2c^2}\right]$$
$$\times\left[-i\frac{g_s}{2}\lambda^\alpha\gamma^\mu\right][\epsilon_{3\mu}^* a_3^{\alpha *}]u(1)c_1 \tag{8.65}$$

(To simplify the already overburdened notation I'll leave the * off the gluon polarization vectors and color states until the end.) Here $q = p_1 - p_3$, so

$$q^2 - m^2c^2 = p_1^2 - 2p_1\cdot p_3 + p_3^2 - m^2c^2 = -2p_1\cdot p_3 \tag{8.66}$$

and hence

$$\mathscr{M}_1 = \frac{-g_s^2}{8}\frac{1}{p_1\cdot p_3}\left\{\bar{v}(2)[\not{\epsilon}_4(\not{p}_1 - \not{p}_3 + mc)\not{\epsilon}_3]u(1)\right\} a_3^\alpha a_4^\beta (c_2^\dagger\lambda^\beta\lambda^\alpha c_1) \tag{8.67}$$

Similarly, for the second diagram:

$$\mathscr{M}_2 = \frac{-g_s^2}{8}\frac{1}{p_1\cdot p_4}\left\{\bar{v}(2)[\not{\epsilon}_3(\not{p}_1 - \not{p}_4 + mc)\not{\epsilon}_4]u(1)\right\} a_3^\alpha a_4^\beta (c_2^\dagger\lambda^\alpha\lambda^\beta c_1) \tag{8.68}$$

Notice that the λ's appear this time in the opposite order. Finally, for the third diagram:

$$\mathscr{M}_3 = i\bar{v}(2)c_2^\dagger\left[-i\frac{g_s}{2}\lambda^\delta\gamma_\sigma\right]u(1)c_1\left[-i\frac{g^{\sigma\lambda}\delta^{\delta\gamma}}{q^2}\right]\{-g_s f^{\alpha\beta\gamma}[g_{\mu\nu}(-p_3+p_4)_\lambda$$
$$+g_{\nu\lambda}(-p_4-q)_\mu + g_{\lambda\mu}(q+p_3)_\nu]\}[\epsilon_3^\mu a_3^\alpha][\epsilon_4^\nu a_4^\beta] \tag{8.69}$$

In this case $q = p_3 + p_4$, so $q^2 = 2p_3\cdot p_4$; simplifying (and using $\epsilon_3\cdot p_3 = \epsilon_4\cdot p_4 = 0$), we find (Problem 8.20):

$$\mathscr{M}_3 = i\frac{g_s^2}{4}\frac{1}{p_3\cdot p_4}\bar{v}(2)[(\epsilon_3\cdot\epsilon_4)(\not{p}_4 - \not{p}_3) + 2(p_3\cdot\epsilon_4)\not{\epsilon}_3 - 2(p_4\cdot\epsilon_3)\not{\epsilon}_4]u(1)$$
$$\times f^{\alpha\beta\gamma}a_3^\alpha a_4^\beta(c_2^\dagger\lambda^\gamma c_1) \tag{8.70}$$

So far, this is all completely general (and rather messy). To make things more manageable, let's assume (as we did in our study of $e^+ e^-$ annihilation) that the initial particles are at rest:

$$p_1 = p_2 = (mc, \mathbf{0}), \quad p_3 = (mc, \mathbf{p}), \quad p_4 = (mc, -\mathbf{p}) \tag{8.71}$$

Then

$$p_1 \cdot p_3 = p_1 \cdot p_4 = (mc)^2, \qquad p_3 \cdot p_4 = 2(mc)^2 \tag{8.72}$$

Meanwhile, in the Coulomb gauge (Equation 8.32)

$$p_3 \cdot \epsilon_4 = -\mathbf{p} \cdot \boldsymbol{\epsilon}_4 = -p_4 \cdot \epsilon_4 = 0 \tag{8.73}$$

(likewise $p_4 \cdot \epsilon_3 = 0$), so two terms in $\mathscr{M}_3$ drop out. Using Equations 7.140 and 7.141 to simplify $\mathscr{M}_1$ and $\mathscr{M}_2$, we find that the total amplitude ($\mathscr{M} = \mathscr{M}_1 + \mathscr{M}_2 + \mathscr{M}_3$) can be written

$$\begin{aligned}\mathscr{M} = -\frac{g_s^2}{8(mc)^2} a_3^\alpha a_4^\beta \bar{v}(2) c_2^\dagger [\not\epsilon_3 \not\epsilon_4 \not p_4 \lambda^\alpha \lambda^\beta + \not\epsilon_4 \not\epsilon_3 \not p_3 \lambda^\beta \lambda^\alpha \\ -i(\epsilon_3 \cdot \epsilon_4)(\not p_4 - \not p_3) f^{\alpha\beta\gamma} \lambda^\gamma] c_1 u(1)\end{aligned} \tag{8.74}$$

We may as well orient our coordinates so that the z axis lies along $\mathbf{p}$; then

$$\not p_3 = mc(\gamma^0 - \gamma^3), \quad \not p_4 = mc(\gamma^0 + \gamma^3), \quad \not p_4 - \not p_3 = 2mc\gamma^3 \tag{8.75}$$

From Equations 7.145 and 7.146 we have

$$\not\epsilon_3 \not\epsilon_4 = -(\boldsymbol{\epsilon}_3 \cdot \boldsymbol{\epsilon}_4) - i(\boldsymbol{\epsilon}_3 \times \boldsymbol{\epsilon}_4) \cdot \boldsymbol{\Sigma}, \quad \not\epsilon_4 \not\epsilon_3 = -(\boldsymbol{\epsilon}_3 \cdot \boldsymbol{\epsilon}_4) + i(\boldsymbol{\epsilon}_3 \times \boldsymbol{\epsilon}_4) \cdot \boldsymbol{\Sigma} \tag{8.76}$$

Putting this into Equation 8.74, and exploiting the commutation relation (Equation 8.35) for the λ's, we obtain

$$\begin{aligned}\mathscr{M} = \frac{g_s^2}{8mc} a_3^\alpha a_4^\beta \bar{v}(2) c_2^\dagger \big[(\boldsymbol{\epsilon}_3 \cdot \boldsymbol{\epsilon}_4)\{\lambda^\alpha, \lambda^\beta\}\gamma^0 \\ +i(\boldsymbol{\epsilon}_3 \times \boldsymbol{\epsilon}_4) \cdot \boldsymbol{\Sigma}([\lambda^\alpha, \lambda^\beta]\gamma^0 + \{\lambda^\alpha, \lambda^\beta\}\gamma^3)\big] c_1 u(1)\end{aligned} \tag{8.77}$$

where curly brackets denote the anticommutator: $\{A, B\} \equiv AB + BA$. You might compare this result with the corresponding expression in QED (Equation 7.146), to which it reduces if you set all the λ's equal to 1, drop the color states a and c, and let $g_s/2 \to g_e$.

Suppose now we put the quarks into a *spin-0* (singlet) state (the triplet state cannot go to two gluons anyway; it needs at least *three*):

$$\mathscr{M} = (\mathscr{M}_{\uparrow\downarrow} - \mathscr{M}_{\downarrow\uparrow})/\sqrt{2} \tag{8.78}$$

For $\mathscr{M}_{\uparrow\downarrow}$ we have (Equations 7.153 and 7.154)

$$\bar{v}(2)\gamma^0 u(1) = \bar{v}(2)\Sigma\gamma^0 u(1) = 0, \quad \bar{v}(2)\Sigma\gamma^3 u(1) = -2mc\hat{z} \tag{8.79}$$

As before, $\mathscr{M}_{\downarrow\uparrow} = -\mathscr{M}_{\uparrow\downarrow}$, and we are left with*

$$\mathscr{M} = -i\sqrt{2}\frac{g_s^2}{4}(\epsilon_3 \times \epsilon_4)_z a_3^\alpha a_4^\beta (c_2^\dagger\{\lambda^\alpha, \lambda^\beta\}c_1) \quad \text{(spin singlet)} \tag{8.80}$$

Once again, we have obtained a result that is identical to the one in QED (Equation 7.158), except that $g_e \to g_s$, and there is a color factor

$$f = \frac{1}{8}a_3^\alpha a_4^\beta (c_2^\dagger\{\lambda^\alpha, \lambda^\beta\}c_1) \tag{8.81}$$

In particular, if the quarks occupy the color singlet state, $(1/\sqrt{3})(r\bar{r} + b\bar{b} + g\bar{g})$, then

$$f = \frac{1}{8}a_3^\alpha a_4^\beta \frac{1}{\sqrt{3}}\left\{(1\ 0\ 0)\{\lambda^\alpha, \lambda^\beta\}\begin{pmatrix}1\\0\\0\end{pmatrix} + (0\ 1\ 0)\{\lambda^\alpha, \lambda^\beta\}\begin{pmatrix}0\\1\\0\end{pmatrix}\right.$$

$$\left. +(0\ 0\ 1)\{\lambda^\alpha, \lambda^\beta\}\begin{pmatrix}0\\0\\1\end{pmatrix}\right\} = \frac{1}{8\sqrt{3}}a_3^\alpha a_4^\beta \mathrm{Tr}\{\lambda^\alpha, \lambda^\beta\}$$

But

$$\mathrm{Tr}\{\lambda^\alpha, \lambda^\beta\} = 2\mathrm{Tr}(\lambda^\alpha\lambda^\beta) = 4\delta^{\alpha\beta} \tag{8.82}$$

(Problem 8.13), so

$$f = \frac{1}{2\sqrt{3}}a_3^\alpha a_4^\alpha \quad \text{(color singlet)} \tag{8.83}$$

Now, the singlet state for two gluons (see Problem 8.22) is

$$|\text{singlet}\rangle = \frac{1}{\sqrt{8}}\sum_{n=1}^{8}|n\rangle_1|n\rangle_2 \tag{8.84}$$

Evidently

$$a_3^\alpha a_4^\alpha = \frac{1}{\sqrt{8}}(8) = 2\sqrt{2} \tag{8.85}$$

* At this stage all terms in $\epsilon_3 \cdot \epsilon_4$ drop out. The fact that $\mathscr{M}_3$ is proportional to $\epsilon_3 \cdot \epsilon_4$ (Equation 8.74) means that the diagram containing a three-gluon vertex makes no contribution, when the quarks are at rest in the spin singlet configuration. Most books simply ignore it from the start, but in principle it should be included (see Problem 8.21).

and hence

$$f = \sqrt{2/3} \tag{8.86}$$

Conclusion: For $q + \bar{q} \to g + g$ in the spin singlet, color singlet configuration, with the quarks at rest, the amplitude is

$$\mathscr{M} = -4\sqrt{2/3}\, g_s^2 \tag{8.87}$$

(compare Equation 7.163), and the cross section is

$$\sigma = \frac{2}{3}\frac{4\pi}{cv}\left(\frac{\hbar\alpha_s}{m}\right)^2 \tag{8.88}$$

(see Equation 7.168). Just as the cross section for $e^+ + e^- \to \gamma + \gamma$ determines the positronium decay rate

$$\Gamma = \sigma v |\psi(0)|^2 \tag{8.89}$$

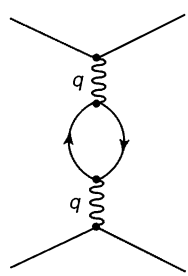
(Equation 7.171), so we can now give a formula for the decay of a spin-0 quarkonium state, such as η_c (note that ψ and Υ themselves carry spin 1, and go to *three* gluons):

$$\Gamma(\eta_c \to 2g) = \frac{8\pi}{3c}\left(\frac{\hbar\alpha_s}{m}\right)^2 |\psi(0)|^2 \tag{8.90}$$

As it stands, this is not terribly useful, since we don't know $\psi(0)$. However, the electromagnetic decay $\eta_c \to 2\gamma$ involves the same factor, and we *can* derive a clean expression for the branching ratio (see Problem 8.23).

8.6 Asymptotic Freedom

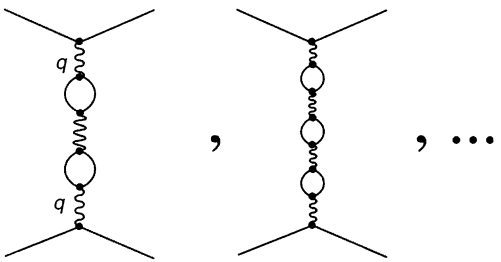
In the last section of Chapter 7 we found that the loop diagram



in QED makes the effective charge of the electron a function of the momentum transfer q:*

$$\alpha(|q^2|) = \alpha(0)\left\{1 + \frac{\alpha(0)}{3\pi}\ln(|q^2|/(mc)^2)\right\} \quad (|q^2| = -q^2 \gg (mc)^2) \tag{8.91}$$

The coupling strength *increases* as the charges get closer together (larger $|q^2|$), a fact that we interpret physically as a consequence of 'vacuum polarization': the vacuum functions as a kind of dielectric medium, partially screening the charge. The closer we approach, the less complete is the screening, and the greater is the effective charge. Of course, Equation 8.91 is valid only to order $\alpha\ (0)^2$. There are higher-order corrections, of which the dominant ones come from *chains* of bubbles:



As it happens, these can be summed explicitly, and the result is†

$$\alpha(|q^2|) = \frac{\alpha(0)}{1 - [\alpha(0)/3\pi]\ln[|q^2|/(mc)^2]} \quad (|q^2| \gg (mc)^2) \tag{8.92}$$

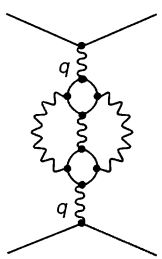
Ostensibly, the coupling *blows up* at $\ln[|q^2|/(mc)^2] = 3\pi/\alpha(0)$. However, this is not to be taken too seriously, since it occurs at an energy of about 10^{280} MeV, which (to put it mildly) is not an accessible region (see Problem 8.24).

* It *also* introduces a *divergent* term, which we soak up in the 'renormalized' charge (Equation 7.189). But that's an entirely different problem, one that (however troublesome you may find it in principle) has no observable consequences, and once the appropriate incantation has been made, is of no further significance. The perfectly *finite* dependence of α on q^2 is the *significant* matter, for it carries direct and measurable implications.

† This is not so surprising. What we have, in effect, is the geometric series

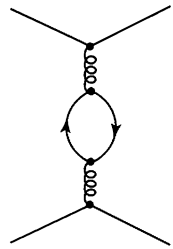
$$1 + x + x^2 + x^3 + \cdots = \frac{1}{1-x}$$

where x is for one bubble, x^2 is for two, and so on. Although Equation 8.92 is correct to all orders in $\alpha(0)$, it is not *exact*, since we are ignoring diagrams such as

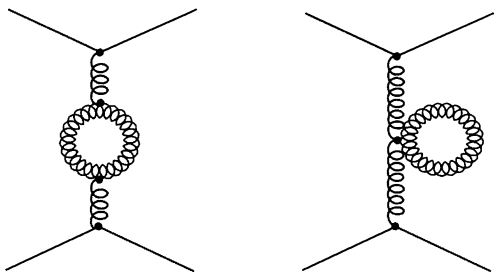


These can be shown to make a much smaller contribution in the limit $|q^2| \gg (mc)^2$. Equation 8.92 is known as the 'leading log' approximation.

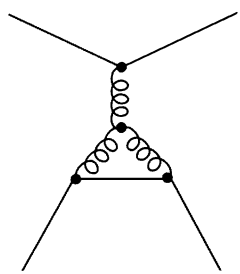
Much the same thing happens in QCD: quark–antiquark bubbles



lead to a screening of the quark color, which (modulo appropriate color factors) is the same as Equation 8.91. However, there is a new twist to the story, for in QCD we also have virtual *gluon* bubbles



as well as diagrams of the form



It turns out [6] that the gluon contribution works in the other direction, producing 'antiscreening' or 'camouflage'. I do not know of a persuasive qualitative explanation of this effect [7] – suffice it to say that the formula for the running coupling constant in QCD (analogous to Equation 8.92) is

$$\alpha_s(|q^2|) = \frac{\alpha_s(\mu^2)}{1 + [\alpha_s(\mu^2)/12\pi](11n - 2f)\ln(|q^2|/\mu^2)} \quad (|q^2| \gg \mu^2) \tag{8.93}$$

where n is the number of colors (3, in the Standard Model), and f is the number of flavors (6, in the Standard Model). In any theory for which $11n > 2f$, antiscreening will dominate, and the coupling constant will *decrease* with increasing $|q^2|$; at short distances the 'strong' force becomes relatively weak. This is the source of *asymptotic freedom*, on which so much of what we can say quantitatively about the hadrons is predicated. Asymptotic freedom is what licenses the use of the Feynman calculus in QCD to calculate interquark potentials; it is a basic ingredient in the theory of quarkonium; and it is presumably responsible for the OZI rule. Chromodynamics would have gone out of business if it had not been for the timely discovery of asymptotic freedom [8].

You may have noticed the appearance of a new parameter, μ, in Equation 8.93 In electrodynamics it is natural to define 'the charge' of a particle as the long-range (fully screened) value – that's what Coulomb and Millikan measured, and it's what an engineer or a chemist or even an atomic physicist (unless he's measuring the Lamb shift) is concerned with. Thus $\alpha(0)$ is the 'good old' fine structure constant, 1/137, and it is the sensible parameter in terms of which to do perturbation expansions. But we don't *have* to do it this way; we *could* work from any other value of q^2 (provided only that we stay well below the singularity in Equation 8.92, where $\alpha(|q^2|)$ runs larger than 1, and perturbation theory breaks down). In QCD, however, we *cannot* work from $q^2 = 0$, because that's where α_s is *large*. We must use as a reference some place where α_s is small enough to justify a perturbation expansion. That's why Equation 8.93 is expressed in terms of $\alpha_s(\mu^2)$, instead of $\alpha_s(0)$. Provided that it's large enough so that $\alpha_s(\mu^2) < 1$, it doesn't matter *what* value of μ you use (see Problem 8.25). Indeed, if we introduce a new variable Λ, defined by

$$\ln \Lambda^2 = \ln \mu^2 - 12\pi/[(11n - 2f)\alpha_s(\mu^2)] \tag{8.94}$$

the running coupling constant can be expressed in terms of a *single* parameter:

$$\alpha_s(|q^2|) = \frac{12\pi}{(11n - 2f)\ln(|q^2|/\Lambda^2)} \quad (|q^2| \gg \Lambda^2) \tag{8.95}$$

(see Problem 8.26). This compact result tells us explicitly the value of the strong coupling at any $|q^2|$, in terms of the constant Λ. Unfortunately, it is hard to determine Λ precisely from experimental data, but Λc appears to lie somewhere in the range

$$100 \text{ MeV} < \Lambda c < 500 \text{ MeV}. \tag{8.96}$$

Notice that whereas the QED coupling varies only minutely over the accessible energy range (Problem 8.24), variation in the QCD coupling is substantial (Problem 8.27).

References

1 Rosenbluth, M. N. (1950) *Physical Review*, **79**, 615.

2 See, for example, Frauenfelder, H. and Henley, E. M. (1991) *Subatomic Physics*, 2nd edn, Prentice-Hall, Englewood Cliffs, N.J. Chapter 6. I should warn you that this subject is a notational nightmare in the literature. There are only two variables in the problem – the incident electron energy (E) and the scattering angle (θ) – but it is common to encounter a more or less random mixture of $E, E', \theta, q^2, Q^2 \equiv -q^2, \tau \equiv -q^2/4M^2c^2, \nu \equiv p \cdot q/Mc, \omega \equiv -2p \cdot q/q^2, W \equiv \sqrt{(q+p)^2}, x \equiv -q^2/2p \cdot q$, and $y \equiv p \cdot q/p \cdot p_1$. Moreover, although there are only two independent form factors involved, there are many different ways to express them. Some authors favor F_1 and F_2, with $K_1 = -q^2(F_1 + KF_2)^2$ and $K_2 = (2Mc)^2F_1^2 - K^2q^2F_2^2$ ($K = 1.7928$ is the 'anomalous' contribution to the proton's magnetic moment); others prefer $G_E = F_1 - K\tau F_2$, $G_M = F_1 + KF_2$. (The latter are related to the Fourier transforms of the charge and magnetic moment distributions, respectively; see; (a) Halzen, F. and Martin, A. D. (1984) *Quarks and Leptons*, John Wiley & Sons, New York. Sect. 8.2.) Anyone can play this game; K_1, and K_2 are my own contributions.

3 The interested reader should consult the classic treatise by Abers, E. S. and Lee, B. W. (1973) *Physics Reports*, **9 C**, 1.

4 People in the know seem to be able to calculate color factors on their fingers; everyone has his or her own tricks. See Perkins, D. H. (1982) *Introduction to High-Energy Physics*, 2nd edn, Addison-Wesley, Reading, M.A. Appendix G (in the 3rd edition (1987) it is Appendix J, and in the 4th (Cambridge, UK: Cambridge University Press, 2000) it has been removed); (a) Halzen, F. and Martin, A. D. (1984) *Quarks and Leptons*, John Wiley & Sons, New York, p. 211. Section 2.15; (b) Quigg, C. (1997) *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions*, Addison Wesley, Reading, M.A., pp. 198–199. I prefer to do it by the more pedestrian method shown here, which is close in spirit to the approach of; (c) Kane, G. L. (1977) *Color Symmetry and Quark Confinement, Proceedings of the 12th Rencontre de Moriond*, vol. **III**, ed. J. Tran Thanh Van, Frontières, Gif-sur-Yvette, France, p. 9.

5 Close, F. (1982) "Demon Nuclei", *Nature*, **296**, 305.

6 Quigg, C. (1997) *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions*, Addison Wesley, Reading, M.A., pp. 198–199. Sect. 8.3.

7 See, however, Quigg, C. *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions*, Addison Wesley, Reading, M.A., p. 223; and (April 1985) *Scientific American*, p. 84.

8 Politzer, H. D. (1973) *Physical Review Letters*, **30**, 1346; (1974) *Physics Reports*, **14 C**, 130; (a) Gross D. J. and Wilczek, F. (1973) *Physical Review Letters*, **30**, 1343.

Problems

8.1
(a) Derive Equation 8.1, from the Feynman rules for QED.
(b) Obtain Equation 8.2 from Equation 8.1
(c) Derive Equation 8.3 from Equation 8.2
(d) Derive Equation 8.4 from Equation 8.3

8.2 Derive Equation 8.5, starting with Equation 8.4

8.3 Why don't we use σ ($e^+e^- \to e^+e^-$) in the denominator, to define R (Equation 8.7)?

8.4 Prove Equation 8.16 [*Hint*: First show that $q_\mu L^{\mu\nu}=0$. Then argue that we *may as well* take $K^{\mu\nu}$ such that $q_\mu K^{\mu\nu}=0$, in the sense that any term in $K^{\mu\nu}$ that does *not* obey $q_\mu K^{\mu\nu}=0$ will contribute nothing to $L^{\mu\nu}K_{\mu\nu}$.] *Comment*: Equation 8.16 actually follows more simply and generally from charge conservation at the proton vertex, but I have not developed the formalism here to make this argument (see Halzen and Martin [2], Sections 8.2 and 8.3).
[One way to proceed is as follows. Take $q^\mu = (0,\ 0,\ 0,\ q)$; then $q_\mu L^{\mu\nu} = 0 \Rightarrow L^{\mu\nu} =$ $\begin{pmatrix} \boxed{} & \begin{matrix}0\\0\\0\end{matrix} \\ \begin{matrix}0&0&0\end{matrix} & 0 \end{pmatrix}$. So $L^{\mu\nu}K_{\mu\nu} =$ $\begin{pmatrix} \boxed{} & \begin{matrix}0\\0\\0\end{matrix} \\ \begin{matrix}0&0&0\end{matrix} & 0 \end{pmatrix}$ $\begin{pmatrix} \boxed{} & \begin{matrix}x\\x\\x\end{matrix} \\ \begin{matrix}x&x&x\end{matrix} & x \end{pmatrix}$, and the x's might as well be zero.]

8.5 Prove Equation 8.17, from Equation 8.16 [*Hint*: First contract $K^{\mu\nu}$ with q_μ, then with p_ν.]

8.6 Find K_1 and K_2, and also G_E and G_M, for a 'Dirac' proton (Equation 8.25).

8.7 Derive Equation 8.19

8.8 Derive Equation 8.20

8.9 Derive Equation 8.21

8.10 Check that the Rosenbluth formula (Equation 8.23) agrees with the Mott formula (Equation 7.131) in the intermediate-energy regime ($mc^2 \ll E \ll Mc^2$). Use the expressions for K_1 and K_2 appropriate to a 'Dirac' proton (Problem 8.6).

8.11 Why can't the 'ninth gluon' be the photon? [*Answer*: The gluon would couple to all baryons with the same strength, not (as the photon does) in proportion to their charge. Since mass and baryon number are approximately proportional in bulk matter, such a force would, in fact, look very much like an extra contribution to gravity. There was a flurry of interest in this possibility in early 1986. (Fischbach, E. *et al.*, (1986) *Physical Review Letters*, **56**, 3. See, however, the comments in *Physical Review Letters*, (1986) **56**, 2423.]

8.12 Color $SU(3)$ transformations relabel 'red', 'blue', and 'green' according to the transformation rule

$$c \to c' = Uc$$

where U is any unitary ($UU^\dagger=1$) 3×3 matrix of determinant 1, and c is a three-element column vector. For example

$$U = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

would take $r \to g$, $g \to b$, $b \to r$. The ninth gluon ($|9\rangle$) is obviously invariant under U, but the octet gluons are *not*. Show that $|3\rangle$ and $|8\rangle$ go into linear combinations of one another:

$$|3'\rangle = \alpha|3\rangle + \beta|8\rangle, \quad |8'\rangle = \gamma|3\rangle + \delta|8\rangle$$

Find the numbers α, β, γ, and δ.

8.13 Show that

$$\mathrm{Tr}(\lambda^\alpha \lambda^\beta) = 2\delta^{\alpha\beta}$$

(Notice that all the λ matrices are traceless.)

8.14 What are the structure constants for $SU(2)$? That is, what are the numbers f^{ijk} in

$$[\sigma^i, \sigma^j] = 2if^{ijk}\sigma^k$$

8.15 (a) Given that $f^{\alpha\beta\gamma}$ is completely antisymmetric (so that $f^{112}=0$ automatically, and having calculated f^{123}, we don't need to bother with f^{213}, f^{231}, etc.) how many *distinct* nontrivial structure constants remain?

$$\left[\textit{Answer}: \frac{8 \cdot 7 \cdot 6}{3 \cdot 2 \cdot 1} = 56\right]$$

Of these, it turns out that only nine are nonzero (those listed in Equation 8.36), and among these there are only three different *numbers*.

(b) Work out $[\lambda^1, \lambda^2]$, and confirm that $f^{12\gamma}=0$ for all γ except 3, while $f^{123}=1$.

(c) Similarly, compute $[\lambda^1, \lambda^3]$ and $[\lambda^4, \lambda^5]$, and determine the resulting structure constants.

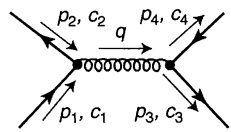
8.16 Calculate the octet $q\bar{q}$ color factor using the state

(a) $b\bar{g}$

(b) $(r\bar{r} - b\bar{b})/\sqrt{2}$

(c) $(r\bar{r} + b\bar{b} - 2g\bar{g})/\sqrt{6}$

8.17 Find the amplitude $\mathscr{M}$ for the diagram



What is the color factor (analogous to Equation 8.47) in this case? Evaluate f in the color singlet configuration. Can you explain this result? [*Answer*: It's *zero*; a singlet cannot couple to an octet (gluon).]

8.18 Calculate the sextet qq color factor using the state $(rb + br)/\sqrt{2}$.

8.19 Color factors always involve expressions of the form $\lambda^\alpha_{ij}\lambda^\alpha_{kl}$ (summed over α). There is a simple formula for this quantity, which shortens the arithmetic:

$$\lambda^\alpha_{ij}\lambda^\alpha_{kl} = 2\delta_{il}\delta_{jk} - \frac{2}{3}\delta_{ij}\delta_{kl}$$

(see Kane [4]). Check this theorem for

(a) $i=j=k=l=1$ (see Equation 8.61)

(b) $i=j=1$, $k=l=2$ (see Equation 8.49)

(c) $i=l=1$, $j=k=2$ (see Equation 8.62)

and

(d) Use it to confirm Equation 8.52

8.20 Derive Equation 8.70, starting from Equation 8.69

8.21 There is a simple test for the gauge invariance of an amplitude ($\mathscr{M}$) in QCD (or QED): Replace any gluon (or photon) polarization vector by its momentum ($\epsilon_3 \to p_3$, say), and you must get zero (see Problem 7.23). Show using this criterion that $\mathscr{M} = \mathscr{M}_1 + \mathscr{M}_2 + \mathscr{M}_3$ is gauge-invariant, but $\mathscr{M}_1 + \mathscr{M}_2$ alone is *not*. [Thus the three-gluon vertex is essential in QCD to preserve gauge invariance. Notice, by contrast, that $\mathscr{M}_1 + \mathscr{M}_2$ alone is gauge-invariant in QED (Example 7.8). The fact that λ matrices do not commute makes the difference.]

8.22 Construct the color singlet combination of two gluons (Equation 8.84). One method is as follows: Let

$$c = \begin{pmatrix} r \\ b \\ g \end{pmatrix}$$

Under $SU(3)$, $c \to c' = Uc$, where U is a unitary matrix of determinant 1. Similarly, let $d^\dagger = (\bar{r}, \bar{b}, \bar{g})$, transforming by the rule $d^\dagger \to d'^\dagger = d^\dagger\, U^\dagger$. Form the matrix

$$M \equiv cd^\dagger = \begin{pmatrix} r\bar{r} & r\bar{b} & r\bar{g} \\ b\bar{r} & b\bar{b} & b\bar{g} \\ g\bar{r} & g\bar{b} & g\bar{g} \end{pmatrix}$$

Note that $M' = c'\, d'^\dagger = UMU^\dagger$. Now remove the trace:

$$N \equiv M - \frac{1}{3}[\mathrm{Tr}(M)], \text{ so that } \mathrm{Tr}(N) = 0$$

[$\mathrm{Tr}(M') = \mathrm{Tr}(M) = (r\bar{r} + b\bar{b} + g\bar{g})$, so this combination is $SU(3)$-invariant; it is the *singlet* combination in $3 \otimes \bar{3} = 1 \oplus 8$, N is the octet.] Note that

$$N' = M' - \frac{1}{3}(\mathrm{Tr}(M')] = UMU^\dagger - \frac{1}{3}[\mathrm{Tr}(M)]UU^\dagger = UNU^\dagger$$

This tells us how the gluons themselves (which are in the octet representation) transform under color $SU(3)$. The question is how to put together *two* octets to make a singlet; that is, how to make something bilinear in N_1 and N_2 which is invariant under U. The solution is

$$s \equiv \mathrm{Tr}(N_1 N_2)$$

for

$$s' = \mathrm{Tr}(N_1' N_2') = \mathrm{Tr}(UN_1 U^\dagger UN_2 U^\dagger) = \mathrm{Tr}(U^\dagger UN_1 N_2) = \mathrm{Tr}(N_1 N_2) = s$$

It remains to figure out what s is in terms of the elements of M_1 and M_2:

$$\begin{aligned}
\mathrm{Tr}(N_1 N_2) &= \mathrm{Tr}\{(M_1 - \frac{1}{3}[\mathrm{Tr}(M_1)])(M_2 - \frac{1}{3}[\mathrm{Tr}(M_2)])\} \\
&= \mathrm{Tr}(M_1 M_2) - \frac{1}{3}[\mathrm{Tr}(M_1)][\mathrm{Tr}(M_2)] \\
&= \frac{2}{3}[(r\bar{r})_1(r\bar{r})_2 + (b\bar{b})_1(b\bar{b})_2 + (g\bar{g})_1(g\bar{g})_2] \\
&\quad -\frac{1}{3}[(r\bar{r})_1(b\bar{b})_2 + (r\bar{r})_1(g\bar{g})_2 + (b\bar{b})_1(r\bar{r})_2 + (b\bar{b})_1(g\bar{g})_2 \\
&\quad +(g\bar{g})_1(r\bar{r})_2 + (g\bar{g})_1(b\bar{b})_2] + [(r\bar{b})_1(b\bar{r})_2 + (r\bar{g})_1(g\bar{r})_2 \\
&\quad +(b\bar{r})_1(r\bar{b})_2 + (b\bar{g})_1(g\bar{b})_2 + (g\bar{r})_1(r\bar{g})_2 + (g\bar{b})_1(b\bar{g})_2]
\end{aligned}$$

$$= |1\rangle_1|1\rangle_2 + |2\rangle_1|2\rangle_2 + |3\rangle_1|3\rangle_2 + |4\rangle_1|4\rangle_2$$

$$+|5\rangle_1|5\rangle_2 + |6\rangle_1|6\rangle_2 + |7\rangle_1|7\rangle_2 + |8\rangle_1|8\rangle_2 = \sum_{n=1}^{8} |n\rangle_1|n\rangle_2$$

(To normalize the state, divide by $\sqrt{8}$.) This – the invariant product of two octets – is the $SU(3)$ analog to the dot product of two 3-vectors in $SU(2)$.

8.23 Determine the branching ratio $\Gamma\ (\eta_c \to 2g)/\Gamma\ (\eta_c \to 2\gamma)$. [*Hint*: Use Equation 8.90 for the numerator, and a suitable modification of Equations 7.168 and 7.171 for the denominator. There are two modifications: (i) the quark charge is Qe and (ii) there is a color factor of 3, for quarks in the singlet state (Equation 8.30). *Answer* : $\frac{9}{8}(\alpha_s/\alpha)^2$.]

8.24 (a) Calculate the energy ($\sqrt{|q^2|c^2}$) at which the QED coupling constant (Equation 8.86) blows up. (Remember, $\alpha(0)=1/137$, the fine structure constant.)

(b) At what energy do we get a 1% departure from $\alpha(0)$? Is this an accessible energy?

8.25 Prove that the value of μ in Equation 9.69 is arbitrary. [That is, suppose physicist A uses the value μ_a, and physicist B uses a different value, μ_b. *Assume* A's version of Equation 9.69 is correct, and *prove* that B's is also correct.]

8.26 Derive Equation 9.71 from Equations 8.93 and 8.94

8.27 Calculate α_s at 10 and 100 GeV. Assume Λc=0.3 GeV. What if Λc=1 GeV? How about Λc=0.1 GeV?

8.28 (Gluon–gluon scattering)

(a) Draw the lowest-order diagrams (there are four of them) representing the interaction of two gluons.

(b) Write down the corresponding amplitudes.

(c) Put the incoming gluons into the color singlet state; do the same for the outgoing gluons. Compute the resulting amplitudes.

(d) Go to the CM frame, in which each gluon has energy E; express all the kinematic factors in terms of E and the scattering angle θ. Add the amplitudes to get the total, $\mathscr{M}$.

(e) Find the differential scattering cross section.

(f) Determine whether the force is attractive or repulsive (if it is the former, this may be a likely glueball configuration).

9

Weak Interactions

This chapter surveys the theory of weak interactions. It relies heavily on Chapter 7, but not on Chapter 8; Section 4.4.1 would be useful background. I begin by stating the Feynman Rules for the coupling of leptons to $W^{\pm}$, and treat three classic problems in some detail: beta decays of the muon, the neutron, and the charged pion. Next, we consider the coupling of quarks to $W^{\pm}$, which brings in the Cabibbo angle, the GIM mechanism, and the Kobayashi–Maskawa matrix. In Section 9.6, I state the Feynman rules for coupling quarks and leptons to the Z^0, and the final section sketches the Glashow–Weinberg–Salam electroweak theory. Throughout this chapter I take the neutrinos to be massless; none of the results are measurably affected if (minute) neutrino masses are included.

9.1
Charged Leptonic Weak Interactions

The mediators of weak interactions (analogous to photons in QED and gluons in QCD) are the W's (W^+ and W^-) and the Z^0. Unlike the photon and gluons, which are massless, these 'intermediate vector bosons' are extremely heavy; experimentally,

$$M_W = 80.40 \pm .03 \text{ GeV}/c^2, \quad M_Z = 91.188 \pm .002 \text{ GeV}/c^2 \tag{9.1}$$

Now, a *massive* particle of spin 1 has *three* allowed polarization states ($m_s = 1, 0, -1$), whereas a free *massless* particle has only *two* (if z is the direction of motion, the 'longitudinal' polarization $m_s = 0$ does not occur). Thus, for photons and gluons, we imposed both the Lorentz condition

$$\epsilon^\mu p_\mu = 0 \tag{9.2}$$

(reducing the number of independent components in ϵ^μ from 4 to 3) and also the Coulomb gauge ($\epsilon^0 = 0$, so that $\boldsymbol{\epsilon}\cdot\mathbf{p} = 0$, which reduces it further from 3 to 2). For the W's and the Z we do *not* impose the latter constraint. As a result, the completeness relation is quite different (see Problem 9.1) and the propagator is

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

no longer simply $-ig_{\mu\nu}/q^2$, but rather,*

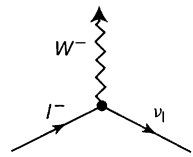
$$\frac{-i(g_{\mu\nu} - q_\mu q_\nu/M^2c^2)}{q^2 - M^2c^2} \quad \text{(propagator for } W \text{ and } Z\text{)} \tag{9.3}$$

where M is M_W or M_Z, as the case may be. In practice, q^2 is ordinarily so much smaller than $(Mc)^2$ that we may safely use

$$\frac{ig_{\mu\nu}}{(Mc)^2} \quad \text{(propagator for } q^2 \ll (Mc)^2\text{)} \tag{9.4}$$

However, when a process involves energies that are comparable to Mc^2 we must, of course, revert to the exact expression.

The theory of 'charged' weak interactions (mediated by the W's) is simpler than that for 'neutral' ones (mediated by the Z), so for the moment I shall concentrate on the former. In this section we consider the coupling of W's to leptons; in the next section we'll discuss their coupling to quarks and hadrons. The fundamental leptonic vertex is



Here an electron, muon, or tau is converted into the associated neutrino, with emission of a W^- (or absorption of W^+). The reverse process ($\nu_l \to l^- + W^+$) is also possible, of course, as well as the 'crossed' reactions involving antileptons. The Feynman rules are the same as for QED (apart from the modifications already mentioned to accommodate the massive mediator), except for the vertex factor, which is

$$\frac{-ig_w}{2\sqrt{2}}\gamma^\mu(1 - \gamma^5) \quad \text{(weak vertex factor)} \tag{9.5}$$

The various 2's are purely conventional, and $g_w = \sqrt{4\pi\alpha_w}$ is the 'weak coupling constant' (analogous to g_e in QED and g_s in QCD). The term $(1 - \gamma^5)$, however, is of profound importance, for γ^μ alone would represent a *vector* coupling (like QED or QCD), whereas $\gamma^\mu\gamma^5$ would be an *axial* vector (see Equation 7.68). A theory that

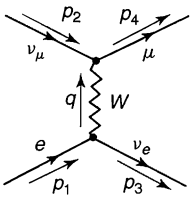
* It might bother you that this does not reduce to the photon propagator as $M \to 0$. For particles of spin 1 (or higher), the massless limit is notoriously treacherous, because in one critical respect it is not a continuous procedure. The number of degrees of freedom (that is, the number of allowed spin orientations) drops abruptly from $2s + 1$ (for $M \neq 0$) to 2 (for $M = 0$). There are ways of formulating the theory that allow a smooth transition to $M = 0$, but only at the cost of introducing spurious nonphysical states.

adds a *vector* to an *axial* vector is bound to violate the conservation of parity, and this is precisely what happens in the weak interactions.*

Example 9.1 Inverse Muon Decay Consider the process

$$\nu_\mu + e^- \to \mu^- + \nu_e$$

represented (in lowest order) by the diagram



Here $q = p_1 - p_3$, and we'll assume $q^2 \ll M_W^2 c^2$, so we can safely use the simplified propagator (Equation 9.4); the amplitude is

$$\mathscr{M} = \frac{g_w^2}{8(M_W c)^2}[\overline{u}(3)\gamma^\mu(1-\gamma^5)u(1)][\overline{u}(4)\gamma_\mu(1-\gamma^5)u(2)] \tag{9.6}$$

Applying Casimir's trick (Equation 7.125), and assuming the neutrino masses are negligible, we find

$$\sum_{\text{spins}} |\mathscr{M}|^2 = \left(\frac{g_w^2}{8(M_W c)^2}\right)^2 \mathrm{Tr}[\gamma^\mu(1-\gamma^5)(\not{p}_1 + m_e c)\gamma^\nu(1-\gamma^5)\not{p}_3]$$
$$\times\ \mathrm{Tr}[\gamma_\mu(1-\gamma^5)\not{p}_2\gamma_\nu(1-\gamma^5)(\not{p}_4 + m_\mu c)] \tag{9.7}$$

The theorems of Section 7.7 yield

$$8[p_1^\mu p_3^\nu + p_1^\nu p_3^\mu - g^{\mu\nu}(p_1 \cdot p_3) - i\epsilon^{\mu\nu\lambda\sigma} p_{1_\lambda} p_{3_\sigma}] \tag{9.8}$$

for the first trace, and

$$8[p_{2_\mu} p_{4_\nu} + p_{2_\nu} p_{4_\mu} - g_{\mu\nu}(p_2 \cdot p_4) - i\epsilon_{\mu\nu\kappa\tau} p_2^\kappa p_4^\tau] \tag{9.9}$$

* In fact, the violation is 'maximal', in the sense that the two terms are equally large. When parity violation was first considered, a factor of the form $(1 + \epsilon\gamma^5)$ was used, but experiments soon dictated that $\epsilon = -1$ (see Problem 9.3). We call it a '$V-A$' ('vector minus axial vector') coupling. Fermi's original theory of beta decay was a pure *vector* theory (like QED), and although others proposed scalar, pseudoscalar, tensor, or pure axial couplings, it was not until 1956 that anyone seriously contemplated *mixing* terms of different parity.

for the second. It follows that*

$$\sum_{\text{spins}} |\mathscr{M}|^2 = 4\left(\frac{g_w}{M_W c}\right)^4 (p_1 \cdot p_2)(p_3 \cdot p_4) \tag{9.10}$$

Actually, we want the sum over final spins but the *average* over initial spins. The electron has *two* spin states, but (massless) neutrinos (as we learned in Section 4.6) have only *one* (if you like, the incident neutrinos are *always* polarized, since they only come 'left-handed'). So

$$\langle|\mathscr{M}|^2\rangle = 2\left(\frac{g_w}{M_W c}\right)^4 (p_1 \cdot p_2)(p_3 \cdot p_4) \tag{9.11}$$

If we now go to the CM frame, and neglect the mass of the electron

$$\langle|\mathscr{M}|^2\rangle = 8\left(\frac{g_w E}{M_W c^2}\right)^4 \left\{1 - \left(\frac{m_\mu c^2}{2E}\right)^2\right\} \tag{9.12}$$

where E is the incident electron (or neutrino) energy. The differential scattering cross section (Equation 6.47) is isotropic (all scattering angles equally likely)

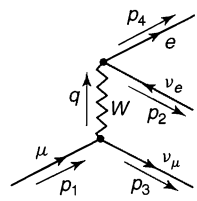
$$\frac{d\sigma}{d\Omega} = \frac{1}{2}\left[\frac{\hbar c g_w^2 E}{4\pi (M_W c^2)^2}\right]^2 \left\{1 - \left(\frac{m_\mu c^2}{2E}\right)^2\right\}^2 \tag{9.13}$$

and the total cross section is

$$\sigma = \frac{1}{8\pi}\left[\left(\frac{g_w}{M_W c^2}\right)^2 \hbar c E\right]^2 \left\{1 - \left(\frac{m_\mu c^2}{2E}\right)^2\right\}^2 \tag{9.14}$$

9.2 Decay of the Muon

Electron–neutrino scattering is not the easiest thing in the world to study experimentally, but the closely related process, muon decay ($\mu \to e + \nu_\mu + \overline{\nu}_e$), is the cleanest of all weak interaction phenomena, theoretically and experimentally. The Feynman diagram



* Note that $\epsilon^{\mu\nu\lambda\sigma}\epsilon_{\mu\nu\kappa\tau} = -2(\delta^\lambda_\kappa\delta^\sigma_\tau - \delta^\lambda_\tau\delta^\sigma_\kappa)$ (Problem 7.35). The traces in Equation 9.7 are special cases of a structure that will occur repeatedly in this chapter; it might be a good idea to pause here and work out the generic result (Problem 9.2).

leads to the amplitude

$$\mathscr{M} = \frac{g_w^2}{8(M_W c)^2}[\overline{u}(3)\gamma^\mu(1-\gamma^5)u(1)][\overline{u}(4)\gamma_\mu(1-\gamma^5)v(2)] \tag{9.15}$$

from which we obtain, as before,

$$\langle|\mathscr{M}|^2\rangle = 2\left(\frac{g_w}{M_W c}\right)^4 (p_1 \cdot p_2)(p_3 \cdot p_4) \tag{9.16}$$

In the muon rest frame, $p_1 = (m_\mu c, \mathbf{0})$, we have

$$p_1 \cdot p_2 = m_\mu E_2 \tag{9.17}$$

and since $p_1 = p_2 + p_3 + p_4$

$$\begin{aligned}(p_3+p_4)^2 &= p_3^2 + p_4^2 + 2p_3 \cdot p_4 = m_e^2 c^2 + 2p_3 \cdot p_4 \\ &= (p_1 - p_2)^2 = p_1^2 + p_2^2 - 2p_1 \cdot p_2 = m_\mu^2 c^2 - 2p_1 \cdot p_2\end{aligned} \tag{9.18}$$

from which it follows that

$$p_3 \cdot p_4 = \frac{(m_\mu^2 - m_e^2)c^2}{2} - m_\mu E_2 \tag{9.19}$$

The algebra will be simpler later on, at no significant cost in accuracy, if we set $m_e = 0$, so that

$$\langle|\mathscr{M}|^2\rangle = \left(\frac{g_w}{M_W c}\right)^4 m_\mu^2 E_2(m_\mu c^2 - 2E_2) = \left(\frac{g_w^2 m_\mu}{M_W^2 c}\right)^2 |\mathbf{p}_2|\left(m_\mu c - 2|\mathbf{p}_2|\right) \tag{9.20}$$

The decay rate is given by Equation 6.21:*

$$\begin{aligned}d\Gamma = &\frac{\langle|\mathscr{M}|^2\rangle}{2\hbar m_\mu}\left(\frac{d^3\mathbf{p}_2}{(2\pi)^3 2|\mathbf{p}_2|}\right)\left(\frac{d^3\mathbf{p}_3}{(2\pi)^3 2|\mathbf{p}_3|}\right)\left(\frac{d^3\mathbf{p}_4}{(2\pi)^3 2|\mathbf{p}_4|}\right) \\ &\times (2\pi)^4\delta^4(p_1 - p_2 - p_3 - p_4)\end{aligned} \tag{9.21}$$

To begin with, we peel apart the delta function:

$$\delta^4(p_1 - p_2 - p_3 - p_4) = \delta\left(m_\mu c - |\mathbf{p}_2| - |\mathbf{p}_3| - |\mathbf{p}_4|\right)\delta^3(\mathbf{p}_2 + \mathbf{p}_3 + \mathbf{p}_4) \tag{9.22}$$

and perform the $\mathbf{p}_3$ integral:

$$d\Gamma = \frac{\langle|\mathscr{M}|^2\rangle}{16(2\pi)^5\hbar m_\mu}\frac{(d^3\mathbf{p}_2)(d^3\mathbf{p}_4)}{|\mathbf{p}_2|\,|\mathbf{p}_2 + \mathbf{p}_4|\,|\mathbf{p}_4|}\delta\left(m_\mu c - |\mathbf{p}_2| - |\mathbf{p}_2 + \mathbf{p}_4| - |\mathbf{p}_4|\right) \tag{9.23}$$

* Note that this is a *three* body decay, so we have to go all the way back to the Golden Rule.

Next we'll do the $\mathbf{p}_2$ integral. Setting the polar axis along $\mathbf{p}_4$ (which is *fixed*, for the purposes of the $\mathbf{p}_2$ integration), we have

$$d^3\mathbf{p}_2 = |\mathbf{p}_2|^2\, d|\mathbf{p}_2|\, \sin\theta\, d\theta\, d\phi \tag{9.24}$$

and

$$|\mathbf{p}_2 + \mathbf{p}_4|^2 = |\mathbf{p}_2|^2 + |\mathbf{p}_4|^2 + 2|\mathbf{p}_2|\,|\mathbf{p}_4|\cos\theta \equiv u^2 \tag{9.25}$$

The ϕ integral is trivial ($\int d\phi = 2\pi$); to carry out the θ integration we change variables ($\theta \to u$):

$$2\,u\,du = -2\,|\mathbf{p}_2|\,|\mathbf{p}_4|\,\sin\theta\, d\theta \tag{9.26}$$

so

$$d\Gamma = \frac{\langle|\mathscr{M}|^2\rangle}{16(2\pi)^4\hbar m_\mu}\frac{d^3\mathbf{p}_4}{|\mathbf{p}_4|^2}d|\mathbf{p}_2|\int_{u_-}^{u_+}\delta\left(m_\mu c - |\mathbf{p}_2| - |\mathbf{p}_4| - u\right)\, du \tag{9.27}$$

where

$$u_\pm \equiv \sqrt{|\mathbf{p}_2|^2 + |\mathbf{p}_4|^2 \pm 2|\mathbf{p}_2|\,|\mathbf{p}_4|} = \Big||\mathbf{p}_2| \pm |\mathbf{p}_4|\Big| \tag{9.28}$$

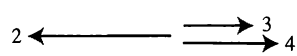
The u integral is 1 if

$$u_- < m_\mu c - |\mathbf{p}_2| - |\mathbf{p}_4| < u_+ \tag{9.29}$$

(and 0 otherwise) – which is to say (Problem 9.4),

$$\left\{\begin{array}{r} |\mathbf{p}_2| < \frac{1}{2}m_\mu c \\ |\mathbf{p}_4| < \frac{1}{2}m_\mu c \\ (|\mathbf{p}_2| + |\mathbf{p}_4|) > \frac{1}{2}m_\mu c \end{array}\right\} \tag{9.30}$$

These constraints make good sense kinematically: particle 2, for example, gets the maximum possible momentum when 3 and 4 emerge diametrically opposite to it:



In this case 2 picks up half the available energy ($\frac{1}{2}m_\mu c^2$), while 3 and 4 share the rest. If there is a nonzero angle between 3 and 4, 2 gets less, and 3 plus 4 correspondingly more. Thus $\frac{1}{2}m_\mu c$ is the maximum momentum for any *individual* outgoing particle, and the *minimum* total for any *pair*.

The θ and ϕ integrals have left us with

$$d\Gamma = \frac{\langle|\mathscr{M}|^2\rangle}{(4\pi)^4\hbar m_\mu}d|\mathbf{p}_2|\frac{d^3\mathbf{p}_4}{|\mathbf{p}_4|^2} \tag{9.31}$$

The inequalities in Equation 9.30 specify the limits on the $|\mathbf{p}_2|$ and $|\mathbf{p}_4|$ integrals: $|\mathbf{p}_2|$ runs from $\frac{1}{2}m_\mu c - |\mathbf{p}_4|$ up to $\frac{1}{2}m_\mu c$, and $|\mathbf{p}_4|$ will then go from 0 to $\frac{1}{2}m_\mu c$. Putting in Equation 9.20* and carrying out the $|\mathbf{p}_2|$ integral, we have

$$\begin{aligned} d\Gamma &= \left(\frac{g_w}{4\pi M_W}\right)^4 \frac{m_\mu}{\hbar c^2}\frac{d^3\mathbf{p}_4}{|\mathbf{p}_4|^2}\int_{(1/2)m_\mu c-|\mathbf{p}_4|}^{(1/2)m_\mu c} |\mathbf{p}_2|(m_\mu c - 2|\mathbf{p}_2|)\, d|\mathbf{p}_2| \\ &= \left(\frac{g_w}{4\pi M_W}\right)^4 \frac{m_\mu}{\hbar c^2}\left(\frac{m_\mu c}{2} - \frac{2}{3}|\mathbf{p}_4|\right) d^3\mathbf{p}_4 \end{aligned} \tag{9.32}$$

Finally, writing

$$d^3\mathbf{p}_4 = 4\pi|\mathbf{p}_4|^2\, d|\mathbf{p}_4|$$

and expressing the answer in terms of the electron energy, $E = |\mathbf{p}_4|c$, we conclude†

$$\frac{d\Gamma}{dE} = \left(\frac{g_w}{M_W c}\right)^4 \frac{m_\mu^2 E^2}{2\hbar(4\pi)^3}\left(1 - \frac{4E}{3m_\mu c^2}\right) \tag{9.33}$$

This tells us the energy distribution of the electrons emitted in muon decay; it nicely matches the experimental spectrum (Figure 9.1). The total decay rate is

$$\begin{aligned} \Gamma &= \left(\frac{g_w}{M_W c}\right)^4 \frac{m_\mu^2}{2\hbar(4\pi)^3}\int_0^{(1/2)m_\mu c^2} E^2\left(1 - \frac{4E}{3m_\mu c^2}\right) dE \\ &= \left(\frac{m_\mu g_w}{M_W}\right)^4 \frac{m_\mu c^2}{12\hbar(8\pi)^3} \end{aligned} \tag{9.34}$$

and hence the lifetime of the muon is

$$\tau = \frac{1}{\Gamma} = \left(\frac{M_W}{m_\mu g_w}\right)^4 \frac{12\hbar(8\pi)^3}{m_\mu c^2} \tag{9.35}$$

Notice that g_w and M_W do not appear *separately*, either in the muon lifetime formula or in the electron–neutrino scattering cross section; only their *ratio* occurs. It is traditional, in fact, to express weak interaction formulas in terms of the 'Fermi coupling constant'

$$G_F \equiv \frac{\sqrt{2}}{8}\left(\frac{g_w}{M_W c^2}\right)^2 (\hbar c)^3 \tag{9.36}$$

* Notice that $\langle\, |\mathscr{M}|^2 \,\rangle$ depends only on the *magnitude* of $\mathbf{p}_2$, not on its *direction*; that's why I was free to ignore it in the θ and ϕ integrations.

† Remember that Equation 9.33 applies *only* up to $E = \frac{1}{2}m_\mu c^2$ (Equation 9.30), at which point it drops abruptly to zero (the corners are softened a bit by the inclusion of particle masses and radiative corrections).

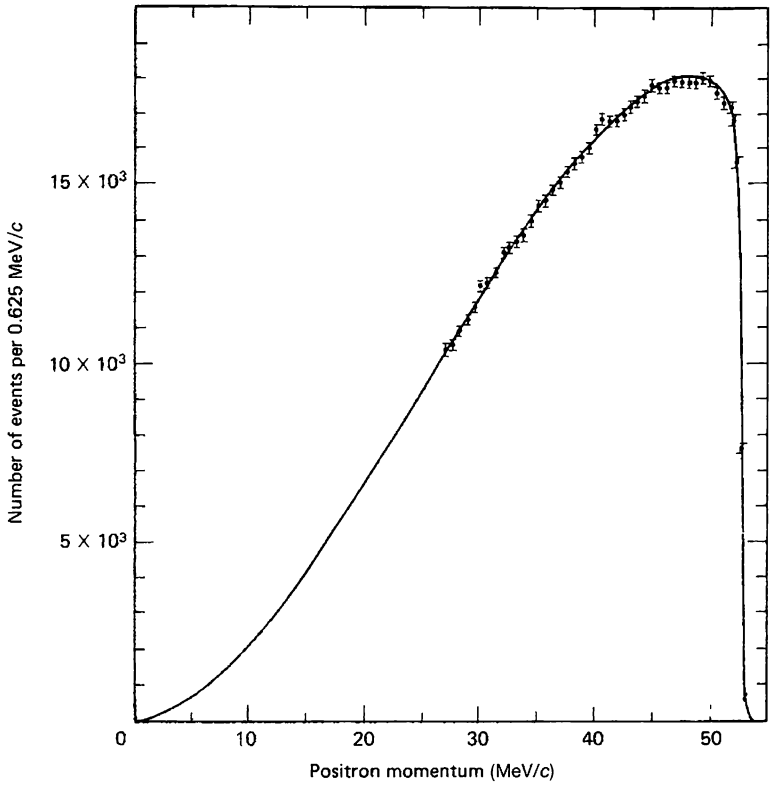
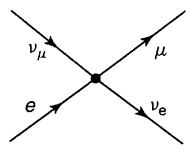


Fig. 9.1 Experimental spectrum of positrons in $\mu^+ \to e^+ + \nu_e + \nu_\mu$. The solid line is the theoretically predicted spectrum based on Equation (9.33), corrected for electromagnetic effects. (*Source*: Bardon, M. *et al.* (1965) *Physical Review Letters*, **14**, 449. For the latest high-precision data on muon decay go to the TWIST collaboration web site at TRIUMF, Vancouver, BC.)

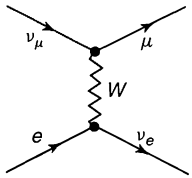
Thus the muon lifetime is written

$$\tau = \frac{192\pi^3\hbar^7}{G_F^2 m_\mu^5 c^4} \tag{9.37}$$

In Fermi's original theory of beta decay (1933) there *was* no W; the interaction was supposed to be a direct four-particle coupling, represented in the Feynman language by a diagram of the form



From the modern perspective, Fermi's theory combined the W propagator with the two vertex factors, in the diagram



to make an effective four-particle coupling constant G_F. It worked, but only because the W is so heavy that Equation 9.4 is a good approximation to the true propagator (Equation 9.3),* and in fact it was recognized already in the 1950s that Fermi's theory could not be valid at high energies. The idea of a weak mediator (analogous to the photon) was suggested by Klein as far back as 1938.

If we put in the observed muon lifetime and mass, we find that

$$G_F/(\hbar c)^3 = \frac{\sqrt{2}}{8}\left(\frac{g_w}{M_W c^2}\right)^2 = 1.166 \times 10^{-5}/\text{GeV}^2 \tag{9.38}$$

The corresponding value of g_w is

$$g_w = 0.653 \tag{9.39}$$

and hence the 'weak fine structure constant' is

$$\alpha_w = \frac{g_w^2}{4\pi} = \frac{1}{29.5} \tag{9.40}$$

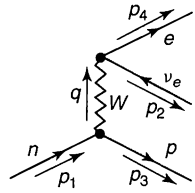
This number should come as something of a shock: it is *larger* than the electromagnetic fine structure constant ($\alpha = \frac{1}{137}$), by a factor of nearly 5! Weak interactions are feeble *not* because the *intrinsic* coupling is small (it *isn't*), but because the *mediators* are so *massive* – or, more precisely, because we typically work at energies so far below the W mass that the denominator in the propagator $|q^2 - M_W^2 c^2|$ is extremely large.

9.3 Decay of the Neutron

The success of the muon decay formula (Equation 9.33) encourages us to apply the same methods to the decay of the neutron, $n \to p + e + \overline{\nu}_e$. Of course, the neutron and proton are composite particles, but just as the Mott and Rutherford cross sections (which treat the proton as an elementary 'Dirac' particle) give a

* Fermi also thought the coupling was pure vector, as I mentioned earlier. Despite these defects (for which Fermi could scarcely be blamed; after all, he invented the theory at a time when the neutrino was a wild speculation and the Dirac equation itself was brand new), Fermi's theory was astonishingly prescient, and all subsequent developments have been relatively small adjustments to it.

good account of low-energy electron–proton scattering, so we might hope that the diagram



(the same as for muon decay, only with $n \to p + W^-$ in place of $\mu \to \nu_\mu + W^-$) will afford a reasonable approximation to neutron beta decay. From a calculational point of view the only new feature is that 3 is now a massive particle (a proton, instead of a neutrino). As it happens (Problem 9.8) this does not change the amplitude:

$$\langle|\mathscr{M}|^2\rangle = 2\left(\frac{g_w}{M_W c}\right)^4 (p_1 \cdot p_2)(p_3 \cdot p_4) \tag{9.41}$$

– same as Equation 9.16. In the rest frame of the neutron, we find

$$\langle|\mathscr{M}|^2\rangle = \frac{m_n}{c}\left(\frac{g_w}{M_W}\right)^4 |\mathbf{p}_2|\left(m_n^2 - m_p^2 - m_e^2 - \frac{2m_n|\mathbf{p}_2|}{c}\right) \tag{9.42}$$

But because the electron rest energy is a substantial fraction of the total energy released, $(m_n - m_p - m_e)c^2$, we *cannot* afford to ignore the electron mass, this time.

The decay rate calculation proceeds as before (with the masses now included):

$$d\Gamma = \frac{\langle|\mathscr{M}|^2\rangle}{2\hbar m_n}\left(\frac{d^3\mathbf{p}_2}{(2\pi)^3 2|\mathbf{p}_2|}\right)\left(\frac{d^3\mathbf{p}_3}{(2\pi)^3 2\sqrt{\mathbf{p}_3^2 + m_p^2c^2}}\right)\left(\frac{d^3\mathbf{p}_4}{(2\pi)^3 2\sqrt{\mathbf{p}_4^2 + m_e^2c^2}}\right)$$
$$\times (2\pi)^4\delta^4(p_1 - p_2 - p_3 - p_4) \tag{9.43}$$

The $\mathbf{p}_3$ integral yields

$$d\Gamma = \frac{\langle|\mathscr{M}|^2\rangle}{16(2\pi)^5\hbar m_n}\frac{d^3\mathbf{p}_2\, d^3\mathbf{p}_4}{|\mathbf{p}_2| u\sqrt{\mathbf{p}_4^2 + m_e^2c^2}}\delta\left(m_n c - |\mathbf{p}_2| - u - \sqrt{\mathbf{p}_4^2 + m_e^2c^2}\right) \tag{9.44}$$

where

$$u \equiv \sqrt{(\mathbf{p}_2 + \mathbf{p}_4)^2 + m_p^2c^2} \tag{9.45}$$

To carry out the $\mathbf{p}_2$ integral, we again set

$$d^3\mathbf{p}_2 = |\mathbf{p}_2|^2 d|\mathbf{p}_2| \sin\theta\, d\theta\, d\phi \tag{9.46}$$

and orient the coordinates so that the z axis lies along $\mathbf{p}_4$ (which is fixed, for purposes of the $\mathbf{p}_2$ integral); then

$$u^2 = |\mathbf{p}_2|^2 + |\mathbf{p}_4|^2 + 2|\mathbf{p}_2||\mathbf{p}_4|\cos\theta + m_p^2c^2 \tag{9.47}$$

and

$$u\,\mathrm{d}u = -|\mathbf{p}_2|\,|\mathbf{p}_4|\,\sin\theta\,\mathrm{d}\theta \tag{9.48}$$

The ϕ and θ (or rather, u) integrals yield

$$\mathrm{d}\Gamma = \frac{\langle|\mathscr{M}|^2\rangle}{(4\pi)^4\hbar m_n}\frac{\mathrm{d}^3\mathbf{p}_4}{|\mathbf{p}_4|\sqrt{\mathbf{p}_4^2 + m_e^2c^2}}\,\mathrm{d}|\mathbf{p}_2|\,I \tag{9.49}$$

where

$$\begin{aligned} I &\equiv \int_{u_-}^{u_+} \delta\left(m_nc - |\mathbf{p}_2| - \sqrt{|\mathbf{p}_4|^2 + m_e^2c^2} - u\right)\mathrm{d}u \\ &= \left\{\begin{array}{ll} 1, & \text{if } u_- < \left(m_nc - |\mathbf{p}_2| - \sqrt{|\mathbf{p}_4|^2 + m_e^2c^2}\right) < u_+ \\ 0, & \text{otherwise} \end{array}\right\} \end{aligned} \tag{9.50}$$

and the limits are

$$u_\pm = \sqrt{(|\mathbf{p}_2| \pm |\mathbf{p}_4|)^2 + m_p^2c^2} \tag{9.51}$$

As before, Equation 9.50 defines the range of the $|\mathbf{p}_2|$ integral; I'll let you work out the algebra (Problem 9.9):

$$p_\pm = \frac{\frac{1}{2}(m_n^2 - m_p^2 + m_e^2)c^2 - m_n\sqrt{|\mathbf{p}_4|^2 + m_e^2c^2}}{m_nc - \sqrt{|\mathbf{p}_4|^2 + m_e^2c^2} \mp |\mathbf{p}_4|} \tag{9.52}$$

With $\langle\,|\mathscr{M}|^2\,\rangle$ from Equation 9.42, the $|\mathbf{p}_2|$ integral becomes

$$\int_{p_-}^{p_+} |\mathbf{p}_2|\left(m_n^2 - m_p^2 - m_e^2 - \frac{2m_n|\mathbf{p}_2|}{c}\right)\mathrm{d}|\mathbf{p}_2| \equiv J \tag{9.53}$$

and since

$$\mathrm{d}^3\mathbf{p}_4 = 4\pi|\mathbf{p}_4|^2\mathrm{d}|\mathbf{p}_4| \tag{9.54}$$

we conclude that

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}E} = \frac{1}{\hbar c^2(4\pi)^3}\left(\frac{g_w}{M_Wc}\right)^4 J(E) \tag{9.55}$$

where $E = c\sqrt{|\mathbf{p}_4|^2 + m_e^2c^2}$ is the electron energy.

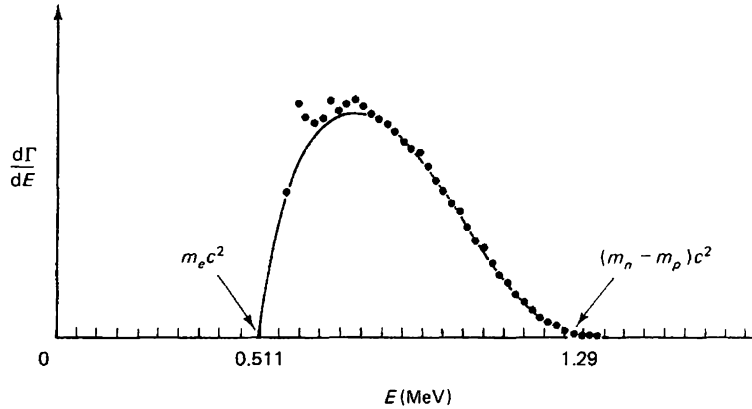


Fig. 9.2 Electron energy distribution from neutron beta decay. (Solid line is the theoretical curve; dots are experimental data.) (*Source*: Christensen, C. J. *et al.* (1972) *Physical Review*, **D5**, 1628. Figure (9.4).)

Equation 9.55 is *exact* (use it, if you like, to rederive Equation 9.33, by setting $m_n \to m_\mu$ and m_p, $m_e \to 0$), but $J(E)$ is a rather cumbersome function:

$$J(E) = \tfrac{1}{2}(m_n^2 - m_p^2 - m_e^2)c^4(p_+^2 - p_-^2) - \tfrac{2m_n c^3}{3}(p_+^3 - p_-^3) \tag{9.56}$$

It pays to approximate, at this stage, recognizing that there are four small numbers here:

$$\epsilon \equiv \frac{m_n - m_p}{m_n} = 0.0014, \qquad \delta \equiv \frac{m_e}{m_n} = 0.0005,$$
$$\eta \equiv \frac{E}{m_n c^2} \quad (\delta < \eta < \epsilon), \qquad \phi \equiv \frac{|\mathbf{p}_4|}{m_n c} \quad (0 < \phi < \eta) \tag{9.57}$$

(The last of these is not independent, of course: $\phi^2 = \eta^2 - \delta^2$.) Expanding to lowest order (Problem 9.9), we obtain

$$J \cong 4m_n^4 c^6 \eta\phi(\epsilon - \eta)^2 = \frac{4}{c^2} E\sqrt{E^2 - m_e^2 c^4}\,[(m_n - m_p)c^2 - E]^2 \tag{9.58}$$

So the distribution of electron energies is given by

$$\frac{d\Gamma}{dE} = \frac{1}{\pi^3 \hbar}\left(\frac{g_w}{2M_W c^2}\right)^4 E\sqrt{E^2 - m_e^2 c^4}\,[(m_n - m_p)c^2 - E]^2 \tag{9.59}$$

The experimental results are shown in Figure 9.2. The electron energies range from $m_e c^2$ up to about $(m_n - m_p)c^2$ (Problem 9.10). Integrating over E, we get the

total decay rate (Problem 9.11):

$$\Gamma = \frac{1}{4\pi^3\hbar}\left(\frac{g_w}{2M_W c^2}\right)^4 (m_e c^2)^5$$
$$\times \left[\frac{1}{15}(2a^4 - 9a^2 - 8)\sqrt{a^2-1} + a\ln(a + \sqrt{a^2-1})\right] \tag{9.60}$$

where

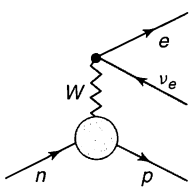
$$a \equiv \frac{m_n - m_p}{m_e} \tag{9.61}$$

Putting in the numbers, I find (Problem 9.12)

$$\tau = \frac{1}{\Gamma} = 1318\ \text{s} \tag{9.62}$$

This is in the ball park, as they say: the experimental neutron lifetime* is 885.7 ± 0.8 seconds, and given that weak decays range from 15 minutes down to 10^{-13} seconds we should perhaps be satisfied to get the right order of magnitude. But why isn't the agreement better?

The main problem is that we have treated the proton and neutron as though they were simple point particles, interacting with the W in exactly the same way leptons do. To be honest about it, we should go back to the beginning, admit that we do not really know how the W couples to composite structures, draw in a blob on the Feynman diagram (to symbolize our ignorance)



and express the amplitude in terms of various unknown 'form factors', whose structure is limited only by Lorentz covariance – just as we did in Chapter 8 for the proton–*photon* vertex. Not until a mature QCD can provide us with the detailed structure of the nucleons will we be in a position to perfect the neutron lifetime calculation.

* This number is from the 2006 *Particle Physics Booklet* (PPB). Free neutrons are hard to work with, and the 'official' neutron lifetime has changed substantially over the years (the first PPB listed it as 1040 ± 130 seconds). Note also that nuclear physicists tend to quote the *half-life* ($t_{1/2} = \tau \ln 2$), and beta-decay specialists often quote the 'comparative half-life' – the so-called '*ft*' value – which has certain kinematic and Coulombic contributions removed (for the neutron the correction factor is about 1.7). This is just to warn you that the numbers given in the literature for the neutron 'lifetime' are all over the map, and it pays to read the fine print and check the date.

And yet, the Mott formula works well for low-energy electron–proton scattering: why does essentially the same procedure give us the right answer in electrodynamics, but not in the weak interactions? In both cases the wavelength of the 'probe' (γ or W, as the case may be) is much larger than the diameter of the 'target' (p or n) (see Problem 9.13); the nucleon's internal structure is not 'resolved', and it behaves as a point particle. The *crucial* question, though, is: *what is the net coupling strength of this object?* Of course, the net *charge* of the proton is still e – it doesn't matter what complicated processes are going on inside – valence quarks emitting virtual gluons, gluons producing quark–antiquark pairs, 'sea' quarks recombining, and so on – because *all this frenzied activity conserves charge*. From the perspective of a long wavelength photon it just looks like a point, and the net charge of the composite nucleon is just the sum of the charges of the valence quarks. But there is no a priori reason to suppose that the same applies to the *weak* coupling; when a gluon splits into a quark–antiquark pair, the net contribution of this pair to the weak coupling may *not* be zero – who knows? To account for this, we make the following replacement in the $n \to p + W$ vertex factor:

$$(1 - \gamma^5) \to (c_V - c_A \gamma^5) \tag{9.63}$$

where c_V is the correction to the vector 'weak charge', and c_A is the correction to the axial vector 'weak charge'. Luckily, the same basic process, $n \to p + e + \overline{\nu}_e$, occurs not only for the *free* neutron, but also within radioactive nuclei, so we have, in principle, many independent opportunities to measure c_V and c_A.* The experimental results are as follows:

$$c_V = 1.000, \quad c_A = 1.270 \pm 0.003 \tag{9.64}$$

Surprisingly, the *vector* weak charge is *not* modified by the strong interactions within the nucleon. Presumably, like electric charge, it is 'protected' by a conservation law; we call this the 'Conserved Vector Current' (CVC) hypothesis.† Even the axial term is not altered *much*; evidently it is 'almost' conserved. We call this the 'Partially Conserved Axial Current' (PCAC) hypothesis.

The effect of this substitution (Equation 9.63) on the neutron lifetime is something you can calculate for yourself, if you have the stamina; to good approximation, the decay rate is increased by a factor of

$$\tfrac{1}{4}(c_V^2 + 3c_A^2) = 1.46 \tag{9.65}$$

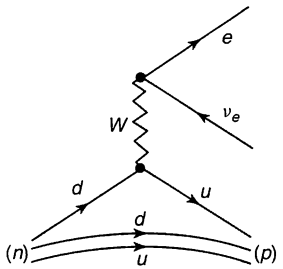
and the lifetime is decreased in the same ratio:

$$\tau = \frac{1316 \text{ s}}{1.46} = 901 \text{ s} \tag{9.66}$$

* A particular favorite is $^{14}\text{O} \to {}^{14}\text{N}$, which is known (from the observed spin and parity of the initial and final states) to involve only *vector* coupling.

† CVC is built into the Standard Model, and nowadays c_V is taken to be 1 exactly; the experiments are interpreted as measurements of the Cabibbo angle (see below) – or, more precisely, of V_{ub}.

This is now within striking distance of the experimental value. Unfortunately, the agreement is deceptive, for there is yet *another* correction to be made. The underlying quark process here is $d \to u + W$ (with two spectators):



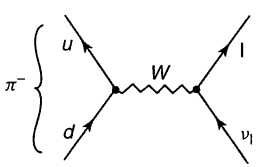
and this quark vertex carries a factor of $\cos\theta_C$, where $\theta_C = 13.15°$ is the 'Cabibbo angle'. I'll have more to say about this in Section 9.5, but the essential point for now is that our theoretical value for the neutron lifetime, corrected for nonconservation of the axial charge and modified by the Cabibbo angle, is

$$\tau = \frac{901\ \text{s}}{\cos^2\theta_C} = 950\ \text{s} \tag{9.67}$$

Two steps forward, one step back!*

9.4 Decay of the Pion

According to the quark model, the decay of a charged pion ($\pi^- \to l^- + \bar{\nu}_l$, where l is a muon or an electron) is really a *scattering* event in which the incident quarks happen to be bound together:

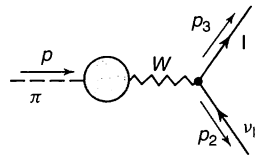


In this sense, it is a weak interaction analog to positronium decay ($e^+ + e^- \to \gamma + \gamma$) or η_c decay ($c + \bar{c} \to g + g$) – electromagnetic and strong processes, respectively. We *could* analyze it this way, following the methods of Example 7.8 and Section 8.5 (see Problem 9.14), but in the end we would be stuck with a factor of $|\psi(0)|^2$, and at this stage we have no idea what the wave function (ψ) of the quarks within a pion

* This isn't the end of the story; there is, for example, a small Coulomb correction, (due to the attraction of the electron and proton in the final state). But we are within 7% of the experimental result, and it is time to move on.

looks like. Given that such a calculation will carry this undetermined multiplicative factor *anyway*, it is simpler to proceed as follows.

Redraw the Feynman diagram, with a blob to represent the coupling of π^- to W^-:



We may not know how the W couples to the pion, but we *do* know how it couples to the leptons, so the amplitude must have the general form

$$\mathscr{M} = \frac{g_w^2}{8(M_W c)^2}[\overline{u}(3)\gamma_\mu(1-\gamma^5)v(2)]F^\mu \tag{9.68}$$

where F^μ is a 'form factor' describing the $\pi \to W$ blob. It has to be a four-vector, to contract with the γ_μ in the lepton term. But the pion has spin zero; the only vector associated with it, out of which we might construct F^μ, is its momentum, p^μ.* (I won't bother with a subscript on the pion's momentum: $p \equiv p_1$.) So F^μ must be some scalar quantity times p^μ:†

$$F^\mu = f_\pi p^\mu \tag{9.69}$$

In principle, f_π is a function of p^2 – the only available scalar – but since the pion is on its mass shell ($p^2 = m_\pi^2 c^2$), f_π is, for our purposes, a fixed number, the 'pion decay constant'.‡

Summing over the outgoing spins, we get

$$\begin{aligned}\langle|\mathscr{M}|^2\rangle &= \left[\frac{f_\pi}{8}\left(\frac{g_w}{M_W c}\right)^2\right]^2 p_\mu p_\nu \mathrm{Tr}[\gamma^\mu(1-\gamma^5)\not{p}_2\gamma^\nu(1-\gamma^5)(\not{p}_3 + m_l c)] \\ &= \frac{1}{8}\left[f_\pi\left(\frac{g_w}{M_W c}\right)^2\right]^2 [2(p\cdot p_2)(p\cdot p_3) - p^2(p_2\cdot p_3)]\end{aligned} \tag{9.70}$$

* Notice that we introduce the (weak) *pion* form factor at the level of $\mathscr{M}$, whereas for the (electromagnetic) proton form factors we waited until the $\langle|\mathscr{M}|^2\rangle$ stage. The reason is that the proton has a spin, and we would have to include that in the roster of available vectors; it is only after we have *averaged* over the spins that the list reduces to two, and the problem becomes manageable. The pion, however, has no spin, so we can afford to introduce the form factor directly in $\mathscr{M}$, where it is only a *vector* quantity, instead of a tensor.

† For reasons that will appear in the next section, it is customary nowadays to factor out the appropriate Cabibbo–Kobayashi–Maskawa (CKM) matrix element in the definition of the meson decay constants: $f_\pi \to V_{du} f_\pi$. To avoid cluttered notation I'll use the older convention.

‡ The corresponding factor for other pseudoscalar mesons will involve a different value of p^2, and a different element in the CKM matrix (see footnote†).

(the trace was already calculated in Equation 9.8). But $p = p_2 + p_3$, so

$$p \cdot p_2 = p_2 \cdot p_3, \qquad p \cdot p_3 = m_l^2 c^2 + p_2 \cdot p_3 \tag{9.71}$$

and

$$p^2 = p_2^2 + p_3^2 + 2p_2 \cdot p_3, \qquad \text{so} \qquad 2p_2 \cdot p_3 = (m_\pi^2 - m_l^2)c^2 \tag{9.72}$$

Thus

$$\langle |\mathscr{M}|^2 \rangle = \left(\frac{g_w}{2M_W} \right)^4 f_\pi^2 m_l^2 (m_\pi^2 - m_l^2) \tag{9.73}$$

(a *constant*).

The decay rate is given by the standard formula (Equation 9.35):

$$\Gamma = \frac{|\mathbf{p}_2|}{8\pi \hbar m_\pi^2 c} \langle |\mathscr{M}|^2 \rangle \tag{9.74}$$

and the outgoing momentum is (see Equation 9.34 or Problem 3.19)

$$|\mathbf{p}_2| = \frac{c}{2m_\pi} (m_\pi^2 - m_l^2) \tag{9.75}$$

So

$$\Gamma = \frac{f_\pi^2}{\pi \hbar m_\pi^3} \left(\frac{g_w}{4M_W} \right)^4 m_l^2 (m_\pi^2 - m_l^2)^2 \tag{9.76}$$

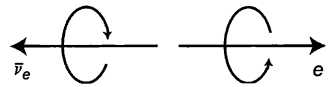
Of course, without knowing the decay constant, f_π, we cannot calculate the pion lifetime.* But we *are* able to determine the *ratio* of the electron and muon decay rates:

$$\frac{\Gamma(\pi^- \to e^- + \bar{\nu}_e)}{\Gamma(\pi^- \to \mu^- + \bar{\nu}_\mu)} = \frac{m_e^2 (m_\pi^2 - m_e^2)^2}{m_\mu^2 (m_\pi^2 - m_\mu^2)^2} = 1.283 \times 10^{-4} \tag{9.77}$$

(The experimental number is $1.230 \pm 0.004 \times 10^{-4}$.) At first glance, this is a very surprising result, for it predicts (correctly) that the pion prefers the *muon* mode, in spite of the fact that the electron is much lighter. Phase space considerations favor decays for which the mass decrease is as large as possible, and unless some conservation law intervenes, we ordinarily find that the lightest final state is the most common one. Pion decay is the notorious exception, and it calls for some special dynamical explanation. A clue is suggested by Equation 9.76: notice that if the electron were *massless*, the $\pi^- \to e^- + \bar{\nu}_e$ mode would be forbidden *altogether*.

* It is a rather striking fact that if you put in $f_\pi = m_\pi c$ (or, better yet, $m_\pi c \cos\theta_C$) you come out very close to the $\pi^- \to \mu^- + \bar{\nu}_\mu$ lifetime, but I know of no persuasive theoretical justification for this ansatz, and it doesn't work for the heavier mesons.

Can we understand this limiting case? Yes: the pion has spin 0, so the electron and the antineutrino must emerge with opposite spins, and hence equal helicities:



The antineutrino is always right-handed, so the electron must be right-handed as well. But if the electron were truly massless, then (like the neutrino) it would only exist as a left-handed particle. More precisely, the $1 - \gamma^5$ in the weak vertex factor would couple only to left-handed electrons, just as it couples only to left-handed neutrinos (see Problem 9.15). That's why, if the electron were massless, the decay $\pi^- \to e^- + \bar{\nu}_e$ could not occur at *all*, and why (the physical electron being very *close* to massless) the decay is so heavily suppressed.

9.5
Charged Weak Interactions of Quarks

In the case of leptons, the coupling to $W^{\pm}$takes place strictly *within a particular generation*:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}, \quad \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}, \quad \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix} \qquad \text{(lepton generations)}$$

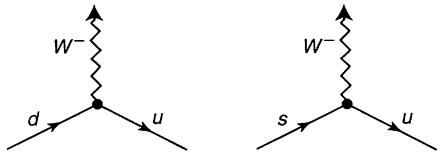
That is, $e^- \to \nu_e + W^-$, $\mu^- \to \nu_\mu + W^-$, $\tau^- \to \nu_\tau + W^-$, but there is no cross-generational coupling, of the form $e^- \to \nu_\mu + W^-$, for example. The coupling of W to *quarks* is not quite so simple, for although the generation structure is similar

$$\begin{pmatrix} u \\ d \end{pmatrix}, \quad \begin{pmatrix} c \\ s \end{pmatrix}, \quad \begin{pmatrix} t \\ b \end{pmatrix} \qquad \text{(quark generations)}$$

the weak interactions do not strictly respect their separate identities. There are, to be sure, interactions of the form $d \to u + W^-$ (the process that underlies neutron decay, $n \to p + e + \bar{\nu}_e$), but there exist as well cross-generational couplings, such as $s \to u + W^-$ (seen, for example, in the decay $\Lambda \to p + e + \bar{\nu}_e$). Indeed, if this were *not* the case, we would have three absolute 'flavor-conservation' laws: conservation of 'upness-plus-downness', 'charm-plus-strangeness', and 'truth-plus-beauty' – analogous to the three lepton number conservation laws. The lightest strange particle (K^-) would be absolutely stable, and so would the B meson (the lightest beautiful particle); our world would be a quite different place.

In 1963 (when u, d, and s were the only quarks known), Cabibbo [1] suggested that the $d \to u + W^-$ vertex carries a factor of $\cos\theta_C$, whereas $s \to u + W^-$

carries a factor of $\sin\theta_C$; apart from that they are identical to the leptonic couplings (Equation 9.5):



$$\frac{-ig_w}{2\sqrt{2}}\gamma^\mu(1-\gamma^5)\cos\theta_C \qquad \frac{-ig_w}{2\sqrt{2}}\gamma^\mu(1-\gamma^5)\sin\theta_C \tag{9.78}$$

The strangeness-changing process ($s \to u + W^-$) is conspicuously *weaker* than the strangeness-conserving one ($d \to u + W^-$), so evidently the 'Cabibbo angle' θ_C is rather small. Experimentally,

$$\theta_C = 13.15^\circ \tag{9.79}$$

The weak interactions *almost* respect quark generations . . . but not *quite*.

Example 9.2 Leptonic Decays Consider the decay $K^- \to l^- + \overline{\nu}_l$, where l is an electron or a muon. This is the analog to π^- decay (Section 9.4), but now the quark vertex is $s + \overline{u} \to W^-$, instead of $d + \overline{u} \to W^-$. From Equation 9.76 we have

$$\Gamma = \frac{f_K^2}{\pi\hbar m_K^3}\left(\frac{g_w}{4M_w}\right)^4 m_l^2(m_K^2 - m_l^2)^2$$

The coupling strength is presumably about the same, except that where f_π contained a factor of $\cos\theta_C$, f_K carries a factor of $\sin\theta_C$. Accordingly,

$$\frac{\Gamma(K^- \to l^- + \overline{\nu}_l)}{\Gamma(\pi^- \to l^- + \overline{\nu}_l)} = \tan^2\theta_C\left(\frac{m_\pi}{m_K}\right)^3\left(\frac{m_K^2 - m_l^2}{m_\pi^2 - m_l^2}\right)^2 \tag{9.80}$$

Putting in the numbers, I get 0.96 for the muon mode ($l = \mu$) and 0.19 for the electron mode ($l = e$). The observed ratios are 1.34 and 0.26, respectively; these decays are pure axial vector, and, as we discovered earlier (Section 9.3), perfect agreement is not to be expected.

Processes of the kind considered in Example 9.2 are called *leptonic* decays. There also exist *semileptonic* decays, such as $\pi^- \to \pi^0 + e^- + \overline{\nu}_e$, $\overline{K}^0 \to \pi^+ + \mu^- + \overline{\nu}_\mu$ (Figure 9.3a), or for that matter the beta decay of the neutron: $n \to p^+ + e^- + \overline{\nu}_e$. Finally, there are *nonleptonic* weak interactions, such as $K^- \to \pi^0 + \pi^-$ or $\Lambda \to p^+ + \pi^-$ (Figure 9.3b). Generally speaking, the latter are the hardest to analyze, because there is strong interaction contamination at *both* ends of the W line [2].

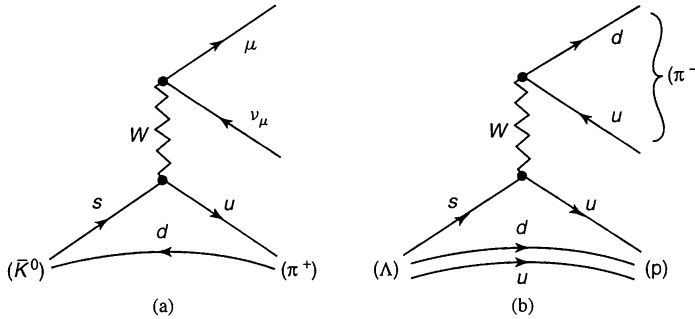


Fig. 9.3 (a) A typical semileptonic decay ($\overline{K}^0 \to \pi^+ + \mu^- + \overline{\nu}_\mu$). (b) A typical nonleptonic weak decay ($\Lambda \to p^+ + \pi^-$).

Example 9.3 Semileptonic Decays In the case of neutron decay ($n \to p + e + \overline{\nu}_e$), the basic quark process is $d \to u + W^-$ (with two spectators). However, there are *two* d quarks in the neutron, and *either one* could couple to the W; the net amplitude for the process is the *sum*. The simplest way to keep track of the numbers is to use the quark wave functions in Section 5.6.1; the flavor states ψ_{12}, for instance, give $n = (ud - du)d/\sqrt{2}$, from which (with $d \to u$) we get $[(uu - uu)d + (ud - du)u]/\sqrt{2} = (ud - du)u/\sqrt{2} = p$. The overall coefficient is then simply $\cos\theta_C$ (as I claimed at the end of Section 9.3). By contrast, in the decay $\Sigma^0 \to \Sigma^+ + e + \overline{\nu}_e$, the quark process is still $d \to u$, but here $\Sigma^0 = [(us - su)d + (ds - sd)u]/2 \to [(us - su)u + (us - su)u]/2 = (us - su)u = \sqrt{2}\Sigma^+$, and hence the amplitude carries a factor of $\sqrt{2}\cos\theta_c$.* The decay rate is given by Equation 9.60, which reduces (in the case $a \gg 1$) to the form

$$\Gamma = \frac{1}{30\pi^3\hbar}\left(\frac{g_w}{2M_W c^2}\right)^4 (\Delta m c^2)^5 X^2$$

where Δm is the baryon mass decrease and X is the Cabibbo factor ($\cos\theta_C$, for neutron decay; $\sqrt{2}\cos\theta_C$, for $\Sigma^0 \to \Sigma^+ + e + \overline{\nu}_e$; etc.). I'll let you work out the numbers for yourself (Problem 9.17).† ▒

Cabibbo's theory was very successful in correlating dozens of decay rates, but there remained a disturbing problem: this picture allowed the K^0 to decay into a $\mu^+\mu^-$ pair (see Figure 9.4). The amplitude should be proportional to $\sin\theta_C \cos\theta_C$, but the calculated rate was far greater than the experimental limit. A solution

* Actually, there is a technical difference here: the active quark is bound to the spectator in a *spin singlet* state. Fortunately, this does not affect the lifetime.

† This procedure includes only the valence quarks, and hence is insensitive to the nonconservation of the axial coupling. As we found in Equation 9.65, PCAC can lead to a correction of nearly 50%, so one does not expect fine precision in the lifetimes. Cabibbo's theory included a way of calculating the axial couplings, but I shall not go into that here.

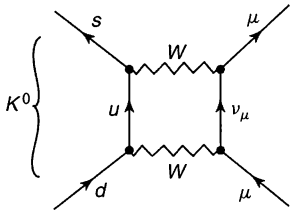


Fig. 9.4 The decay $K^0 \to \mu^+ + \mu^-$.

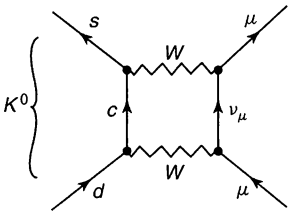
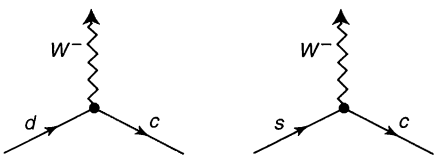


Fig. 9.5 The GIM mechanism. This diagram cancels Figure 9.4. Note the virtual c quark replacing the u.

to this dilemma was proposed in 1970 by Glashow, Iliopoulos, and Maiani (GIM) [3]. They introduced a fourth quark, c (note that this was four years before the 'November Revolution' produced the first direct experimental evidence for charm) whose couplings to s and d carry factors of $\cos\theta_C$ and $-\sin\theta_C$, respectively:



$$\frac{-ig_w}{2\sqrt{2}}\gamma^\mu(1-\gamma^5)(-\sin\theta_C) \qquad \frac{-ig_w}{2\sqrt{2}}\gamma^\mu(1-\gamma^5)\cos\theta_C \tag{9.81}$$

In the 'GIM mechanism', the diagram in Figure 9.4 is canceled by the corresponding diagram with c in place of u (Figure 9.5), for this time the amplitude is proportional to $-\sin\theta_C\cos\theta_C$.*

The Cabibbo–GIM scheme invites a simple and beautiful interpretation: instead of the physical quarks d and s, the 'correct' states to use in the weak interactions are d' and s', given by

$$d' = d\cos\theta_C + s\sin\theta_C, \quad s' = -d\sin\theta_C + s\cos\theta_C \tag{9.82}$$

* The cancellation is not *perfect*, because the mass of the c is not the same as the mass of the u. However, these virtual particles are so far off the mass shell that both propagators are essentially just $i\not{p}/q^2$. (In calculating $\mathscr{M}$ we shall be integrating over the one remaining internal momentum which is not fixed by the conservation laws. This is essentially the momentum 'circulating around the loop'. Because of the two W propagators, the main contribution will come in the region of the W mass, which is so much greater than the c or u mass that the latter can, to good approximation, be neglected. Actually, the decay *does* occur, it's just extremely slow, and if you *include* the effects of u/c mass difference, the calculation is consistent with the observed rate.)

or, in matrix form

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_C & \sin\theta_C \\ -\sin\theta_C & \cos\theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} \tag{9.83}$$

The W's couple to the 'Cabibbo-rotated' states

$$\begin{pmatrix} u \\ d' \end{pmatrix}, \begin{pmatrix} c \\ s' \end{pmatrix}$$

in exactly the same way that they couple to lepton pairs, $\begin{pmatrix} \nu_e \\ e \end{pmatrix}$ and $\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}$; their couplings to the *physical* particles (states of specific flavor) are then given by

$$\begin{pmatrix} u \\ d' \end{pmatrix} = \begin{pmatrix} u \\ d\cos\theta_C + s\sin\theta_C \end{pmatrix}, \qquad \begin{pmatrix} c \\ s' \end{pmatrix} = \begin{pmatrix} c \\ -d\sin\theta_C + s\cos\theta_C \end{pmatrix} \tag{9.84}$$

That is, $d \to u + W^-$ carries a factor $\cos\theta_C$, $s \to u + W^-$ carries a factor $\sin\theta_C$, and so on.*

At the time, the GIM mechanism seemed a little extravagant – introducing a new quark just to fix a rather esoteric technical defect in a largely untested theory. But the skeptics were silenced by the discovery of the $\psi(c\bar{c})$ in 1974. Meanwhile, Kobayashi and Maskawa [4] had generalized the Cabibbo–GIM scheme to handle *three* generations of quarks.† the 'weak interaction generations',

$$\begin{pmatrix} u \\ d' \end{pmatrix}, \begin{pmatrix} c \\ s' \end{pmatrix}, \begin{pmatrix} t \\ b' \end{pmatrix} \tag{9.85}$$

* It is purely conventional that we 'rotate' d and s, rather than u and c; we could accomplish the same purpose by introducing $u' = u\cos\theta_C - c\sin\theta_C$ and $c' = u\sin\theta_C + c\cos\theta_C$. Incidentally, you might be wondering whether a similar rotation occurs in the *lepton* sector. If all neutrinos were massless, any linear combination of them would *still* be massless, and there would be no 'tag' to identify the 'unrotated' states. But, if neutrinos have mass (as we now know they *do*), there is no reason to suppose that the 'mass eigenstates' are the same as the weak interaction states, and the same rotation story plays out – only in reverse, since the 'familiar' neutrinos are the ones paired with the charged leptons in the weak interactions and we need to rotate *back* to get the 'physical' states (see Chapter 11).

† It is interesting to note that Kobayashi and Maskawa proposed a third quark generation before the *second* was complete, and long before there was any experimental evidence for a third. They were motivated by a desire to explain *CP* violation within the Cabibbo–GIM scheme. It turned out that for this purpose they needed a complex number in the 'rotation' matrix (Equation 9.83), but such a term could always be eliminated by suitable redefinition of the quark phases, unless they went to a 3×3 matrix, and hence to three generations (Problem 9.18).

are related to the physical quark states by the CKM matrix:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{9.86}$$

where V_{ud}, for example, specifies the coupling of u to d ($d \to u + W^-$).

There are nine (complex) elements in the CKM matrix, but they are not all independent (see Problem 9.18); V can be reduced to a kind of 'canonical form', in which there remain just three 'generalized Cabibbo angles', ($\theta_{12}, \theta_{23}, \theta_{13}$) and one phase factor ($\delta$) [5] :

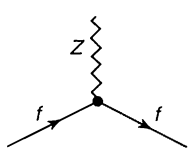
$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{9.87}$$

Here c_{ij} stands for $\cos\theta_{ij}$, and s_{ij} for $\sin\theta_{ij}$. If $\theta_{23} = \theta_{13} = 0$, the third generation does not mix with the other two, and we recover the Cabibbo–GIM picture, with $\theta_{12} = \theta_C$. However, there is compelling evidence (namely, the observed decay of the $B^-(b\bar{u})$ meson) for *some* third-generation mixing, although it must be fairly small in order to account for the success of the original Cabibbo–GIM scheme. The Standard Model offers no insight into the CKM matrix (indeed, this is one of its most conspicuous weaknesses); for the moment, we simply take the values of the matrix elements from experiment. The magnitudes are [6]:

$$|V_{ij}| = \begin{pmatrix} 0.9738 & 0.2272 & 0.0040 \\ 0.2271 & 0.9730 & 0.0422 \\ 0.0081 & 0.0416 & 0.9991 \end{pmatrix} \tag{9.88}$$

9.6 Neutral Weak Interactions

In 1958, Bludman [7] suggested that there might exist *neutral* weak interactions, mediated by an uncharged partner of the W's – the Z^0:



Here f stands for any lepton or any quark. Notice that the *same* fermion comes out as went in (just as in QED and QCD). We do not allow couplings of the form $\mu^- \to e^- + Z^0$, for example (this would violate conservation of muon and electron number), nor of the form $s \to d + Z^0$ (such a strangeness-changing neutral process would lead to $K^0 \to \mu^+ + \mu^-$, which, as I have already remarked, is strongly

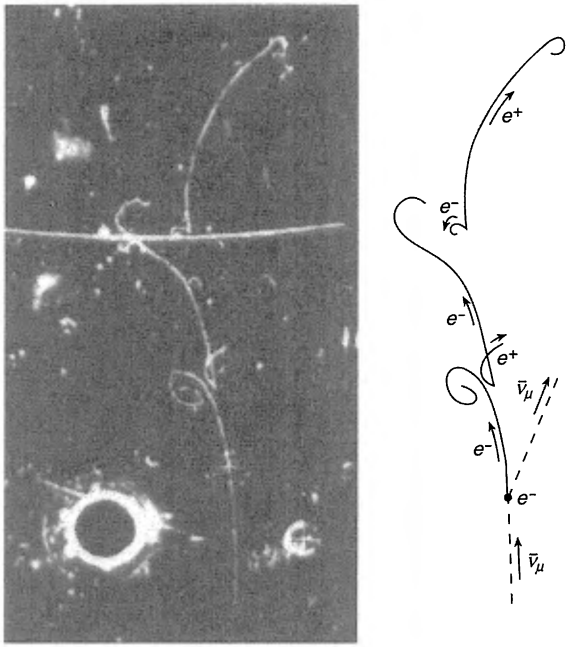
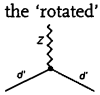


Fig. 9.6 The first picture of a neutral weak process ($\bar{\nu}_\mu + e^- \to \bar{\nu}_\mu + e^-$). The neutrino enters from below (leaving no track), and strikes an electron, which moves off (upward), emitting two photons (which show up in the figure only when they subsequently produce electron–positron pairs) as it slows down and spirals inward in the superimposed magnetic field (the big circle in the lower left is a lamp). (*Source*: CERN.)

suppressed).* In 1961, Glashow [8] published the first paper on unification of weak and electromagnetic interactions; his theory required the existence of neutral weak processes, and specified their structure (see Section 9.7). In 1967, Weinberg and Salam [9] formulated Glashow's model as a 'spontaneously broken gauge theory', and in 1971, 't Hooft [10] demonstrated that the Glashow–Weinberg–Salam(GWS) scheme is renormalizable. Thus there were increasingly persuasive theoretical reasons for thinking that neutral weak interactions occur in nature, but for a long time there were no *experimental* data to support this hope. Finally, in 1973 [11],

* In the case of *neutral* processes, it doesn't matter whether you use the physical states (d, s, b) or the 'rotated' states (d', s', b'). Schematically, the argument runs as follows:



gives $\mathscr{M} \sim \bar{d}'d' = \bar{d}d\cos^2\theta_C + \bar{s}s\sin^2\theta_C + (\bar{d}s + \bar{s}d)\sin\theta_C\cos\theta_C$,

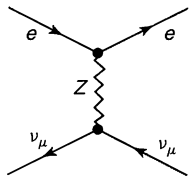
gives $\mathscr{M} \sim \bar{s}'s' = \bar{d}d\sin^2\theta_C + \bar{s}s\cos^2\theta_C + (\bar{d}s + \bar{s}d)\sin\theta_C\cos\theta_C$.

So the *sum* of the two is $\mathscr{M} \sim \bar{d}'d' + \bar{s}'s' = \bar{d}d + \bar{s}s$. Thus the *net* amplitude, once both diagrams are combined, is the same whichever states we use. (The same argument generalizes to three generations, as long as the CKM matrix is unitary.)

a bubble chamber photograph at CERN (Figure 9.6) revealed the first convincing evidence for the reaction

$$\bar{\nu}_\mu + e \to \bar{\nu}_\mu + e$$

suggesting mediation by the Z^0:

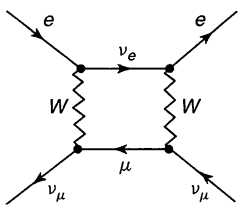


The same series of experiments also witnessed the corresponding neutrino–*quark* process, in the form of inclusive neutrino–nucleon scattering:

$$\bar{\nu}_\mu + N \to \bar{\nu}_\mu + X$$

$$\nu_\mu + N \to \nu_\mu + X$$

Their cross sections were about a third as large as those of the related *charged* events ($\bar{\nu}_\mu + N \to \mu^+ + X$ and $\nu_\mu + N \to \mu^- + X$), indicating that this was indeed a new kind of weak interaction, and not simply the higher-order process,



(which would yield a far smaller cross section). The CERN results came as welcome encouragement to electroweak theorists, who had been out on a limb now for several years [12].

As we have seen, the coupling of quarks and leptons to $W^\pm$ is a universal '*V-A*' form; the vertex factor is always

$$\frac{-ig_w}{2\sqrt{2}}\gamma^\mu(1-\gamma^5) \qquad (W^\pm \text{ vertex factor}) \tag{9.89}$$

(It is true that the axial coupling to *composite* structures, such as the proton, is modified, but that is a result of strong interaction contamination – the underlying quark process is pure *V-A*.) The coupling of the Z^0 is not so simple:

$$\frac{-ig_z}{2}\gamma^\mu(c_V^f - c_A^f\gamma^5) \qquad (Z^0 \text{ vertex factor}) \tag{9.90}$$

Table 9.1 Neutral vector and axial vector couplings in the GWS model

f	c_V	c_A
ν_e, ν_μ, ν_τ	$\frac{1}{2}$	$\frac{1}{2}$
e^-, μ^-, τ^-	$-\frac{1}{2} + 2\sin^2\theta_w$	$-\frac{1}{2}$
u, c, t	$\frac{1}{2} - \frac{4}{3}\sin^2\theta_w$	$\frac{1}{2}$
d, s, b	$-\frac{1}{2} + \frac{2}{3}\sin^2\theta_w$	$-\frac{1}{2}$

where g_Z is the neutral coupling constant, and the coefficients c_V^f and c_A^f depend on the particular quark or lepton (f) involved. In the GWS model, *all* these numbers are determined by a single fundamental parameter θ_w, called the 'weak mixing angle' (or 'Weinberg angle'), as indicated in Table 9.1. The weak and electromagnetic coupling constants are related:

$$g_w = \frac{g_e}{\sin\theta_w}, \qquad g_Z = \frac{g_e}{\sin\theta_w \cos\theta_w} \tag{9.91}$$

where g_e, remember, is essentially the charge of the electron ($g_e = e\sqrt{4\pi/\hbar c}$). Finally, the $W^\pm$and Z^0 masses are related by

$$M_W = M_Z \cos\theta_w \tag{9.92}$$

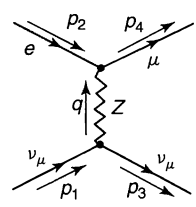
Equations 9.90–9.92 are the basic predictions of the GWS model; you'll see how they were obtained in the next section.

The Standard Model provides no way to calculate θ_w itself; like the CKM matrix, its value is taken from experiment:

$$\theta_w = 28.75^\circ \qquad (\sin^2\theta_w = 0.2314) \tag{9.93}$$

But *given* the value of θ_w, we can calculate the W and Z masses (see Problem 9.20). Their discovery by Rubbia at CERN in 1983, at $M_W = 82$ GeV/ c^2 and $M_Z = 92$ GeV/ c^2 (as predicted) was persuasive evidence for the GWS model [13].

Example 9.4 Elastic Neutrino–Electron Scattering In Example 9.1 we calculated the cross section for the W-mediated process $\nu_\mu + e \to \nu_e + \mu$. We now consider the related Z^0-mediated reaction $\nu_\mu + e \to \nu_\mu + e$.



The Z^0 propagator is (Equation 9.3)

$$\frac{-i(g_{\mu\nu} - q_\mu q_\nu / M_Z^2 c^2)}{q^2 - M_Z^2 c^2} \tag{9.94}$$

At low energies ($q^2 \ll M_Z^2 c^2$) it reduces to

$$\frac{i g_{\mu\nu}}{(M_Z c)^2} \tag{9.95}$$

With this approximation, the amplitude is

$$\mathscr{M} = \frac{g_z^2}{8(M_Z c)^2}[\overline{u}(3)\gamma^\mu(1-\gamma^5)u(1)][\overline{u}(4)\gamma_\mu(c_V - c_A\gamma^5)u(2)] \tag{9.96}$$

and hence (Problem 9.2)

$$\begin{aligned}\langle|\mathscr{M}|^2\rangle &= 2\left(\frac{g_z}{4M_Z c}\right)^4 \mathrm{Tr}\{\gamma^\mu(1-\gamma^5)\not{p}_1\gamma^\nu(1-\gamma^5)\not{p}_3\} \\ &\quad \times \mathrm{Tr}\{\gamma_\mu(c_V - c_A\gamma^5)(\not{p}_2 + mc)\gamma_\nu(c_V - c_A\gamma^5)(\not{p}_4 + mc)\} \\ &= \frac{1}{2}\left(\frac{g_z}{M_Z c}\right)^4 \{(c_V + c_A)^2(p_1 \cdot p_2)(p_3 \cdot p_4) \\ &\quad + (c_V - c_A)^2(p_1 \cdot p_4)(p_2 \cdot p_3) - (mc)^2(c_V^2 - c_A^2)(p_1 \cdot p_3)\}\end{aligned} \tag{9.97}$$

where m is the mass of the electron, and c_V and c_A are the neutral weak couplings for the electron. If we now go to the CM frame, and ignore the electron mass ($m \to 0$), we find

$$\langle|\mathscr{M}|^2\rangle = 2\left(\frac{g_z E}{M_Z c^2}\right)^4 \left[(c_V + c_A)^2 + (c_V - c_A)^2 \cos^4\frac{\theta}{2}\right] \tag{9.98}$$

where E is the electron (or neutrino) energy, and θ is the scattering angle (Figure 9.7). The differential scattering cross section (Equation 9.47) is

$$\frac{d\sigma}{d\Omega} = 2\left(\frac{\hbar c}{\pi}\right)^2 \left(\frac{g_z}{4M_Z c^2}\right)^4 E^2 \left[(c_V + c_A)^2 + (c_V - c_A)^2 \cos^4\frac{\theta}{2}\right] \tag{9.99}$$

and the total cross section (integrating over all angles) is

$$\sigma = \frac{2}{3\pi}(\hbar c)^2 \left(\frac{g_z}{2M_Z c^2}\right)^4 E^2(c_V^2 + c_A^2 + c_V c_A) \tag{9.100}$$

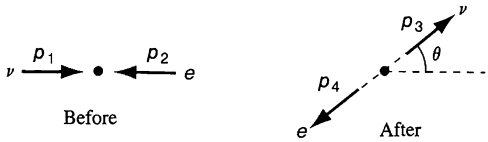


Fig. 9.7 Elastic neutrino–electron scattering in the *CM*.

Putting in the GWS values for c_V and c_A (from Table 9.1), and comparing the result of Example 9.1 (Equation 9.14), we find that for energies substantially above the muon mass

$$\frac{\sigma(\nu_\mu + e^- \to \nu_\mu + e^-)}{\sigma(\nu_\mu + e^- \to \nu_e + \mu^-)} = \frac{1}{4} - \sin^2\theta_w + \frac{4}{3}\sin^4\theta_w = 0.0900 \tag{9.101}$$

The current experimental value [14] is 0.11, which, given the 10% uncertainties in the measurements, is reasonable agreement.

You might well ask why it took so long for neutral weak interactions to be detected in the laboratory; after all, 15 years separate Bludman's original speculations from the definitive experiments at CERN. The reason is that *most* neutral processes are 'masked' by competing electromagnetic ones. For example, $e^+ + e^- \to \mu^+ + \mu^-$ can occur *either* by a virtual Z^0 *or* by a virtual γ (Figure 9.8); at low energies the photon mechanism overwhelmingly dominates.* That's why *neutrino* scattering was originally used to confirm the existence of neutral weak interactions; neutrinos have no electromagnetic coupling, so the weak effects are not obscured. But neutrino experiments are notoriously difficult – hence the long delay. An alternative is to work at high energy – specifically, in the neighborhood of the Z^0 mass, where the denominator of the Z^0 propagator is small, and the 'weak' interaction is correspondingly large. In the early days it was hard to estimate θ_w, and hence the Z^0 mass was quite uncertain. But by the late seventies, a variety of experimental data pointed to $\theta_w \approx 29°$, and hence to $M_Z = 90\ \text{GeV}/c^2$ (see Problem 9.20). This prediction was stunningly confirmed in 1983 [13], and inspired the

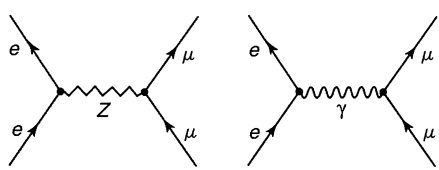


Fig. 9.8 Weak and electromagnetic contributions to $e^+ + e^- \to \mu^+ + \mu^-$.

* In principle, there is weak contamination in *every* electromagnetic process, since the Z^0 couples to everything the γ does (and then some). For example, the Coulomb potential binding the electrons to the nucleus in an atom is slightly modified by Z^0 exchange, and this is observable in atomic spectra. Similarly, there is a weak contribution to electron–proton scattering. Although these effects are minute, they leave a tell-tale fingerprint: parity violation [15].

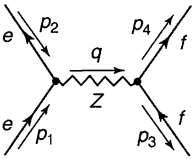


Fig. 9.9 Electron–positron scattering near the Z^0 pole.

construction of electron–positron colliders designed to operate at the Z^0 peak: SLC at SLAC, and LEP at CERN.

Example 9.5 Electron–Positron Scattering Near the Z^0 Pole Consider the process $e^+ + e^- \to f + \bar{f}$ (Figure 9.9), where f is any quark or lepton.* This time we shall *not* use the approximate form of the Z^0 propagator (Equation 9.95), for we are interested precisely in the regime $q^2 \approx (M_Z c)^2$. The amplitude is

$$\mathscr{M} = -\frac{g_z^2}{4\left[q^2 - (M_Z c)^2\right]}[\overline{u}(4)\gamma^\mu(c_V^f - c_A^f\gamma^5)v(3)] \times \left(g_{\mu\nu} - \frac{q_\mu q_\nu}{(M_Z c)^2}\right)[\overline{v}(2)\gamma^\nu(c_V^e - c_A^e\gamma^5)u(1)] \tag{9.102}$$

where $q = p_1 + p_2 = p_3 + p_4$. Since we are working in the vicinity of 90 GeV, we can afford to ignore the lepton and quark masses.† In this case the second term in the propagator contributes nothing, for q_μ contracts with γ^μ to give

$$\overline{u}(4)\not{q}(c_V - c_A\gamma^5)v(3)$$

but $\not{q} = \not{p}_3 + \not{p}_4$ and $\overline{u}(4)\not{p}_4 = 0$ (Equation 9.96 with $m = 0$), and

$$\not{p}_3(c_V - c_A\gamma^5)v(3) = (c_V + c_A\gamma^5)\not{p}_3 v(3) = 0$$

for the same reason. Thus

$$\mathscr{M} = -\frac{g_z^2}{4[q^2 - (M_Z c)^2]}[\overline{u}(4)\gamma^\mu(c_V^f - c_A^f\gamma^5)v(3)][\overline{v}(2)\gamma_\mu(c_V^e - c_A^e\gamma^5)u(1)] \tag{9.103}$$

and it follows that

$$\langle|\mathscr{M}|^2\rangle = \left[\frac{g_z^2}{8(q^2 - (M_Z c)^2)}\right]^2 \mathrm{Tr}\{\gamma^\mu(c_V^f - c_A^f\gamma^5)\not{p}_3\gamma^\nu(c_V^f - c_A^f\gamma^5)\not{p}_4\} \times \mathrm{Tr}\{\gamma_\mu(c_V^e - c_A^e\gamma^5)\not{p}_1\gamma_\nu(c_V^e - c_A^e\gamma^5)\not{p}_2\} \tag{9.104}$$

* *Not* an electron, however, for then we would have to include the rotated diagram.

† I assume $m_f \ll M_Z$, which excludes the top quark. But the t cannot be produced anyway, at these energies.

Now, the first trace is (Problem 9.2)

$$4[(c_V^f)^2 + (c_A^f)^2][p_3^\mu p_4^\nu + p_3^\nu p_4^\mu - g^{\mu\nu}(p_3 \cdot p_4)] - 8ic_V^f c_A^f \epsilon^{\mu\nu\lambda\sigma} p_{3\lambda} p_{4\sigma} \tag{9.105}$$

and there is the corresponding expression for the second trace, so

$$\begin{aligned}\langle|\mathscr{M}|^2\rangle = \frac{1}{2}\left[\frac{g_z^2}{q^2 - (M_Z c)^2}\right]^2 &\{[(c_V^f)^2 + (c_A^f)^2][(c_V^e)^2 + (c_A^e)^2] \\ &\times [(p_1 \cdot p_3)(p_2 \cdot p_4) + (p_1 \cdot p_4)(p_2 \cdot p_3)] \\ &+ 4c_V^f c_A^f c_V^e c_A^e[(p_1 \cdot p_3)(p_2 \cdot p_4) - (p_1 \cdot p_4)(p_2 \cdot p_3)]\}\end{aligned} \tag{9.106}$$

In the CM frame this reduces to

$$\begin{aligned}\langle|\mathscr{M}|^2\rangle &= \left[\frac{g_z^2 E^2}{(2E)^2 - (M_Z c^2)^2}\right]^2 \\ &\times \left\{[(c_V^f)^2 + (c_A^f)^2][(c_V^e)^2 + (c_A^e)^2](1 + \cos^2\theta) - 8c_V^f c_A^f c_V^e c_A^e \cos\theta\right\}\end{aligned} \tag{9.107}$$

where E is the energy of each particle and θ is the angle between $\mathbf{p}_1$ and $\mathbf{p}_3$. The differential scattering cross section (Equation 9.47) is, therefore,

$$\frac{d\sigma}{d\Omega} = \left(\frac{\hbar c g_z^2 E}{16\pi[(2E)^2 - (M_Z c^2)^2]}\right)^2$$

$$\times \left\{[(c_V^f)^2 + (c_A^f)^2][(c_V^e)^2 + (c_A^e)^2](1 + \cos^2\theta) - 8c_V^f c_A^f c_V^e c_A^e \cos\theta\right\} \tag{9.108}$$

and the total cross section is

$$\sigma = \frac{1}{3\pi}\left(\frac{\hbar c g_z^2 E}{4[(2E)^2 - (M_Z c^2)^2]}\right)^2 [(c_V^f)^2 + (c_A^f)^2][(c_V^e)^2 + (c_A^e)^2] \tag{9.109}$$

As it stands, σ blows up at the Z^0 pole – that is, when the total energy ($2E$) hits the value $M_Z c^2$ (just right to put the Z^0 on its mass shell). The problem is that we have treated the Z^0 as a *stable* particle, which it is not. Its lifetime is finite, and this has the effect of 'smearing out' its mass. We can account for this by modifying the propagator [16]

$$\frac{1}{q^2 - (M_Z c)^2} \to \frac{1}{q^2 - (M_Z c)^2 + i\hbar M_Z \Gamma_Z} \tag{9.110}$$

where Γ_Z is the decay rate (experimentally, $\Gamma_Z = 3.791 \pm 0.003 \times 10^{24}$/s). With this adjustment, the cross section becomes

$$\sigma = \frac{(\hbar c g_Z^2 E)^2}{48\pi} \frac{[(c_V^f)^2 + (c_A^f)^2][(c_V^e)^2 + (c_A^e)^2]}{[(2E)^2 - (M_Z c^2)^2]^2 + (\hbar M_Z c^2 \Gamma_Z)^2} \tag{9.111}$$

Because $\hbar\Gamma_Z \ll M_Z c^2$, the correction for finite Z^0 lifetime is negligible except in the immediate vicinity of the Z^0 pole, where it has the effect of softening the infinite spike.

In Chapter 8 we calculated the cross section for the same process when mediated by a *photon* (Equation 8.6):

$$\sigma = \frac{(\hbar c g_e^2)^2}{48\pi} \frac{(Q^f)^2}{E^2} \tag{9.112}$$

(where Q^f is the charge of f, in units of e). Thus the ratio of weak to electromagnetic rates in (for example) muon production, is

$$\frac{\sigma(e^+e^- \to Z^0 \to \mu^+\mu^-)}{\sigma(e^+e^- \to \gamma \to \mu^+\mu^-)} = \left\{ \frac{\left[\frac{1}{2} - 2\sin^2\theta_w + 4\sin^4\theta_w\right]^2}{(\sin\theta_w \cos\theta_w)^4} \right\} \times \frac{E^4}{[(2E)^2 - (M_Z c^2)^2]^2 + (\hbar\Gamma_Z M_Z c^2)^2} \tag{9.113}$$

The factor in curly brackets is approximately 2. Substantially *below* the Z^0 pole ($2E \ll M_Z c^2$), then,

$$\frac{\sigma_Z}{\sigma_\gamma} \cong 2\left(\frac{E}{M_Z c^2}\right)^4 \tag{9.114}$$

and the electromagnetic route dominates (at $2E = \frac{1}{2}M_Z c^2$, for instance, the weak contribution is less than 1%). But right *on* the Z^0 pole ($2E = M_Z c^2$),

$$\frac{\sigma_Z}{\sigma_\gamma} \cong \frac{1}{8}\left(\frac{M_Z c^2}{\hbar\Gamma_Z}\right) \approx 200 \tag{9.115}$$

At the Z^0 pole, therefore, the *weak* mechanism is favored, by a factor of around 200 (Figure 9.10).*

* Equally interesting is the electromagnetic-weak 'interference' that occurs when the two amplitudes are combined: $|\mathscr{M}_\gamma + \mathscr{M}_Z|^2 = |\mathscr{M}_\gamma|^2 + |\mathscr{M}_Z|^2 + 2\,\mathrm{Re}(\mathscr{M}_\gamma \mathscr{M}_Z)$. We have calculated $|\mathscr{M}_Z|^2$ and (in Chapter 8) $|\mathscr{M}_\gamma|^2$, but the cross term provides a sensitive test of the GWS theory, even at energies substantially below the Z^0 pole. (See Halzen and Martin, ref. [11], Section 13.6, and ref. [15].) Indeed, it was the success of the electroweak interference experiments in 1978 that convinced most theorists that the GWS model is correct. For a contemporary account, see *Physics Today*, September 1978, p. 17.

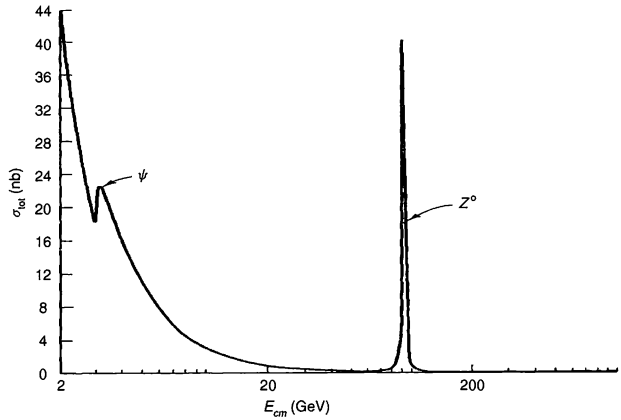


Fig. 9.10 Electron-positron scattering in the neighborhood of the Z^0 pole.

9.7 Electroweak Unification

9.7.1 Chiral Fermion States

All the cards are now on the table;* it remains only to explain where the GWS parameters in Table 9.1 and Equations 9.90–9.92 *come* from. Glashow's original aim was to *unify* the weak and electromagnetic interactions – to combine them into a single theoretical system, in which they would appear not as unrelated phenomena, but rather as different manifestations of one fundamental 'electroweak' interaction. This was a bold proposition, in 1961 [17]. In the first place, there was the enormous disparity in strength between weak and electromagnetic forces. However, as Glashow and others recognized, this could be accounted for if the weak interactions were mediated by extremely massive particles. Of course, this immediately begs the second question: if it's really all one basic interaction, how come the electromagnetic mediator (γ) is massless, when the weak mediators ($W^{\pm}$ and Z^0) are so heavy? Glashow had no particularly good answer ('It is a stumbling block we must overlook', he said coyly). The solution was provided by Weinberg and Salam, in 1967 (see refs. [8] and [9]) in the form of the 'Higgs mechanism' (Chapter 10). Finally, there is a structural difference between the electromagnetic and weak vertex factors, which at first glance would seem to preclude any possibility of unification: the former are purely vectorial (γ^{μ}), whereas the latter contain vector and axial vector parts. In particular, the $W^{\pm}$coupling is 'maximally' mixed $V-A$ in character: $\gamma^{\mu}(1-\gamma^5)$.

* I have not discussed the couplings of W's and Z^0's to one another (or of W's to the photon). The rules are similar to those for gluon–gluon coupling in QCD, and are listed in Appendix D.

This last difficulty is overcome by the ingenious device of absorbing the matrix $(1-\gamma^5)$ into the particle spinor itself. Specifically, we define

$$u_L(p) \equiv \frac{(1-\gamma^5)}{2} u(p) \tag{9.116}$$

The subscript (L) stands for 'left-handed', and is supposed to make you think 'helicity -1'. However, this is seriously misleading, since u_L is not, in general, a helicity eigenstate. In fact, for solutions to the Dirac equation,

$$\gamma^5 u(p) = \begin{pmatrix} \dfrac{c(\mathbf{p}\cdot\boldsymbol{\sigma})}{E+mc^2} & 0 \\ 0 & \dfrac{c(\mathbf{p}\cdot\boldsymbol{\sigma})}{E-mc^2} \end{pmatrix} u(p) \tag{9.117}$$

(Problem 9.26). *If the particle in question is massless*, then $E = |\mathbf{p}|c$, and

$$\gamma^5 u(p) = (\hat{\mathbf{p}}\cdot\boldsymbol{\Sigma}) u(p) \tag{9.118}$$

where

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} \tag{9.119}$$

as before. Remember that $(\hbar/2)\boldsymbol{\Sigma}$ is the spin matrix for a Dirac particle, and hence $(\hat{\mathbf{p}}\cdot\boldsymbol{\Sigma})$ is the helicity, with eigenvalues ± 1. Accordingly

$$\frac{1}{2}(1-\gamma^5)u(p) = \begin{Bmatrix} 0, & \text{if } u(p) \text{ carries helicity } +1 \\ u(p), & \text{if } u(p) \text{ carries helicity } -1 \end{Bmatrix} \quad (\text{for } m=0) \tag{9.120}$$

(If $u(p)$ is *not* a helicity eigenstate, $\frac{1}{2}(1-\gamma^5)$ functions as a 'projection operator', picking out the helicity -1 component.) On the other hand, if the particle is *not* massless, it is only in the ultrarelativistic regime ($E \gg mc^2$) that Equation 9.118 holds (approximately), and hence only in this limit that u_L (as defined by Equation 9.116 carries helicity -1. Nevertheless, everybody *calls* u_L a 'left-handed' state, and I shall stick to the customary language.*

Meanwhile, for *antiparticles* we define†

$$v_L(p) \equiv \frac{(1+\gamma^5)}{2} v(p) \tag{9.121}$$

* Please understand: Equation 9.116 is the *definition* of u_L – nobody's arguing about that. I'm only warning you that the *name* is misleading: 'left-handed' does *not* mean 'helicity -1', except in contexts where the particle's mass is negligible.

† If the sign of γ^5 seems strange, refer to the footnote following Equation 7.30.

Table 9.2 Chiral spinors

Particles	Antiparticles
$u_L = \frac{1}{2}(1-\gamma^5)u$	$v_L = \frac{1}{2}(1+\gamma^5)v$
$u_R = \frac{1}{2}(1+\gamma^5)u$	$v_R = \frac{1}{2}(1-\gamma^5)v$
$\bar{u}_L = \bar{u}\frac{1}{2}(1+\gamma^5)$	$\bar{v}_L = \bar{v}\frac{1}{2}(1-\gamma^5)$
$\bar{u}_R = \bar{u}\frac{1}{2}(1-\gamma^5)$	$\bar{v}_R = \bar{v}\frac{1}{2}(1+\gamma^5)$

R and L correspond to helicity $+1$ and -1 if $m = 0$, and *approximately* so if $E \gg mc^2$.

The corresponding 'right-handed' spinors are

$$u_R(p) = \frac{(1+\gamma^5)}{2}u(p), \quad v_R(p) \equiv \frac{(1-\gamma^5)}{2}v(p) \tag{9.122}$$

As for the adjoint spinors, since γ^5 is Hermitian ($\gamma^{5\dagger} = \gamma^5$), and it anticommutes with γ^μ ($\gamma^\mu\gamma^5 = -\gamma^5\gamma^\mu$),

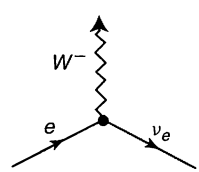
$$\bar{u}_L = u_L^\dagger\gamma^0 = u^\dagger\frac{(1-\gamma^5)}{2}\gamma^0 = u^\dagger\gamma^0\frac{(1+\gamma^5)}{2} = \bar{u}\frac{(1+\gamma^5)}{2} \tag{9.123}$$

Similarly

$$\bar{v}_L = \bar{v}\frac{(1-\gamma^5)}{2}, \quad \bar{u}_R = \bar{u}\frac{(1-\gamma^5)}{2}, \quad \bar{v}_R = \bar{v}\frac{(1+\gamma^5)}{2} \tag{9.124}$$

We call these various spinors (summarized in Table 9.2) 'chiral' fermion states (from the Greek word for 'hand' – same root as 'chiropractor').

I emphasize that this is nothing but notation; it is *useful* because it allows us to recast the weak and electromagnetic interactions in a form that facilitates their unification. Consider, to begin with, the coupling of an electron and a neutrino to the W^- (as it occurs, say, in inverse beta decay, Example 9.1):



The contribution to $\mathscr{M}$ from this vertex is given by

$$j_\mu^- = \bar{\nu}\gamma_\mu\left(\frac{1-\gamma^5}{2}\right)e \tag{9.125}$$

(here e and ν stand for the particle spinors; for a while we need to keep careful track of the different particle species, and u_e, u_{ν_e}, etc. just gets too cumbersome). This quantity is called the weak 'current'; as we shall see, it plays a role analogous to the electric current in QED. Now

$$\left(\frac{1-\gamma^5}{2}\right)^2 = \frac{1}{4}[1-2\gamma^5+(\gamma^5)^2] = \left(\frac{1-\gamma^5}{2}\right) \tag{9.126}$$

and

$$\gamma_\mu\left(\frac{1-\gamma^5}{2}\right) = \left(\frac{1+\gamma^5}{2}\right)\gamma_\mu \tag{9.127}$$

so

$$\gamma_\mu\left(\frac{1-\gamma^5}{2}\right) = \left(\frac{1+\gamma^5}{2}\right)\gamma_\mu\left(\frac{1-\gamma^5}{2}\right) \tag{9.128}$$

This may not look like much of an improvement, but it enables us to write Equation 9.125 more neatly, in terms of the chiral spinors:

$$j_\mu^- = \bar{\nu}_L\gamma_\mu e_L \tag{9.129}$$

The weak vertex factor is now *purely vectorial* – but it couples *left-handed* electrons to *left-handed* neutrinos. In this sense it is still structurally different from the fundamental vertex in QED; however, we can play a similar game there, too. Notice that

$$u = \left(\frac{1-\gamma^5}{2}\right)u + \left(\frac{1+\gamma^5}{2}\right)u = u_L + u_R \tag{9.130}$$

(similarly, $\bar{u} = \bar{u}_L + \bar{u}_R$), so the electromagnetic analog can itself be written in terms of chiral spinors:

$$j_\mu^{em} = -\bar{e}\gamma_\mu e = -(\bar{e}_L + \bar{e}_R)\gamma_\mu(e_L + e_R) = -\bar{e}_L\gamma_\mu e_L - \bar{e}_R\gamma_\mu e_R \tag{9.131}$$

(For future purposes, it is best to build in a factor of -1, to account for the negative charge of the electron). Observe that the 'cross terms' vanish:

$$\bar{e}_L\gamma_\mu e_R = \bar{e}\left(\frac{1+\gamma^5}{2}\right)\gamma_\mu\left(\frac{1+\gamma^5}{2}\right)e = \bar{e}\gamma_\mu\left(\frac{1-\gamma^5}{2}\right)\left(\frac{1+\gamma^5}{2}\right)e \tag{9.132}$$

but

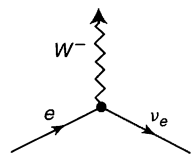
$$(1-\gamma^5)(1+\gamma^5) = 1-(\gamma^5)^2 = 0 \tag{9.133}$$

Equations 9.129 and 9.131 are beginning to look like the stuff of which one might build a unified theory. It is true that the weak current only couples left-handed states, whereas the electromagnetic current couples both types, but apart from that they are strikingly similar. So attractive is this formulation that physicists have come to regard left- and right-handed fermions almost as different particles.* In this view, the factor $(1-\gamma^5)/2$ in the charged weak vertex characterizes the participating particles, rather than the interaction itself; the latter is vectorial in *all* cases – strong, electromagnetic, and weak alike.

9.7.2
Weak Isospin and Hypercharge

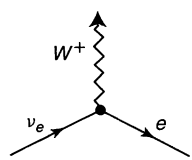
In addition to the (negatively charged) weak current

$$j_\mu^- = \bar{\nu}_L \gamma_\mu e_L$$



describing the process $e^- \to \nu_e + W^-$, there is also, of course, a positively charged current

$$j_\mu^+ = \bar{e}_L \gamma_\mu \nu_L$$



* There is a danger in carrying this *too* far. You may find yourself wondering, for example, whether the left-handed electron necessarily has the same *mass* as the right-handed electron; or, noting that no vector interaction can couple a left-handed particle to a right-handed one (see Equations 9.132 and 9.133), you may ask how the two 'worlds' communicate at all. Both questions are based on a misunderstanding of u_L and u_R. The problem is that, useful as it is in describing particle *interactions*, handedness is not *conserved* in the propagation of a free particle (unless its mass is zero). Formally, γ^5 does not commute with the free-particle Hamiltonian. In fact, u_L and u_R do not satisfy the Dirac equation (see Problem 9.26). A particle that starts out left-handed will soon evolve a right-handed component. (By contrast, *helicity is* conserved in free-particle propagation.) Only for *massless* fermions can left- and right-handed species be considered distinct particles in the full sense of the word.

representing the process $\nu_e \to e^- + W^+$. We can express them both in a more compact notation by introducing the left-handed doublet

$$\chi_L = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \tag{9.134}$$

and the 2×2 matrices

$$\tau^+ \equiv \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \tau^- \equiv \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{9.135}$$

so that

$$j_\mu^\pm = \overline{\chi}_L \gamma_\mu \tau^\pm \chi_L \tag{9.136}$$

The matrices $\tau^\pm$are linear combinations of the first two Pauli spin matrices (Equation 4.26):

$$\tau^\pm = \tfrac{1}{2}(\tau^1 \pm i\tau^2) \tag{9.137}$$

(I use the letter τ here, instead of σ, to avoid any possible confusion with ordinary spin.) This is all very reminiscent of *isospin*: in Section 4.5, we put the proton and neutron into a doublet similar to Equation 9.134. Indeed, we could contemplate a full 'weak isospin'' symmetry, if only there were a *third* weak current, corresponding to $\frac{1}{2}\tau^3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$:

$$j_\mu^3 = \overline{\chi}_L \gamma_\mu \tfrac{1}{2}\tau^3 \chi_L = \tfrac{1}{2}\overline{\nu}_L \gamma_\mu \nu_L - \tfrac{1}{2}\overline{e}_L \gamma_\mu e_L \tag{9.138}$$

'Perfect!' (I hear you exclaim), 'There's the *neutral* weak current!' Not so fast. *This* current only couples *left*-handed particles; in the older language, it is pure *V-A*, whereas the neutral weak interaction involves right-handed components as well. But hang on – we're *almost* there.

Building on the parallel with isospin, we are led to consider a weak analog of *hypercharge* (Y),* which is related to electric charge (Q, in units of e) and the third component of isospin (I^3), by the Gell-Mann–Nishijima formula (Equation 4.37):

$$Q = I^3 + \tfrac{1}{2}Y \tag{9.139}$$

We introduce, then, the 'weak hypercharge' current

$$j_\mu^Y = 2j_\mu^{em} - 2j_\mu^3 = -2\overline{e}_R \gamma_\mu e_R - \overline{e}_L \gamma_\mu e_L - \overline{\nu}_L \gamma_\mu \nu_L \tag{9.140}$$

* You have probably forgotten this word, but *hypercharge* is essentially the same as strangeness, only shifted, in the case of baryons, so that the center row of Eightfold Way diagrams will always carry $Y = 0$. Specifically, $Y = S + A$, where A is the baryon number.

This is an *invariant* construct, as far as weak isospin is concerned, for the latter does not touch *right*-handed components at all, and the combination

$$\bar{e}_L\gamma_\mu e_L + \bar{\nu}_L\gamma_\mu \nu_L = \bar{\chi}_L\gamma_\mu \chi_L$$

is itself invariant.* The underlying symmetry group is called $SU(2)_L \otimes U(1)$; $SU(2)_L$ refers to the weak isospin (with a subscript to remind us that it involves left-handed states only), and $U(1)$ refers to weak hypercharge (involving both chiralities).

I have developed all this in terms of the electron and its neutrino, but it is a simple matter to extend it to the other leptons and quarks. From the left-handed doublets (Cabibbo-rotated, in the case of the quarks)

$$\chi_L \to \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, \begin{pmatrix} u \\ d' \end{pmatrix}_L, \begin{pmatrix} c \\ s' \end{pmatrix}_L, \begin{pmatrix} t \\ b' \end{pmatrix}_L \tag{9.141}$$

we construct three weak isospin currents

$$\mathbf{j}_\mu = \tfrac{1}{2}\bar{\chi}_L\gamma_\mu \boldsymbol{\tau}\chi_L \tag{9.142}$$

and a weak hypercharge current

$$j^Y_\mu = 2j^{em}_\mu - 2j^3_\mu \tag{9.143}$$

where j^{em}_μ is the electric current:

$$j^{em}_\mu = \sum_{i=1}^{2} Q_i(\bar{u}_{iL}\gamma_\mu u_{iL} + \bar{u}_{iR}\gamma_\mu u_{iR}) \tag{9.144}$$

(summed over the particles in the doublet, with Q_i the electric charge).†

* If you care to think of it this way, what we have done is to combine two weak isospin doublets to make an isotriplet, $\bar{\nu}_L e_L, (\bar{\nu}_L\nu_L - \bar{e}_L e_L), \bar{e}_L\nu_L$ (analogous to Equation 5.38), and an isosinglet $(\bar{\nu}_L\nu_L + e_L e_L)$ (analogous to Equation 5.39). The first three go to make the weak isospin currents $j^\pm$ and j^3; the last, together with a right-handed piece, makes the weak hypercharge current, j^Y.

† You might ask what the *difference* is between *weak* isospin (and hypercharge) and their ordinary ('strong') counterparts. The question is particularly pertinent when you come to the light quarks: the weak isospin doublet is $\begin{pmatrix} u \\ d' \end{pmatrix}_L$, whereas the strong isospin doublet is $\begin{pmatrix} u \\ d \end{pmatrix}$. Pretty similar ... is there anything to this? Nope. After all, (i) weak isospin applies to leptons as well as quarks (and to all three quark generations); (ii) weak isospin involves only the *left-handed* chiralities, (all right-handed states are *singlets* – that is, invariant – as far as weak isospin is concerned); (iii) weak isodoublets are Cabibbo-rotated. To put it plainly, strong isospin and weak isospin have nothing to do with one another, save for a common mathematical structure (which, for that matter, they share with many other systems, such as ordinary spin $\frac{1}{2}$) and the (perhaps unfortunate) similarity in their names.

9.7.3
Electroweak Mixing

Now, the GWS model asserts that the three weak isospin currents couple, with strength g_w, to a weak isotriplet of vector bosons, $\mathbf{W}$, whereas the weak hypercharge current couples with strength $g'/2$ to an isosinglet, B:

$$-i\left[g_w \mathbf{j}_\mu \cdot \mathbf{W}^\mu + \frac{g'}{2} j_\mu^Y B^\mu\right] \tag{9.145}$$

(These four particles correspond, ultimately, to the weak and electromagnetic mediators: $W^\pm$, Z^0, and γ – but with a twist, as we shall soon see.) I use bold face here to denote a three-vector in weak isospin space; the dot product can be written out explicitly:

$$\mathbf{j}_\mu \cdot \mathbf{W}^\mu = j_\mu^1 W^{\mu^1} + j_\mu^2 W^{\mu^2} + j_\mu^3 W^{\mu^3} \tag{9.146}$$

or, in terms of the charged currents, $j_\mu^\pm = j_\mu^1 \pm i j_\mu^2$:

$$\mathbf{j}_\mu \cdot \mathbf{W}^\mu = (1/\sqrt{2}) j_\mu^+ W^{\mu+} + (1/\sqrt{2}) j_\mu^- W^{\mu^-} + j_\mu^3 W^{\mu^3} \tag{9.147}$$

where

$$W_\mu^\pm \equiv (1/\sqrt{2})(W_\mu^1 \mp i W_\mu^2) \tag{9.148}$$

are the wave functions representing the $W^\pm$ particles.

The couplings to $W^\pm$ can be read off immediately, from the coefficients of $W_\mu^\pm$ in Equation 9.147 For example, in the process $e^- \to \nu_e + W^-$ we have $j_\mu^- = \overline{\nu}_L \gamma_\mu e_L = \overline{\nu}\gamma_\mu[(1-\gamma^5)/2]e$, giving a term

$$-ig_w(1/\sqrt{2}) j_\mu^- W^{\mu^-} = -\frac{ig_w}{2\sqrt{2}}[\overline{\nu}\gamma_\mu(1-\gamma^5)e] W^{\mu^-} \tag{9.149}$$

The vertex factor is

$$\frac{-ig_w}{2\sqrt{2}}\gamma_\mu(1-\gamma^5) \tag{9.150}$$

which is exactly what we started with (Equation 9.5).

But the two *neutral* states (W^3 and B) 'mix', in Glashow's theory, producing one massless linear combination (the photon), and an orthogonal massive combination (the Z^0):

$$\begin{aligned} A_\mu &= B_\mu \cos\theta_w + W_\mu^3 \sin\theta_w \\ Z_\mu &= -B_\mu \sin\theta_w + W_\mu^3 \cos\theta_w \end{aligned} \tag{9.151}$$

(You see why θ_w is called the 'weak mixing angle'.) In terms of the *physical* fields (A^μ and Z^μ), the neutral portion of the electroweak interaction (Equation 9.145) reads as follows:

$$-i\left[g_w j^3_\mu W^{\mu 3} + \frac{g'}{2} j^Y_\mu B^\mu\right] = -i\left\{\left[g_w \sin\theta_w j^3_\mu + \frac{g'}{2}\cos\theta_w j^Y_\mu\right] A^\mu + \left[g_w \cos\theta_w j^3_\mu - \frac{g'}{2}\sin\theta_w j^Y_\mu\right] Z^\mu\right\} \tag{9.152}$$

Of course, we *know* the electromagnetic coupling; in the present language it is

$$-ig_e j^{em}_\mu A^\mu \tag{9.153}$$

Meanwhile, from Equation 9.143, $j^{em}_\mu = j^3_\mu + \frac{1}{2} j^Y_\mu$. Consistency of the unified electroweak theory with ordinary QED requires

$$g_w \sin\theta_w = g' \cos\theta_w = g_e \tag{9.154}$$

Evidently the weak and electromagnetic coupling constants are *not independent.*

There remain the weak couplings to the Z^0. Using Equations 9.143, 9.152, and 9.154, we obtain

$$-ig_z(j^3_\mu - \sin^2\theta_w j^{em}_\mu)Z^\mu \tag{9.155}$$

where

$$g_z = \frac{g_e}{\sin\theta_w \cos\theta_w} \tag{9.156}$$

From Equation 9.155 we can pick out the neutral weak couplings. For example, the process $\nu_e \to \nu_e + Z^0$ comes exclusively from the j^3_μ term; referring back to Equation 9.138, we have

$$-i\frac{g_z}{2}(\overline{\nu}_L \gamma_\mu \nu_L)Z^\mu = -\frac{ig_z}{2}\left[\overline{\nu}\gamma_\mu\left(\frac{1-\gamma^5}{2}\right)\nu\right]Z^\mu \tag{9.157}$$

and hence the vector and axial vector couplings (Equation 9.90) are $c^\nu_V = c^\nu_A = \frac{1}{2}$. I'll leave it for you to work out the other entries in Table 9.1 (Problem 9.28).*

All this raises some obvious questions: by what mechanism is the underlying $SU(2)_L \otimes U(1)$ symmetry of the electroweak interactions 'broken'? Why do the B and W^3 states 'mix' to form the Z^0 and the photon? If weak and electromagnetic interactions are, deep down, both manifestations of a single electroweak force, how come the weak mediators ($W^\pm$and Z^0) are so heavy, while the electromagnetic mediator (γ) is massless? I'll address these questions in the next chapter.

* Since the weak mixing angle is undetermined in the GWS model, there remain, in effect, *two* independent coupling constants (g_e and g_w, say, or g_e and g_z); in this sense, it is not a completely *unified* theory, but rather an *integrated* theory of weak and electromagnetic interactions.

References

1 Cabibbo, N. (1963) *Physical Review Letters*, **10**, 531.

2 For more detailed calculations in weak interaction theory consult the classic treatise by Marshak, R. E., Riazuddin and Ryan, C. P. (1969) *Theory of Weak Interactions in Particle Physics*, John Wiley & Sons, New York; or the briefer account by (a) Commins, E. D. (1973) *Weak Interactions*, McGraw-Hill, New York. For a comprehensive review of weak interactions in the quark model, see (b) Donoghue, J. F., Golowich, E. and Holstein, B. (1986) *Physics Reports*, **131**, 319; or (c) Commins, E. D. and Bucksbaum, P. H. (1983) *Weak Interactions of Leptons and Quarks*, Cambridge University Press, Cambridge.

3 Glashow, S. L., Iliopoulos, J. and Maiani, L. (1970) *Physical Review*, **D2**, 1585. This and the other fundamental papers on weak interaction theory are reprinted in (a) Lai, C. H. (ed) (1981) *Gauge Theory of Weak and Electromagnetic Interactions*, World Scientific, Singapore.

4 Kobayashi, M. and Maskawa, K. (1973) *Progress in Theoretical Physics*, **49**, 652.

5 Different authors use different parameterizations; I follow the *Review of Particle Physics* convention.

6 *Particle Physics Booklet* (2006). For a useful discussion of the CKM matrix, see Cheng, T. -P. and Li, L. -F. (1984) *Gauge Theory of Elementary Particle Physics*, Oxford, New York. Sect. 12.2.

7 Bludman, S. A. (1958) *Nuovo Cimento*, **9**, 443.

8 Glashow, S. L. (1961) *Nuclear Physics*, **22**, 579. Reprinted in (a) Lai, C. H. (ed) (1981) *Gauge Theory of Weak and Electromagnetic Interactions*, World Scientific, Singapore.

9 Weinberg, S. (1967) *Physical Review Letters*, **19**, 1264; (a) Salam, A. (1968) *Elementary Particle Theory*, (eds N. Svartholm), Almquist and Wiksell, Stockholm, reprinted in Lai, C. H. (ed) (1981) *Gauge Theory of Weak and Electromagnetic Interactions*, World Scientific, Singapore. See also (b) Weinberg's Nobel Prize lecture, reprinted in (1980) *Science*, **210**, 1212.

10 't Hooft, G. (1971) *Nuclear Physics*, **B33**, 173; and (1971) **B35**, 167. Reprinted in (a) Lai, C. H. (ed) (1981) *Gauge Theory of Weak and Electromagnetic Interactions*, World Scientific, Singapore.

11 Hasert, F. J. et al. (1973) *Physics Letters*, **45B**, 138; (1974) *Nuclear Physics*, **B73**, 1.

12 Meanwhile a series of deep inelastic neutrino–proton scattering experiments (also at CERN) not only supported the basic structure of charged and neutral weak interactions but also helped to confirm the quark model itself. See, for example, Perkins, D. H. (2000) *Introduction to High Energy Physics*, 4th edn, Cambridge University Press, Cambridge, UK. Sect. 8.7; (a) Halzen, F. and Martin, A. D. (1984) *Quarks and Leptons*, John Wiley & Sons, New York. Sects. 12.7 and 12.10; (b) Close, F. E. (1979) *An Introduction to Quarks and Partons*, Academic, London. Sect. 11.3.

13 Arnison, G. et al. (1983) *Physics Letters*, **122B**, 103; and (1983) **126B**, 398. For a review of these discoveries, see (a) Radermacher, E. (1985) *Progress in Particle and Nuclear Physics*, **14**, 231.

14 Data on $\nu_\mu + e^- \rightarrow \nu_e + \mu^-$ are from Vilain, P. et al. (1995) *Physics Letters B*, **364**, 121; data on $\nu_\mu + e^- \rightarrow \nu_\mu + e^-$ are from (a) Ahrens, L. A. et al. (1990) *Physical Review D*, **41**, 3297. Earlier data by (b) Alibran, P. et al. (1978) *Physics Letters*, **74B**, 422; which were inconsistent with the GWS model, turned out to be wrong, although they caused some consternation at the time; they were corrected by (c) Armenise, N. et al. (1979) *Physics Letters*, **86B**, 225.

15 Fortson, E. N. and Wilets, L. (1980) *Advances in Atomic and Molecular Physics*, **16**, 319; (a) Prescott, C. Y. et al. (1978) *Physics Letters*, **77B**, 347; and (1979) **84B**, 524; (b) Wu, S. L. (1984) *Physical Reports*, **107**, 229; (c) Wood, C. S. et al. (1997) *Science*, **275**, 1759.
16 See, for example, Frauenfelder, H. and Henley, E. M. (1991) *Subatomic Physics*, 2nd edn, Prentice-Hall, Englewood Cliffs, N.J. Sect. 5.7.
17 Even bolder in 1957, when Schwinger laid the essential groundwork for such a theory: Schwinger, J. (1957) *Annals of Physics (New York)*, **2**, 407. Reprinted in (a) Lai, C. H. (ed) (1981) *Gauge Theory of Weak and Electromagnetic Interactions*, World Scientific, Singapore.

Problems

9.1 Derive the completeness relation for a massive particle of spin 1 (see Problem 9.27 for the massless analog). [*Hint*: Let the z axis point along $\mathbf{p}$. First construct three mutually orthogonal polarization vectors $(\epsilon_\mu^{(1)}, \epsilon_\mu^{(2)}, \epsilon_\mu^{(3)})$ that satisfy $p^\mu \epsilon_\mu = 0$ and $\epsilon_\mu \epsilon^\mu = -1$.]

$$\left[\textit{Answer}: \sum_{s=1,2,3} \epsilon_\mu^{(s)} \epsilon_\nu^{(s)*} = -g_{\mu\nu} + \frac{p_\mu p_\nu}{(Mc)^2} \right] \tag{9.158}$$

9.2 Calculate the trace

$$\mathrm{Tr}\left[\gamma^\mu (c_V - c_A \gamma^5)(\not{p}_1 + m_1 c)\gamma^\nu (c_V - c_A \gamma^5)(\not{p}_2 + m_2 c)\right]$$

for arbitrary (real) numbers c_V and c_A.

$$\Big[\textit{Answer}: 4(c_V^2 + c_A^2)[p_1^\mu p_2^\nu + p_1^\nu p_2^\mu - p_1 \cdot p_2\, g^{\mu\nu}]$$

$$+4(c_V^2 - c_A^2) m_1 m_2 c^2 g^{\mu\nu} - 8ic_V c_A\, \epsilon^{\mu\nu\lambda\sigma} p_{1\lambda} p_{2\sigma}\Big] \tag{9.159}$$

9.3 (a) Calculate $\langle\, |\mathscr{M}|^2 \,\rangle$ for $\nu_\mu + e^- \to \mu^- + \nu_e$ using the more general coupling $\gamma^\mu (1 + \epsilon\, \gamma^5)$. Check that your answer reduces to Equation 9.11 in the case $\epsilon = -1$.

$$\Big[\textit{Answer}: \frac{1}{4}\left(\frac{g_w}{M_W c}\right)^4 [(1-\epsilon^2)^2 (p_1 \cdot p_4)(p_2 \cdot p_3)$$

$$+(1 + 6\epsilon^2 + \epsilon^4)(p_1 \cdot p_2)(p_3 \cdot p_4)]\Big]$$

(b) Let $m_e = m_\mu = 0$, and calculate the CM differential scattering cross section. Also, find the *total* cross section.

(c) If you had accurate experimental data on this reaction, how could you determine ϵ?

9.4 Show that Equation 9.30 is equivalent to Equation 9.29.

9.5 By making the appropriate changes in Equation 9.35, determine the lifetime of the τ lepton, pretending the decay is purely leptonic. (Assume also that the muon mass can be neglected, in comparison with m_τ.) Compare the experimental value.

9.6 Suppose the weak interaction were *pure vector* (no γ^5 in Equation 9.5). Would you still get the same shape for the graph in Figure 9.1?

9.7 What is the *average* value of the electron energy in muon decay?

$$[Answer: (7/20)m_\mu c^2]$$

9.8 Using the coupling $\gamma^\mu(1+\epsilon\gamma^5)$ for $n \to p + W$, but $\gamma^\mu(1-\gamma^5)$ for the leptons, calculate the spin-averaged amplitude for neutron beta decay. Show that your result reduces to Equation 9.41 when $\epsilon = -1$.

$$\left[\ Answer: \langle|\mathscr{M}|^2\rangle = \frac{1}{2}\left(\frac{g_W}{M_W c}\right)^4 [(p_1 \cdot p_2)(p_3 \cdot p_4)(1-\epsilon)^2 + (p_1 \cdot p_4)(p_2 \cdot p_3)(1+\epsilon)^2 - (1-\epsilon^2)m_p m_n c^2 (p_2 \cdot p_4)]\right]$$

9.9 (a) Derive Equation 9.52. (b) Derive Equation 9.58.

9.10 In the text, I said that electron energies in neutron decay range up to about $(m_n - m_p)c^2$. This is not *exact*, since it ignores the *kinetic* energy of the proton and the neutrino. What kinematic configuration gives the maximum electron energy? Apply conservation of energy and momentum to determine the *exact* maximum electron energy.

$$[Answer: (m_n^2 - m_p^2 + m_e^2)c^2/2m_n.]$$

How far off is the approximate answer (give the percent error)?

9.11 (a) Integrate Equation 9.59 to get Equation 9.60.
(b) Approximate as suitable for $m_e \ll \Delta\, m = (m_n - m_p)$. Note that m_e now drops out.

9.12 Obtain Equation 9.62.

9.13 Find the minimum de Broglie wavelength ($\lambda = h/p$) of the W in neutron decay, and compare it with the diameter of the neutron ($\sim 10^{-13}$ cm). [*Answer*: maximum $|\mathbf{p}| = 1.18$ MeV/c, occurring when p and e emerge back to back, so the minimum $\lambda = 10^{-10}$ cm]

9.14 Analyze π^- decay as a scattering process, using the methods of Example 7.8 and Section 8.5. Calculate the decay rate, and, by comparing your answer with the one in the text, obtain the formula for f_π in terms of $|\psi(0)|^2$. Take the quarks to be massless.

$$\left[Answer: f_\pi^2 = \frac{2\hbar^3}{3c}\frac{(2m_\pi^2 + m_l^2)}{m_\pi m_l^2}\cos^2\theta_C|\psi(0)|^2\right]$$

9.15 Show that if $mc^2 \ll E$

$$\gamma^5 u \cong \begin{pmatrix} \boldsymbol{\sigma}\cdot\hat{\mathbf{p}} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\hat{\mathbf{p}} \end{pmatrix} u$$

where u is a particle spinor satisfying the Dirac equation:

$$u = \begin{pmatrix} u_A \\ \dfrac{c(\mathbf{p}\cdot\boldsymbol{\sigma})}{E+mc^2}u_A \end{pmatrix}$$

(Eqs. 7.35 and 7.41). Show therefore that the *projection matrix*

$$P_\pm \equiv \tfrac{1}{2}(1 \pm \gamma^5)$$

picks out the helicity ± 1 component of u:

$$\boldsymbol{\Sigma} \cdot \hat{\mathbf{p}}(P_{\pm}u) = \pm(P_{\pm}u)$$

9.16 Calculate the ratio of the decay rates $K^- \to e^- + \overline{\nu}_e$ and $K^- \to \mu^- + \overline{\nu}_\mu$. Compare the observed branching ratios.

9.17 Calculate decay rates for the following processes: (a) $\Sigma^0 \to \Sigma^+ + e + \overline{\nu}_e$, (b) $\Sigma^- \to \Lambda + e + \overline{\nu}_e$, (c) $\Xi^- \to \Xi^0 + e + \overline{\nu}_e$, (d) $\Lambda \to p + e + \overline{\nu}_e$, (e) $\Sigma^- \to n + e + \overline{\nu}_e$; (f) $\Xi^0 \to \Sigma^+ + e + \overline{\nu}_e$. Assume the coupling is always $\gamma^\mu(1 - \gamma^5)$ – that is, ignore the strong interaction corrections to the axial coupling – but do not forget the Cabibbo factor. Compare the experimental data.

9.18 (a) Show that as long as the CKM matrix is *unitary** ($V^{-1} = V^\dagger$), the GIM mechanism for eliminating $K^0 \to \mu^+\mu^-$ works for three (or any number of) generations. [*Note:* $u \to d + W^+$ carries a CKM factor V_{ud}; $d \to u + W^-$ carries a factor V^*_{ud}.

(b) How many independent real parameters are there in the general 3×3 unitary matrix? How about $n \times n$? [*Hint:* It helps to know that any unitary matrix (U) can be written in the form $U = e^{iH}$, where H is a hermitian matrix. So an equivalent question is, how many independent real parameters are there in the general *hermitian* matrix?] We are free to change the *phase* of each quark wave function (normalization of u really only determines $|N|^2$; see Problem 7.3), so $2n$ of these parameters are arbitrary – or rather, $(2n - 1)$, since changing the phase of *all* quark wave functions by the same amount has no effect on V. *Question*: Can we thus reduce the CKM matrix to a *real* matrix (if it is real and unitary, then it is *orthogonal*: $V^{-1} = \tilde{V}$).

(c) How many independent real parameters are there in the general 3×3 (real) orthogonal matrix? How about $n \times n$?

(d) So, what is the answer? *Can* you reduce the CKM matrix to real form? How about for only two generations ($n = 2$)?

9.19 Show that the CKM matrix (Equation 9.87) is unitary for any (real) numbers θ_{12}, θ_{23}, θ_{13}, and δ.

9.20 Using the experimental values of the Fermi constant G_F (Equation 9.38) and the weak mixing angle θ_w (Equation 9.93), 'predict' the mass of the $W^\pm$and the Z^0, in GWS theory. Compare the experimental values.

9.21 In Example 9.4 I used *muon* neutrinos, rather than *electron* neutrinos. As a matter of fact, ν_μ and $\overline{\nu}_\mu$ beams are easier to produce than ν_e and $\overline{\nu}_e$, but there is also a *theoretical* reason why $\nu_\mu + e^- \to \nu_\mu + e^-$ is simpler than $\nu_e + e^- \to \nu_e + e^-$ or $\overline{\nu}_e + e^- \to \overline{\nu}_e + e^-$. Explain.

9.22 (a) Calculate the differential and total cross section for $\overline{\nu}_\mu + e^- \to \overline{\nu}_\mu + e^-$ in the GWS model. [*Answer*: Same as Equation 9.100, only with the sign of $c_A c_V$ reversed.]

(b) Find the ratio $\sigma(\overline{\nu}_\mu + e^- \to \overline{\nu}_\mu + e^-)/\sigma(\nu_\mu + e^- \to \nu_\mu + e^-)$. Assume the energy is high enough that you can set $m_e = 0$.

9.23 (a) Calculate the decay rate for $Z^0 \to f + \bar{f}$, where f is any quark or any lepton. Assume f is so light (compared to the Z) that its mass can be neglected. (You'll need the completeness relation for the Z^0 – see Problem 9.1.)

$$\left[\textit{Answer}: \Gamma(Z^0 \to f + \bar{f}) = \frac{g_Z^2 M_Z c^2}{48\pi\hbar}(|c_V^f|^2 + |c_A^f|^2).\right]$$

(b) Assuming these are the dominant decay modes, find the branching ratio for each species of quark and lepton (remember that the quarks come in three colors). Should you include the top quark among the allowed decays? [*Answer*: 3% each for e, μ, τ; 7% each for ν_e, ν_μ, ν_τ; 12% each for u, c; 15% each for d, s, b.]

* For experimental confirmation see Problem 9.33.

(c) Calculate the lifetime of the Z^0. Quantitatively, how would it change if there were a fourth generation (quarks and leptons)? (Notice that an accurate measurement of the Z^0 lifetime tells us how many quarks and leptons there can be with masses less than 45 GeV/c^2.)

9.24 Estimate R (the total ratio of quark pair production to muon pair production in e^+e^- scattering), when the process is mediated by Z^0. For the sake of argument, pretend the top quark is light enough so that Equation 9.109 can be used. Don't forget color.

9.25 Graph the ratio in Equation 9.113 as a function of $x \equiv 2E/M_Zc^2$. Use $\Gamma_Z = 7.3(g_z^2/48\pi)(M_Zc^2/\hbar)$ (Problem 9.23).

9.26 (a) If $u(p)$ satisfies the (momentum space) Dirac equation (Equation 7.49), show that u_L and u_R (Table 9.2) do *not* (unless $m = 0$).

(b) Find the eigenvalues and eigenspinors of the matrices $P_\pm \equiv \frac{1}{2}(1 \pm \gamma^5)$.

(c) Can there exist spinors that are simultaneously eigenstates of P_+ (say) *and* of the Dirac operator $(\not{p} - mc)$?

[*Answer* : No; these operators do not commute.]

9.27 Work out the weak isospin currents $j_\mu^\pm$ and j_μ^3 for the light quark doublet u and d'. Also, construct the electromagnetic current (j_μ^{em}) and the weak hypercharge current (j_μ^Y). (Leave your answers in terms of d'.)

9.28 From Equation 9.155, determine the vector and axial vector couplings in Table 9.1.

9.29 In Problem 9.5 you found the decay rate Γ for $\tau \to e + \nu_\tau + \overline{\nu}_e$, and for $\tau \to \mu + \nu_\tau + \overline{\nu}_\mu$ (which is essentially the same). How about for the hadronic modes ($\tau \to d + \nu_\tau + \overline{u}$ and $\tau \to s + \nu_\tau + \overline{u}$)? Estimate the lifetime of the τ (including both leptonic and hadronic modes) and the branching ratios for the electron, muon, and hadron modes. Compare the experimental values. [*Partial answer:* $\Gamma_{\text{tot}} = 5\,\Gamma$]

9.30 (a) Estimate the lifetime of the charmed quark. (First decide what modes dominate, and then make the appropriate modifications in the muon decay formula, Equation 9.35) [*Hint:* Refer to Problem (9.29)]

(b) On the basis of (a), estimate the lifetime of the D meson ($D^0 = c\overline{u}$ and $D^+ = c\overline{d}$), treating the light quark as a spectator. Also estimate the branching ratios for the various semileptonic modes and for the hadronic mode. Compare the experimental values.

(c) In the same way, estimate the lifetime of the B meson ($B^0 = b\overline{d}$ and $B^- = b\overline{u}$). Note that more decay modes are available to the b quark. Find the branching ratios, and compare the experimental values.

(d) According to Equation 9.35, the decay rate goes like the *fifth power* of the mass. The bottom quark is almost four times as massive as the charmed quark. Why, then, isn't the lifetime of the D meson 1000 times longer than that of the B? In *fact*, their lifetimes are quite comparable, but this is something of a coincidence. Explain.

9.31 Calculate the lifetime of the top quark. Note that because $m_t > m_b + m_W$, the top can decay into a *real* W ($t \to b + W^+$), whereas all other quarks must go via a *virtual* W. As a consequence, its lifetime is much shorter, and that's why it does not form bound states ('truthful' mesons and baryons). Take the b quark to be massless (compared to t and W). [*Answer:* 4×10^{-25} s]

9.32 The radical new [your name] theory of weak interactions asserts that the W actually has spin 0 (not 1), and the coupling is 'scalar/pseudo-scalar', instead of 'vector/axial-vector'. Specifically, in your theory the W propagator is

$$\frac{-i}{q^2 - (M_Wc)^2} \approx \frac{i}{(M_Wc)^2}$$

(replacing Equation 9.4), and the vertex factor is

$$\frac{-ig_w}{2\sqrt{2}}(1-\gamma^5)$$

(replacing Equation 9.5). Consider 'inverse muon decay' ($\nu_\mu + e \to \mu + \nu_e$), in this theory:

(a) Draw the Feynman diagram, and construct the amplitude, $\mathscr{M}$.
(b) Determine the spin-averaged quantity, $\langle |\mathscr{M}|^2 \rangle$.
(c) Find the differential scattering cross section, in the CM frame, in terms of the electron energy E and the scattering angle θ. Assume $E \gg m_\mu c^2 \gg m_e c^2$, so you can safely neglect the masses of both the electron and the muon (and, of course, the neutrinos).
(d) Calculate the total cross section, under the same conditions.
(e) By comparing the orthodox predictions for this process, instruct the experimentalists how best to confirm your theory (and demolish the Standard Model). [*Note:* There is no reason to suppose that the weak coupling constant (g_w) in your theory has the same value as it does in the Standard Model, so a test that depends on this number is not going to be very persuasive.]

9.33 The rows (and columns) of a unitary matrix are orthonormal. This suggests a number of tests of the CKM model, as the values of the matrix elements are measured with increasing precision. For example, orthogonality of the first and third columns implies (Equation 9.86)

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$$

or (dividing by the middle term)

$$1 + z_1 + z_2 = 0, \quad \text{where} \quad z_1 \equiv \frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}, \quad z_2 \equiv \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}$$

Plotted in the complex plane, the numbers 1, z_1, and z_2 must add up to form a closed loop, called the 'unitarity triangle'. Look up the best current values for the CKM matrix elements, and plot 1, z_1, and z_2. Does their sum in fact form a closed triangle?

9.34 Find the threshold ν_μ energy for inverse muon decay (Example 9.1), assuming the target electron is at rest. Why is the answer so huge, when all we're doing is producing a muon?

10

Gauge Theories

This chapter introduces the 'gauge theories' that describe all elementary particle interactions. I begin with the Lagrangian formulation of classical mechanics, and proceed to Lagrangian field theory, the principle of local gauge invariance, the notion of spontaneous symmetry-breaking, and the Higgs mechanism (which accounts for the mass of the W's and the Z). This material is quite abstract (in contrast to previous chapters); it concerns the fundamental quantum field theories from which the Feynman rules derive. It will not help you to calculate any cross sections or lifetimes. On the other hand, the ideas discussed here constitute the foundation on which virtually all modern theories are predicated. To understand this chapter it will help to have studied some Lagrangian mechanics, but more essential is the relativistic notation in Chapter 3, the taste of group theory in Chapter 4, the Feynman calculus from Chapter 6, and the Dirac equation from Chapter 7.

10.1
Lagrangian Formulation of Classical Particle Mechanics

According to Newton's second law of motion, a particle of mass m, subjected to a force $\mathbf{F}$, undergoes an acceleration $\mathbf{a}$ given by

$$\mathbf{F} = m\mathbf{a} \tag{10.1}$$

If the force is *conservative*, it can be expressed as the gradient of a scalar potential energy function U:

$$\mathbf{F} = -\nabla U \tag{10.2}$$

and Newton's law reads

$$m\frac{d\mathbf{v}}{dt} = -\nabla U \tag{10.3}$$

where $\mathbf{v}$ is the velocity [1].

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

An alternative formulation of classical mechanics begins with the 'Lagrangian'

$$L = T - U \tag{10.4}$$

where T is the kinetic energy of the particle:

$$T = \frac{1}{2}m\mathrm{v}^2 \tag{10.5}$$

The Lagrangian is a function of the coordinates q_i (say, $q_1 = x$, $q_2 = y$, $q_3 = z$) and their time derivatives $\dot{q}_i$ ($\dot{q}_1 = v_x$, $\dot{q}_2 = v_y$, $\dot{q}_3 = v_z$). In the Lagrangian formulation, the fundamental law of motion is the Euler–Lagrange equation [2]:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i} \quad (i = 1, 2, 3) \tag{10.6}$$

Thus in Cartesian coordinates we have

$$\frac{\partial L}{\partial \dot{q}_1} = \frac{\partial T}{\partial v_x} = mv_x \tag{10.7}$$

$$\frac{\partial L}{\partial q_1} = -\frac{\partial U}{\partial x} \tag{10.8}$$

and the Euler–Lagrange equation (for $i = 1$) reproduces the x component of Newton's law, in the form of Equation 10.3. The Lagrangian formulation is thus equivalent to Newton's (at least, for conservative systems), but it has certain theoretical advantages, as we shall see in the following sections (see also Problem 10.1).

10.2 Lagrangians in Relativistic Field Theory

A *particle*, by its nature, is a *localized* entity; in classical particle mechanics we are typically interested in calculating its *position* as a function of time: $x(t)$, $y(t)$, $z(t)$. A *field*, on the other hand, occupies some *region* of space; in field theory our concern is to calculate one or more functions of *position and time*: $\phi_i(x, y, z, t)$. The field variables ϕ_i might be, for example, the temperature at each point in a room, or the electric potential V, or the three components of the magnetic field $\mathbf{B}$. In particle mechanics, we introduced a Lagrangian L that was a function of the coordinates, q_i, and their time derivatives, $\dot{q}_i$; in field theory we start with a Lagrangian (technically, a Lagrangian *density*) $\mathscr{L}$, which is a function of the fields ϕ_i and their x, y, z, and t derivatives:

$$\partial_\mu \phi_i \equiv \frac{\partial \phi_i}{\partial x^\mu} \tag{10.9}$$

In the former case, the left side of the Euler–Lagrange Equation 10.6 involves only the *time* derivative; a *relativistic* theory must treat space and time coordinates on

an equal footing, and the Euler–Lagrange equations generalize in the simplest possible way, to:

$$\partial_\mu \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i)} \right) = \frac{\partial \mathscr{L}}{\partial \phi_i} \qquad (i = 1, 2, 3, \ldots) \tag{10.10}$$

Example 10.1 The Klein–Gordon Lagrangian for a Scalar (Spin-0) Field Suppose we have a single, scalar field variable ϕ, and the Lagrangian is

$$\mathscr{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2}\left(\frac{mc}{\hbar}\right)^2 \phi^2 \tag{10.11}$$

In this case,

$$\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)} = \partial^\mu \phi \tag{10.12}$$

(If this confuses you, write out the Lagrangian 'longhand':

$$\mathscr{L} = \frac{1}{2}[\partial_0\phi\partial_0\phi - \partial_1\phi\partial_1\phi - \partial_2\phi\partial_2\phi - \partial_3\phi\partial_3\phi] - \frac{1}{2}\left(\frac{mc}{\hbar}\right)^2 \phi^2$$

In this form, it is clear that

$$\frac{\partial \mathscr{L}}{\partial(\partial_0\phi)} = \partial_0\phi = \partial^0\phi, \qquad \frac{\partial \mathscr{L}}{\partial(\partial_1\phi)} = -\partial_1\phi = \partial^1\phi,$$

and so on.) Meanwhile

$$\frac{\partial \mathscr{L}}{\partial \phi} = -\left(\frac{mc}{\hbar}\right)^2 \phi$$

and hence the Euler–Lagrange formula leads to

$$\partial_\mu \partial^\mu \phi + \left(\frac{mc}{\hbar}\right)^2 \phi = 0 \tag{10.13}$$

which is the Klein–Gordon equation (Equation 7.9), describing (in quantum field theory) a particle of spin 0 and mass m. ▩

Example 10.2 The Dirac Lagrangian for a Spinor (Spin-$\frac{1}{2}$) Field Consider now a spinor field ψ, and the Lagrangian

$$\mathscr{L} = i(\hbar c)\overline{\psi}\gamma^\mu \partial_\mu \psi - (mc^2)\overline{\psi}\psi \tag{10.14}$$

We treat ψ and the adjoint spinor $\overline{\psi}$ as independent field variables.* Applying the Euler–Lagrange equation to $\overline{\psi}$, I find

$$\frac{\partial \mathscr{L}}{\partial(\partial_\mu \overline{\psi})} = 0, \qquad \frac{\partial \mathscr{L}}{\partial \overline{\psi}} = i\hbar c \gamma^\mu \partial_\mu \psi - mc^2 \psi$$

so that

$$i\gamma^\mu \partial_\mu \psi - \left(\frac{mc}{\hbar}\right)\psi = 0 \tag{10.15}$$

This is the Dirac equation (Equation 7.20), describing (in quantum field theory) a particle of spin $\frac{1}{2}$ and mass m. Meanwhile, if we apply the Euler–Lagrange equation to ψ, we obtain

$$\frac{\partial \mathscr{L}}{\partial(\partial_\mu \psi)} = i\hbar c \overline{\psi} \gamma^\mu, \qquad \frac{\partial \mathscr{L}}{\partial \psi} = -mc^2 \overline{\psi}$$

and hence

$$i\partial_\mu \overline{\psi} \gamma^\mu + \left(\frac{mc}{\hbar}\right)\overline{\psi} = 0$$

which is the adjoint of the Dirac equation (see Problem 7.15).

Example 10.3 The Proca Lagrangian for a Vector (Spin-1) Field Finally, suppose we take a vector field, A^μ, with the Lagrangian

$$\mathscr{L} = \frac{-1}{16\pi}(\partial^\mu A^\nu - \partial^\nu A^\mu)(\partial_\mu A_\nu - \partial_\nu A_\mu) + \frac{1}{8\pi}\left(\frac{mc}{\hbar}\right)^2 A^\nu A_\nu \tag{10.16}$$

Here

$$\frac{\partial \mathscr{L}}{\partial(\partial_\mu A_\nu)} = \frac{-1}{4\pi}(\partial^\mu A^\nu - \partial^\nu A^\mu) \tag{10.17}$$

(see Problem 10.2) and

$$\frac{\partial \mathscr{L}}{\partial A_\nu} = \frac{1}{4\pi}\left(\frac{mc}{\hbar}\right)^2 A^\nu \tag{10.18}$$

so the Euler–Lagrange equation yields

$$\partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) + \left(\frac{mc}{\hbar}\right)^2 A^\nu = 0 \tag{10.19}$$

* Since ψ is a complex spinor, there are actually *eight* independent fields here (i runs from 1 to 8): the real and imaginary parts of each of the four components of ψ. But in applying the Euler–Lagrange equations any linear combinations of these eight will do just as well, and we choose to use the four components of ψ plus the four components of $\overline{\psi}$.

This is called the Proca equation; it describes a particle of spin 1 and mass m. Incidentally, since the combination $(\partial^\mu A^\nu - \partial^\nu A^\mu)$ occurs repeatedly in this theory, it is useful to introduce the shorthand

$$F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu \tag{10.20}$$

Then, the Lagrangian reads

$$\mathscr{L} = -\frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} + \frac{1}{8\pi}\left(\frac{mc}{\hbar}\right)^2 A^\nu A_\nu \tag{10.21}$$

and the field equation becomes

$$\partial_\mu F^{\mu\nu} + \left(\frac{mc}{\hbar}\right)^2 A^\nu = 0 \tag{10.22}$$

If the notation is beginning to remind you of electrodynamics, it's no accident, for the electromagnetic field *is* precisely a massless vector field; if you set $m = 0$ in Equation 10.22 you're left with Maxwell's equations for empty space.*

The Lagrangians in these examples came out of thin air (or rather, they were concocted in such a way as to reproduce the desired field equations). In classical particle mechanics, L is *derived* ($L = T - U$), but in relativistic field theory $\mathscr{L}$ is usually taken as *axiomatic* – you have to start somewhere. The Lagrangian for a particular system is not unique; you can always multiply $\mathscr{L}$ by a constant, or add a constant – or for that matter the divergence of an arbitrary vector function ($\partial_\mu M^\mu$, where M^μ is any function of ϕ_i and $\partial_\mu \phi_i$); such terms cancel out when you apply the Euler–Lagrange equations, so they do not affect the field equations. In this sense, the factors of $\frac{1}{2}$ in the Klein–Gordon Lagrangian, for example, are purely conventional.† Apart from that, however, what we have here are *the* Lagrangians for spin 0, spin $\frac{1}{2}$, and spin 1. So far, however, we are talking only of *free* fields, with no sources or interactions.

* Notice that in this formulation A^μ is the fundamental quantity and $F^{\mu\nu}$ is just convenient notation (Equation 10.20) – the reverse of the perspective taken in classical electrodynamics, where **E** and **B** (hence $F^{\mu\nu}$) are fundamental and the potentials are constructs. In particular, for purposes of the Euler–Lagrange equations the 'fields' are the components of A^μ, *not* $F^{\mu\nu}$.

† The Lagrangian (L) carries units of energy (Equation 10.4), and the Lagrangian *density* ($\mathscr{L}$) has the units of energy *per unit volume*. The *fields* carry dimensions as follows:

ϕ (scalar field): $\sqrt{ML}/T$

ψ (spinor field): $L^{-3/2}$

A^μ (vector field): $\sqrt{ML}/T$

These are chosen so that ψ will go over to the Schrodinger wave function (in the nonrelativistic limit) and A^μ to the Maxwell vector potential (in the nonquantum limit). By the way, in Heaviside–Lorentz units the Proca and Maxwell Lagrangians are conventionally multiplied by 4π.

Example 10.4 The Maxwell Lagrangian for a Massless Vector Field with Source J^μ
Suppose

$$\mathscr{L} = \frac{-1}{16\pi} F^{\mu\nu} F_{\mu\nu} - \frac{1}{c} J^\mu A_\mu \tag{10.23}$$

where $F^{\mu\nu}$ (again) stands for $(\partial^\mu A^\nu - \partial^\nu A^\mu)$ and J^μ is some specified function. The Euler–Lagrange equations yield

$$\partial_\mu F^{\mu\nu} = \frac{4\pi}{c} J^\nu \tag{10.24}$$

which (as we found in Section 7.4) is the tensor form of Maxwell's equations, describing the electromagnetic fields produced by a current J^μ. Incidentally, it follows from Equation 10.24 that

$$\partial_\nu J^\nu = 0 \tag{10.25}$$

That is, the internal consistency of the Maxwell Lagrangian (Equation 10.23) requires that the current satisfy the continuity equation (Equation 7.74); you can't just put in *any* old function for J^μ – it's got to respect conservation of charge. ▪

10.3 Local Gauge Invariance

Notice that the Dirac Lagrangian

$$\mathscr{L} = i\hbar c \overline{\psi} \gamma^\mu \partial_\mu \psi - mc^2 \overline{\psi} \psi$$

is invariant under the transformation

$$\psi \to e^{i\theta} \psi \qquad \text{(global phase transformation)} \tag{10.26}$$

(where θ is any real number), for then $\overline{\psi} \to e^{-i\theta} \overline{\psi}$, and in the combination $\overline{\psi}\psi$ the exponential factors cancel out. (Already in *non*relativistic quantum mechanics, of course, the overall phase of the wave function is arbitrary.) But what if the phase factor is different at different space–time points; that is, *what if θ is a function of x^μ*:

$$\psi \to e^{i\theta(x)} \psi \qquad \text{(local phase transformation)} \tag{10.27}$$

Is the Lagrangian invariant under such a '*local*' phase transformation? The answer is *no*, for now we pick up an extra term from the derivative of θ:

$$\partial_\mu (e^{i\theta} \psi) = i(\partial_\mu \theta) e^{i\theta} \psi + e^{i\theta} \partial_\mu \psi \tag{10.28}$$

so that

$$\mathscr{L} \to \mathscr{L} - \hbar c(\partial_\mu \theta)\overline{\psi}\gamma^\mu \psi \tag{10.29}$$

For what follows, it is convenient to pull a factor of $-(q/\hbar c)$ out of θ, letting

$$\lambda(x) \equiv -\frac{\hbar c}{q}\theta(x) \tag{10.30}$$

where q is the charge of the particle involved. In terms of λ, then,

$$\mathscr{L} \to \mathscr{L} + (q\overline{\psi}\gamma^\mu \psi)\partial_\mu \lambda \tag{10.31}$$

under the local phase transformation

$$\psi \to e^{-iq\lambda(x)/\hbar c}\psi \tag{10.32}$$

So far, there is nothing particularly new or deep about this. The crucial point comes when we *demand that the complete Lagrangian be invariant under local phase transformations.** Since the *free* Dirac Lagrangian (Equation 10.14) is *not* locally phase invariant, we are obliged to *add* something, in order to soak up the extra term in Equation 10.31. Specifically, suppose

$$\mathscr{L} = [i\hbar c\overline{\psi}\gamma^\mu \partial_\mu \psi - mc^2\overline{\psi}\psi] - (q\overline{\psi}\gamma^\mu \psi)A_\mu \tag{10.33}$$

where A_μ is some *new* field, which changes (in coordination with the local phase transformation of ψ) according to the rule

$$A_\mu \to A_\mu + \partial_\mu \lambda \tag{10.34}$$

This 'new, improved' Lagrangian *is* now locally invariant – the $\partial_\mu \lambda$ in Equation 10.34 exactly compensates for the 'extra' term in Equation 10.31. The price we have to pay is the introduction of a new vector field that couples to ψ through the last term in Equation 10.33 (see Problem 10.6). But Equation 10.33 isn't the whole story; the *full* Lagrangian must include a 'free' term for the field A^μ itself. Since it's a vector, we look to the Proca Lagrangian (Equation 10.21)

$$\mathscr{L} = \frac{-1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \frac{1}{8\pi}\left(\frac{m_A c}{\hbar}\right)^2 A^\nu A_\nu$$

But there is a problem here, for whereas $F^{\mu\nu} \equiv (\partial^\mu A^\nu - \partial^\nu A^\mu)$ is invariant under Equation 10.34 (as you should check for yourself), $A^\nu A_\nu$ is *not. Evidently the new field must be massless* ($m_A = 0$), otherwise the invariance will be lost.

* I know of no compelling physical argument for insisting that a global invariance *should* hold locally. If you believe that phase transformations are in some sense 'fundamental', then I suppose one should be able to carry them out independently at spacelike separated points (which are, after all, out of communication with one another). But I think this begs the question. Better, for the moment at least, to take the requirement of local phase invariance as a new principle of physics in its own right.

Conclusion: If we start with the Dirac Lagrangian, and demand *local* phase invariance, we are forced to introduce a massless vector field (A^μ), and the complete Lagrangian becomes

$$\mathscr{L} = [i\hbar c\overline{\psi}\gamma^\mu\partial_\mu\psi - mc^2\overline{\psi}\psi] - \left[\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu}\right] - (q\overline{\psi}\gamma^\mu\psi)A_\mu \tag{10.35}$$

As you will have guessed, A^μ is nothing but the electromagnetic potential; the transformation rule for A^μ (Equation 10.34) is precisely the *gauge* invariance* we found back in Chapter 7 (Equation 7.81), and the last two terms in Equation 10.35 reproduce the Maxwell Lagrangian (Equation 10.23), with the current density

$$J^\mu = cq(\overline{\psi}\gamma^\mu\psi) \tag{10.36}$$

Thus the requirement of local phase invariance, applied to the free Dirac Lagrangian, generates all of electrodynamics and specifies the current produced by Dirac particles.

This is a truly breathtaking accomplishment. The critical step was the added term in Equation 10.33. How was this obtained? The difference between global and local phase transformations arises when we calculate *derivatives* of the fields (Equation 10.28):

$$\partial_\mu\psi \to e^{-iq\lambda/\hbar c}\left[\partial_\mu - i\frac{q}{\hbar c}(\partial_\mu\lambda)\right]\psi \tag{10.37}$$

Instead of a simple phase factor, we pick up an extra piece involving $\partial_\mu\lambda$. *If in the original (free) Lagrangian we replace every derivative* (∂_μ) *by the so-called 'covariant derivative'*

$$\mathscr{D}_\mu \equiv \partial_\mu + i\frac{q}{\hbar c}A_\mu \tag{10.38}$$

(and every ∂^μ by $\mathscr{D}^\mu$) the gauge transformation of A_μ (Equation 10.34) will cancel the offending term in Equation 10.37

$$\mathscr{D}_\mu\psi \to e^{-iq\lambda/\hbar c}\mathscr{D}_\mu\psi \tag{10.39}$$

and the invariance of $\mathscr{L}$ is restored. The substitution of $\mathscr{D}_\mu$ for ∂_μ, then, is a beautifully simple device for converting a *globally* invariant Lagrangian into a *locally* invariant one; we call it the *'minimal coupling rule'*.† But the covariant

* Because of the connection with gauge invariance in classical electrodynamics, we now call Equations 10.34 *and* 10.26 'gauge transformations', A^μ is called the 'gauge field', and the entire strategy is called 'gauge theory'.

† The minimal coupling rule is much older than the principle of local gauge invariance. In terms of momentum ($p_\mu \leftrightarrow i\hbar\partial_\mu$) it reads $p_\mu \to p_\mu - i(q/c)A_\mu$, and is a well-known trick in classical electrodynamics for obtaining the equation of motion for a charged particle in the presence of electrodynamic fields. It amounts, in this sense, to a sophisticated formulation of the Lorentz force law. In modern particle theory we prefer to regard local gauge invariance as fundamental and minimal coupling as the vehicle for achieving it.

derivative introduces a new vector field (A_μ), which requires its own *free* Lagrangian; if the latter is not to spoil local gauge invariance, we must take the gauge field to be *massless*. This leads to the final expression (Equation 10.35), which people in the know would immediately recognize as the Lagrangian for quantum electrodynamics – Dirac fields (electrons and positrons) interacting with Maxwell fields (photons).

The idea of local gauge invariance goes back to Hermann Weyl in 1918 [3]. However, its power and generality were not fully appreciated until the early 1970s. Our starting point (the global phase transformation in Equation 10.26) may be thought of as multiplication of ψ by a unitary 1×1 matrix:

$$\psi \to U\psi, \quad \text{where} \quad U^\dagger U = 1 \tag{10.40}$$

(here, $U = e^{i\theta}$). The group of *all* such matrices is $U(1)$ (see Table 4.2), and hence the symmetry involved is called '$U(1)$ gauge invariance'. This terminology is extravagant for the case at hand (a 1×1 matrix is a *number*, so why not leave it at that?), but in 1954 Yang and Mills [4] applied the same strategy (insisting that a *global* invariance hold *locally*) to the group $SU(2)$, and later on the idea was extended to color $SU(3)$, producing chromodynamics. In the Standard Model, *all* of the fundamental interactions are generated in this way.

10.4 Yang–Mills Theory

Suppose now that we have *two* spin-$\frac{1}{2}$ fields, ψ_1 and ψ_2. The Lagrangian, in the absence of any interactions, is

$$\mathscr{L} = [i\hbar c\overline{\psi}_1\gamma^\mu\partial_\mu\psi_1 - m_1c^2\overline{\psi}_1\psi_1] + [i\hbar c\overline{\psi}_2\gamma^\mu\partial_\mu\psi_2 - m_2c^2\overline{\psi}_2\psi_2] \tag{10.41}$$

It's just the *sum* of the two Dirac Lagrangians. (Apply the Euler–Lagrange equations to this $\mathscr{L}$, and you'll find that both ψ_1 and ψ_2 obey the Dirac equation, with the appropriate mass.) But we can write Equation 10.41 more compactly by combining ψ_1 and ψ_2 into a two-component column vector:

$$\psi \equiv \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \tag{10.42}$$

(Of course, ψ_1 and ψ_2 are themselves four-component Dirac spinors, and you might prefer a double-index notation: $\psi_{\alpha,i}$, where $\alpha = 1, 2$ identifies the *particle* and $i = 1, 2, 3, 4$ labels the spinor component. However, in the present context we are *only* concerned with the particle index, although the Dirac matrices, of course, act on the spinor indices.) The adjoint spinor is

$$\overline{\psi} = (\overline{\psi}_1\ \overline{\psi}_2) \tag{10.43}$$

and the Lagrangian becomes

$$\mathscr{L} = i\hbar c\overline{\psi}\gamma^{\mu}\partial_{\mu}\psi - c^2\overline{\psi}M\psi \tag{10.44}$$

where

$$M = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \tag{10.45}$$

is the 'mass matrix'. In particular, if the two masses happen to be *equal* Equation 10.44 reduces to

$$\mathscr{L} = i\hbar c\overline{\psi}\gamma^{\mu}\partial_{\mu}\psi - mc^2\overline{\psi}\psi \tag{10.46}$$

This *looks* just like the one-particle Dirac Lagrangian. However, ψ is now a two-element column vector, and $\mathscr{L}$ admits a more general global invariance than before:

$$\psi \to U\psi \tag{10.47}$$

where U is any 2×2 unitary matrix

$$U^{\dagger}U = 1 \tag{10.48}$$

For under the transformation in Equation 10.47,

$$\overline{\psi} \to \overline{\psi}U^{\dagger} \tag{10.49}$$

and hence the combination $\overline{\psi}\psi$ is invariant. Now, just as any complex number of modulus 1 can be written in the form $e^{i\theta}$, with real θ, so any unitary matrix can be written in the form [5]

$$U = e^{iH} \tag{10.50}$$

where H is *Hermitian* ($H^{\dagger} = H$).* Moreover, the most general Hermitian 2×2 matrix can be expressed in terms of four real numbers, a_1, a_2, a_3, and θ (Problem 10.10):

$$H = \theta 1 + \boldsymbol{\tau}\cdot\mathbf{a} \tag{10.51}$$

* In matrix theory the natural generalization of complex conjugation (*) is Hermitian conjugation (†) – *transpose* conjugation. Of course, there's no distinction in the case of 1×1 matrices (complex numbers), but for higher dimensions it is the Hermitian conjugate that shares the most useful properties of ordinary complex conjugation. In this sense the closest analog to a *real* number ($a = a^*$) is a *Hermitian* matrix ($A = A^{\dagger}$), and the analog to a number of modulus 1 ($a^*a = 1$) is a *unitary* matrix ($A^{\dagger}A = 1$).

where 1 is the 2×2 unit matrix, τ_1, τ_2, τ_3 are the Pauli matrices (Equation 4.26), and the dot product is a convenient shorthand for $\tau_1 a_1 + \tau_2 a_2 + \tau_3 a_3$. Thus any unitary 2×2 matrix can be expressed as a product:

$$U = e^{i\theta} e^{i\boldsymbol{\tau} \cdot \mathbf{a}} \tag{10.52}$$

We have already explored the implications of *phase* transformations ($e^{i\theta}$); in this section we shall concentrate on transformations of the form

$$\psi \to e^{i\boldsymbol{\tau} \cdot \mathbf{a}} \psi \qquad \text{(global } SU(2) \text{ transformation)} \tag{10.53}$$

The matrix $e^{i\boldsymbol{\tau} \cdot \mathbf{a}}$ has determinant 1 (see Problem 4.22), and therefore belongs to the group $SU(2)$. Generalizing the terminology of Section 10.3, we say that the Lagrangian is invariant under global $SU(2)$ gauge transformations.* What Yang and Mills did was to promote this *global* invariance to the status of a *local* invariance.

The inspiration and the strategy were similar to Weyl's, but the implementation is more subtle; in fact, it's quite remarkable that it works at all. The first step is to let the parameters $(\mathbf{a})$ be functions of x^μ (as before, I'll let $\boldsymbol{\lambda}(x) \equiv -(\hbar c/q)\mathbf{a}(x)$, where q is a coupling constant analogous to electric charge):

$$\psi \to S\psi, \text{ where } S \equiv e^{-iq\boldsymbol{\tau} \cdot \boldsymbol{\lambda}(x)/\hbar c} \text{ (local } SU(2) \text{ transformation)} \tag{10.54}$$

As it stands, $\mathscr{L}$ is *not* invariant under such a transformation, for the derivative picks up an extra term:

$$\partial_\mu \psi \to S\partial_\mu \psi + (\partial_\mu S)\psi \tag{10.55}$$

The remedy, again, is to replace the derivative in $\mathscr{L}$ by a 'covariant derivative', modeled on Equation 10.38, but taking into account the structure of Equation 10.55:

$$\mathscr{D}_\mu \equiv \partial_\mu + i\frac{q}{\hbar c}\boldsymbol{\tau} \cdot \mathbf{A}_\mu \tag{10.56}$$

and assign to the gauge fields $\mathbf{A}_\mu$ (it takes *three* of them this time) a transformation rule such that

$$\mathscr{D}_\mu \psi \to S(\mathscr{D}_\mu \psi) \tag{10.57}$$

for then the Lagrangian (Equation 10.46) will clearly be invariant.

It is not a trivial matter to deduce the transformation rule for $\mathbf{A}_\mu$ from Equation 10.57 [6]. I'll leave it for you to show (Problem 10.11) that $\mathbf{A}_\mu \to \mathbf{A}'_\mu$, where $\mathbf{A}'_\mu$ is given by

$$\boldsymbol{\tau} \cdot \mathbf{A}'_\mu = S(\boldsymbol{\tau} \cdot \mathbf{A}_\mu)S^{-1} + i\left(\frac{\hbar c}{q}\right)(\partial_\mu S)S^{-1} \tag{10.58}$$

* It is *also* invariant under the *larger* group $U(2)$. But Equation 10.52 shows that any element of $U(2)$ can be expressed as an element of $SU(2)$ times an appropriate phase factor (in group-theoretical language, $U(2) = U(1) \otimes SU(2)$), and since we have already studied $U(1)$ invariance, the only thing *new* here is the $SU(2)$ symmetry.

This much is relatively straightforward. But S and S^{-1} in the first term cannot be brought together, because they do not commute with $\boldsymbol{\tau} \cdot \mathbf{A}_\mu$. Nor is the gradient of S simply $-i(q\boldsymbol{\tau} \cdot \partial_\mu \boldsymbol{\lambda}/\hbar c)S$, because S does not commute with $\boldsymbol{\tau} \cdot \partial_\mu \boldsymbol{\lambda}$. You can work out the exact result (using Problems 4.20 and 4.21), if you have the energy, but the answer is not particularly illuminating. For our purposes it will suffice to know the *approximate* transformation rule, in the limiting case of very small $|\boldsymbol{\lambda}|$, for which we may expand S and keep only the first-order terms:

$$S \cong 1 - \frac{iq}{\hbar c}\boldsymbol{\tau} \cdot \boldsymbol{\lambda}, \quad S^{-1} \cong 1 + \frac{iq}{\hbar c}\boldsymbol{\tau} \cdot \boldsymbol{\lambda}, \quad \partial_\mu S \cong -\frac{iq}{\hbar c}\boldsymbol{\tau} \cdot \partial_\mu \boldsymbol{\lambda} \tag{10.59}$$

In this approximation Equation 10.58 yields

$$\boldsymbol{\tau} \cdot \mathbf{A}'_\mu \cong \boldsymbol{\tau} \cdot \mathbf{A}_\mu + \frac{iq}{\hbar c}[\boldsymbol{\tau} \cdot \mathbf{A}_\mu, \boldsymbol{\tau} \cdot \boldsymbol{\lambda}] + \boldsymbol{\tau} \cdot \partial_\mu \boldsymbol{\lambda} \tag{10.60}$$

and hence (using Problem 4.20, to evaluate the commutator)

$$\mathbf{A}'_\mu \cong \mathbf{A}_\mu + \partial_\mu \boldsymbol{\lambda} + \frac{2q}{\hbar c}(\boldsymbol{\lambda} \times \mathbf{A}_\mu) \tag{10.61}$$

The resulting Lagrangian

$$\mathscr{L} = i\hbar c\overline{\psi}\gamma^\mu \mathscr{D}_\mu \psi - mc^2\overline{\psi}\psi = [i\hbar c\overline{\psi}\gamma^\mu \partial_\mu \psi - mc^2\overline{\psi}\psi] - (q\overline{\psi}\gamma^\mu \boldsymbol{\tau}\psi) \cdot \mathbf{A}_\mu \tag{10.62}$$

is invariant under local gauge transformations (Equations 10.54 and 10.58), but we have been obliged to introduce three new vector fields $\mathbf{A}^\mu = (A_1^\mu, A_2^\mu, A_3^\mu)$, and they will require their own *free* Lagrangian:

$$\mathscr{L}_A = -\frac{1}{16\pi}F_1^{\mu\nu}F_{\mu\nu 1} - \frac{1}{16\pi}F_2^{\mu\nu}F_{\mu\nu 2} - \frac{1}{16\pi}F_3^{\mu\nu}F_{\mu\nu 3} = -\frac{1}{16\pi}\mathbf{F}^{\mu\nu} \cdot \mathbf{F}_{\mu\nu} \tag{10.63}$$

(Again, the three-vector notation pertains to the *particle* indices.) The Proca mass term

$$\frac{1}{8\pi}\left(\frac{m_A c}{\hbar}\right)^2 \mathbf{A}^\nu \cdot \mathbf{A}_\nu \tag{10.64}$$

is excluded by local gauge invariance; as before, the gauge fields must be massless. But this time the old association $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ must itself be modified, for with this definition the gauge field Lagrangian (Equation 10.63) is not invariant *either* (see Problem 10.12). Rather, we take*

$$\mathbf{F}^{\mu\nu} \equiv \partial^\mu \mathbf{A}^\nu - \partial^\nu \mathbf{A}^\mu - \frac{2q}{\hbar c}(\mathbf{A}^\mu \times \mathbf{A}^\nu) \tag{10.65}$$

* This definition is not as arbitrary as it may seem. The point is that with *three*-vector fields there is a second antisymmetric tensor form available $(\mathbf{A}^\mu \times \mathbf{A}^\nu)$, and the coefficient, $-2q/\hbar c$, is chosen precisely to *make* $\mathscr{L}_A$ invariant. Notice that when the coupling constant q goes to zero we are left with the free Dirac Lagrangian for each spinor field and the free (massless) Proca Lagrangian for each of the three gauge fields.

Under infinitesimal local gauge transformations (Equation 10.61),

$$\mathbf{F}^{\mu\nu} \to \mathbf{F}^{\mu\nu} + \frac{2q}{\hbar c}(\boldsymbol{\lambda} \times \mathbf{F}^{\mu\nu}) \tag{10.66}$$

(Problem 10.13), and hence $\mathscr{L}_A$ is invariant. (See Problem 10.14 for a proof that the invariance extends to *finite* gauge transformations.)

Conclusion: The complete Yang–Mills Lagrangian is

$$\mathscr{L} = [i\hbar c\overline{\psi}\gamma^\mu\partial_\mu\psi - mc^2\overline{\psi}\psi] - \frac{1}{16\pi}\mathbf{F}^{\mu\nu}\cdot\mathbf{F}_{\mu\nu} - (q\overline{\psi}\gamma^\mu\boldsymbol{\tau}\psi)\cdot\mathbf{A}_\mu \tag{10.67}$$

with $\mathbf{F}^{\mu\nu}$ defined by Equation 10.65; it is invariant under local $SU(2)$ gauge transformations (Equations 10.54 and 10.58), and describes two equal-mass Dirac fields in interaction with three massless vector gauge fields. All this results from insisting that the *global* $SU(2)$ invariance of the original free Lagrangian (Equation 10.46) shall hold *locally*. Borrowing the language of electrodynamics, we say that the Dirac fields generate three *currents*

$$\mathbf{J}^\mu \equiv cq(\overline{\psi}\gamma^\mu\boldsymbol{\tau}\psi) \tag{10.68}$$

which act as *sources* for the gauge fields; the Lagrangian for the gauge fields alone

$$\mathscr{L} = -\frac{1}{16\pi}\mathbf{F}^{\mu\nu}\cdot\mathbf{F}_{\mu\nu} - \frac{1}{c}\mathbf{J}^\mu\cdot\mathbf{A}_\mu \tag{10.69}$$

is reminiscent of the Maxwell Lagrangian (Equation 10.23), and gives rise to a rich and interesting *classical* field theory [7] (see Problem 10.15).

Although Yang–Mills theory was inspired by the same *idea* as Weyl's (namely: a global invariance should hold locally), the implementation was more subtle at two points: (i) the local transformation rule for gauge fields, and (ii) the expression for $F^{\mu\nu}$ in terms of A^μ. Both complications derive from the fact that the symmetry group in question is non-Abelian (2×2 matrices do not commute, whereas 1×1 matrices – obviously – do). To emphasize the distinction, we refer to the Weyl case as an *Abelian* gauge theory and Yang–Mills as a *non-Abelian* gauge theory. In contemporary elementary particle physics, many symmetry groups have been explored; we shall encounter a few in the remaining sections of this book. However, the *hard* work is over: extending non-Abelian gauge theory to higher symmetry groups is a straightforward procedure, once the Yang–Mills model is on the table.

Curiously, though, Yang–Mills theory in its original form turned out to be of little use. After all, it starts from the premise that there exist two elementary spin-$\frac{1}{2}$ particles of equal mass, and as far as we know there are no such pairs in nature. Yang and Mills themselves had the nucleon system (proton and neutron) in mind, and thought of their model as a way of implementing Heisenberg's isospin invariance in the strong interactions. The small mass difference between proton and neutron, 1.29 MeV/c^2, would be attributed to electromagnetic symmetry-breaking. For the

theory to succeed there had to exist a massless isotriplet of vector (spin-1) particles. The only candidates in sight are the ρ mesons; but they are hardly *massless* ($M_\rho = 770\ \mathrm{MeV}/c^2$), and this is not a minor discrepancy that can be plausibly blamed on electromagnetic contamination. A number of attempts were made to doctor up Yang–Mills theory to accommodate massive gauge bosons, but by the time they finally bore fruit (through the Higgs mechanism) it was pretty clear that p, n, and ρ are composite particles anyway, and that isospin is just one component of a larger flavor symmetry that is too drastically broken to play any fundamental role in the strong interactions. When non-Abelian gauge theory finally came into its own, it was in the context of *color* $SU(3)$ symmetry in the strong interactions and (weak) isospin-hypercharge $SU(2)_L \otimes U(1)$ symmetry in the weak interactions. Meanwhile, for more than a decade after 1954 the Yang–Mills model languished – a lovely idea that nature had evidently chosen not to exploit.

10.5 Chromodynamics

According to the Standard Model, each flavor of quark comes in three colors – red, blue, and green. Although the various *flavors* carry different masses (Table 4.4), the three colors of a given flavor are all supposed to weigh the same. Thus the free Lagrangian for a particular flavor reads

$$\begin{aligned}\mathscr{L} &= [i\hbar c\overline{\psi}_r\gamma^\mu\partial_\mu\psi_r - mc^2\overline{\psi}_r\psi_r] + [i\hbar c\overline{\psi}_b\gamma^\mu\partial_\mu\psi_b - mc^2\overline{\psi}_b\psi_b] \\ &\quad + [i\hbar c\overline{\psi}_g\gamma^\mu\partial_\mu\psi_g - mc^2\overline{\psi}_g\psi_g]\end{aligned} \tag{10.70}$$

As before, we can simplify the notation by introducing

$$\psi \equiv \begin{pmatrix}\psi_r \\ \psi_b \\ \psi_g\end{pmatrix}, \qquad \overline{\psi} = (\overline{\psi}_r\ \overline{\psi}_b\ \overline{\psi}_g) \tag{10.71}$$

so that

$$\mathscr{L} = i\hbar c\overline{\psi}\gamma^\mu\partial_\mu\psi - mc^2\overline{\psi}\psi \tag{10.72}$$

This looks just like the original Dirac Lagrangian, only ψ now stands for a three-component column vector (each element of which is itself a four-component Dirac spinor). Just as the one-particle Dirac Lagrangian (Equation 10.14) has (global) $U(1)$ phase invariance, and the (equal mass) two-particle Lagrangian (Equation 10.41) admits $U(2)$ invariance, so this (equal mass) three-particle Lagrangian exhibits $U(3)$ symmetry. That is to say, it is invariant under transformations of the form

$$\psi \to U\psi \qquad (\overline{\psi} \to \overline{\psi}U^\dagger) \tag{10.73}$$

where U is any unitary 3×3 matrix:

$$U^\dagger U = 1 \tag{10.74}$$

But remember (Equation 10.50), any unitary matrix can be written as an exponentiated Hermitian matrix:

$$U = e^{iH}, \qquad \text{with } H^\dagger = H \tag{10.75}$$

Moreover, any 3×3 Hermitian matrix can be expressed in terms of nine real numbers, $a_1, a_2, \ldots, a_8$, and θ (Problem 10.16):

$$H = \theta 1 + \boldsymbol{\lambda} \cdot \mathbf{a} \tag{10.76}$$

where 1 is the 3×3 unit matrix, $\lambda_1, \lambda_2, \ldots, \lambda_8$ are the Gell-Mann matrices (Equation 8.34), and the dot product now denotes a sum from 1 to 8:

$$\boldsymbol{\lambda} \cdot \mathbf{a} \equiv \lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_8 a_8 \tag{10.77}$$

Thus,

$$U = e^{i\theta} e^{i\boldsymbol{\lambda} \cdot \mathbf{a}} \tag{10.78}$$

We have already explored *phase* transformations ($e^{i\theta}$); what is *new* is the second term. The matrix $e^{i\boldsymbol{\lambda} \cdot \mathbf{a}}$ has determinant 1 (see Problem 10.17); it belongs to the group $SU(3)$.* So what we are interested in is the invariance of the Lagrangian (Equation 10.72) under $SU(3)$ transformations, a global symmetry that we now propose to make local.

That is, we modify $\mathscr{L}$ in such a way as to render it invariant under *local* $SU(3)$ gauge transformations:

$$\psi \to S\psi, \quad \text{where} \quad S \equiv e^{-iq\boldsymbol{\lambda} \cdot \boldsymbol{\phi}(x)/\hbar c} \tag{10.79}$$

(again, I let $\boldsymbol{\phi} \equiv -(\hbar c/q)\mathbf{a}$, with the coupling constant q playing a role analogous to electric charge in QED). As always, the trick is to replace the ordinary derivative, ∂_μ, by the 'covariant derivative' $\mathscr{D}_\mu$:

$$\mathscr{D}_\mu \equiv \partial_\mu + i\frac{q}{\hbar c}\boldsymbol{\lambda} \cdot \mathbf{A}_\mu \tag{10.80}$$

and assign to the gauge fields $\mathbf{A}_\mu$ (there are *eight* of them, notice) a transformation rule such that

$$\mathscr{D}_\mu \psi \to S(\mathscr{D}_\mu \psi) \tag{10.81}$$

* In the language of group theory, $U(3) = U(1) \otimes SU(3)$.

Again (see Equation 10.58), this entails

$$\boldsymbol{\lambda} \cdot \mathbf{A}'_\mu = S(\boldsymbol{\lambda} \cdot \mathbf{A}_\mu)S^{-1} + i\left(\frac{\hbar c}{q}\right)(\partial_\mu S)S^{-1} \tag{10.82}$$

which, for the infinitesimal case, yields a formula identical to Equation 10.61:

$$\mathbf{A}'_\mu \cong \mathbf{A}_\mu + \partial_\mu \boldsymbol{\phi} + \frac{2q}{\hbar c}(\boldsymbol{\phi} \times \mathbf{A}_\mu) \tag{10.83}$$

However, this time the cross product notation is shorthand for

$$(\mathbf{B} \times \mathbf{C})_i = \sum_{j,k=1}^{8} f_{ijk} B_j C_k \tag{10.84}$$

where f_{ijk} are the structure constants of $SU(3)$ (Equation 8.35), analogous to ϵ_{ijk} for $SU(2)$ (Problem 10.18).

The modified Lagrangian

$$\mathscr{L} = i\hbar c\overline{\psi}\gamma^\mu \mathscr{D}_\mu \psi - mc^2\overline{\psi}\psi = [i\hbar c\overline{\psi}\gamma^\mu \partial_\mu \psi - mc^2\overline{\psi}\psi] - (q\overline{\psi}\gamma^\mu \boldsymbol{\lambda}\psi) \cdot \mathbf{A}_\mu \tag{10.85}$$

is invariant under local $SU(3)$ gauge transformations (Equations 10.79 and 10.82), but as usual the cost is the introduction of gauge fields $\mathbf{A}^\mu$ (eight of them, this time). In particle language, these correspond to the eight gluons, just as the $U(1)$ gauge field in Weyl's theory represents the photon.* To finish the job, we must adjoin the free gluon Lagrangian

$$\mathscr{L}_{\text{gluons}} = -\frac{1}{16\pi}\mathbf{F}^{\mu\nu} \cdot \mathbf{F}_{\mu\nu} \tag{10.86}$$

where, as in the Yang–Mills case

$$\mathbf{F}^{\mu\nu} \equiv \partial^\mu \mathbf{A}^\nu - \partial^\nu \mathbf{A}^\mu - \frac{2q}{\hbar c}(\mathbf{A}^\mu \times \mathbf{A}^\nu) \tag{10.87}$$

(with the $SU(3)$ 'cross product' defined by Equation 10.84).

Conclusion: The complete Lagrangian for chromodynamics is

$$\mathscr{L} = [i\hbar c\overline{\psi}\gamma^\mu \partial_\mu \psi - mc^2\overline{\psi}\psi] - \frac{1}{16\pi}\mathbf{F}^{\mu\nu} \cdot \mathbf{F}_{\mu\nu} - (q\overline{\psi}\gamma^\mu \boldsymbol{\lambda}\psi) \cdot \mathbf{A}_\mu \tag{10.88}$$

$\mathscr{L}$ is invariant under local $SU(3)$ gauge transformations and describes three equal-mass Dirac fields (the three colors of a given quark flavor) in interaction with eight massless vector fields (the gluons). It derives from the requirement that the

* Remember that a 'ninth gluon', coupling universally to all quarks, is excluded by experiment (see Problem 8.11).

global $SU(3)$ symmetry of the original Lagrangian (Equation 10.70) should hold *locally*. The Dirac fields constitute eight color *currents*

$$\mathbf{J}^{\mu} \equiv cq(\overline{\psi}\gamma^{\mu}\lambda\psi) \tag{10.89}$$

which act as *sources* for the color fields ($\mathbf{A}_{\mu}$), in the same way that the *electric* current acts as the source for *electromagnetic* fields. The theory described here is very close in structure to that of Yang and Mills; in this case, however, we believe it to be the correct description of a phenomenon realized in nature: the strong interaction. (Of course, we need six replicas of ψ, in Equation 10.88, each with the appropriate mass, to handle the six quark flavors.)

10.6 Feynman Rules

Up to this point, the Lagrangians we have considered might just as well describe *classical* fields as *quantum* ones; indeed, the Maxwell Lagrangian will be found in any textbook on classical electrodynamics. The passage from a classical field theory to the corresponding quantum field theory does not involve modification of the Lagrangian or the field equations, but rather a *reinterpretation* of the field variables; the fields are 'quantized,' and *particles* emerge as quanta of the associated fields. Thus, the photon is the quantum of the electrodynamic field, A^{μ}; leptons and quarks are quanta of Dirac fields; gluons are quanta of the eight $SU(3)$ gauge fields; and $W^{\pm}$ and Z^0 are quanta of the corresponding Proca fields. The quantization procedure itself is recondite, and this is not the place to go into it [8]; for our purposes the essential point is that *each Lagrangian prescribes a particular set of Feynman rules*. What we need, then, is a protocol for obtaining the Feynman rules dictated by a given Lagrangian.

To begin with, notice that $\mathscr{L}$ consists of two kinds of terms: the *free* Lagrangian for each participating field, plus various *interaction* terms ($\mathscr{L}_{\text{int}}$). The former – Klein–Gordon, for spin 0; Dirac, for spin $\frac{1}{2}$; Proca, for spin 1; or something more exotic, for a theory with higher spin – determines the *propagator*; the latter – obtained by invoking local gauge invariance, or by some other means – determine the *vertex factors*:

Free Lagrangian $\Rightarrow$ propagator
Interaction terms $\Rightarrow$ vertex factors

Let us consider the propagators first.

Application of the Euler–Lagrange equation to the free Lagrangian yields the free field equations (Eqs. 10.13, 10.15, and 10.22):

$$\left[\partial^{\mu}\partial_{\mu} + \left(\frac{mc}{\hbar}\right)^2\right]\phi = 0 \quad \text{(Klein–Gordon, for spin 0)}$$

$$\left[i\gamma^\mu\partial_\mu - \left(\frac{mc}{\hbar}\right)\right]\psi = 0 \quad \left(\text{Dirac, for spin } \frac{1}{2}\right)$$

$$\left[\partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) + \left(\frac{mc}{\hbar}\right)^2 A^\nu\right] = 0 \quad \text{(Proca, for spin 1)}$$

The corresponding 'momentum–space' equations are obtained by the standard prescription $p_\mu \leftrightarrow i\hbar\partial_\mu$:

$$[p^2 - (mc)^2]\phi = 0 \quad \text{(spin 0)} \tag{10.90}$$

$$[\not{p} - (mc)]\psi = 0 \quad \left(\text{spin } \frac{1}{2}\right) \tag{10.91}$$

$$[(-p^2 + (mc)^2)g_{\mu\nu} + p_\mu p_\nu]A^\nu = 0 \quad \text{(spin 1)} \tag{10.92}$$

The propagator is simply (i times) the *inverse* of the factor in square brackets:

$$\text{Spin-0 propagator: } \frac{i}{p^2 - (mc)^2} \tag{10.93}$$

$$\text{Spin-}\tfrac{1}{2}\text{ propagator: } \frac{i}{\not{p} - mc} = i\frac{(\not{p} + mc)}{p^2 - (mc)^2} \tag{10.94}$$

$$\text{Spin-1 propagator: } \frac{-i}{p^2 - (mc)^2}\left[g_{\mu\nu} - \frac{p_\mu p_\nu}{(mc)^2}\right] \tag{10.95}$$

Note that in the second case this factor is a 4×4 *matrix* and we want the *matrix* inverse; in the third case the factor is a second-rank tensor ($T_{\mu\nu}$) and we want the *tensor* inverse $(T^{-1})_{\mu\nu}$, such that $T_{\mu\lambda}(T^{-1})^{\lambda\nu} = \delta_\mu^\nu$ (Problem 10.19). These are precisely the propagators we used in Chapters 6, 7, and 9.* Since we obviously cannot set $m \to 0$ in the Proca propagator (Equation 10.95), we must go back to the free field equation (Equation 10.22) to work out the photon propagator:

$$\partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) = 0 \quad \text{(massless spin 1)} \tag{10.96}$$

As I have remarked before, this equation does not uniquely determine A^μ; if we impose the Lorentz condition (Equation 7.82)

$$\partial_\mu A^\mu = 0$$

then (Equation 10.96) reduces to

$$\partial^2 A^\nu = 0 \tag{10.97}$$

* Actually, this procedure only determines the propagator up to a multiplicative constant, since the field equations can always be multiplied by such a factor. In the 'canonical' form of these equations, the coefficient of mc or $(mc)^2$ is taken to be ± 1, with the sign matching that of the mass term in $\mathscr{L}$. Other conventions lead to a slightly different set of Feynman rules, but do not, of course, change the calculated reaction amplitudes.

which, in momentum space, can be written as

$$(-p^2 g_{\mu\nu})A^\nu = 0 \tag{10.98}$$

So the photon propagator is

$$\text{Massless spin-1 propagator: } -i\frac{g_{\mu\nu}}{p^2} \tag{10.99}$$

To get the vertex factors, first write down $i\,\mathscr{L}_{int}$ in momentum space ($i\hbar\partial_\mu \to p_\mu$) and examine the *fields* involved; these determine the qualitative structure of the interaction. For example, in the case of the QED Lagrangian (Equation 10.35)

$$i\mathscr{L}_{\text{int}} = -i(q\overline{\psi}\gamma^\mu\psi)A_\mu \tag{10.100}$$

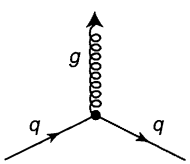
there are *three* fields involved ($\overline{\psi}$, ψ, and A_μ) and this defines a vertex in which three lines are joined – an incoming fermion, an outgoing fermion, and a photon. To obtain the vertex factor itself, *simply rub out the field variables*:

$$-i\sqrt{\frac{4\pi}{\hbar c}}q\gamma^\mu = ig_e\gamma^\mu \qquad \text{(QED vertex factor)} \tag{10.101}$$

(In the case of the photon, what we actually rub out is $\sqrt{\hbar c/4\pi}A_\mu$; the extra factor is due to our use of cgs units which are, for this purpose, a little cumbersome.) The same goes for chromodynamics (Equation 10.88): the quark–gluon coupling

$$\mathscr{L}_{\text{int}} = -(q\overline{\psi}\gamma^\mu\boldsymbol{\lambda}\psi)\cdot\mathbf{A}_\mu \tag{10.102}$$

yields a vertex of the form



with the vertex factor

$$-i\frac{g_s}{2}\gamma^\mu\boldsymbol{\lambda} \tag{10.103}$$

(The strong coupling constant is traditionally defined with a factor of 2: $g_s \equiv 2\sqrt{4\pi/\hbar c}q$, where q is the 'strong charge' appearing in the Lagrangian). However, there are also direct gluon–gluon couplings, coming from the $\mathbf{F}^{\mu\nu}\cdot\mathbf{F}_{\mu\nu}$ term in $\mathscr{L}$,

since $\mathbf{F}^{\mu\nu}$ contains not only the 'free' part, $\partial^\mu \mathbf{A}^\nu - \partial^\nu \mathbf{A}^\mu$, but also an interaction term, $-2q/\hbar c(\mathbf{A}^\mu \times \mathbf{A}^\nu)$ (Equation 10.87). Squaring it out, we find

$$\begin{aligned}\mathscr{L}_{\text{int}} &= \left(\frac{q}{8\pi\hbar c}\right)[(\partial^\mu \mathbf{A}^\nu - \partial^\nu \mathbf{A}^\mu)\cdot(\mathbf{A}_\mu \times \mathbf{A}_\nu) + (\mathbf{A}^\mu \times \mathbf{A}^\nu)\cdot(\partial_\mu \mathbf{A}_\nu - \partial_\nu \mathbf{A}_\mu)] \\ &\quad - \frac{q^2}{4\pi(\hbar c)^2}(\mathbf{A}^\mu \times \mathbf{A}^\nu)\cdot(\mathbf{A}_\mu \times \mathbf{A}_\nu) \qquad (10.104)\end{aligned}$$

The first term carries three factors of $\mathbf{A}^\mu$, and leads to the three-gluon vertex (Equation 8.42); the second term carries *four* factors of $\mathbf{A}^\mu$, and gives the four-gluon vertex (Equation 8.43). (For practice in extracting Feynman rules from Lagrangians, see Problems 10.20 and 10.21.)

10.7 The Mass Term

The principle of local gauge invariance works beautifully for the strong and electromagnetic interactions. In the first place, it gives us a *machine* for determining the couplings (in the 'old days' the construction of $\mathscr{L}_{int}$ was a purely *ad hoc* guess). Moreover, as 't Hooft and others proved in the early 1970s, [9] gauge theories are renormalizable. But the application to *weak* interactions was stymied by the fact that gauge fields have to be massless. Remember, the mass term in the Proca Lagrangian is not locally gauge invariant, and while the photon and the gluons *are* massless, the W's and the Z^0 certainly are *not*. So the question arises: can we doctor up gauge theory in such a way as to accommodate massive gauge fields? The answer is *yes*, but the procedure – exploiting spontaneous symmetry-breaking and the Higgs mechanism – is diabolically subtle, and it pays to begin by thinking very carefully about how one identifies the mass term in a Lagrangian.

Suppose, for instance, you were given the following Lagrangian for a scalar field ϕ:

$$\mathscr{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) + e^{-(\alpha\phi)^2} \qquad (10.105)$$

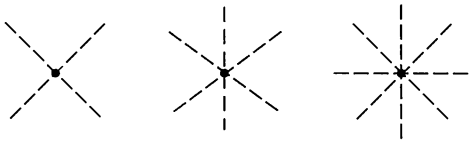
where α is some (real) constant. Where is the mass term here? At first glance there's no sign of one, and you might conclude that this is a *massless* field. But that is incorrect, for if you expand the exponential, $\mathscr{L}$ takes the form

$$\mathscr{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) + 1 - \alpha^2\phi^2 + \frac{1}{2}\alpha^4\phi^4 - \frac{1}{6}\alpha^6\phi^6 + \cdots \qquad (10.106)$$

The 1 is irrelevant (a constant term in $\mathscr{L}$ has no effect on the field equations), but the second term looks just like the mass term in the Klein–Gordon Lagrangian (Equation 10.11), with $\alpha^2 = \frac{1}{2}(mc/\hbar)^2$. Evidently, this Lagrangian describes a particle of mass

$$m = \sqrt{2}\alpha\hbar/c \qquad (10.107)$$

The higher-order terms represent couplings, of the form



and so on. This is not supposed to be a realistic theory, of course – I offer it only as an example of how the mass term in a Lagrangian may be 'disguised'. To expose it, we expand $\mathscr{L}$ in powers of ϕ and pick out the term proportional to ϕ^2 (in general, it's the term of *second order* in the fields – ϕ, ψ, A^μ, or whatever).

But there is a deeper subtlety lurking here, which I illustrate with the following Lagrangian:

$$\mathscr{L} = \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) + \frac{1}{2}\mu^2\phi^2 - \frac{1}{4}\lambda^2\phi^4 \tag{10.108}$$

Here μ and λ are (real) constants. The second term looks like a mass (and the third like an interaction). But wait! The *sign* is wrong (compare Equation 10.11) – if that's a *mass* term, then m is *imaginary*, which is nonsense. How, then, *should* we interpret this Lagrangian?* To answer this question, we must understand that the Feynman calculus is really a perturbation procedure, in which we start from the ground state (the 'vacuum') and treat the fields as fluctuations about that state. For the Lagrangians we have considered so far, the ground state – the field configuration of minimum energy – has always been the trivial one: $\phi = 0$. But for the Lagrangian in Equation 10.108, $\phi = 0$ is *not* the ground state. To determine the *true* ground state, we write $\mathscr{L}$ as a 'kinetic' term ($\frac{1}{2}\partial_\mu\phi\partial^\mu\phi$) minus a 'potential' term (inspired by the classical Lagrangian, Equation 10.4):

$$\mathscr{L} = \mathscr{T} - \mathscr{U} \tag{10.109}$$

and look for the minimum of $\mathscr{U}$. In the present case,

$$\mathscr{U}(\phi) = -\tfrac{1}{2}\mu^2\phi^2 + \tfrac{1}{4}\lambda^2\phi^4 \tag{10.110}$$

* I like to imagine that God has a giant computer–controlled factory, which takes Lagrangians as input and delivers the universes they represent as output. Usually God's computer has no difficulty – when you feed in the Maxwell Lagrangian, Equation 10.35, for example, it immediately creates an electromagnetic universe of interacting electrons, positrons, and photons. Sometimes it takes a little longer – the Lagrangian in Equation 10.105, for instance, confuses it at first, until it deciphers the 'hidden' mass term. And occasionally it returns an error message: 'this Lagrangian does not describe a possible universe; please check for syntax errors or incorrect signs'. That's what it would do, for example, if you fed it the Lagrangian in Equation 10.108 without the λ term.

and the minimum occurs at

$$\phi = \pm\mu/\lambda \tag{10.111}$$

(see Figure 10.1). The Feynman calculus must be formulated in terms of deviations from one or the other of these ground states. This suggests that we introduce a *new* field variable, η, defined by

$$\eta \equiv \phi \pm \frac{\mu}{\lambda} \tag{10.112}$$

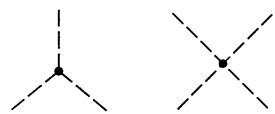
In terms of η, the Lagrangian reads

$$\mathscr{L} = \tfrac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - \mu^2\eta^2 \pm \mu\lambda\eta^3 - \tfrac{1}{4}\lambda^2\eta^4 + \tfrac{1}{4}(\mu^2/\lambda)^2 \tag{10.113}$$

The second quantity is now a mass term with the *correct* sign, and we discover (comparing Equation 10.11) that the mass of the particle is

$$m = \sqrt{2}\mu\hbar/c \tag{10.114}$$

Meanwhile, the third and fourth terms represent couplings of the form



(the last term is a constant, signifying nothing).

I emphasize that these Lagrangians (Equations 10.108 and 10.113) represent *exactly the same physical system*; all we have done is to change the *notation* (Equation 10.112). But the first version is not suited to the Feynman calculus (technically, a perturbation series in ϕ would not converge, because it is an expansion about an unstable point); only in the second formulation can we read off the mass and the vertex factors.

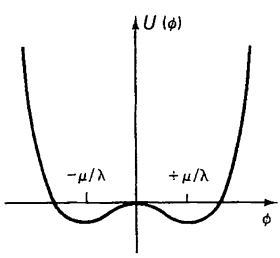


Fig. 10.1 Graph of $\mathscr{U}(\phi)$ (Equation 10.110).

Conclusion: To identify the mass term in a Lagrangian, we first locate the ground state (the field configuration for which $\mathscr{U}(\phi)$ is a minimum) and re-express $\mathscr{L}$ as a function of the deviation, η, from this minimum. Expanding in powers of η, we obtain the mass from the coefficient of the η^2 term.

10.8 Spontaneous Symmetry-breaking

The example we have just considered illustrates another phenomenon of importance: *spontaneous symmetry-breaking*. The original Lagrangian (Equation 10.108) is *even* in ϕ: it is invariant as $\phi \to -\phi$. But the reformulated Lagrangian (Equation 10.113) is *not* even in η; the symmetry has been 'broken'. How did this happen? It happened because the 'vacuum' (whichever of the two ground states we care to work with) does not share the symmetry of the Lagrangian. (The collection of *all* ground states, of course, *does*, but to set up the Feynman formalism we are obliged to work with one or the other of them, and that spoils the symmetry.) We call this 'spontaneous' symmetry-breaking because no *external* agency is responsible (as occurs, for example, when gravity breaks the three-dimensional symmetry in this room, making 'up' and 'down' quite different from 'left' and 'right'). To put it the other way around, the true symmetry of the system is 'concealed' by the arbitrary selection of a particular (asymmetrical) ground state. There are examples of spontaneous symmetry-breaking in many branches of physics. Take, for instance, a thin plastic strip (say, a short ruler): if you squeeze the ends together, it will snap into a curved configuration, but it can just as well buckle to the left as to the right – both are ground states for the system, and either one breaks the left–right symmetry (Figure 10.2).

But the spontaneously broken symmetry we have just considered was a *discrete* symmetry, with just *two* ground states. More interesting things happen when we consider *continuous* symmetries. (Replace the plastic *strip* in Figure 10.2 with a

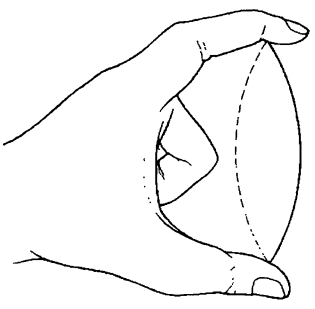


Fig. 10.2 Spontaneous symmetry-breaking in a plastic strip.

plastic *rod* – say, a knitting needle. Then it can buckle in *any* direction, not just left or right.*) It is easy to construct a Lagrangian with spontaneously broken continuous symmetry. For example,

$$\mathscr{L} = \tfrac{1}{2}(\partial_\mu \phi_1)(\partial^\mu \phi_1) + \tfrac{1}{2}(\partial_\mu \phi_2)(\partial^\mu \phi_2) + \tfrac{1}{2}\mu^2(\phi_1^2 + \phi_2^2) - \tfrac{1}{4}\lambda^2(\phi_1^2 + \phi_2^2)^2 \tag{10.115}$$

This is identical to Equation 10.108, except that now there are *two* fields, ϕ_1 and ϕ_2, and because $\mathscr{L}$ involves only the sum of the squares, it is invariant under *rotations* in ϕ_1, ϕ_2 space.†

This time the 'potential energy' function is

$$\mathscr{U} = -\tfrac{1}{2}\mu^2(\phi_1^2 + \phi_2^2) + \tfrac{1}{4}\lambda^2(\phi_1^2 + \phi_2^2)^2 \tag{10.116}$$

and the minima lie on a *circle* of radius μ/λ:

$$\phi_{1\text{min}}^2 + \phi_{2\text{min}}^2 = \mu^2/\lambda^2 \tag{10.117}$$

(Figure 10.3). To apply the Feynman calculus, we have to expand about a *particular* ground state ('the vacuum') – we may as well pick

$$\phi_{1\text{min}} = \mu/\lambda; \quad \phi_{2\text{min}} = 0 \tag{10.118}$$

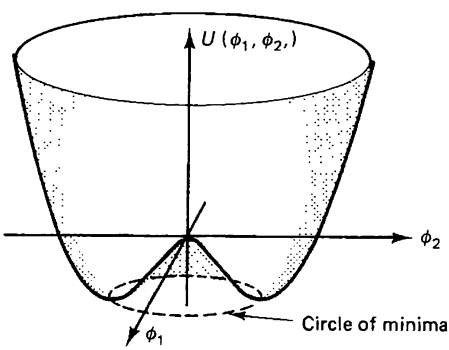


Fig. 10.3 The potential function (Equation 10.116).

* A more sophisticated example is the ferromagnet: in the ground state all the electron spins are aligned, but the *direction* of alignment is an accident of history. The *theory* is symmetrical, but a given piece of iron has to pick a particular orientation, and that ('spontaneously') breaks the symmetry.

† Group theoretically, it is invariant under $SO(2)$: $\phi_1 \to \phi_1 \cos\theta + \phi_2 \sin\theta$; $\phi_2 \to -\phi_1 \sin\theta + \phi_2 \cos\theta$, for any 'rotation angle' θ (see Problem 4.6).

As before, we introduce new fields, η and ξ, which are the fluctuations about this vacuum state:

$$\eta \equiv \phi_1 - \mu/\lambda; \quad \xi \equiv \phi_2 \tag{10.119}$$

Rewriting the Lagrangian in terms of these new field variables, we find (Problem 10.22):

$$\begin{aligned}\mathscr{L} &= \left[\frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \mu^2\eta^2\right] + \left[\frac{1}{2}(\partial_\mu \xi)(\partial^\mu \xi)\right] \\ &\quad - \left[\mu\lambda(\eta^3 + \eta\xi^2) + \frac{\lambda^2}{4}(\eta^4 + \xi^4 + 2\eta^2\xi^2)\right] + \frac{\mu^4}{4\lambda^2}\end{aligned} \tag{10.120}$$

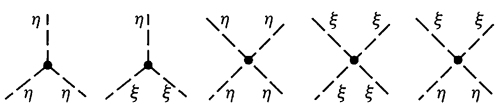
The first term is a free Klein–Gordon Lagrangian (Equation 10.11) for the field η, which evidently carries a mass

$$m_\eta = \sqrt{2}\,\mu\hbar/c \tag{10.121}$$

(same as before, Equation 10.114); the second term is a free Lagrangian for the field ξ, which is evidently *massless*:

$$m_\xi = 0 \tag{10.122}$$

and the third term defines five couplings:



(the final constant, of course, is irrelevant). In this form the Lagrangian doesn't look symmetrical at all; the symmetry of Equation 10.115 has been broken (or rather, 'hidden') by the selection of a particular vacuum state.

The important thing to notice here is that one of the fields (ξ) is *automatically massless*. This is no accident. It can be shown (Goldstone's theorem [10]) that spontaneous breaking of a *continuous* global symmetry is *always* accompanied by the appearance of one or more massless scalar (spin-0) particles (we call them 'Goldstone bosons').* Well, this is a disaster; we were hoping to use the mechanism of spontaneous symmetry-breaking to account for the mass of the weak interaction gauge fields, but now we find that this introduces a massless scalar boson, and there is no such thing on the roster of known elementary particles.† But hold on, for there is one final incredible twist in the story. It comes

* Intuitively, this is related to the fact that there is no resistance to excitations in the ξ direction. Flick the bent knitting needle and it will spin freely about the axis, whereas *radial* excitations encounter a restoring force, and the system oscillates.

† It is hard to imagine that such a particle could have escaped detection. With *heavy* particles, this is always a possibility – maybe you just didn't have enough energy to *produce* it – but a *massless* particle would surely have shown up somewhere, if only in the form of 'missing' energy and momentum.

when we apply the idea of spontaneous symmetry-breaking to the case of local gauge invariance.

10.9 The Higgs Mechanism

The Lagrangian we studied in Section 10.8 can be written more neatly if we combine the two real fields, ϕ_1 and ϕ_2, into a single complex field:

$$\phi \equiv \phi_1 + i\phi_2 \tag{10.123}$$

so that

$$\phi^*\phi = \phi_1^2 + \phi_2^2 \tag{10.124}$$

In this notation (and it *is* nothing but notation), the Lagrangian (Equation 10.115) reads

$$\mathscr{L} = \tfrac{1}{2}(\partial_\mu\phi)^*(\partial^\mu\phi) + \tfrac{1}{2}\mu^2(\phi^*\phi) - \tfrac{1}{4}\lambda^2(\phi^*\phi)^2 \tag{10.125}$$

and the rotational $SO(2)$ symmetry that was spontaneously broken becomes invariance under $U(1)$ phase transformations:

$$\phi \to e^{i\theta}\phi \tag{10.126}$$

This is precisely the kind of symmetry we considered back in Section 10.3, except that now we are working with scalar fields instead of with spinors. We can make the system invariant under *local* gauge transformations

$$\phi \to e^{i\theta(x)}\phi \tag{10.127}$$

by the usual device of introducing a massless gauge field A^μ, and replacing the derivatives in Equation 10.125 with *covariant* derivatives (Equation 10.38):

$$\mathscr{D}_\mu = \partial_\mu + i\frac{q}{\hbar c}A_\mu \tag{10.128}$$

Thus

$$\begin{aligned}\mathscr{L} &= \frac{1}{2}\left[\left(\partial_\mu - \frac{iq}{\hbar c}A_\mu\right)\phi^*\right]\left[\left(\partial^\mu + \frac{iq}{\hbar c}A^\mu\right)\phi\right] \\ &\quad + \frac{1}{2}\mu^2(\phi^*\phi) - \frac{1}{4}\lambda^2(\phi^*\phi)^2 - \frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu}\end{aligned} \tag{10.129}$$

Now we simply retrace our steps in Section 10.8, applying them to the *locally* invariant Lagrangian (Equation 10.129). Defining the new fields

$$\eta \equiv \phi_1 - \mu/\lambda, \quad \xi \equiv \phi_2 \tag{10.130}$$

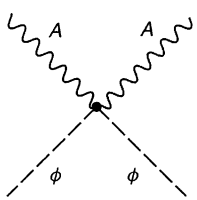
(compare Equation 10.119), the Lagrangian becomes (Problem 10.25):

$$\begin{aligned}\mathscr{L} &= \left[\frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \mu^2\eta^2\right] + \left[\frac{1}{2}(\partial_\mu \xi)(\partial^\mu \xi)\right] \\ &+ \left[-\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}\left(\frac{q}{\hbar c}\frac{\mu}{\lambda}\right)^2 A_\mu A^\mu\right] \\ &+ \left\{\frac{q}{\hbar c}[\eta(\partial_\mu \xi) - \xi(\partial_\mu \eta)]A^\mu + \frac{\mu}{\lambda}\left(\frac{q}{\hbar c}\right)^2 \eta(A_\mu A^\mu)\right. \\ &\left. + \frac{1}{2}\left(\frac{q}{\hbar c}\right)^2(\xi^2+\eta^2)(A_\mu A^\mu) - \lambda\mu(\eta^3+\eta\xi^2) - \frac{1}{4}\lambda^2(\eta^4+2\eta^2\xi^2+\xi^4)\right\} \\ &+ \left(\frac{\mu}{\lambda}\frac{q}{\hbar c}\right)(\partial_\mu \xi)A^\mu + \left(\frac{\mu^2}{2\lambda}\right)^2\end{aligned} \tag{10.131}$$

The first line is the same as before (Equation 10.120); it represents a scalar particle (η), of mass $\sqrt{2}\mu\hbar/c$, and a massless Goldstone boson (ξ). The second line describes the free gauge field A^μ, but – *mirabile dictu!* – it has acquired a mass:

$$m_A = 2\sqrt{\pi}\left(\frac{q\mu}{\lambda c^2}\right) \tag{10.132}$$

(compare the Proca Lagrangian, Equation 10.121). The term in curly brackets specifies various couplings of ξ, η, and A^μ (Problem 10.26). It is interesting to see where the mass of A^μ came from: the original Lagrangian (Equation 10.129) contains a term of the form $\phi^*\phi\ A_\mu A^\mu$, which – absent spontaneous symmetry-breaking – would represent a *coupling*:

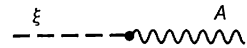


But when the ground state moves 'off center', and the field ϕ_1 picks up a constant (Equation 10.130), this piece of the Lagrangian emerges as a Proca mass term.

However, we still have that unwanted Goldstone boson (ξ). Moreover, there is a suspicious-looking quantity in $\mathscr{L}$:

$$\left(\frac{\mu}{\lambda}\frac{q}{\hbar c}\right)(\partial_\mu \xi)A^\mu \tag{10.133}$$

What are we to make of *this*? If we read it as an *interaction*, it leads to a vertex of the form



in which the ξ turns into an A. Any such term, bilinear in two different fields, indicates that we have incorrectly identified the fundamental particles in the theory (see Problem 10.23). Both difficulties involve the field $\xi = \phi_2$, and both can be resolved exploiting the local gauge invariance of $\mathscr{L}$ (in the original form, Equation 10.129) to *transform this field away entirely!* Writing Equation 10.126 in terms of its real and imaginary parts,

$$\begin{aligned}\phi \to \phi' &= (\cos\theta + i\sin\theta)(\phi_1 + i\phi_2) \\ &= (\phi_1\cos\theta - \phi_2\sin\theta) + i(\phi_1\sin\theta + \phi_2\cos\theta)\end{aligned} \tag{10.134}$$

we see that picking

$$\theta = -\tan^{-1}(\phi_2/\phi_1) \tag{10.135}$$

will render ϕ' real, which is to say that $\phi'_2 = 0$. The gauge field A^μ will transform accordingly (Equation 10.34), but the Lagrangian will take the same form in terms of the new field variables as it did in terms of the old ones (that's what it *means* to say that $\mathscr{L}$ is invariant). The only difference is that ξ is now zero. *In this particular gauge*, then, the Lagrangian (Equation 10.131) reduces to

$$\begin{aligned}\mathscr{L} = &\left[\frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - \mu^2\eta^2\right] + \left[-\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}\left(\frac{q}{\hbar c}\frac{\mu}{\lambda}\right)^2 A_\mu A^\mu\right] \\ &+ \left\{\frac{\mu}{\lambda}\left(\frac{q}{\hbar c}\right)^2\eta(A_\mu A^\mu) + \frac{1}{2}\left(\frac{q}{\hbar c}\right)^2\eta^2(A_\mu A^\mu) - \lambda\mu\eta^3 - \frac{1}{4}\lambda^2\eta^4\right\} \\ &+ \left(\frac{\mu^2}{2\lambda}\right)^2\end{aligned} \tag{10.136}$$

By an astute choice of gauge, we have eliminated the Goldstone boson and the offending term in $\mathscr{L}$; we are left with a single *massive* scalar η (the 'Higgs' particle) and a *massive* gauge field A^μ.

Please understand: The Lagrangians in Equations 10.129 and 10.136 describe *exactly the same physical system*; all we have done is to select a convenient gauge (Equation 10.135) and rewrite the fields in terms of fluctuations about a particular ground state (Equation 10.130). We have sacrificed manifest symmetry in favor of notation that makes the physical content more transparent, and allows us to extract the Feynman rules more directly. But it's still the same Lagrangian. There is an illuminating way to think about this: a *massless* vector field carries *two* degrees of freedom (transverse polarizations); when A^μ acquires mass, it picks up a third (longitudinal polarization). Where did this extra degree of freedom come from? *Answer:* It came from the Goldstone boson, which meanwhile disappeared from

the theory. The gauge field 'ate' the Goldstone boson, thereby acquiring both a mass and a third polarization state.* This is the famous *Higgs mechanism*, the remarkable offspring of the marriage of local gauge invariance and spontaneous symmetry-breaking [11].

According to the Standard Model, the Higgs mechanism is responsible for the masses of the weak interaction gauge bosons ($W^{\pm}$ and Z^0). The details are still matters of speculation – the Higgs particle has never been seen in the laboratory (presumably it is just too heavy to make with any existing accelerator), and the Higgs 'potential', $\mathscr{U}(\phi)$, is unknown (I used $\mathscr{U} = -\frac{1}{2}\mu^2(\phi^*\phi) + \frac{1}{4}\lambda^2(\phi^*\phi)^2$ just for the sake of argument).† There may in fact be many Higgs particles, or it may be a composite structure, but never mind: the important thing is that we have found a way *in principle* of imparting mass to the gauge fields,‡ and that is our license to believe that all the fundamental interactions – weak as well as strong and electromagnetic – can be described by local gauge theories [12].

References

1 For an introduction to the Lagrangian formulation of classical mechanics, see Taylor, J. R. (2005) *Classical Mechanics*, University Science Books, Sausalito, C.A. or; (a) Marion, J. B. and Thornton, S. T. (1995) *Classical Dynamics of Particles and Systems*, 4th edn, Harcourt, Orlando, F.L.

2 The Euler–Lagrange equation can be obtained, in turn, from the Principle of Least Action. See Poole, C. P., Sapko, J. L. and Goldstein, H. (2002) *Classical Mechanics*, 3rd edn, Addison-Wesley, Reading, M.A. or; (a) Lanczos, C. (1986) *The Variational Principles of Mechanics*, Dover, New York.

3 Weyl, H. (1919) *Annals of Physics*, 59, 101. See also; (a) Al-Kuwari, H. and Taha, M. O. (1991) *American Journal of Physics*, 59, 4. or; (b) Moriyasu, K. (1983) *An Elementary Primer for Gauge Theory*, World Scientific, Singapore.

4 Yang, C. N. and Mills, R. L. (1954) *Physics Review*, 96, 191. For historical commentary see; (a) Mills, R. (1989) *American Journal of Physics*, 57, 493. or; (b) 't Hooft, G. (2005) *Fifty Years of Yang–Mills Theory*, Singapore: World Scientific.

5 See, for example, Chevalley, C. (1946) *Theory of Lie Groups*, Princeton University Press, Princeton, N.J.

* We don't *have* to adopt any particular gauge. However, if we do not, the theory will contain a nonphysical 'ghost' particle, and it is simplest to eliminate it explicitly from the start.

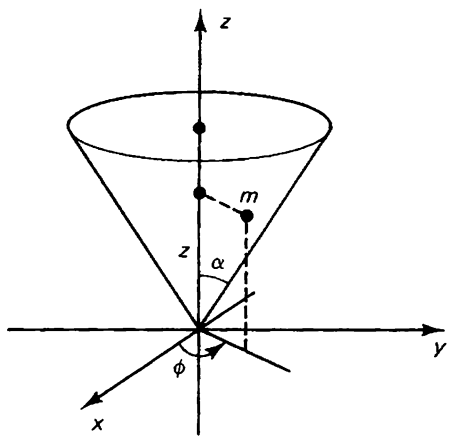
† Actually, for the theory to be renormalizable the potential has to be quartic in the fields.

‡ In the Standard Model, the Higgs particle is *also* responsible for the masses of quarks and leptons; they are initially taken to be massless, but are assumed to have Yukawa couplings (see Problem 10.21) to the Higgs particle. When the latter is 'shifted' by spontaneous symmetry-breaking (Equation 10.130), the Yukawa coupling splits into two parts, one of which is a true interaction and the other a *mass* term for the field ψ. This is a nice *idea*, but it does not help us to *calculate* the fermion masses, since the Yukawa coupling constants themselves are unknown. Only when (and *if*) the Higgs particle is actually *found* will it be possible to confirm all this empirically. See Chapter 12.

6 Local gauge theories are treated with extraordinary clarity and beauty in the unpublished research notes Wheeler, N. A. (1981) *'Classical Chromodynamics' and 'Bare Bones of the Classical Theory of Gauge Fields'*, Reed College, Portland, O.R.
7 Actor, A. (1979) *Reviews of Modern Physics*, **51**, 461.
8 The details will be found in any treatise on quantum field theory, such as those cited in the Introduction.
9 't Hooft, G. (1971) *Nuclear Physics*, **B33**, 173; (1971) **B35**, 167. Reprinted in (a) Lai, C. H. (ed) (1981)*Gauge Theory of Weak and Electromagnetic Interactions*, World Scientific, Singapore.
10 Goldstone, J. (1961) *Nuovo Cimento*, **19**, 154. Goldstone, J., Salam, A. and Weinberg, S. (1962) *Physics Review*, **127**, 965.
11 Higgs, P. W. (1964) *Physics Letters*, **12**, 132; Englert, F. and Brout, R. (1964) *Physical Review Letters*, **13**, 321.
12 For detailed application of the Higgs mechanism in the electroweak theory of Glashow, Weinberg, and Salam, see, for instance, Chapter 15 of Halzen, F. and Martin, A. D. 1984) *Quarks and Leptons* John Wiley & Sons, Ltd, New York.

Problems

10.1 One advantage of the Lagrangian formulation is that it does not commit us to any particular coordinate system – the q's in Equation 10.6 could be Cartesian coordinates, or polar coordinates, or any other variables we might use to designate the particle's position. Suppose, for example, we want to analyze the motion of a particle that slides frictionlessly on the inside surface of a cone mounted with its axis pointing upward, as shown.



(a) Express T and U in terms of the variables z and ϕ and the constants α (the opening angle of the cone), m (the mass of the particle), and g (the acceleration of gravity).
(b) Construct the Lagrangian, and apply the Euler–Lagrange equations to obtain differential equations for $z(t)$ and ϕ (t).
(c) Show that $L = (m \tan^2\alpha)z^2\phi$ is a constant of the motion. What *is* this quantity, physically?

(d) Use the result in (c) to eliminate ϕ from the z equation. (You are left with a *second*-order differential equation for $z(t)$; if you want to pursue the problem further, it is easiest to invoke conservation of *energy*, which yields a *first*-order equation for z.)

10.2 Derive Equation 10.17

10.3 Starting with Equation 10.19, show that $\partial_\mu A^\mu = 0$, and hence that each component of A^μ satisfies the Klein–Gordon equation: $\Box A^\nu + (mc/\hbar)^2 A^\nu = 0$.

10.4 As it stands, the Dirac Lagrangian (Equation 10.14) treats ψ and $\overline{\psi}$ asymmetrically. Some people prefer to deal with them on an equal footing, using the modified Lagrangian

$$\mathscr{L} = \frac{i\hbar c}{2}\left[\overline{\psi}\gamma^\mu(\partial_\mu\psi) - (\partial_\mu\overline{\psi})\gamma^\mu\psi\right] - (mc^2)\overline{\psi}\psi$$

Apply the Euler–Lagrange equations to this $\mathscr{L}$, and show that you get the Dirac equation (Equation 10.15) and its adjoint.

10.5 The Klein–Gordon Lagrangian for a *complex* field would be

$$\mathscr{L} = \tfrac{1}{2}(\partial_\mu\phi)^*(\partial^\mu\phi) - \tfrac{1}{2}(mc/\hbar)^2\phi^*\phi$$

Treating ϕ and ϕ^* as independent field variables, deduce the field equations for each, and show that these field equations are consistent (i.e. one is the complex conjugate of the other).

10.6 Apply the Euler–Lagrange equations to Equation 10.33 to obtain the Dirac equation with electromagnetic coupling.

10.7 Show that the Dirac current (Equation 10.36) satisfies the continuity equation (Equation 10.25).

10.8 The complex Klein–Gordon Lagrangian (Problem 10.5) is invariant under the global gauge transformation $\phi \to e^{i\theta}\phi$. Impose local gauge invariance to construct the complete gauge-invariant Lagrangian, and determine the current density J^μ. Using the Euler-Lagrange equation for ϕ, show that this current obeys the continuity equation (Equation 10.25). [*Warning:* The current is defined by Equation 10.24, *not* by Equation 10.23. It is true that the former follows (ordinarily) from the latter, but not when J^μ depends explicitly on A^μ. In this (rare) circumstance you cannot just pick off the term in $\mathscr{L}$ that is proportional to A^μ; rather, you must use the Euler–Lagrange equations to determine $\partial_\mu F^{\mu\nu}$, and get the current from that.]

10.9 (a) Suppose the field variables (ϕ_i) are subjected to an infinitesimal global transformation $\delta\phi_i$. Show that the Lagrangian $\mathscr{L}(\phi_i, \partial_\mu\phi_i)$ changes by an amount

$$\delta\mathscr{L} = \partial_\mu\left\{\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_i)}\delta\phi_i\right\}$$

In particular, if the Lagrangian is *invariant* under the transformation in question, then $\delta\mathscr{L} = 0$, and the term in curly brackets constitutes a conserved current (that is, it obeys the continuity equation). More precisely, if the transformation $\delta\phi_i$ is specified by a parameter $\delta\theta$, the Noetherian current is

$$J^\mu = \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_i)}\frac{\delta\phi_i}{\delta\theta}$$

(up to an overall constant, chosen for convenience in the particular context). This is the essence of Noether's theorem [3], relating symmetries of the Lagrangian to conservation laws.

(b) Apply Noether's theorem to the Dirac Lagrangian (Equation 10.14), to construct the conserved current associated with global phase invariance (Equation 10.26). Compare the electric current (Equation 10.36).

(c) Do the same for the complex Klein–Gordon Lagrangian in Problem 10.8

10.10 Derive Equation 10.51

10.11 Deduce Equation 10.58 from Equation 10.57, using Equations 10.54–10.56.

10.12 Suppose we were to define

$$\mathbf{F}^{\mu\nu} \equiv \partial^\mu \mathbf{A}^\nu - \partial^\nu \mathbf{A}^\mu$$

in Yang–Mills theory.

(a) Find the transformation rule for this $\mathbf{F}^{\mu\nu}$, under infinitesimal gauge transformations (Equation 10.61).

(b) Determine the infinitesimal transformation rule for $\mathscr{L}_A$ (Equation 10.63), in this case. Is the Lagrangian invariant?

$$\left[\textit{Answers}: \text{(a) } \mathbf{F}^{\mu\nu} \to \mathbf{F}^{\mu\nu} + \frac{2g}{\hbar c}[\boldsymbol{\lambda} \times \mathbf{F}^{\mu\nu} + \mathbf{A}^\mu \times \partial^\nu \boldsymbol{\lambda} - \mathbf{A}^\nu \times \partial^\mu \boldsymbol{\lambda}]\right.$$

$$\left.\text{(b) } \mathbf{F}^{\mu\nu} \cdot \mathbf{F}_{\mu\nu} \to \mathbf{F}^{\mu\nu} \cdot \mathbf{F}_{\mu\nu} + \frac{8g}{\hbar c}(\mathbf{A}_\nu \times \mathbf{F}^{\mu\nu}) \cdot \partial_\mu \boldsymbol{\lambda}\right]$$

10.13 Derive Equation 10.66, starting with Equations 10.61 and 10.65.

10.14 Prove that gauge field Lagrangian (Equation 10.63) is invariant under finite local gauge transformations, as follows:

(a) Using Equations 10.58 and 10.65, show that

$$\boldsymbol{\tau} \cdot \mathbf{F}^{\mu\nu\prime} = S(\boldsymbol{\tau} \cdot \mathbf{F}^{\mu\nu})S^{-1}$$

[Note that $\partial_\mu(S^{-1}S) = 0 \Rightarrow (\partial_\mu S^{-1})S = -S^{-1}(\partial_\mu S)$.]

(b) Show, therefore, that

$$\mathrm{Tr}[(\boldsymbol{\tau} \cdot \mathbf{F}^{\mu\nu})(\boldsymbol{\tau} \cdot \mathbf{F}_{\mu\nu})]$$

is invariant.

(c) Using Problem 4.20(c), show that the trace in (b) is equal to $2\mathbf{F}^{\mu\nu} \cdot \mathbf{F}_{\mu\nu}$.

10.15 Apply the Euler–Lagrange equations to the Lagrangian in Equation 10.69. Using the standard associations (Equations 7.71, 7.72, and 7.79), obtain 'Maxwell's equations' for classical Yang–Mills theory. [Note that there are *three* charge densities, three current densities, three scalar potentials, three vector potentials, three 'electric' fields, and three 'magnetic' fields, in this theory.] (Unlike electrodynamics, your expressions for the divergence and curl of the **E**'s and **B**'s will inevitably involve the *potentials*.)

10.16 Show that any Hermitian 3×3 matrix can be written as a linear combination of the unit matrix and the eight Gell-Mann matrices (Equation 10.76)

10.17 (a) Show that $\det(e^A) = e^{\mathrm{Tr}(A)}$, for any matrix A. [*Hint*: Check it first for a *diagonal* matrix. Then extend the proof to any *diagonalizable* matrix ($S^{-1}AS = D$, where D is diagonal, for some matrix S) – show that $\mathrm{Tr}(A) = \mathrm{Tr}(D)$ and $S^{-1}e^A S = e^D$, so that $\det(e^A) = \det(e^D)$. Of course, not all matrices *are* diagonalizable; however, every matrix can be brought into *Jordan canonical form* ($S^{-1}AS = J$, where J is diagonal except for some 1's immediately below the main diagonal). Take it from there.]

(b) Show that $e^{i\boldsymbol{\lambda}\cdot\mathbf{a}}$ (in Equation 10.78) has determinant 1.

10.18 Starting with Equation 10.81, derive Equations 10.82 and 10.83.

10.19 Confirm that the Proca propagator (Equation 10.95) is the inverse of the tensor in Equation 10.92, in the sense explained in the text.

10.20 Construct the Lagrangian for ABC theory (Chapter 6).

10.21 Give a physical interpretation of the Yukawa Lagrangian:

$$\begin{aligned}\mathscr{L} &= [i\hbar c\overline{\psi}\gamma^\mu\partial_\mu\psi - m_1c^2\overline{\psi}\psi] \\ &\quad + \left[\frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \frac{1}{2}\left(\frac{m_2c}{\hbar}\right)^2\phi^2\right] - \alpha_Y\overline{\psi}\psi\phi \end{aligned} \tag{10.137}$$

What are the spins and masses of the particles? What are their propagators? Draw the Feynman diagram for their interaction and determine the vertex factor.

10.22 Derive Equation 10.120

10.23 Suppose we took

$$\psi_1 \equiv (\eta + \xi)/\sqrt{2} \quad \text{and} \quad \psi_2 \equiv (\eta - \xi)/\sqrt{2}$$

as the fundamental fields, instead of Equation 10.119 Express the Lagrangian (Equation 10.120) in terms of ψ_1 and ψ_2.

[*Comment*: Offhand, it looks as though we have two *massive* fields here, and thus escape Goldstone's theorem. Unfortunately, there is *also* a term of the form $-\mu^2\psi_1\psi_2$. If you interpret this as an *interaction*, it converts ψ_1 into ψ_2, and vice versa, but that means neither one exists as an independent free particle. Rather, such an expression should be interpreted as an off-diagonal term in the mass matrix (Equation 10.45), indicating that we have incorrectly identified the fundamental fields in the theory. The *physical* fields are those for which M is *diagonal* and for which no direct transitions from one to the other can occur. We have encountered this situation once before, in Section 4.4.3: we found that $K^0 \leftrightarrow \overline{K}^0$, and hence that these are not the physical particle states; instead, the linear combinations K_1 and K_2, in terms of which the mass matrix is diagonal, are the 'true' particles.]

10.24 Generalize the argument following Equation 10.115 to *three* fields (ϕ_1, ϕ_2, ϕ_3). What are the masses of the three particles? How many Goldstone bosons are there in this case?

10.25 Starting from Equations 10.129 and 10.130, derive Equation 10.131

10.26 Draw the primitive vertices for all the interactions in curly brackets in Equation 10.131 Circle the ones that survive in Equation 10.136

11

Neutrino Oscillations

Recent experiments have shown that neutrinos can convert from one flavor to another (for instance, $\nu_e \leftrightarrow \nu_\mu$). This means that neutrinos have nonzero mass, and that the lepton numbers (electron, muon, and tau) are not separately conserved. Neutrino oscillations resolve the solar neutrino problem, and suggest modest changes in the Standard Model. The treatment here is largely self-contained, and could even be read immediately after Chapter 2.

11.1
The Solar Neutrino Problem

The story begins [1] in the middle of the nineteenth century, when Lord Rayleigh undertook to calculate the age of the sun. He assumed (as everyone did, at the time) that the source of the sun's energy was gravity – the energy accumulated when all that matter 'fell down' from infinity is liberated over time in the form of radiation. On the basis of the known rate of solar radiation (which he took to be constant), Rayleigh showed that the maximum possible age of the sun was substantially shorter than the age of the earth as estimated by geologists, and, more to the point, shorter than Darwin's theory of evolution required. This pleased Lord Rayleigh, who was opposed to evolution on quaint religious grounds, but it worried Darwin, who removed his own estimate from subsequent editions of his book.

In 1896, Becquerel discovered radioactivity. In subsequent studies he and the Curies noted that radioactive substances such as radium give off prodigious amounts of heat. This suggested that nuclear fission, not gravity, might be the source of the sun's energy, and this would allow for a much longer lifetime. The only trouble was that there didn't appear to *be* any radioactive stuff in the sun, which is made almost entirely of hydrogen (plus a small amount of light elements, but certainly not uranium or radium).

By 1920, Aston completed a series of meticulous measurements of atomic weights, and Eddington noticed that four hydrogen atoms weigh slightly more than one atom of helium-4. This implied (in view of Einstein's $E = mc^2$) that the fusion of four hydrogens would be energetically favorable, and would release a substantial

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

amount of energy. Eddington suggested that this process (nuclear fusion) powers the sun, and in essence he was right. Of course, Eddington didn't know what the *mechanism* for binding the hydrogens together might be; this had to await the development of nuclear physics in the 1930s – in particular, Chadwick's discovery of the neutron and Pauli's invention of the neutrino.

In 1938, Hans Bethe worked out the details, which turn out to be quite complicated. In heavy stars the dominant mechanism is the CNO (Carbon–Nitrogen –Oxygen) cycle, in which the fusion process is 'catalyzed' by small amounts of those three elements. But in the sun (and other relatively light stars) the dominant route is the so-called *pp* chain (Figure 11.1). To begin with, a pair of protons (hydrogen nuclei) combine to make a deuteron, a positron, and a neutrino. (The deuteron is a proton and a neutron, so what really happened here is that a proton converted into a neutron, a positron, and a neutrino – the reverse of neutron decay.) Alternatively, the outgoing positron could be replaced by an incoming electron. Either way, we have produced deuterons (along with some neutrinos) from protons. The deuteron soon picks up another proton to make a helium-3 nucleus (two protons and a neutron), releasing energy in the form of a photon. Helium-3 has three options: it can join with another loose proton to make an alpha particle – the nucleus of helium-4 (two protons and two neutrons). Once again, a proton has been converted into a neutron (with emission of a positron and a neutrino). Or two helium-3s can get together to make an alpha particle and two leftover protons. Or the helium-3 can combine with an alpha particle (produced in one of the previous reactions) to make beryllium-7, with the emission of a photon. Finally, the beryllium can either

The pp Chain

Step 1: Two protons make a deuteron

$$\begin{aligned} p + p &\rightarrow d + e^+ + \nu_e \\ p + p + e^- &\rightarrow d + \nu_e \end{aligned}$$

Step 2: Deuteron plus proton makes ^{3}He.

$$d + p \rightarrow {}^3\mathrm{He} + \gamma$$

Step 3: Helium-3 makes alpha particle or ^{7}Be.

$$\begin{aligned} {}^3\mathrm{He} + p &\rightarrow \alpha + e^+ + \nu_e \\ {}^3\mathrm{He} + {}^3\mathrm{He} &\rightarrow \alpha + p + p \\ {}^3\mathrm{He} + \alpha &\rightarrow {}^7\mathrm{Be} + \gamma \end{aligned}$$

Step 4: Berillium makes alpha particles.

$$\begin{aligned} {}^7\mathrm{Be} + e^- &\rightarrow {}^7\mathrm{Li} + \nu_e \\ {}^7\mathrm{Li} + p &\rightarrow \alpha + \alpha \\ {}^7\mathrm{Be} + p &\rightarrow {}^8\mathrm{B} + \gamma \\ {}^8\mathrm{B} &\rightarrow {}^8\mathrm{Be}^* + e^+ + \nu_e \\ {}^8\mathrm{Be}^* &\rightarrow \alpha + \alpha \end{aligned}$$

Fig. 11.1 The *pp* chain: how protons make alpha particles in the sun.

absorb an electron, making lithium, which picks up a proton, yielding two alpha particles, or else it absorbs a proton, making boron, which goes to an excited state of beryllium-8, and from there to two alpha particles.

The details are not so important; the *point* is that it all starts out as hydrogen (protons), and it all ends up as α particles (helium-4 nuclei) – precisely Eddington's reaction – plus some electrons, positrons, photons ... and neutrinos. But is this complicated story really *true*? How can we tell what is going on inside the sun? Photons take athousand years to work their way out from the center to the surface, and what we see from earth doesn't tell us much about the interior. But neutrinos – because they interact so weakly, emerge virtually unscathed by passage through the sun. Neutrinos, therefore, are the perfect probes for studying the interior of the sun.

In the *pp* chain there are five reactions that yield neutrinos, and for each one the neutrinos come out with a characteristic energy spectrum, as shown in Figure 11.2 The overwhelming majority come from the initial reaction $p + p \rightarrow d + e + \nu_e$. Unfortunately, they carry relatively low energy, and most detectors are insensitive in this regime. For that reason, even though the boron-8 neutrinos are far less abundant, most experiments actually work with them.

There are certainly plenty of neutrinos coming from the sun. John Bahcall, who was responsible for most of the calculations of solar neutrino abundances, liked to say that 100 billion neutrinos pass through your thumbnail every second; and yet they are so ethereal that you can look forward to only one or two neutrino-induced

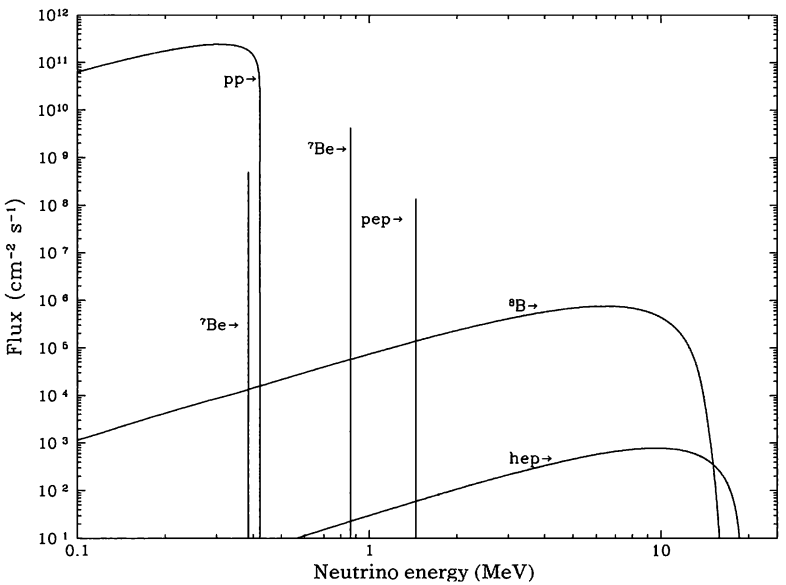


Fig. 11.2 The calculated energy spectra for solar neutrinos. (*Source*: J. N. Bahcall, A.M. Serenelli, and S. Basu, *Astrophysical Journal* **621**, L85 (2005).)

reactions in your body during your entire lifetime. In 1968, Ray Davis *et al.* [2] reported the first experiments to measure solar neutrinos, using a huge tank of chlorine (actually, cleaning fluid) in the Homestake mine in South Dakota (you have to do it deep underground to eliminate background from cosmic rays). Chlorine can absorb a neutrino and convert to argon by the reaction $\nu_e + {}^{37}\mathrm{Cl} \to {}^{37}\mathrm{Ar} + e$ (essentially again $\nu_e + n \to p + e$). The Davis experiment – for which he was finally awarded Nobel Prize in 2002 – collected argon atoms for several months (they were produced at a rate of about one atom every two *days*). The total accumulation was only about a third of what Bahcall predicted [3]. Thus was born the famous *solar neutrino problem.*

11.2 Oscillations

At the time, most physicists assumed the experiments were wrong. After all, Davis claimed to have flushed a total of 33 argon atoms out of a tank containing 615 metric tons of tetrachloroethylene – it was not hard to imagine that he might have missed a few. On the theory side, Bahcall's calculations required an audacious confidence in the so-called Standard Solar Model of the interior of the sun. But gradually the community came to take the solar neutrino problem seriously – especially when other experiments, using quite different detection methods, confirmed the deficit.

In 1968, Bruno Pontecorvo suggested a beautifully simple explanation for the solar neutrino problem. He proposed that the electron neutrinos produced by the sun are transformed in flight into a different species (muon neutrinos, say, or even antineutrinos), to which Davis' experiment was insensitive [4]. This is the mechanism we now call *neutrino oscillation.* The theory is quite simple – it is basically the quantum mechanics of mixed states, which itself is almost identical to the classical theory of coupled oscillators [5]. Consider the case of just *two* neutrino types – say, ν_e and ν_μ. If one can spontaneously convert into the other, it means that neither is an eigenfunction of the Hamiltonian. The *true* stationary states for the system are evidently some orthogonal linear combinations:

$$\nu_1 = \cos\theta\, \nu_\mu - \sin\theta\, \nu_e; \quad \nu_2 = \sin\theta\, \nu_\mu + \cos\theta\, \nu_e \tag{11.1}$$

(Writing the coefficients as sines and cosines is just a cute way of enforcing normalization.)

According to the Schrödinger equation, these eigenstates have the simple time dependence $e^{-iE_1 t/\hbar}$:

$$\nu_1(t) = \nu_1(0)e^{-iE_1 t/\hbar}; \quad \nu_2(t) = \nu_2(0)e^{-iE_2 t/\hbar} \tag{11.2}$$

Suppose, for example, that the particle starts out as an electron neutrino:

$$\nu_e(0) = 1,\ \nu_\mu(0) = 0, \quad \text{so} \quad \nu_1(0) = -\sin\theta,\ \nu_2(0) = \cos\theta \tag{11.3}$$

In that case

$$\nu_1(t) = -\sin\theta e^{-iE_1 t/\hbar} \quad \text{and} \quad \nu_2(t) = \cos\theta e^{-iE_2 t/\hbar} \tag{11.4}$$

Solving Equation 11.1 for ν_μ,

$$\nu_\mu(t) = \cos\theta\, \nu_1(t) + \sin\theta\, \nu_2(t) = \sin\theta\cos\theta\left(-e^{-iE_1 t/\hbar} + e^{-iE_2 t/\hbar}\right) \tag{11.5}$$

The probability that the electron neutrino has converted into a muon neutrino, after a time t, is evidently

$$\begin{aligned}
|\nu_\mu(t)|^2 &= (\sin\theta\cos\theta)^2\left(e^{-iE_2 t/\hbar} - e^{-iE_1 t/\hbar}\right)\left(e^{iE_2 t/\hbar} - e^{iE_1 t/\hbar}\right) \\
&= \frac{\sin^2(2\theta)}{4}\left(1 - e^{i(E_2-E_1)t/\hbar} - e^{-i(E_2-E_1)t/\hbar} + 1\right) \\
&= \frac{\sin^2(2\theta)}{4}\left(2 - 2\cos\frac{(E_2-E_1)t}{\hbar}\right) = \frac{\sin^2(2\theta)}{4} 4\sin^2\left(\frac{E_2-E_1}{2\hbar}t\right)
\end{aligned}$$

or

$$P_{\nu_e\to\nu_\mu} = \left[\sin(2\theta)\sin\left(\frac{E_2-E_1}{2\hbar}t\right)\right]^2 \tag{11.6}$$

You see why they are called neutrino *oscillations*: ν_e will convert to ν_μ, and then *back again*, sinusoidally, just as coupled oscillators go back and forth between the normal modes. In this theory the electron and muon neutrinos themselves do not have well-defined energies – or masses; the 'mass eigenstates' are ν_1 and ν_2, with masses m_1 and m_2.* What is the energy of a highly relativistic particle of mass m and momentum $\mathbf{p}$? Well, $E^2 - |\mathbf{p}|^2c^2 = m^2c^4$, so

$$\begin{aligned}
E^2 &= |\mathbf{p}|^2c^2 + m^2c^4 = |\mathbf{p}|^2c^2\left(1 + \frac{m^2c^2}{|\mathbf{p}|^2}\right) \\
E &\approx |\mathbf{p}|c\left(1 + \frac{1}{2}\frac{m^2c^2}{|\mathbf{p}|^2}\right) = |\mathbf{p}|c + \frac{m^2c^3}{2|\mathbf{p}|}
\end{aligned}$$

Evidently, then,†

$$E_2 - E_1 \approx \frac{m_2^2c^3 - m_1^2c^3}{2|\mathbf{p}|} \approx \frac{(m_2^2 - m_1^2)}{2E}c^4 \tag{11.7}$$

* In particular, it is literally *nonsense* to speak of the 'mass' of an electron neutrino (for example) – it *has* no mass, any more than a three-note chord has a (single) pitch.

† I follow here the standard derivation, in which $\mathbf{p}$, not E, is held constant. Kayser [6] notes that this is 'technically incorrect', but a 'harmless error', since it leads (much more simply) to the right answer.

and hence*

$$P_{\nu_e \to \nu_\mu} = \left\{ \sin(2\theta) \sin\left[\frac{(m_2^2 - m_1^2)c^4}{4\hbar E} t \right] \right\}^2 \tag{11.8}$$

Write the answer, if you prefer, in terms of the distance $z \approx ct$ the neutrinos have traveled:

$$P_{\nu_e \to \nu_\mu} = \left\{ \sin(2\theta) \sin\left[\frac{(m_2^2 - m_1^2)c^3}{4\hbar E} z \right] \right\}^2 \tag{11.9}$$

In particular, after a distance

$$L = \frac{2\pi \hbar E}{(m_2^2 - m_1^2)c^3} \tag{11.10}$$

the probability of conversion hits a maximum, $\sin^2(2\theta)$, and at $2L$ they are all back to electron neutrinos.

Notice that two things are necessary, in order for neutrino oscillations to occur: There must be mixing (θ), and the masses must be unequal – in particular, they cannot both be zero. It is sometimes asserted that the Standard Model requires that neutrinos be massless, but I don't agree. It's true that some of the calculations are *easier* if you make this assumption, but there is no *fundamental* reason why neutrinos should have zero mass (whereas for the *photon* this is absolutely essential). Cross-generation mixing is a more significant change, though it already happens in the quark sector, and in a way it would be more surprising if it did *not* occur for the leptons.†

11.3 Confirmation

In 2001, the Super-Kamiokande collaboration presented its results on solar neutrinos [9]. Unlike the Davis experiment, SuperK uses *water* as the detector (Figure 11.3), and it is sensitive to muon and tau neutrinos as well as electron neutrinos. The process is elastic neutrino–electron scattering: $\nu + e \to \nu + e$; the outgoing electron is detected by the Cerenkov radiation it emits in water. The neutrino can be of any type, but the detection efficiency is 6.5 times greater

* Exactly the same formalism applies to neutral kaon mixing (Section 4.8.1) – see Problem 11.2.

† For neutrinos passing through *matter* (as opposed to vacuum) there are additional effects, due to elastic scattering of electron neutrinos ($\nu_e + e \to \nu_e + e$, by exchange of a W) and the Z^0-mediated interaction of neutrinos of any flavor with e, p, and n. This possibility, first noted by Wolfenstein, Mikheyev, and Smirnov [7] (hence known originally as the MSW effect), does not alter the functional form of Equation 11.9, but it does modify the effective mixing angle and mass splitting in a manner that depends on the density of the matter and the energy of the beam [8].

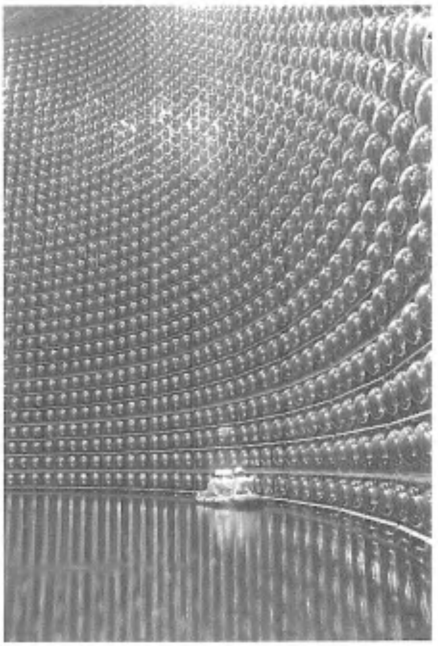


Fig. 11.3 The Super-Kamiokande detector (note the people in the rubber raft).

for electron neutrinos than for the other two kinds.* They recorded 45% of the predicted number ... *assuming* all of these neutrinos were still electron neutrinos. But remember that their detector is less efficient in counting μ and τ neutrinos. If some of the ν_e's had converted to ν_μ's or ν_τ's, then the *actual* flux would be higher – but how *much* higher they could not say, because they had no way of knowing what *fraction* of the neutrinos had in fact converted. You could look back at the Homestake data (remember, Homestake counted only *electron* neutrinos), but the conditions were sufficiently different that the comparison was not persuasive.

Meanwhile, at the Sudbury Neutrino Observatory (SNO) a very similar experiment was under way, using *heavy* water (D_2O) instead of ordinary water. The virtue of heavy water is that the *neutrons* present admit two *other* reactions (in addition to elastic scattering off electrons), and these enable one to measure separately the *electron* neutrino flux and the *total* neutrino flux (Figure 11.4). In the summer of 2001 the SNO collaboration published their first results [10], reporting on the neutrino absorption process (which applies only to electron neutrinos). They got 35% of the predicted flux. If you compare this with the SuperK data (45%) it appears that 10% of the neutrinos detected at SuperK must, in fact, have been ν_μ's or ν_τ's. But we know that the detector is 6.5 times more efficient for electron neutrinos, so if they *had* been ν_e's, they would have accounted for

* Elastic neutrino–electron scattering can proceed via Z^0 exchange for all three neutrino species, but for electron neutrinos there is an extra diagram, mediated by the W (see Problem 11.3).

Detection Methods

Homestake experiment (1968):

$$\nu_e + {}^{37}\mathrm{Cl} \rightarrow {}^{37}\mathrm{Ar} + e$$

Super-Kamiokande experiment (1998):

$$\nu + e \rightarrow \nu + e$$

Solar neutrino observatory (2002):

$$\begin{aligned} \nu_e + d &\rightarrow p + p + e \\ \nu + d &\rightarrow n + p + \nu \\ \nu + e &\rightarrow \nu + e \end{aligned}$$

Fig. 11.4 Detection mechanisms at Homestake, SuperK, and SNO.

$6.5 \times 10 = 65\%$, and $35 + 65 = 100$ – right on the money! This was just *too* perfect to be an accident, and many people concluded right then that the solar neutrino problem was solved, and neutrino oscillations confirmed. Still, not everyone was convinced, because this argument involves an awkward concatenation of data from different instruments, taken under different conditions. To nail it down definitively, the two measurements – the total flux and the electron-neutrino flux – had to be taken under identical conditions.* *Those* results were finally provided by the SNO collaboration in April 2002 [12]. Suffice it to say that they perfectly confirmed the tentative conclusions of the previous summer, with

$$\theta_{\mathrm{sol}} \approx \pi/6, \qquad \Delta(m^2)_{\mathrm{sol}} \approx 8 \times 10^{-5}\ (\mathrm{eV/c^2})^2 \tag{11.11}$$

(for the conversion of electron neutrinos to muon and/or tau neutrinos).

Of course, the sun is not the only supplier of neutrinos. There are also terrestrial sources (radioactive materials, nuclear reactors, and particle accelerators), atmospheric sources (cosmic rays), and astronomical sources (supernovae). In fact, the first strong evidence for neutrino oscillations was obtained at Kamiokande [13] (predecessor to SuperK) in the early 1990s using atmospheric neutrinos. Atmospheric neutrinos come mainly from the decay of pions and muons produced when cosmic rays (high-energy protons from outer space) hit air molecules in the upper atmosphere:

$$\begin{aligned} \pi^+ &\rightarrow \mu^+ + \nu_\mu, \qquad & \mu^+ &\rightarrow e^+ + \nu_e + \overline{\nu}_\mu, \\ \pi^- &\rightarrow \mu^- + \overline{\nu}_\mu, \qquad & \mu^- &\rightarrow e^- + \overline{\nu}_e + \nu_\mu \end{aligned} \tag{11.12}$$

* Incidentally, you may have been wondering whether neutrinos don't simply *decay* – that would certainly account for the deficit. But what could they decay *into*? Maybe some even lighter fermion we never noticed before. This was actually a viable (if implausible) option until the SNO experiments demonstrated conclusively not only that electron neutrinos are *missing*, but that the other flavors are *appearing* in their place. Kayser calls this the 'smoking gun' evidence for neutrino oscillations. It may well be true that the heaviest neutrino, at least, is unstable, but its lifetime is presumably too long to affect current experiments [11].

Evidently, there should be twice as many muon neutrinos (and antineutrinos) as electron neutrinos.* In fact, however, Kamiokande found roughly *equal* numbers of electron and muon neutrinos. This suggests that the muon neutrinos are converting to a different flavor. Indeed, the Kamiokande detector was able to sense the *direction* from which the neutrinos came; those from directly overhead, which had traveled only 10 km or so, arrived in the expected ratio (2: 1), but as the zenith angle increased (and with it the distance to the source), the ratio decreased (see Problem 11.4). These results were confirmed and improved by SuperK in 1998 [14]. It seems that the muon neutrinos convert into tau neutrinos, with

$$\theta_{\text{atm}} \approx \pi/4, \qquad \Delta(m)^2_{\text{atm}} \approx 3 \times 10^{-3}\ (\text{eV}/c^2)^2 \tag{11.13}$$

The atmospheric neutrino experiments (involving *muon* neutrino oscillations) tell us nothing about the solar neutrino problem (which involves *electron* neutrinos), but it is comforting to see the same phenomenon play out in two different contexts.

The ideal test of neutrino oscillations would involve a fixed source (a reactor or an accelerator) and a movable detector. As the separation increases, one would monitor the sinusoidal variation predicted by Equation 11.9. Unfortunately, neutrino detectors tend to be huge, and oscillation lengths are typically in the range of hundreds of kilometers (while the flux from a point source falls off like $1/r^2$). So one must make do with fixed targets and extremely intense sources, and study the variation with *energy*. The KamLAND experiment [15] uses a new detector at the SuperK site and looks at neutrinos from several power reactors 150–200 km away; the MINOS experiment [16] uses a detector in a mine in Soudan, Minnesota, to monitor accelerator-generated neutrinos from Fermilab, 750 km away in Illinois.

11.4 Neutrino Masses

With three neutrinos there are three mass splittings:

$$\Delta_{21} = m_2^2 - m_1^2, \quad \Delta_{32} = m_3^2 - m_2^2, \quad \Delta_{31} = m_3^2 - m_1^2 \tag{11.14}$$

Only two of them are independent ($\Delta_{31} = \Delta_{21} + \Delta_{32}$).† The oscillation measurements (Equations 11.11 and 11.13) indicate that one splitting is quite small, and

* Of course, not all pions decay to muons, and not all muons decay before reaching ground level; moreover, kaons as well as pions are produced by cosmic rays. So the factor of two is not *exact*, but it should be pretty close.

† The LSND experiment at Los Alamos reported a third mass splitting incompatible with this constraint [17], and was for a while interpreted as evidence of a fourth neutrino. Since, however, it was already established (see Section 11.9) that there are exactly three light neutrinos participating in the weak interactions, the 'extra' neutrino was taken to be 'sterile' (noninteracting, except for gravity). At any rate, the MiniBooNE experiment at Fermilab has pretty decisively repudiated the LSND result [18], and with it the notion of sterile neutrinos.

the others relatively large; we call ν_1 and ν_2 the closely-spaced pair (with $m_2 > m_1$), and ν_3 the loner. This structure is somewhat reminiscent of the charged leptons (e and μ fairly close in mass, τ much higher), and the quarks (d and s close, b higher; u and c relatively close, t much higher), so it is natural to assume that ν_3 is *heavier* than the other two – but it is possible that the neutrino spectrum is 'inverted', with ν_3 much *lighter* than ν_1 and ν_2 (Figure 11.5).

Unfortunately, oscillations are only sensitive to *differences* in (the squares of) neutrino masses, and one would like to measure the individual neutrino masses directly. This is not easy [19]. The standard method is to study the high energy cut-off (analogous to 9.2) in the beta-decay spectrum of tritium, but while these experiments set *upper bounds* on the neutrino mass, no measurement to date has established an actual mass. Meanwhile, an independent upper bound was provided serendipitously by the supernova SN1987A: 19 neutrinos, with a range of energies, were detected in the burst, which lasted only 10 seconds. For massive particles the speed is (of course) a function of energy, and the fact that these arrived so close together puts a limit of about $20\,\mathrm{eV}/c^2$ on the neutrino mass (see Problem 11.5). On the other hand, the atmospheric neutrino oscillations (Equation 11.13) imply that at least one of the neutrino masses must exceed $0.04\,\mathrm{eV}/c^2$. From all available evidence the best we can say today (2008) is that the heaviest neutrino mass lies somewhere between $0.04\,\mathrm{eV}/c^2$ and $0.4\,\mathrm{eV}/c^2$.*

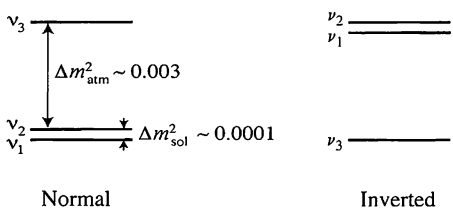


Fig. 11.5 'Normal' and 'inverted' neutrino mass spectrum. The units are $(\mathrm{eV}/c^2)^2$.

* Alone among the quarks and leptons, neutrinos could conceivably be their own antiparticles – 'Majorana' as opposed to 'Dirac' neutrinos (Problem 7.51). In Section 1.5 I mentioned the Davis and Harmer experiment, which appears to demonstrate that ν_e is distinct from $\overline{\nu}_e$. But it could be the *helicity* of the (anti)neutrino that forbids Equation 1.13. The ultimate test is *neutrinoless double beta decay*, in which a nucleus with atomic number Z goes to a nucleus of atomic number $Z + 2$, with the emission of two electrons and *no neutrinos* – in effect, the decay of two neutrons with annihilation of the accompanying neutrinos. This should be possible if $\overline{\nu}_e \equiv \nu_e$, but it has never been observed. One reason for interest in this scenario is that Majorana neutrinos are required by the so-called 'See-Saw' mechanism, which accounts for the extraordinary smallness of the neutrino masses by postulating that they are paired with extremely *heavy* neutrinos in a scheme whereby their masses are inversely proportional [20]. In any event, neutrino flavor oscillations work the same for Dirac and Majorana neutrinos.

11.5 The Mixing Matrix

In Section 11.2 I discussed oscillations between *two* neutrino species (ν_e and ν_μ, for the sake of argument). Of course, there are actually *three* kinds, and this complicates the algebra a bit.* But the essential point is unchanged: neutrinos *interact* as *flavor eigenstates* (ν_e is the particle that goes with the electron, ν_μ with the muon, and ν_τ with the tau), but they *propagate* as eigenstates of the free-particle Hamiltonian – the *mass eigenstates* ν_1, ν_2, and ν_3. The flavor eigenstates evolve in a complicated, oscillatory manner, because really they carry three different masses that are playing off against each other, like the beats of a coupled oscillator.

The same mixing happens with quarks, except that for them the familiar flavors (d, s, and b) *are* the mass eigenstates, and it is the 'weak eigenstates' (d', s', and b', Equation 9.85) that are 'rotated'; they are the ones that correspond to the neutrinos.† The CKM matrix (Equation 9.86) relates the weak eigenstates to the mass eigenstates in the quark sector; the analogous construct for leptons is sometimes called the 'MNS matrix' [22]:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{11.15}$$

As before (Equation 9.87), it can be expressed in terms of three angles (θ_{12}, θ_{23}, θ_{13}) and one phase factor (δ):

$$U = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{11.16}$$

($c_{ij} \equiv \cos\theta_{ij}$, $s_{ij} \equiv \sin\theta_{ij}$). But whereas the mixing angles for quarks are all rather small (so the CKM matrix is not far from diagonal, and the cross-generational couplings are suppressed), two of the leptonic mixing angles ($\theta_{12} \approx \theta_{sol}$ and $\theta_{23} \approx \theta_{atm}$) are large. Experimentally, $\theta_{sol} = 34 \pm 2°$ and $\theta_{atm} = 45 \pm 8°$. On the other hand, θ_{13} is known [23] to be less than 10°.

* As it turns out, if one of the three masses is substantially different from the others (which is in fact the case, as we have seen), then 'quasi-two-neutrino oscillation' (described by Equation 11.9) remains an excellent approximation [21].

† There is nothing *deep* here. Quarks interact dominantly by the strong interactions, which are agnostic – you could use *either* set of states; for them it is natural to let flavor coincide with mass. But neutrinos only interact weakly, so for them it seems more natural to use the weak eigenstates to define flavor. In retrospect, it would be better to speak uniformly of 'mass eigenstates' and 'weak eigenstates'; the standard flavors coincide with mass eigenstates for quarks, but with weak eigenstates for leptons.

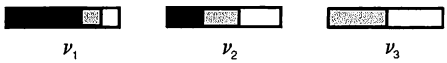


Fig. 11.6 Flavor contents of the neutrino mass eigenstates. Black is ν_e, gray is ν_μ, white is ν_τ. (The electron-neutrino component of ν_3 is too small to show on this scale).

Because U is a unitary matrix ($U^{-1} = U^\dagger$), it is easy to invert Equation 11.15, expressing the mass eigenstates in terms of the flavor states:

$$\begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} = \begin{pmatrix} U_{e1}^* & U_{\mu 1}^* & U_{\tau 1}^* \\ U_{e2}^* & U_{\mu 2}^* & U_{\tau 2}^* \\ U_{e3}^* & U_{\mu 3}^* & U_{\tau 3}^* \end{pmatrix} \begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} \tag{11.17}$$

It appears that ν_3 is an almost perfect 50–50 blend of ν_μ and ν_τ (with a tiny admixture of ν_e); ν_2 is a roughly equal combination of all three flavors; and ν_1 is mostly ν_e (Figure 11.6). But it will be several years before we have accurate numbers for the elements of the MNS matrix, and who knows how long before we can actually *calculate* them.

References

1 For a collection of the major papers on the solar neutrino problem, see Bahcall, J. N. (ed) et al. (2002) *Solar Neutrinos: The First Thirty Years*, Westview, Boulder, CO. For useful reviews of neutrino oscillations, see (a) Bilendy, S. M. and Petcov, S. T. (1987) *Reviews Modern Physics*, **59**, 671; (b) Haxton, W. C. and Holstein, B. R. (2000) *American Journal of Physics*, **68**, 15; (c) McDonald, A. B., Klein, J. R. and Wark, D. L. (April 2003) *Scientific American*, 40; (d) Waltham, C. (2004) *American Journal of Physics*, **72**, 742.

2 Davis, R. Jr., Harmer, D. S. and Hoffman, K. C. (1968) *Physical Review Letters*, **20**, 1205.

3 Bahcall, J. N., Bahcall, N. A. and Shaviv, G. (1968) *Physical Review Letters*, **20**, 1209.

4 Pontecorvo first discussed the possibility of neutrino oscillations – by analogy with $K^0/\overline{K}^0$ oscillations – in 1957: Pontecorvo, B. (1958) *Soviet Physics JETP*, **6**, 429. He revived the idea as a possible resolution to the solar neutrino problem in 1968: (a) Gribov, V. N. and Pontecorvo, B. M. (1969) *Physics Letters B*, **28**, 493.

5 For a charming discussion, seeLyons, L. (June 1999) *CERN Courier*, p. 32. For the relativistic treatment see (a) Sassaroli, E. (1999) *American Journal of Physics*, **67**, 869.

6 For accessible, authoritative, and well-written material on neutrino oscillations, see practically anything by Kayser, B. (2004) *This Quote is From his Lecture*, at the SLAC Summer Institute; which includes a very careful treatment of the kinematics. See also (a) Burkhardt, H. et al. (2003) *Physics Letters B*, **566**, 137.

7 Wolfenstein, L. (1978) *Physical Review*, **D17**, 2369; (a) Mikheyev, S. and Smirnov, A. (1986) *Soviet Journal of Nuclear Physics*, **42**, 913; (1986) *JETP*, **64**, 4; and (1986) *Nuovo Cimento*, **9 C**, 17.

8 SeeKayser, B.,Burkhardt, H. et al. (2003) *Physics Letters B*, **566**, 137.

9 Fukuda, S. et al. (2001) *Physical Review Letters*, **86**, 5651.

10 Ahmad, Q. R. et al. (2001) *Physical Review Letters*, **87**, 071301.

11 Barger, V. etal. (1999) *Physics Letters B*, **462**, 109; (a) Beacom, J. F. and Bell, N. F. (2002) *Physical Review D*, **65**, 113009.
12 Ahmad, Q. R. et al. (2002) *Physical Review Letters*, **89**, 011301. For useful commentary see (a) Schwarzschild, B. (July 2002) *Physics Today*, 13.
13 Hirata, K. S. et al. (1992) *Physics Letters B*, **280**, 146.
14 Eguchi, K. et al. (2003) *Physical Review Letters*, **90**, 021802; (a) Kearns, E., Kajita, T. and Totsuka, Y. (August 1999) *Scientific American*, 68.
15 Fukuda, Y. et al. (1998) *Physical Review Letters*, **81**, 1562. For an illuminating discussion of the SuperKamiokande atmospheric neutrino experiments see (a) Schwarzschild, B. (August 1998) *Physics Today*, 17.
16 Michael, D. G. et al. (2006) *Physical Review Letters*, **97**, 191801.
17 Athanassopoulos, C. et al. (1995) *Physical Review Letters*, **75**, 2650.
18 See Schwarzschild, B. (June2007) *Physics Today*, 18. For delightful commentary see (a) Cole, K. C. (June 2, 2003), *The New Yorker*, p. 48.
19 For a survey of direct neutrino mass measurements see Haxton, W. C. and Holstein, B. R. (2000) *American Journal of Physics*, **68**, 15.
20 For interesting commentary on the distinction between neutrinos and antineutrinos seeBoas, M. (1994) *American Journal of Physics*, **62**, 972; (a) Hammond, R. (1995) *American Journal of Physics*, **63**, 489; (b)Wagner, R. G. (1997) *American Journal of Physics*, **65**, 105; (c) Fewell, M. P. (1998) *American Journal of Physics*, **66**, 751; (d) Holstein, B. R. (1998) *American Journal of Physics*, **66**, 1045; (e) Fewell, M. P. (1998) *American Journal of Physics*, **66**, 751; (f) Boya, L. J. (2000) *American Journal of Physics*, **68**, 193. On neutrinoless double beta decay ('$\beta\beta(0\nu)$) see (g) Elliott, S. R. and Vogel, P. (2002) *Annual Review of Nuclear and Particle Science*, **52**, 115.
21 See Kayser, B., in *Review of Particle Physics* (2006) p. 156.
22 This is in honor ofMaki, Z., Nakagawa, M. and Sakata, S. (1962) whose pioneering work *Progress in Theoretical Physics*, **28**, 870, long predates the discovery of neutrino oscillations (or of the tau).
23 For planned experiments to measure θ_{13}, see Feder, T. (November 2006) *Physics Today*, 31.

Problems

11.1 Estimate the lifetime of the sun, assuming (as Lord Kelvin did) that the source of the energy radiated is gravity. Look up any empirical numbers (the power radiated by the sun, the mass, and radius of the sun).

11.2 (a) What is the period of $K^0 \leftrightharpoons \overline{K}^0$ oscillations (Section 4.4.3)? [*Hint:* The mass eigenstates are K_S^0 and K_L^0. In the neutrino case (Eq. 11.7) the particles were highly relativistic; for the K's, assume on the contrary that the kinetic energy is substantially *less* than the rest energy.]

(b) Compare your result in **(a)** to the lifetimes of K_S^0 and K_L^0. Notice that the K_S^0 component of the beam dies out – leaving pure K_L^0 – well before significant oscillation can occur.

11.3 Draw the lowest-order diagrams for elastic neutrino–electron scattering, **(a)** for electron neutrinos, **(b)** for muon neutrinos, **(c)** for tau neutrinos.

11.4 **(a)** Suppose atmospheric neutrinos are produced at an altitude h, and the detector is at sea level. Find the distance x from the source to the detector, as a function of the zenith angle Θ (directly overhead is $\Theta = 0$, the horizon is $\Theta = 90°$, straight down is $\Theta = 180°$). Let R be the radius of the earth.

(b) Suppose 95% of the 'upper' neutrinos (overhead to horizon) reach the detector, but only 50% of the 'lower' neutrinos (below the horizon) do. Using the oscillation formula Equation 11.9 (but this time for muon neutrinos converting to tau neutrinos), determine θ and Δm^2. Assume $h = 10$ km and $E = 1$GeV. [This problem was posed by Waltham [1]. You'll need a computer to get the numerical answer.]

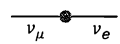
11.5 **(a)** Show that the velocity of an ultra-relativistic particle (mass m) with energy E is approximately

$$v \approx c\left[1 - \frac{1}{2}\left(\frac{mc^2}{E}\right)^2\right]; \quad \frac{1}{v} \approx \frac{1}{c}\left[1 + \frac{1}{2}\left(\frac{mc^2}{E}\right)^2\right]$$

(b) Supernova SN1987A occurred in the Large Magellanic Cloud (1.7×10^5 light years from Earth). Neutrinos from this explosion, with energies ranging from 20 MeV to 30 MeV, were detected within a 10 s time interval. What upper bound on the neutrino mass does this imply? [Assume the neutrinos all started out at the same instant.]

11.6 With neutrino oscillations, the individual lepton numbers (L_e, L_μ, and L_τ) are no longer conserved, and this means that the decay $\mu \to e + \gamma$ (the *absence* of which suggested these conservation laws in the first place – see Eq. 1.16) is *possible*, in principle.

(a) Draw a Feynman diagram for this process. *Note*: neutrino oscillations can be represented by a blob:



(b) In this process you must 'borrow' the energy necessary to make the virtual W. According to the uncertainty principle (see Problem 1.2), how soon must you 'repay' the debt? How far could a neutrino get in this time? Given that neutrino oscillations occur over distance scales of many *kilometers*, does it seem likely that you could 'borrow' the energy long enough for $\mu \to e + \gamma$ to occur?

12

Afterword: What's Next?

So far, I have talked almost exclusively about established 'facts'. With the possible exception of the Higgs mechanism, any future theory will have to include all of this. But the Standard Model is certainly not the last word on the subject. Already there are intriguing theoretical speculations and tantalizing experimental indications of what the future may hold. Increasingly, the impetus is coming from observations in astrophysics and cosmology, rather than traditional collider experiments. In this chapter, I'll explore some of the directions in which future discoveries seem most likely. I'll start (Section 12.1) with the hunt for the Higgs, which tops the agenda for the Large Hadron Collider (LHC) (and for the remaining lifetime of the Tevatron), and may lead to the explanation for all particle masses. Next (Section 12.2), I'll discuss Grand Unification, which was the 'natural' next step 30 years ago, but hit a brick wall when the predicted decay of the proton was not observed; it nevertheless sets the context for all subsequent theoretical developments. Then (Section 12.3) I'll consider CP violation and its implications for the matter/antimatter asymmetry of the universe. Section 12.4 is a scandalously brief introduction to supersymmetry, extra dimensions, and string theory, ideas that have dominated theoretical particle physics since 1984 and for which the first experimental support may come from the LHC. Finally, in Section 12.5 we'll study Dark Matter and Dark Energy, which by current estimates account for 95% of the matter in the universe, leaving only a paltry 5% for the 'ordinary' particles we encountered in the first 11 chapters.*

12.1
The Higgs Boson

In the Higgs mechanism, a gauge symmetry is spontaneously broken by a two-component scalar field ϕ, whose ground state is not zero (Section 10.9). One component of ϕ is reincarnated as the third (longitudinal) polarization state

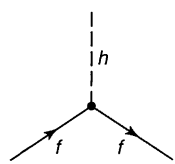
* In retrospect, one might call the period from the early 1930s to 1954 the era of cosmic rays, and from the Cosmotron to the Large Hadron Collider (LHC) – let's say 2010 – the era of accelerator physics; in this sense, we are now entering the era of particle astrophysics (1). Part of the reason is simple economics: to reach ever higher energies, accelerators have become so huge and so expensive that it is hard to imagine anything beyond the International Linear Collider (ILC) now on the drawing boards. Astrophysics offers a relatively inexpensive window into vastly higher energy regimes.

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

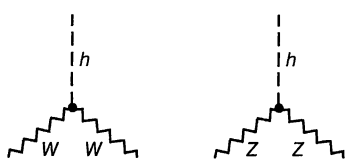
for a now *massive* gauge field, but the other remains, representing a neutral particle of spin 0: the Higgs boson [2].

Most particle physicists believe in the Higgs mechanism because it seems to be the only way (certainly it is the *cleanest* way) to account for the mass of the W and the Z, in the context of local gauge theory. But if there really *is* a Higgs field, permeating all of space, with a nonzero value even in 'vacuum', it could account as well for the masses of the quarks and leptons, whose interaction with the primordial Higgs field has been likened to wading through deep water, imparting an effective inertia to (almost) everything that moves. In this more exalted vision, the Higgs particle becomes the source of *all* mass.* The quarks and leptons are 'born' massless,† but with Yukawa couplings (Problem 10.21) to the ϕ: $\mathcal{L}_{\text{int}} = -\alpha_f \overline{\psi}_f \psi_f \phi$, where f denotes the particular quark or lepton. When ϕ is 'shifted' by spontaneous symmetry-breaking (Equation 10.130), $\mathcal{L}_{\text{int}}$ splits into two pieces, one of which is a Yukawa coupling to the physical Higgs field and the other a pure Fermion mass term, $-m_f c^2 \overline{\psi}_f \psi_f$ (in the notation of Section 10.9, $m_f c^2 = (\mu/\lambda)\alpha_f$). Unfortunately, this doesn't help us to *calculate* the particle masses – it simply trades one unknown parameter (m_f) for another (α_f). But it does suggest that the strength of the coupling to the Higgs is proportional to mass.

In the simplest theory (the 'Minimal Standard Model', MSM), there are *four* scalar fields to begin with – two charged and two neutral. Three of these are 'eaten' by the $W\pm$ and Z^0 (which thereby acquire mass) and the fourth remains as the neutral Higgs field. More complicated schemes have been proposed, involving multiple or composite Higgs particles,‡ but the MSM provides a useful roadmap for experimental and theoretical exploration of the Higgs sector. In this model, the Higgs (h) interacts with quarks and leptons by the diagram



(vertex factor $-im_f c^2/v$), and with the weak mediators by

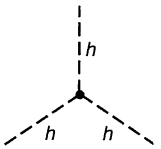


* Leon Lederman famously called it *'The God Particle'* (New York: Delta, 1993).

† In the Standard Model Lagrangian, fermion mass terms ($\overline{\psi}\psi$) are not invariant under the electroweak symmetry $SU(2)_L \times U(1)$, so the 'starting' masses of the quarks and leptons have to be zero, and the physical masses arise only when the symmetry is (spontaneously) broken.

‡ In supersymmetric theories, for example, there are at least five Higgs bosons and in technicolor the role of the Higgs is played by a bound state of two fermions.

(vertex factor $2iM_m^2c^2g^{\mu\nu}/(\hbar^2 v)$, where the subscript m stands for W or Z). There is, as well, a direct Higgs–Higgs coupling*



(vertex factor $-3im_h^2c^2/(\hbar^2 v)$). Here, v is the 'vacuum expectation value' of ϕ_1 (μ/λ, for the potential in Section 10.8). It can be calculated from the mass of the W (see Problem 12.1)

$$\sqrt{\hbar c}\, v = \frac{2M_W c^2}{g_w} = 246 \text{ GeV} \tag{12.1}$$

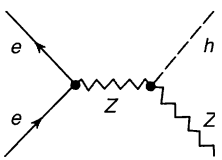
The mass of the Higgs itself is not determined by the theory.†

It would be nice to know whether this story (or some variation on it) is actually *true*. The Higgs particle is the only element in the Standard Model for which there is as yet no compelling experimental evidence. It *may* have been seen at LEP (CERN), in the last months before it shut down (to make way for the LHC) [4], it could still be found at the Tevatron (Fermilab), and unless current theories are wildly off it will certainly be observed at the LHC. Various constraints – experimental and theoretical – suggest that its mass must lie in the range

$$114 \text{ GeV}/c^2 < m_h < 250 \text{ GeV}/c^2 \tag{12.2}$$

with a most probable value around 120 GeV/c^2 [5]. The LHC will explore the entire region up to 1 TeV and beyond.

At LEP (an electron–positron collider) the Higgs was sought in the Z-'bremsstrahlung' reaction $e^+ + e^- \to Z + h$:

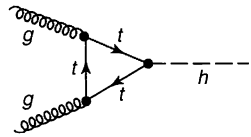


At hadron colliders (the Tevatron and LHC), the dominant production mechanism is gluon 'fusion', $g + g \to h$ via a quark loop (mainly the top, since it's the heaviest,

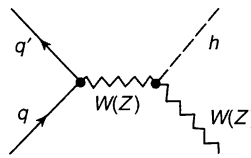
* There are also 'four-point' couplings $hh \to ZZ$, $hh \to WW$, and $hh \to hh$ (3).

† In Equation 10.121, m_η involves only μ, not μ/λ, so it is sensitive to the shape of the potential, not just the vacuum expectation value of ϕ_1.

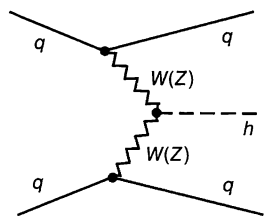
and hence has the strongest coupling to the Higgs):



but several other modes are expected to contribute, notably W/Z-bremsstrahlung:

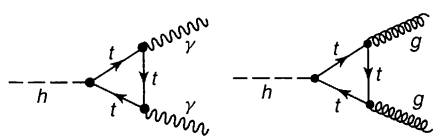


and W/Z fusion:



(direct quark fusion, $q + \bar{q} \to h$, does not contribute much in the MSM, because the only readily available quarks are u's and d's, which, since they are very light, couple weakly to the h).

How do we expect the Higgs to decay? Because Higgs couplings are proportional to mass (or mass *squared*, in the case of the W and Z), heavy daughters are favored, if they are kinematically allowed. The branching ratios depend a lot on the mass of the Higgs (see Figure 12.1). If m_h is less than about 140 GeV/c^2, the dominant mode is $h \to b\bar{b}$, but above that $h \to W^+W^-$ takes over (with a virtual W up to 160 GeV/c^2 and real W's from then on); $h \to ZZ$ is close behind (especially above 180 GeV/c^2), and in the unlikely event that the Higgs is heavy enough to make two tops ($m_h > 360$ GeV/c^2) $h \to t\bar{t}$ assumes third place. More exotic decays are also possible, such as a photon or gluon pair:



These decay rates have all been calculated in great detail [6] (you can do some of them yourself – see Problems 12.2 and 12.3).

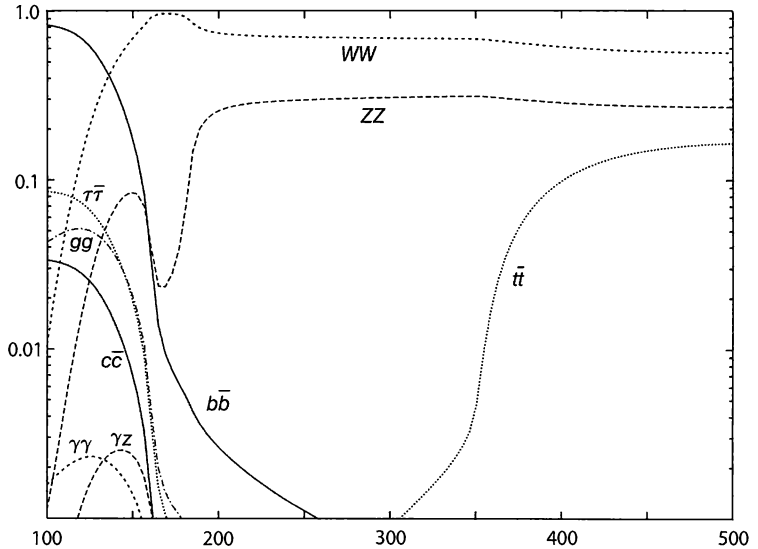


Fig. 12.1 Branching ratios for Higgs decay, as functions of the Higgs mass (in GeV/c^2). (*Source*: Gunion, J. F. *etal.* (1990) *The Higgs Hunter's Guide*, Addison-Wesley, Redwood City, CA.)

As soon as the Higgs mass is established, one will be able to draw a vertical line at the appropriate point in Figure 12.1 and read off the branching ratios. If the measurements disagree (as they probably *will*), then the Higgs sector is more interesting than the MSM contemplates. And of course if no Higgs particle is found at *all*, then we have a revolution on our hands.

12.2 Grand Unification

With the success of electroweak unification, in the 1960s, the logical next step was to include the strong interactions, in a 'Grand Unified Theory' (GUT) that would identify all three forces as different manifestations of a single underlying interaction. Of course, the strong forces are enormously more powerful than the others; but the same could be said of electromagnetic versus weak forces, and we now understand *that* disparity as an artifact of the huge mass of the W and Z – their intrinsic strengths are quite similar, but it is only at energies well above $M_W c^2$ that the unity becomes manifest.

Moreover, as we saw in Sections 7.9 and 8.6, the coupling 'constants' themselves are functions of energy – the strong and weak couplings go *down*, while the electromagnetic coupling goes *up*. It is irresistible to suppose that they coalesce at some point (Figure 12.2); above the *grand unification scale* ($\approx 10^{16}$ GeV) there is just

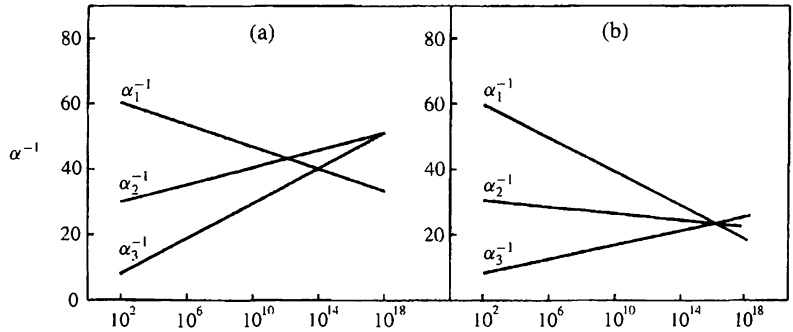


Fig. 12.2 Convergence of the coupling constants at the GUT scale (a) in the Minimal Standard Model, (b) with supersymmetry. The horizontal axis is energy, in GeV.

one universal coupling constant, and the strong, electromagnetic, and weak forces are identical in strength.*

The first (and simplest) GUT was introduced by Georgi and Glashow in 1974 [7]. It led to a spectacular prediction: the proton is unstable, decaying (for example) into a positron and a pion

$$p \to e^+ + \pi^0 \tag{12.3}$$

The lifetime is reassuringly long – at least 10^{30} years, which is 10^{20} times the age of the universe – though (since we have easy access to a lot of protons) not beyond the range of measurement. In 30 years of increasingly precise experiments, however, proton decay has never been observed [8]. The current lower bound is

$$\tau_{\text{proton}} > 10^{33} \text{ years} \tag{12.4}$$

(which probably vetoes the Georgi–Glashow model). More elaborate GUTs have been proposed, but almost all of them require proton decay at *some* level.

Although there is no direct experimental evidence in support of grand unification, belief in it is an uncontested article of faith among theorists. In a way, the 'natural' evolution of particle physics was rudely interrupted by the failure to detect proton decay. Had proton decay been discovered in – say – 1985, one can easily imagine that enormous efforts, theoretical and experimental, would have been devoted to fleshing out the details of grand unification, just as the previous two decades had fleshed out the Standard Model. But that's not what happened, and today grand unification simmers, half-cooked, on a back burner.† What are its essential features, and why should we take it seriously? [9].

* Awkwardly, it is now clear that they do *not* (quite) meet at a single point, in the MSM; one of the attractions of supersymmetry is that it makes perfect convergence possible.

† Testable predictions of grand unification, at the relatively low energies presently accessible, are few and far between. Proton decay, if it exists, is the best probe available, but we are fast approaching the practical limit on proton lifetime measurements (see Problem 12.4).

Table 12.1 Fermion States in the $SU(5)$ GUT

Quintet	$e_L, \nu_e, \overline{d_R^r}, \overline{d}_R^b, \overline{d}_R^g$
Decuplet	$e_R, u_R^r, u_R^b, u_R^g, \overline{u}_L^r, \overline{u}_L^b, \overline{u}_L^g, \overline{d}_L^r, \overline{d}_L^b, \overline{d}_L^g$

Grand unification contemplates an overarching symmetry group ($SU(5)$, in the Glashow–Georgi version) that contains as subgroups the (color) $SU(3)$ and $SU(2)_L \otimes U(1)$ symmetries of the Standard Model. The fundamental fermions (quarks and leptons) are assigned to representations of this group, much as the Eightfold Way assigned baryons and mesons to (octet, nonet, and decuplet) representations of (flavor) $SU(3)$. The first generation comprises 15 particle states: u and d, each in three colors and two chiralities (L and R), e (L and R), and ν_e (L only *). In the $SU(5)$ GUT, they constitute a quintet and a decuplet† (Table 12.1); in the absence of symmetry-breaking (presumably by the Higgs mechanism), the states in each multiplet share the same mass and interact identically. (The same goes, of course, for the other two generations.) There are 24 mediators‡ (Table 12.2): the 8 gluons, the photon, W^+, W^-, and Z, and 12 new ones – the X (charge $\pm 4/3$, 3 colors, hence 6 in all) and the Y (charge $\pm 1/3$, 3 colors, for another 6). They couple leptons to (anti)quarks,§ and hence are known as *leptoquarks*. For instance, $\overline{d} \to e + X$ and $\overline{u} \to e + Y$:

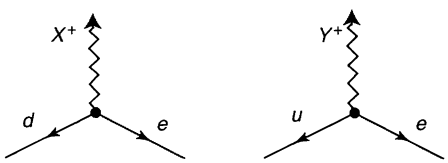


Table 12.2 Gauge Bosons in the $SU(5)$ GUT

	Charge	Mass
8 gluons	0	0
1 photon	0	0
3 $W\pm$, Z	1, −1, 0	$\sim 10^2$ GeV/c^2
6 X	4/3, −4/3	$\sim 10^{16}$ GeV/c^2
6 Y	1/3, −1/3	$\sim 10^{16}$ GeV/c^2

* In 1974, it was assumed that the neutrino is massless, and the fact that there was no natural place for ν_R was taken as a virtue of the theory. For massive neutrinos ν_R must be assigned, awkwardly, to a singlet representation of $SU(5)$ – or, in the case of Majorana neutrinos, the Higgs sector must be expanded.

† The fact that they don't all fit into a single irreducible representation is an unattractive feature of the $SU(5)$ model; the $SO(10)$ GUT assigns all 15, plus ν_R, to a 16-dimensional representation.

‡ In general, $SU(n)$ has $n^2 - 1$ mediators (eight gluons for color $SU(3)$, three intermediate vector bosons for $SU(2)_L$); $U(n)$ has n (hence one photon).

§ Notice that it is the *anti-d* that lies in the same multiplet as the electron.

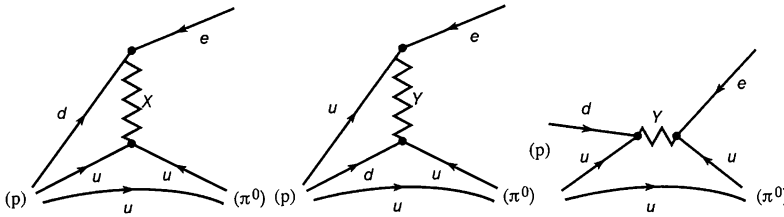
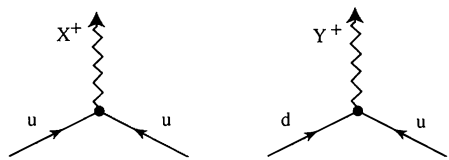


Fig. 12.3 Proton decay in the $SU(5)$ GUT.

They also couple quarks to antiquarks (in this context, they are sometimes called *diquarks*), as in $u \to \overline{u} + X$ and $d \to \overline{u} + Y$:*



This larger symmetry is badly broken, obviously (quarks and leptons do *not* have the same mass, and the strong interactions are – well – *stronger* than the others). Just as electroweak symmetry becomes apparent at energies well above the W/Z mass, GUT symmetry prevails at energies above the (huge) grand unification scale. That's why it is so difficult to test grand unification in the laboratory – even though its implications are, in principle, dramatic. The leptoquark couplings allow for nonconservation of lepton and baryon number, and hence license the decay of the proton, via diagrams such those in Figure (12.3). But because these mediators are so heavy (presumably in the neighborhood of the GUT scale: $M_X \sim M_Y \sim 10^{16}$ GeV/c^2), the decay rate is extremely small (Problem 12.5).

Apart from the largely aesthetic attraction of unifying the fundamental forces of particle physics, grand unification purports to 'explain' the relation between quark and lepton charges (and beyond that the quantization of charge itself). For technical reasons the sum of the charges in a multiplet must be zero, and putting quarks and leptons into the *same* multiplet forces (in the case of the $SU(5)$ quintet)

$$q_e - 3q_d = 0 \tag{12.5}$$

* Ostensibly these reactions do not conserve color, but remember that the 'cross product' of two color states carries a single color (Equation 10.84), and it is such a combination that is implied here.

Our world would be a radically different place if the electron and proton did not have precisely opposite charge, but short of grand unification there is no reason of principle why this had to be so.*

12.3 Matter/Antimatter Asymmetry

Everyone assumes that the Big Bang created matter and antimatter in exactly equal amounts. If this is the case, how come we are surrounded by electrons, protons, and neutrons, with no positrons, antiprotons, or antineutrons in sight? Of course, if a positron (for example) *does* show its face, it doesn't last long: as soon as it encounters an electron, they annihilate. But this doesn't explain the preponderance of leftover electrons. Perhaps it's a local phenomenon – our matter-dominated corner of the universe is balanced by an antimatter region somewhere out there. However, there is no evidence for this – on the contrary, astrophysical observations indicate that the known universe, at least, is all matter (if there *were* an antimatter zone, the border would be an extremely violent place, and it is hard to imagine that the cosmic microwave background would show no sign of the disturbance) [12]. Alternatively, some process must have favored matter over antimatter in the course of cosmic evolution. What sort of mechanism might do the job?

In 1967, Sakharov [13] identified the necessary ingredients. Obviously, there must be an interaction that violates conservation of baryon and lepton number (something grand unification could supply). There must have been a period when the universe was substantially out of thermal equilibrium (otherwise any reaction $i \to f$ would go just as often the other way, $f \to i$, and there would be no net change in baryon number). And, crucially, there must be CP violation – some reaction $i \to f$ whose rate is different from its CP-conjugate, $\tilde{i} \to \tilde{f}$ (otherwise, again, there would be no net change in baryon number). Conveniently, CP violation had recently been discovered by Cronin and Fitch in the $K^0/\overline{K}^0$ system.

To this day, the underlying nature of CP violation is not well understood. Parity violation was very easy to incorporate into the theory of weak interactions: one simply replaced vector couplings, γ^μ, by vector/axial vector couplings, $\gamma^\mu(1-\gamma^5)$ (Section 9.1). But the only known source of CP violation is the residual phase δ in the CKM matrix (Equation 9.87), and it is hardly obvious why this breaks CP invariance. Consider a process $i \to f$, and the CP-reversed process $\tilde{i} \to \tilde{f}$ (for instance, if i includes a left-handed electron, $\tilde{i}$ includes a right-handed positron); CP violation means that the rate for $\tilde{i} \to \tilde{f}$ is not the same as for $i \to f$ (for

* A more problematic implication of grand unification is the existence of super heavy 't Hooft–Polyakov magnetic monopoles (10), which should be present in large numbers (left over from the Big Bang), but have never been detected in the laboratory (well ... maybe *once* (11)). Inflationary cosmology can account for a dilution in the number, but the prediction of (unobserved) monopoles in grand unification – and for that matter in other theories as well – remains a troubling problem.

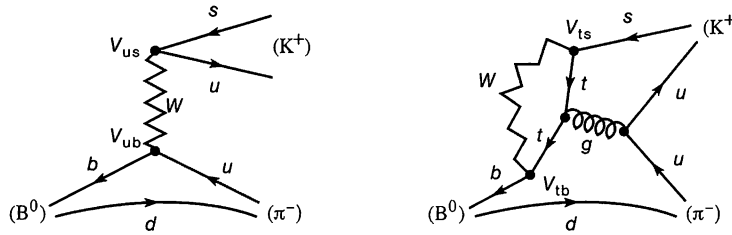


Fig. 12.4 Two diagrams for $B^0 \to K^+ + \pi^-$. The second is a 'penguin' (Problem 4.40).

example, $B^0 \to K^+ + \pi^-$ is 13% more common than $\overline{B}^0 \to K^- + \pi^+$). Now, the amplitude, $\mathscr{M}$, is a complex number, and ordinarily it is the same for $\tilde{i} \to \tilde{f}$ as for $i \to f$ – *except* that any CKM element gets conjugated. Thus

$$\mathscr{M} = |\mathscr{M}|e^{i\phi}e^{i\theta}, \qquad \tilde{\mathscr{M}} = |\mathscr{M}|e^{i\phi}e^{-i\theta} \tag{12.6}$$

where θ is the 'conjugating' phase and ϕ the 'ordinary' phase.* On the other hand, reaction rates are proportional to $|\mathscr{M}|^2$, so there is no CP violation, even though the amplitudes themselves are different.

But suppose that the process $(i \to f)$ can proceed by *two different routes* (for example, B^0 can go to $K^+ + \pi^-$ in several distinct ways – see Figure 12.4). Then $\mathscr{M} = \mathscr{M}_1 + \mathscr{M}_2$, with

$$\mathscr{M}_1 = |\mathscr{M}_1|e^{i\phi_1}e^{i\theta_1}, \quad \mathscr{M}_2 = |\mathscr{M}_2|e^{i\phi_2}e^{i\theta_2} \tag{12.7}$$

and $\tilde{\mathscr{M}} = \tilde{\mathscr{M}}_1 + \tilde{\mathscr{M}}_2$, with

$$\tilde{\mathscr{M}}_1 = |\mathscr{M}_1|e^{i\phi_1}e^{-i\theta_1}, \quad \tilde{\mathscr{M}}_2 = |\mathscr{M}_2|e^{i\phi_2}e^{-i\theta_2} \tag{12.8}$$

It follows (Problem 12.6) that

$$|\mathscr{M}|^2 - |\tilde{\mathscr{M}}|^2 = -4|\mathscr{M}_1||\mathscr{M}_2|\sin(\phi_1 - \phi_2)\sin(\theta_1 - \theta_2). \tag{12.9}$$

In this case the rates are *not* the same, and CP is violated. Notice that there has to be a conjugating phase (from the CKM matrix) as well as a nonconjugating phase – and these have to be *different* for the two contributing routes.

From the experiments, we know that CP violation occurs in the weak interactions of quarks, and is attributable to the phase factor in the CKM matrix.† Unfortunately, this is *nowhere near* enough to account for the matter dominance of the universe [15],

* In the literature, they are sometimes called 'weak' and 'strong' phases, respectively. The distinction is subtle, but in practice θ comes exclusively from the CKM matrix element, and ϕ typically involves strong interaction effects (14).

† In fact, all such CP violating effects are proportional to the height of the 'unitarity triangle' (Problem 9.33).

so one is forced to speculate about other mechanisms of CP violation. With massive neutrinos and the leptonic analog to the CKM matrix (Section 11.5), the same phenomenon should occur in the lepton sector, where it would reveal itself, for example, in unequal probabilities for $\nu_e \to \nu_\mu$ and $\overline{\nu}_e \to \overline{\nu}_\mu$. This has not been observed (yet), but it is conceivable that it would provide a mechanism (sometimes called *leptogenesis**) for the observed matter/antimatter asymmetry. Another possibility is CP violation in the *strong* interactions (in this case the 'smoking gun' would be a nonzero electric dipole moment for the neutron). CP violation has never been observed in strong processes, but there does not seem to be any fundamental theoretical prohibition.† At this point, the matter/antimatter asymmetry of the universe remains an uncompleted puzzle; the essential missing piece is the nature of the CP violation responsible. It is far from clear how this story will resolve itself.

12.4 Supersymmetry, Strings, Extra Dimensions

12.4.1 Supersymmetry

The classic symmetries of quantum mechanics involve different states of the same system. Rotational invariance, for instance, requires that the theory be unchanged when the state ψ is replaced by its rotated version $U(\boldsymbol{\theta})\psi$ (Equation 4.27) – or, more precisely, the Lagrangian is unchanged (in first order)‡ when the wave function is incremented by the infinitesimal amount $\delta\psi = (-i/\hbar)[\delta\boldsymbol{\theta} \cdot \mathbf{S}]\psi$ (Equation 4.28). Particle physics long ago generalized the idea to 'internal symmetries' involving closely related particles (flavor multiplets, for example). In 1974, Wess and Zumino [17] introduced a more radical symmetry that stirred together fermions and bosons. For example, a scalar field ϕ could mix with a spinor field ψ

$$\delta\phi = 2\overline{\epsilon}\psi, \quad \delta\psi = -\left(\frac{i}{\hbar c}\right)\gamma^\mu \epsilon(\partial_\mu \phi) \tag{12.10}$$

where ϵ is an infinitesimal spinor describing the transformation (analogous to $\delta\boldsymbol{\theta}$ for rotations), and $\overline{\epsilon} \equiv \epsilon^\dagger \gamma^0$ is its adjoint. What if we insist that the theory be

* The terminology is not entirely consistent: *baryogenesis* is the generic word for the origin of matter dominance, so leptogenesis is actually one possible mechanism for baryogenesis.

† Indeed, it is something of a mystery why strong CP violation does *not* occur. One possible explanation was suggested by Peccei and Quinn (16) in 1977: a neutral spin-0 particle (the *axion*) couples to the quarks in such a way as to cancel dynamically any strong CP violation. Axions have not been observed, but they remain among the viable candidates for dark matter.

‡ It is generally simpler to work with infinitesimal transformations, and there is no loss of generality since a finite transformation can be built up as a sequence of infinitesimal ones (Problem 12.7).

invariant under such a transformation? It is not hard to construct a Lagrangian with this property; the combined free Klein–Gordon and Dirac Lagrangians are invariant

$$\mathscr{L} = \frac{1}{2}\left[\partial^\mu \phi^* \partial_\mu \phi - \left(\frac{mc}{\hbar}\right)^2 \phi^* \phi\right] + i(\hbar c)\overline{\psi}\gamma^\mu \partial_\mu \psi - (mc^2)\overline{\psi}\psi \tag{12.11}$$

as long as the boson ϕ and its fermion partner ψ carry the same mass (Problem 12.8). A similar game can be played joining a particle of spin 1/2 to a particle of spin 1 – and in general pairing particles whose spins differ by 1/2. Invariance of this kind, linking fermions and bosons, is called 'supersymmetry'.

Over the past 30 years an enormous amount of work has been done on supersymmetry [18], and I think it is fair to say that most particle physicists are convinced (without as yet any supporting experimental evidence) that it is a fundamental symmetry of nature. Supersymmetry carries the stupendous implication that every fermion has a bosonic partner (identified by putting an 's' in front of the name – thus 'squark', 'slepton', 'selectron', 'sneutrino', etc.) and every boson has a fermionic partner (identified by putting an 'ino' after the name – thus 'photino', 'gluino', 'wino', 'higgsino', etc.). Where *are* all these particles? If supersymmetry were unbroken, they would share the masses of their 'ordinary' twins – the photino would be a massless particle of spin 1/2, and the selectron a spin-0 particle with a mass of 0.511 MeV/c^2. This is nonsense, obviously – no such particles exist. So the symmetry must be badly broken (perhaps spontaneously, but there are other possibilities, especially if gravity is brought into the picture). Presumably the supersymmetric particles are much heavier – too heavy to be produced by any existing machine, though there are strong indications that at least some of them should be accessible to the LHC.

Hmm . . . . Why should we take such an outlandish scheme seriously? Supersymmetry has the potential to solve several thorny problems, among them the following:

1. By introducing a number of new particles, it modifies the energy dependence of the three running coupling constants (see Equations 7.191 and 8.94), making possible their perfect convergence at the GUT scale (Figure 12.2).
2. It offers a 'natural' solution to the so-called *hierarchy* problem. The Higgs mass is renormalized by various loop diagrams (Section 6.3.3), which drive it way out of acceptable range unless there are magical cancellations ('fine tuning'). But loop corrections are of opposite sign for bosons and fermions, so supersymmetry, by pairing particles with 'sparticles', makes the cancellation exact and automatic.
3. In most models, the lightest supersymmetric particle is colorless, neutral, and stable, making it an attractive candidate for Dark Matter (Section 12.5).

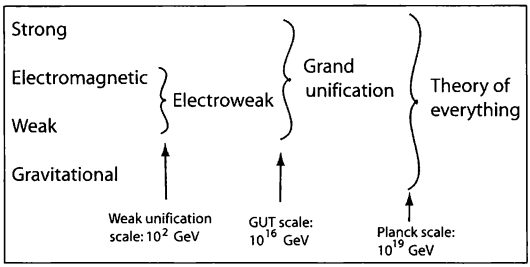


Fig. 12.5 Unification of the four interactions.

Moreover, attempts to formulate a quantum theory of gravity seem to require supersymmetry. On the other hand, the minimal supersymmetric models involve at least 124 independent parameters [19] – five times the (already embarrassing) number in the Standard Model – and they do not easily accommodate neutrino masses. If supersymmetric particles are discovered* at the LHC, it will be a spectacular triumph of inspired audacity. But I wouldn't bet your last dollar on it.

12.4.2 Strings

For decades, a fundamental challenge in theoretical physics has been the formulation of a quantum theory of gravity – the quantized version of General Relativity (analogous to QED, the quantized version of electrodynamics). Generations of physicists have tried, and failed – for point masses the theory seems to be incorrigibly nonrenormalizable. While this is embarrassing, it has not, so far, been catastrophic for particle physics, where gravity is much too feeble to play a significant role. But at extremely close range (which is to say, at very high energy – specifically, the Planck scale: 10^{19} GeV) quantum gravity is bound to come into the picture. Moreover, the old dream of unifying the forces of nature leads inexorably to a putative 'theory of everything' that would include gravity along with the strong, electromagnetic, and weak interactions (Figure 12.5).

String theory proposes to solve these problems (and more) [22]. In string theory the basic units of matter are not (zero-dimensional) particles, but rather one-dimensional 'strings' (or higher-dimensional 'branes'), of which 'particles' are various vibrational modes. The theory underwent an extraordinary evolution between the 1970s, when a few lonely visionaries took up the cause, and 2000, by which time it was well established as the dominant paradigm. Early versions contained only bosons, and consistency required 25 space dimensions. This seemed a trifle extravagant, but it was possible to imagine that 22 of them are 'curled up'

* There was a flurry of excitement in 2001, when discrepancies between the measured and calculated values of the anomalous magnetic moment of the muon seemed to suggest a contribution from supersymmetric particles (20). But it turned out that the calculations were in error – when a sign mistake in one term was corrected the disagreement largely evaporated (21).

(*compactified*), and hence irrelevant on the macroscopic scale.* Fermions were later incorporated via supersymmetry (hence 'superstrings') and the number of space dimensions dropped to 9 or 10. Meanwhile, it was realized that the theory automatically includes the graviton, making it a natural candidate for quantum gravity.

In the early days, one of the great attractions of superstring theory was that it appeared to be uniquely determined – we live, it seemed, in the *only mathematically possible* world. Physics would no longer be a matter of discovering contingent laws by experimental observation, but of working out the inescapable implications of the one allowed theory. Sadly, this particular hope has turned inside out, and 'M-theory' now suggests that there is in fact a whole 'landscape' of permissible models (10^{500} of them, by some estimates), and no way (short of the anthropic principle†) to choose the correct one.

At this point an entire generation of theoretical physicists is way out on a limb. Superstring theory still holds out the best hope for ultimate unification of all four interactions, and it is probably the most promising candidate for quantum gravity. But it has proved diabolically difficult to extract verifiable (or falsifiable) predictions about the low-energy world we inhabit. The discovery of supersymmetric particles, or indications of extra dimensions [23], would lend some support, but anything approaching a confirmation of superstring theory seems, at this point, a very long way off [24].

12.5 Dark Matter/Dark Energy

Persuasive astronomical evidence now indicates that the matter we know about – described by the Standard Model – represents a measly 5% of the mass/energy content of the universe. The rest is *Dark Matter* (about 20%) and *Dark Energy* (75%). The implications for particle physics are humbling: we have only seen the tip of the iceberg. What *is* all this other stuff, and how has it managed to elude us?

12.5.1 Dark Matter

In 1933, Fritz Zwicky measured the velocities of galaxies in the Coma cluster (from the Doppler shift of their atomic spectra), and used this information to determine

* The idea of extra dimensions was not new. T. Kaluza first introduced the notion in 1919, in an effort to unify electrodynamics and gravity, and in 1926 O. Klein suggested compactification as a device for 'hiding' extra dimensions. (If you want to specify the location of an ant on a clothes line, you would probably just report its distance z from one end – only for much smaller bugs would the azimuthal position ϕ be of interest or importance.)

† The anthropic principle holds that the laws and parameters of physics are what they are 'because' (if that's the right word) if they were different, human life would be impossible and we wouldn't be here to discover them.

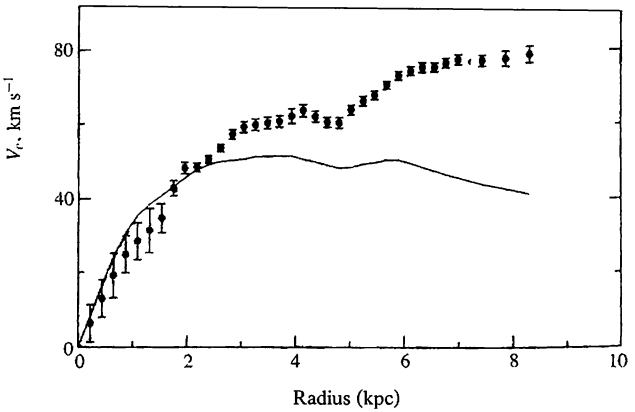


Fig. 12.6 Rotation curve for the galaxy NGC 1560. The solid line represents the curve to be expected on the basis of all observed matter (stars plus gas). [*Source*: A. H. Broeils *Astron. and Astrophys.* **256** 19 (1992).]

the mass of the cluster. The result was surprising: 400 times larger than the *visible* stars in the cluster. Evidently the galaxies contain a lot of matter that does not radiate (and is called, therefore, *dark matter*) [25]. More recently, *rotation curves* have been measured for a number of galaxies (including our own). These plot the (tangential) velocity v as a function of distance r from the galactic center. Newton's law of universal gravitation says that for stars well away from the core v should decrease as $1/\sqrt{r}$ (Problem 12.10); instead, it typically *increases* (Figure 12.6). This suggests that the dark matter permeates a spherical 'halo' extending well outside the galactic nucleus.* Today it is even possible to map out the distribution of dark matter, using gravitational lensing (the bending of light as it passes through).

So far, though, our only evidence for dark matter comes from its large-scale gravitational effects, and it is natural to wonder whether perhaps Newton's laws (and also General Relativity) are incorrect on some scale, and there is actually *no* dark matter out there [26]. Short of such a radical alternative, the question remains: what *is* this stuff? Could it be ordinary cold matter – sand and gravel, perhaps, the remnants of extinct stars or dead planets. Almost certainly not. Cosmological models that are convincingly corroborated by the observed abundances of light elements do not allow for anywhere near enough baryons to account for dark matter [27]. What about neutrinos? Probably not – even though there are enormous numbers of them, they are much too light to contribute more than a small fraction of the observed dark matter.† Evidently we are looking for something much more

* The dark matter discussed here is not to be confused with the 'missing mass' required to 'close' the universe. We will talk about that in the next section.

† Moreover, neutrinos would constitute 'hot' dark matter – they are by nature highly relativistic, and it is hard to imagine that they could be confined to galactic halos (or to the primordial aggregates from which galaxies emerged).

massive than neutrinos, but (like neutrinos) weakly interacting; Bahcall called them WIMPs (Weakly Interacting Massive Particles).* Their mass is tentatively estimated to lie in the range 100–200 GeV/c^2; they are certainly neutral (otherwise they would radiate) and stable (left over from the Big Bang). No such particle, of course, is known to the Standard Model. But supersymmetry does suggest a candidate: the lightest supersymmetric particle (probably a mixture of the photino and the higgsino – or possibly the Zino – called the 'neutralino' ... obviously, this terminology is getting out of hand) is presumably absolutely stable. Large numbers might be left over from the Big Bang. Another possibility is the axion – the hypothetical particle introduced to account for the absence of strong CP violation. But surely the most exciting possibility would be something entirely new and unanticipated.

How is all this going to be decided? Since the late 1980s a number of WIMP searches have been under way. They are based on the realization that the solar system orbits around the galactic center at 220 km s^{-1},† and the earth orbits the sun at 30 km s^{-1}, so we face a 'dark matter headwind' – 235 km s^{-1} in (northern) summer and 205 km s^{-1} in winter. (The seasonal variation is a lucky thing, for it should enable experimentalists to filter the signal out of a much larger – but constant – background due to natural radioactivity and cosmic rays.) Several different detection mechanisms have been tried [27], but it is only recently that their greatly improved sensitivity has approached the requisite level. There have already been some (questionable) events [28], and convincing evidence may well come in the next few years. Meanwhile, the LHC should be in a position to *create* dark matter, and at that point the remaining task will be to demonstrate that the three approaches (galactic, terrestrial, and accelerator) are all talking about the same particle [29].

12.5.2
Dark Energy

Before 1998, it was taken for granted that the expansion of the universe is slowing down, due to the gravitational attraction of all matter; the only question was whether the energy density of the universe is great enough to reverse the expansion completely, leading to a 'big crunch' (see Problem 12.10). Visible matter and dark matter together amount to about a third of the 'critical density', so for those who believed the expansion 'should' reverse‡ there was a second 'missing mass' paradox, unrelated to the dark matter problem: where is all that 'extra' energy?

* In principle, dark matter might interact *only* gravitationally, but as Cline (27) remarks coyly, 'If that is really the case, physicists have no hope of ever detecting it' (that is, as individual particles). For this reason, at least, it is generally assumed that dark matter participates in the weak interaction.

† The dark matter halo (since it is only very weakly coupled to matter) does not (one assumes) share in the galactic rotation.

‡ The widely accepted inflationary cosmology *requires* that the total density of the universe have exactly the critical value.

This problem was turned inside out by the astonishing discovery that the expansion of the universe is not slowing down at all, but rather *accelerating*. Evidently Newtonian gravity (universal attraction) is not right on the largest scale – either that or there is some *new* force that is repulsive in nature and overwhelms gravity in this case. In General Relativity, there is a (sort of) natural place for an extra term that could account for the phenomenon: the *cosmological constant*, Λ. Einstein's original theory (with *no* cosmological constant) implied that the universe expands – something he regarded as absurd. He was able to rescue the theory by introducing an *ad hoc* source term, whose strength (Λ) could be adjusted to stabilize the universe. (Mathematically, the cosmological constant introduces a kind of primordial repulsion, or negative pressure, that balances the universal attraction on a cosmic scale.) Later, when Hubble discovered that the universe is *in fact* expanding, a chagrined Einstein disowned the cosmological constant, calling it 'my greatest blunder'. But when the accelerated expansion was discovered, the obvious remedy was to resurrect the cosmological constant [30].

There is, however, a subtle distinction between the original notion of a cosmological constant and its contemporary reincarnation. Einstein conceived of Λ as an unexplained fundamental constant of nature – analogous to Planck's constant or Boltzmann's constant; there were *two* distinct sources of gravitation: matter (actually, the *stress tensor*, incorporating energy, momentum, and stress of all forms), and Λ. In the modern version Λ is taken to have a *dynamical* origin, in the form of *dark energy* associated with the vacuum expectation value of some quantum field. It is, in effect, a constant term in the stress tensor, pervading all space uniformly,* that we choose to peel off and treat separately. But what the nature of this field (or fields) might be is at this stage a mystery. *Worse* than a mystery, because attempts to construct model theories tend to yield values of Λ that are 120 *orders of magnitude* too great![31] Obviously, we have a lot to learn.

12.6 Conclusion

Most particle physicists anticipate that the LHC will produce Higgs bosons. Many believe it will create the first supersymmetric particles. Some think it will yield evidence of extra dimensions. Perhaps. But there is another possibility that very few take seriously: substructure – the idea that quarks and leptons (and maybe also the mediators) are composite particles, made of even more elementary constituents. This would change everything, just as the quark model changed everything 40 years ago, and Rutherford's atomic model changed everything a century ago. In any event, we almost certainly stand at the threshold of a fundamental revolution in elementary particle physics [32].

* This is in contrast to dark *matter*, which is concentrated in galactic halos.

References

1 See, for example, the special section in (2007) *Science*, **315**, 55.

2 The Higgs mechanism was proposed independently by Englert, F. and Brout, R. (1964) *Physical Review Letters*, **13**, 321; and (a) Higgs, P. W. (1964) *Physics Letters*, **12**, 132; (1964) *Physical Review Letters*, **13**, 508.

3 The 'bible' for this material is Gunion, J. F. et al. (1990) *The Higgs Hunter's Guide*, Addison-Wesley, Redwood City, CA.

4 LEP had already excluded a Higgs boson with mass below 114 GeV/c^2; the statistically marginal observation in 2000 suggests a mass of 115 Gev/c^2. See, for example, Renton, P. (2004) *Nature*, **428**, 141.

5 Estimates of the Higgs mass are sensitive to the exact value of the top quark mass, and of course they depend on the model – supersymmetry puts an upper limit on the mass of the lightest Higgs particle at 140 GeV/c^2. See, for instance, Schwarzschild, B. (August 2004), *Physics Today*, 26.

6 For an accessible introduction see Peskin, M. (2006) Physics 450 lecture notes, Fall 2006, Stanford University, available on-line.

7 Georgi, H. and Glashow, S. L. (1974) *Physical Review Letters*, **32**, 438.

8 Shiozawa, M. et al. (1998) *Physical Review Letters*, **81**, 3319.

9 For a comprehensive review see Ross, G. (1985) *Grand Unified Theories*, Perseus, Reading, MA; (a) Langacker, P. (1981) *Physics Reports*, **72**, 185.

10 't Hooft, G. (1974) *Nuclear Physics*, **B79**, 276; (a) Polyakov, A. M. (1974) *JETP Letters*, **20**, 194.

11 Cabrera, B. (1982) *Physical Review Letters*, **48**, 1378.

12 For an account of the search for stray antihelium from some distant antigalaxy see Barry, P. (2007) *Science News*, **171**, 296.

13 Sakharov, A. D. (1967) *JETP Letters*, 5, 24.

14 See, for example, the article by Kirkby, D. and Nir, Y. (2006) *Review of Particle Physics*, 146.

15 See, for instance, Peskin, M. (2002) *Nature*, **419**, 24; or (a) Harrison, P. (July 2003) *Physics World*, 27.

16 Peccei, R. D. and Quinn, H. R. (1977) *Physics Review Letters*, **38**, 1440; and (1977) *Physical Review D*, **16**, 1791. For a delightful account of the Quinn/Peccei mechanism see (a) Sikivie, P. (December 1996) *Physics Today*, 22; for the current state of axion searches, see (b) van Bibber, K. and Rosenberg, L. J. (August 2006) *Physics Today*, 30.

17 Wess, J. and Zumino, B. (1974) *Physics Letters B*, **49**, 52.

18 For an excellent semi-popular account of supersymmetry, see Kane, G. (2000) *Supersymmetry: Unveiling the Ultimate Laws of Nature*, Perseus, Cambridge, MA; for a reasonably accessible introduction to the technical details see *A Supersymmetry Primer*, by S. P. Martin, on the web at hep-ph/9709356.

19 See the article by Haber, H. E. (2006) RPP, p. 1105.

20 Brown, H. N. et al. (2001) *Physical Review Letters*, **86**, 2227.

21 Schwarzschild, B. (February 2002) *Physics Today*, 18.

22 For a thrilling semi-popular history of string theory, see Greene, B. R. (1999) *The Elegant Universe*, W. W. Norton, New York; for an accessible introduction to the theory itself, see (a) Zwiebach, B. (2004) *A First Course in String Theory*, Cambridge University Press, Cambridge, UK; the standard graduate textbook is (b) Polchinski, J. (1998) *String Theory*, Cambridge University Press, Cambridge, UK.

23 For an accessible account of extra dimensions see Randall, L. (July 2007) *Physics Today*, 80. For the status of searches for extra dimensions see the article by (a) Giudice, G. F. and Wells, J. D. (2006) *Review of Particle Physics*, 1165.

24 In 2006, there were signs of a backlash against string theory and its pervasive influence on particle physics: Richter, B. (October **2006**) *Physics Today*, **8**; (a) Smolin, L. (**2006**) *The Trouble With Physics*, Houghton Mifflin, New York; (b) Woit, P. (**2006**) *Not Even Wrong*, Basic Books, New York.
25 Zwicky, F. (**1933**) *Helvetica Physics Acta*, **6**, 124. For a masterful review of the first 55 years, see (a) Trimble, V. (**1987**) *Annual Review of Astronomy and Astrophysics*, **25**, 425.
26 Milgrom, M. (August **2002**) *Scientific American*, 42; (March **2007**) *Science News*, 206.
27 For a beautifully clear survey of dark matter candidates, and experimental searches for dark matter, see Cline, D. B. (March **2003**) *Scientific American*, 51.
28 Schwarzschild, B. (August **2007**) *Physics Today*, 16.
29 Peskin, M. E. (**2007**), *Journal of the Physical Society of Japan*, **76**, 111017.
30 This is not the only possibility; for an outstanding review see the article by Turner, M. S. and Huterer, D. (**2007**), *Journal of the Physical Society of Japan*, **76**, 111015.
31 For an accessible explanation see Adler, R. J., Casey, B. and Jacob, O. C. (**1995**) *American Journal of Physics*, **63**, 620.
32 Ellis, J. (**2007**) *Nature*, **448**, 297.
33 Planck, M. (**1899**) *Sitzungsber. Dtsch. Akad. Wiss. Berlin–Math-Phys. Tech. Kl.*, 440.

Problems

12.1 (a) Use Equation 10.132 to determine the mass of the W, in terms of $v = \mu/\lambda$ and $q = g_w\sqrt{\hbar c/4\pi}$. Thus, confirm Equation (12.1)

(b) Use Problem 10.21 and Equation 10.130 to determine the vertex factor for the coupling of the Higgs to a quark or lepton.

(c) Use Equation 10.136 to determine the vertex factors for the couplings hWW, hZZ, and hhh.

12.2 (a) Calculate the decay rate for $h \to f + \bar{f}$ (where f is a quark or lepton), in the MSM.

$$\left[\textit{Answer}: \frac{\alpha_w}{8\hbar} m_h c^2 \left(\frac{m_f}{M_W}\right)^2 \left[1 - \left(\frac{2m_f}{m_h}\right)^2\right]^{3/2}\right]$$

(b) If $m_h = 120\ \text{GeV}/c^2$, what are the branching ratios $\Gamma(b\bar{b})/\Gamma(c\bar{c})$ and $\Gamma(b\bar{b})/\Gamma(\tau^+\tau^-)$? [Include a factor of 3, for color, in the case of quarks.]

12.3 (a) Calculate the decay rates for $h \to W^+ + W^-$ and $h \to Z + Z$, in the MSM.

[*Answer:*

$$\Gamma(W^+W^-) = \frac{\alpha_w m_h c^2}{16\hbar}\left(\frac{m_h}{M_W}\right)^2 \left(1 - 4\frac{M_W^2}{m_h^2} + 12\frac{M_W^4}{m_h^4}\right)\left[1 - 4\frac{M_W^2}{m_h^2}\right]^{1/2}$$

$$\Gamma(ZZ) = \frac{\alpha_w m_h c^2}{32\hbar}\left(\frac{m_h}{M_W}\right)^2 \left(1 - 4\frac{M_Z^2}{m_h^2} + 12\frac{M_Z^4}{m_h^4}\right)\left[1 - 4\frac{M_Z^2}{m_h^2}\right]^{1/2}\Bigg]$$

(b) If $m_h = 120\ \text{GeV}/c^2$, what is the ratio $\Gamma(W^+W^-)/\Gamma(ZZ)$?

12.4 Estimate the longest proton lifetime that could be measured in a realistic laboratory experiment. [*Hint:* How many protons could you sample, in a practical experiment (Super-K, for instance)? How long would you (or – more to the point – your funding agency) be prepared to wait?]

12.5 Estimate the lifetime of the proton, in the Glashow/Georgi model.
[*Hint:* Don't try to *calculate* anything here – you don't have anywhere near enough information. The real question is how the lifetime formula depends on the various masses. Study other decays – the muon, the neutron, the pion – and exploit dimensional analysis, if it helps.]

12.6 Derive Equation (12.9), from Equations (12.7) and (12.8)

12.7 Consider vectors in the xy plane:

(a) Show that a (counterclockwise) rotation θ carries a vector $\mathbf{a} = (a_x, a_y)$ into $\mathbf{a}' = (a'_x, a'_y)$ given by

$$a'_x = \cos\theta\, a_x - \sin\theta\, a_y, \quad a'_y = \sin\theta\, a_x + \cos\theta\, a_y$$

(b) Show that the dot product of two vectors is invariant under such a rotation: $\mathbf{a}' \cdot \mathbf{b}' = \mathbf{a} \cdot \mathbf{b}$.

(c) Now consider an *infinitesimal* rotation $d\theta$. Expand the transformation rule in (a) to first order in $d\theta$.

(d) Show that the dot product is invariant (to first order) under infinitesimal rotations. [Of course, if you already know it's invariant under *finite* transformations, the proof for infinitesimal transformations is redundant. The point is that the infinitesimal case is typically much simpler.]

12.8 The purpose of this problem is to prove that the action described by the Lagrangian in Eq. (12.11) is invariant under the supersymmetry transformations in Eq. (12.10).

(a) Show that $\delta\phi^* = 2\overline{\psi}\,\epsilon$ and $\delta\overline{\psi} = (i/\hbar c)\overline{\epsilon}\gamma^\mu(\partial_\mu\phi^*)$.

(b) Consider first the scalar 'kinetic' term, $\mathcal{L}_1 = \frac{1}{2}(\partial^\mu\phi^*)(\partial_\mu\phi)$; show that $\delta\mathcal{L}_1 = (\partial^\mu\phi)(\partial_\mu\overline{\psi})\epsilon + \overline{\epsilon}(\partial^\mu\phi^*)(\partial_\mu\psi)$.

(c) Next treat the spinor 'kinetic' term, $\mathcal{L}_2 = i\hbar c\,\overline{\psi}\gamma^\mu(\partial_\mu\psi)$. Show that $\delta\mathcal{L}_2 = -\delta\mathcal{L}_1 + \partial_\mu Q^\mu$, where $Q^\mu \equiv \overline{\psi}(\partial^\mu\phi)\epsilon + \frac{1}{2}\overline{\epsilon}\sigma^{\mu\nu}[\phi^*(\partial_\nu\psi) - (\partial_\nu\phi^*)\psi]$, where $\sigma^{\mu\nu}$ is defined in Eq. 7.69.

(d) Now examine the mass terms, $\mathcal{L}_3 = -\frac{1}{2}(mc/\hbar)^2\phi^*\phi$ and $\mathcal{L}_4 = -mc^2\,\overline{\psi}\psi$. Show that $\delta\mathcal{L}_3 = -(mc/\hbar)^2(\overline{\psi}\epsilon\phi + \phi^*\overline{\epsilon}\psi)$ and $\delta\mathcal{L}_4 = i(mc/\hbar)[-\overline{\epsilon}\gamma^\mu(\partial_\mu\phi^*)\psi + \overline{\psi}\gamma^\mu\epsilon(\partial_\mu\phi)]$.

(e) Finally, invoke the Dirac equation, which follows from the Euler-Lagrange equations (Eq. 10.15), to show that $\delta\mathcal{L}_4 = -\delta\mathcal{L}_3 + \partial_\mu R^\mu$, where $R^\mu = i(mc/\hbar)[-\overline{\epsilon}\gamma^\mu\phi^*\psi + \overline{\psi}\gamma^\mu\epsilon\phi]$.

Although the full Lagrangian ($\mathscr{L} = \mathscr{L}_1 + \mathscr{L}_2 + \mathscr{L}_3 + \mathscr{L}_4$) is not invariant, it changes only by a total divergence, $\delta\mathscr{L} = \partial_\mu(Q^\mu + R^\mu)$, so the action and the equations of motion *are* invariant. Notice, however, that the scalar and the spinor have to carry the *same* mass for this to work.

12.9 (a) From c, $\hbar$, and G (Newton's constant of universal gravitation), construct a quantity l_P with the dimensions of length, a quantity t_P with the dimensions of time, and a quantity m_P with the dimensions of mass. These are known as the *Planck length*, the *Planck time*, and the *Planck mass*, respectively, after Max Planck, who first published them in 1899 – the year *before* the eponymous constant itself [33]. Work out the actual *numbers*, in meters, seconds, and kilograms. Also calculate the *Planck energy* ($E_P = m_P c^2$), in GeV. [These quantities set the scale at which quantum gravity is expected to be relevant.]

(b) What is the gravitational analog to the fine structure constant? Find the actual number, using (i) the mass of the electron, (ii) the Planck mass.

12.10 Find the velocity v as a function of orbital radius r, for an object in a circular trajectory around a fixed center of mass M (for example, a planet about the sun).

12.11 A quick naive way to calculate the critical density is to picture the universe as a uniform sphere of radius R, and set the escape velocity for a particle at the surface equal to the expansion velocity (from Hubble's law), $v = HR$. On this basis, show that

$$\rho_c = \frac{3H^2}{8\pi G}$$

Look up the value of Hubble's constant (H), and determine the critical density, in $\mathrm{kg/m^{-3}}$.

Appendix A

The Dirac Delta Function

Introduction to the Dirac delta function

The Dirac delta function, $\delta(x)$, is an infinitely high, infinitesimally narrow spike at the origin, with area 1 (Figure A.1). Specifically

$$\delta(x) = \begin{cases} 0, & \text{if } x \neq 0 \\ \infty, & \text{if } x = 0 \end{cases} \quad \text{and} \quad \int_{-\infty}^{\infty} \delta(x)\,\mathrm{d}x = 1 \tag{A.1}$$

Technically, it's not a function at all, since its value is not finite at $x = 0$. In the mathematical literature it is known as a *generalized* function, or *distribution*. It is, if you like, the *limit* of a *sequence* of functions, such as rectangles of height n and width $1/n$, or isosceles triangles of height n and base $2/n$ (Figure A.2), or any other shape you might wish to use.

If $f(x)$ is some 'ordinary' function (that is, *not* another delta function – in fact, just to be on the safe side let's say that $f(x)$ is *continuous*) then the *product* $f(x)\delta(x)$ is zero everywhere except at $x = 0$. It follows that

$$f(x)\delta(x) = f(0)\delta(x) \tag{A.2}$$

(This is the most important fact about the delta function, so make sure you understand why it is true. The point is that since the product is zero anyway *except* at $x = 0$, we may as well replace $f(x)$ by the value it assumes at the origin.) In particular

$$\int_{-\infty}^{\infty} f(x)\delta(x)\,\mathrm{d}x = f(0)\int_{-\infty}^{\infty} \delta(x)\,\mathrm{d}x = f(0) \tag{A.3}$$

Under an integral, the delta function 'picks out' the value of $f(x)$ at $x = 0$. (Here and below, the integral need not run from $-\infty$ to $+\infty$; it is sufficient that the domain extends across the delta function, and $-\epsilon$ to $+\epsilon$ would do just as well.)

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

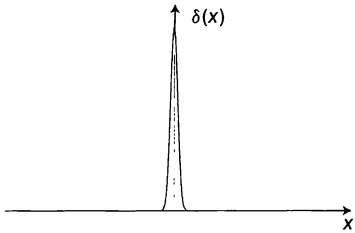


Fig. A.1 The Dirac delta function (you must imagine, however, that the spike is infinitely high and infinitesimally narrow).

Of course, we can move the spike from $x = 0$ to some other point, $x = a$:

$$\delta(x-a) = \begin{cases} 0, & \text{if } x \neq a \\ \infty, & \text{if } x = a \end{cases} \quad \text{and} \quad \int_{-\infty}^{\infty} \delta(x-a)\,\mathrm{d}x = 1 \tag{A.4}$$

(see Figure A.3). Equation A.2 generalizes to

$$f(x)\delta(x-a) = f(a)\delta(x-a) \tag{A.5}$$

and Equation A.3 becomes

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)\,\mathrm{d}x = f(a) \tag{A.6}$$

How should we interpret the expression $\delta(kx)$, if k is some nonzero (real) number? Suppose we multiply by an 'ordinary' function $f(x)$ and integrate:

$$\int_{-\infty}^{\infty} f(x)\delta(kx)\,\mathrm{d}x$$

We may change variables, letting $y \equiv kx$, so that $x = y/k$, and $\mathrm{d}x = 1/k\,\mathrm{d}y$. If k is positive, the integration still runs from $-\infty$ to $+\infty$, but if k is *negative*, then $x = \infty$

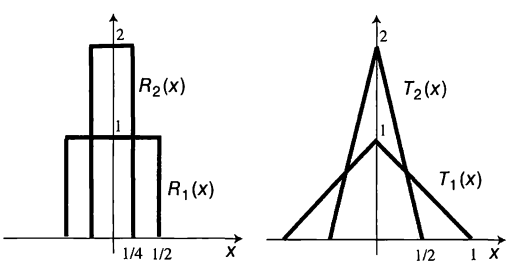


Fig. A.2 Two sequences of functions whose limit is $\delta(x)$.

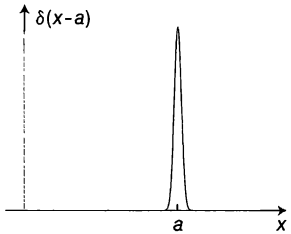


Fig. A.3 'Graph' of $\delta(x - a)$.

implies $y = -\infty$, and vice versa, so the limits are reversed – restoring the "proper" order costs a minus sign. Thus

$$\int_{-\infty}^{\infty} f(x)\delta(kx)\,\mathrm{d}x = \pm\int_{-\infty}^{\infty} f(y/k)\delta(y)\frac{\mathrm{d}y}{k}$$
$$= \pm\frac{1}{k}f(0) = \frac{1}{|k|}f(0) \tag{A.7}$$

(The lower signs apply when k is negative, and we account for this neatly by putting absolute value bars around the k, as indicated.) In this context, then, $\delta(kx)$ serves the same purpose as $(1/|k|)\delta(x)$:

$$\int_{-\infty}^{\infty} f(x)\delta(kx)\,\mathrm{d}x = \int_{-\infty}^{\infty} f(x)\left[\frac{1}{|k|}\delta(x)\right]\mathrm{d}x \tag{A.8}$$

Because this holds for *any* $f(x)$, it follows that the delta function expressions are equal:*

$$\delta(kx) = \frac{1}{|k|}\delta(x) \tag{A.9}$$

What we have just analyzed is really a special case of the general form $\delta(g(x))$, where $g(x)$ is some function of x. In general, $\delta(g(x))$ has spikes at the zeros, x_1, x_2, $x_3, \ldots,$ of $g(x)$:

$$g(x_i) = 0 \quad (i = 1, 2, 3, \ldots, n) \tag{A.10}$$

* You ought to ponder that last step for a moment. Ordinarily, the equality of two integrals certainly does *not* imply equality of the integrands. The crucial point here is that the integrals are equal for *any* $f(x)$. Suppose the delta function expressions $\delta(kx)$ and $(1/|k|)\delta(x)$ actually *differed*, say, in the neighborhood of the point $x = 17$. Then I would pick a function $f(x)$ that was sharply peaked about $x = 17$, and the integrals would not be equal. Since, on the contrary, the integrals *must* be equal, it follows that the delta function expressions are themselves equal. Well, technically they might still differ at isolated points, provided these contribute nothing to the integral, but we can silence this objection by noting that both sides of Equation A.9 are clearly zero except at $x = 0$.

In the neighborhood of the ith zero, we may expand $g(x)$ as a Taylor series:

$$g(x) = g(x_i) + (x - x_i)g'(x_i) + \tfrac{1}{2}(x - x_i)^2 g''(x_i) + \cdots \cong (x - x_i)g'(x_i) \tag{A.11}$$

In view of Equation A.9, the spike at x_i has the form

$$\delta(g(x)) = \frac{1}{|g'(x_i)|}\delta(x - x_i) \qquad (x \cong x_i) \tag{A.12}$$

The factor $|g'(x_i)|^{-1}$ tells us the 'strength' of the delta function at x_i. Putting this together with the spikes at the other zeros, we conclude that

$$\delta(g(x)) = \sum_{i=1}^{n} \frac{1}{|g'(x_i)|}\delta(x - x_i) \tag{A.13}$$

Thus, any expression of the form $\delta(g(x))$ can be reduced to a sum of simple delta functions.*

Example A.1 Simplify the expression $\delta(x^2 + x - 2)$.

Solution: Here $g(x) = x^2 + x - 2 = (x - 1)(x + 2)$; there are two zeros, at $x_1 = 1$ and $x_2 = -2$. Differentiating, $g'(x) = 2x + 1$, so $g'(x_1) = 3$ and $g'(x_2) = -3$. Thus

$$\delta(x^2 + x - 2) = \frac{1}{3}\delta(x - 1) + \frac{1}{3}\delta(x + 2)$$

The Dirac delta function can be thought of as the *derivative* of the Heaviside step function (Figure A.4):†

$$\theta(x) \equiv \begin{cases} 0, & (x < 0) \\ 1, & (x > 0) \end{cases} \tag{A.14}$$

Obviously, $d\theta/dx$ is zero everywhere *except* at the origin, while

$$\int_{-\infty}^{\infty} \frac{d\theta}{dx}\, dx = \theta(\infty) - \theta(-\infty) = 1 - 0 = 1 \tag{A.15}$$

so $d\theta/dx$ satisfies the defining conditions (Equation A.1) for $\delta(x)$.

It is an easy matter to generalize the delta function to three (or more) dimensions:

$$\delta^3(\mathbf{r}) = \delta(x)\delta(y)\delta(z) \tag{A.16}$$

* Equation A.13 is *exact*, notwithstanding the truncated Taylor series (Equation A.11) I used in its derivation. At x_i, the 'extra' terms are zero, since they contain powers of $(x - x_i)$.

† The value at the discontinuity seldom matters, but if it worries you, define $\theta(0) \equiv 1/2$.

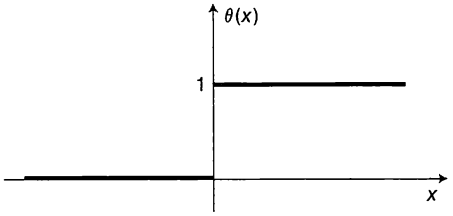


Fig. A.4 The Heaviside theta ('step') function.

This three-dimensional delta function is zero everywhere except at the origin, where it blows up. The triple integral over $\delta^3(\mathbf{r})$ is 1:

$$\int \delta^3(\mathbf{r})\,\mathrm{d}^3r = \int \delta(x)\delta(y)\delta(z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = 1 \tag{A.17}$$

and

$$\int f(\mathbf{r})\delta^3(\mathbf{r}-\mathbf{r}_0)\,\mathrm{d}^3r = f(\mathbf{r}_0) \tag{A.18}$$

For example, the charge density (charge per unit volume) of a point charge q located at the point $\mathbf{r}_0$ can be written as

$$\rho(\mathbf{r}) = q\delta^3(\mathbf{r}-\mathbf{r}_0) \tag{A.19}$$

Problems

A.1 (a) $\int_0^3 (2x^2+7x+3)\delta(x-1)\,\mathrm{d}x = ?$
(b) $\int_0^3 \ln(1+x)\delta(\pi - x)\,\mathrm{d}x = ?$

A.2 Use Equation A.13 to simplify the expression $\delta(\sqrt{x^2+1}-x-1)$.

A.3 Use Equation A.13 to simplify the expression $\delta(\sin x)$. Sketch this function.

A.4 Let $f(y) = \int_0^2 \delta(y - x(2-x))\,\mathrm{d}x$. Find $f(y)$, and plot it from $y=-2$ to $y=+2$.

A.5 $\int_{-1}^5 x^4 \left[\frac{\mathrm{d}^2}{\mathrm{d}x^2}\delta(x-3)\right]\mathrm{d}x = ?$ [*Hint*: Integrate by parts.]

A.6 Evaluate the integral (to 5 significant digits)

$$\int_{-1}^{5} \theta(2x-4)e^{-3x}\,\mathrm{d}x$$

A.7 Evaluate $\int \mathbf{r}\cdot(\mathbf{a}-\mathbf{r})\delta^3(\mathbf{r}-\mathbf{b})\,\mathrm{d}^3r$, if $\mathbf{a} = (1,\ 2,\ 3)$, $\mathbf{b} = (3,\ 2,\ 1)$, and the integration is over a sphere of radius 1.5 centered at (2, 2, 2).

Appendix B

Decay Rates and Cross Sections

Summary of formulas for decay rates and scattering cross sections.

B.1
Decays

Suppose particle 1 decays into particles 2, 3, 4, . . . , n:

$$1 \to 2 + 3 + 4 + \cdots + n$$

The decay rate is given by the formula

$$\begin{aligned} d\Gamma &= |\mathscr{M}|^2 \frac{S}{2\hbar m_1} \left\{ \left[\frac{c\, d^3\mathbf{p}_2}{(2\pi)^3 2E_2} \right] \left[\frac{c\, d^3\mathbf{p}_3}{(2\pi)^3 2E_3} \right] \cdots \left[\frac{c\, d^3\mathbf{p}_n}{(2\pi)^3 2E_n} \right] \right\} \\ &\quad \times (2\pi)^4 \delta^4(p_1 - p_2 - p_3 - \cdots - p_n) \end{aligned} \tag{B.1}$$

where $p_i = (E_i/c, \mathbf{p}_i)$ is the 4-momentum of the ith particle (which carries mass m_i, so $E_i = c\sqrt{\mathbf{p}_i^2 + m_i^2 c^2}$). The decaying particle is presumed to be at rest: $p_1 = (m_1 c; \mathbf{0})$; S is a product of statistical factors: $1/j!$ for each group of j identical particles in the final state.

B.1.1
Two-body Decays

If there are just *two* particles in the final state, the integrals can be performed explicitly. The total decay rate is

$$\Gamma = \frac{S|\mathbf{p}|}{8\pi\hbar m_1^2 c} |\mathscr{M}|^2 \tag{B.2}$$

where $|\mathbf{p}|$ is the magnitude of either outgoing momentum:

$$|\mathbf{p}| = \frac{c}{2m_1} \sqrt{m_1^4 + m_2^4 + m_3^4 - 2m_1^2 m_2^2 - 2m_1^2 m_3^2 - 2m_2^2 m_3^2} \tag{B.3}$$

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

In particular, if the outgoing particles are *massless*, then $|\mathbf{p}| = m_1 c/2$, and

$$\Gamma = \frac{S}{16\pi\hbar m_1}|\mathscr{M}|^2 \tag{B.4}$$

B.2 Cross Sections

Suppose particles 1 and 2 collide, producing particles 3, 4, ..., n:

$$1 + 2 \rightarrow 3 + 4 + \cdots + n$$

The cross section is given by the formula

$$\begin{aligned} d\sigma &= |\mathscr{M}|^2 \frac{\hbar^2 S}{4\sqrt{(p_1 \cdot p_2)^2 - (m_1 m_2 c^2)^2}} \\ &\times \left\{ \left[\frac{c\, d^3\mathbf{p}_3}{(2\pi)^3 2E_3} \right] \left[\frac{c\, d^3\mathbf{p}_4}{(2\pi)^3 2E_4} \right] \cdots \left[\frac{c\, d^3\mathbf{p}_n}{(2\pi)^3 2E_n} \right] \right\} \\ &\times (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4 - \cdots - p_n) \end{aligned} \tag{B.5}$$

where (as before) $p_i = (E_i/c, \mathbf{p}_i)$ is the 4-momentum of particle i (mass m_i), $E_i = c\sqrt{m_i^2 c^2 + \mathbf{p}_i^2}$, and S is a statistical factor ($1/j!$ for each group of j identical particles in the final state).

B.2.1 Two-body Scattering

If there are just *two* particles in the final state, the integrals can be performed explicitly.

(a) *In the center-of-momentum frame*

$$\sqrt{(p_1 \cdot p_2)^2 - (m_1 m_2 c^2)^2} = (E_1 + E_2)|\mathbf{p}_1|/c \tag{B.6}$$

and

$$\frac{d\sigma}{d\Omega} = \left(\frac{\hbar c}{8\pi}\right)^2 \frac{S|\mathscr{M}|^2}{(E_1 + E_2)^2} \frac{|\mathbf{p}_f|}{|\mathbf{p}_i|} \tag{B.7}$$

where $|\mathbf{p}_i|$ is the magnitude of either incoming momentum, and $|\mathbf{p}_f|$ is the magnitude of either outgoing momentum. In

particular, for *elastic* scattering ($A + B \to A + B$), $|\mathbf{p}_i| = |\mathbf{p}_f|$, so, letting $E \equiv (E_1 + E_2)/2$:

$$\frac{d\sigma}{d\Omega} = \left(\frac{\hbar c}{16\pi}\right)^2 \frac{S|\mathscr{M}|^2}{E^2} \tag{B.8}$$

(b) *In the lab frame (particle 2 at rest)*

$$\sqrt{(p_1 \cdot p_2)^2 - (m_1 m_2 c^2)^2} = m_2 c|\mathbf{p}_1| \tag{B.9}$$

In the case of *elastic* scattering ($A + B \to A + B$),

$$\frac{d\sigma}{d\Omega} = \left(\frac{\hbar}{8\pi}\right)^2 \frac{\mathbf{p}_3^2 S|\mathscr{M}|^2}{m_2|\mathbf{p}_1|\,\left|\mathbf{p}_3|(E_1 + m_2c^2) - |\mathbf{p}_1|E_3\cos\theta\right|} \tag{B.10}$$

If, in particular, the incident particle is *massless* ($m_1 = 0$), this reduces to

$$\frac{d\sigma}{d\Omega} = \left(\frac{\hbar E_3}{8\pi m_2 c E_1}\right)^2 S|\mathscr{M}|^2 \tag{B.11}$$

If the target recoil is negligible ($m_2c^2 \gg E_1$), then Equation B.10 reduces to

$$\frac{d\sigma}{d\Omega} = \left(\frac{\hbar}{8\pi m_2 c}\right)^2 |\mathscr{M}|^2 \tag{B.12}$$

If the outgoing particles are massless ($m_3 = m_4 = 0$), Equation B.5 yields

$$\frac{d\sigma}{d\Omega} = \left(\frac{\hbar}{8\pi}\right)^2 \frac{S|\mathscr{M}|^2|\mathbf{p}_3|}{m_2|\mathbf{p}_1|(E_1 + m_2c^2 - |\mathbf{p}_1|c\cos\theta)} \tag{B.13}$$

Appendix C

Pauli and Dirac Matrices

Pauli and Dirac matrices.

C.1
Pauli Matrices

The Pauli matrices are three Hermitian, unitary, traceless 2×2 matrices:

$$\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{C.1}$$

(Often we use *numerical* indices: $\sigma_1 = \sigma_x, \sigma_2 = \sigma_y, \sigma_3 = \sigma_z$; $\boldsymbol{\sigma}$ is not part of a four-vector, and we do not distinguish upper and lower indices: $\sigma_1 = \sigma^1, \sigma_2, = \sigma^2$, $\sigma_3 = \sigma^3$.)

(a) *Product rules.*

$$\sigma_i \sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k \tag{C.2}$$

(A 2×2 unit matrix is implied in the first term, and summation over k in the second). Thus, in particular:

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1 \tag{C.3}$$

$$\sigma_x\sigma_y = i\sigma_z, \qquad \sigma_y\sigma_z = i\sigma_x, \qquad \sigma_z\sigma_x = i\sigma_y \tag{C.4}$$

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k \quad \text{(commutator)} \tag{C.5}$$

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij} \quad \text{(anticommutator)} \tag{C.6}$$

and for any two vectors $\mathbf{a}$ and $\mathbf{b}$,

$$(\mathbf{a} \cdot \boldsymbol{\sigma})(\mathbf{b} \cdot \boldsymbol{\sigma}) = \mathbf{a} \cdot \mathbf{b} + i\boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b}) \tag{C.7}$$

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

(b) *Exponentials.*

$$e^{i\theta\cdot\sigma} = \cos\theta + i\hat{\theta}\cdot\sigma\sin\theta \tag{C.8}$$

C.2 Dirac Matrices

The Dirac matrices are four unitary traceless 4×4 matrices:

$$\gamma^0 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \tag{C.9}$$

(Here 1 is the 2×2 unit matrix, and 0 is the 2×2 matrix of zeros; σ^i are the Pauli matrices. Lowering indices changes the sign of the 'spatial' components: $\gamma_0 = \gamma^0$, $\gamma_i = -\gamma^i$.) We introduce as well the auxiliary matrices

$$\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{C.10}$$

$$\Sigma \equiv \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix} \tag{C.11}$$

$$\sigma^{\mu\nu} = \frac{i}{2}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu) \tag{C.12}$$

For any four-vector a^μ, we define the 4×4 matrix $\not{a}$ as follows:

$$\not{a} \equiv a_\mu\gamma^\mu \tag{C.13}$$

(a) *Product rules.* In terms of the metric

$$g^{\mu\nu} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{C.14}$$

(note that $g^{\mu\nu}g_{\mu\nu} = 4$), we have:

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}, \qquad \not{a}\not{b} + \not{b}\not{a} = 2a\cdot b \tag{C.15}$$

$$\gamma_\mu\gamma^\mu = 4 \tag{C.16}$$

$$\gamma_\mu\gamma^\nu\gamma^\mu = -2\gamma^\nu, \qquad \gamma_\mu\not{a}\gamma^\mu = -2\not{a} \tag{C.17}$$

$$\gamma_\mu\gamma^\nu\gamma^\lambda\gamma^\mu = 4g^{\nu\lambda}, \qquad \gamma_\mu\not{a}\not{b}\gamma^\mu = 4a\cdot b \tag{C.18}$$

$$\gamma_\mu\gamma^\nu\gamma^\lambda\gamma^\sigma\gamma^\mu = -2\gamma^\sigma\gamma^\lambda\gamma^\nu, \quad \gamma_\mu\not{a}\not{b}\not{c}\gamma^\mu = -2\not{c}\not{b}\not{a} \tag{C.19}$$

(b) *Trace theorems.* The trace of the product of an *odd* number of gamma matrices is *zero.*

$$\mathrm{Tr}(1) = 4 \tag{C.20}$$

$$\mathrm{Tr}(\gamma^\mu\gamma^\nu) = 4g^{\mu\nu}, \qquad \mathrm{Tr}(\not{a}\not{b}) = 4a\cdot b \tag{C.21}$$

$$\mathrm{Tr}(\gamma^\mu\gamma^\nu\gamma^\lambda\gamma^\sigma) = 4(g^{\mu\nu}g^{\lambda\sigma} - g^{\mu\lambda}g^{\nu\sigma} + g^{\mu\sigma}g^{\nu\lambda}),$$

$$\mathrm{Tr}(\not{a}\not{b}\not{c}\not{d}) = 4\left[(a\cdot b)(c\cdot d) - (a\cdot c)(b\cdot d) + (a\cdot d)(b\cdot c)\right] \tag{C.22}$$

Since γ^5 is the product of an *even* number of γ matrices, it follows that $\mathrm{Tr}(\gamma^5\gamma^\mu) = 0$ and $\mathrm{Tr}(\gamma^5\gamma^\mu\gamma^\nu\gamma^\lambda) = 0$. When γ^5 is multiplied by an *even* number of γ's, we find

$$\mathrm{Tr}(\gamma^5) = 0 \tag{C.23}$$

$$\mathrm{Tr}(\gamma^5\gamma^\mu\gamma^\nu) = 0, \qquad \mathrm{Tr}(\gamma^5\not{a}\not{b}) = 0 \tag{C.24}$$

$$\mathrm{Tr}(\gamma^5\gamma^\mu\gamma^\nu\gamma^\lambda\gamma^\sigma) = 4i\epsilon^{\mu\nu\lambda\sigma},$$

$$\mathrm{Tr}(\gamma^5\not{a}\not{b}\not{c}\not{d}) = 4i\epsilon^{\mu\nu\lambda\sigma}a_\mu b_\nu c_\lambda \mathrm{d}_\sigma \tag{C.25}$$

where $\epsilon^{\mu\nu\lambda\sigma} = -1$, if $\mu\nu\lambda\sigma$ is an even permutation of 0123, $+1$ for an odd permutation, and 0 if any two indices are the same. Note that

$$\epsilon^{\mu\nu\lambda\sigma}\epsilon_{\mu\nu\kappa\tau} = -2(\delta^\lambda_\kappa\delta^\sigma_\tau - \delta^\lambda_\tau\delta^\sigma_\kappa) \tag{C.26}$$

(c) *Anticommutation relations.*

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}, \qquad \{\gamma^\mu, \gamma^5\} = 0 \tag{C.27}$$

Appendix D

Feynman Rules (Tree Level)

Feynman rules for QED, QCD, and weak interactions.

D.1
External Lines

$$\text{Spin } 0 : \text{(nothing)}$$

$$\text{Spin } \tfrac{1}{2} : \begin{cases} \text{Incoming particle} : \ u \\ \text{Incoming antiparticle} : \ \overline{v} \\ \text{Outgoing particle} : \ \overline{u} \\ \text{Outgoing antiparticle} : \ v \end{cases}$$

$$\text{Spin } 1 : \begin{cases} \text{incoming} : \ \epsilon_\mu \\ \text{outgoing} : \ \epsilon_\mu^* \end{cases}$$

D.2
Propagators

$$\text{Spin } 0 : \frac{i}{q^2 - (mc)^2}$$

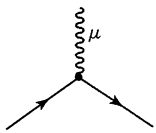
$$\text{Spin } \tfrac{1}{2} : \frac{i(\not{q} + mc)}{q^2 - (mc)^2}$$

$$\text{Spin } 1 : \begin{cases} \text{Massless} : \dfrac{-ig_{\mu\nu}}{q^2} \\ \\ \text{Massive} : \dfrac{-i[g_{\mu\nu} - q_\mu q_\nu/(mc)^2]}{q^2 - (mc)^2} \end{cases}$$

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

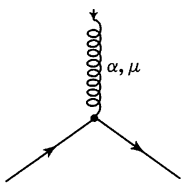
D.3 Vertex Factors

QED:

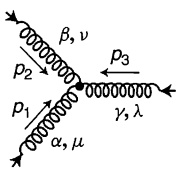


$$ig_e\gamma^\mu \quad (g_e = \sqrt{4\pi\alpha})$$

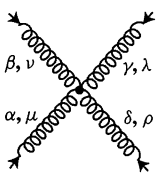
QCD:



$$\frac{-ig_s}{2}\lambda^\alpha\gamma^\mu$$

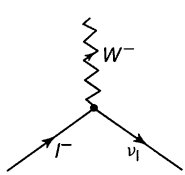


$$-g_s f^{\alpha\beta\gamma}[g_{\mu\nu}(q_1 - q_2)_\lambda + g_{\nu\lambda}(q_2 - q_3)_\mu + g_{\lambda\mu}(q_3 - q_1)_\nu]$$

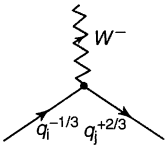


$$-ig_s^2[f^{\alpha\beta\eta}f^{\gamma\delta\eta}(g_{\mu\lambda}g_{\nu\rho} - g_{\mu\rho}g_{\nu\lambda}) + f^{\alpha\delta\eta}f^{\beta\gamma\eta}(g_{\mu\nu}g_{\lambda\rho} - g_{\mu\lambda}g_{\nu\rho}) + f^{\alpha\gamma\eta}f^{\delta\beta\eta}(g_{\mu\rho}g_{\nu\lambda} - g_{\mu\nu}g_{\lambda\rho})]$$

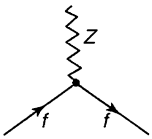
GWS:



$$\frac{-ig_w}{2\sqrt{2}}\gamma^\mu(1-\gamma^5)$$ (Here l is any lepton, and ν_l the corresponding neutrino.)

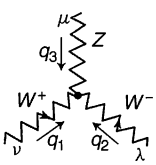


$\frac{-ig_w}{2\sqrt{2}}\gamma^\mu(1-\gamma^5)V_{ij}$ (Here $i = u$, c, or t, and $j = d$, s, or b; V is the CKM matrix.)

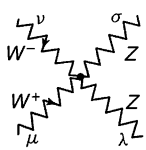


$\frac{-ig_Z}{2}\gamma^\mu(c_V^f - c_A^f\gamma^5)$ Here f is any quark or lepton; c_V and c_A are given in the following table:

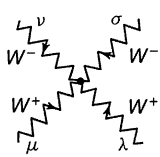
f	c_V	c_A
ν_e, ν_μ, ν_τ	$\frac{1}{2}$	$\frac{1}{2}$
e^-, μ^-, τ^-	$-\frac{1}{2} + 2\sin^2\theta_w$	$-\frac{1}{2}$
u, c, t	$\frac{1}{2} - \frac{4}{3}\sin^2\theta_w$	$\frac{1}{2}$
d, s, b	$-\frac{1}{2} + \frac{2}{3}\sin^2\theta_w$	$-\frac{1}{2}$



$$ig_w\cos\theta_w[g_{\nu\lambda}(q_1-q_2)_\mu + g_{\lambda\mu}(q_2-q_3)_\nu + g_{\mu\nu}(q_3-q_1)_\lambda]$$



$$-ig_w^2\cos^2\theta_w(2g_{\mu\nu}g_{\lambda\sigma} - g_{\mu\lambda}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\lambda})$$

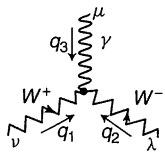


$$ig_w^2(2g_{\mu\lambda}g_{\nu\sigma} - g_{\mu\nu}g_{\lambda\sigma} - g_{\mu\sigma}g_{\nu\lambda})$$

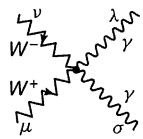
The weak coupling constants are related to the electromagnetic coupling constant:

$$g_w = \frac{g_e}{\sin\theta_w}; \quad g_z = \frac{g_e}{\sin\theta_w \cos\theta_w}.$$

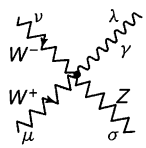
There are also 'mixed' couplings of the photon to the W and Z:



$$ig_e[g_{\nu\lambda}(q_1 - q_2)_\mu + g_{\lambda\mu}(q_2 - q_3)_\nu + g_{\mu\nu}(q_3 - q_1)_\lambda]$$



$$-ig_e^2(2g_{\mu\nu}g_{\lambda\sigma} - g_{\mu\lambda}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\lambda})$$



$$-ig_e g_w \cos\theta_w(2g_{\mu\nu}g_{\lambda\sigma} - g_{\mu\lambda}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\lambda})$$

Index

a

A: *see* Baryon number
ABC theory 211–223
Abelian 118, 365
  – *See also* Gauge invariance; Groups
Accelerator 4–7, 401
Additive quantum number 141
Adjoint 236, 361
Allowed 120, 127, 161–162
α: *see* Fine structure constant
α_s 173
α_w 308, 315
α particle 5, 14, 156, 388–389
Amplitude 132–133, 156, 203–204, 211–214, 244
Anderson, C. D. 19, 21, 31
Angular momentum 120–128
  – addition 122–125, 154
  – eigenfunctions: *see* Spherical harmonics
  – eigenvalues 120–121
  – matrices: *see* Spin matrices
  – orbital 120–121, 161–163, 182
  – spin 120–121
Annihilation: *see* Pair annihilation
Anomalous magnetic moment
  – electron 18, 165, 167, 246, 266
  – muon 413
  – proton 168
Anthropic principle 414
Antibaryon 37, 40
Anticommutator 155, 228, 237, 252, 435
Antielectron: *see* Positron
Antimatter 21, 23
Antimeson 37, 41
Antineutrino 24, 27, 128
Antineutron 21–22
Antiparticle 3, 20–23, 37, 39, 61–62, 230–234
Antiproton 21, 33, 106–107
Antiquark 39, 49
Antiscreening 300–301
Antisymmetric state 122, 125, 183–184
Antisymmetric tensor 110–111, 253
Antisymmetrization 245
Associated Laguerre polynomial 162
Associated production 34
Astrophysics 7, 394, 396, 400–401
Asymptotic freedom 68–70, 83, 220, 298–301
Atmospheric neutrino 394–395, 399–400
Atomic number 24
Atomic weight 24, 387
Axial vector 139–141, 308
Axion 411, 416
Azimuthal angle 201

b

B: *see* Bottom
B factory 47, 148
B meson 47, 148–149
b quark 47
Bahcall, J. N. 389–390, 416
Bare: *see* Charge; Coupling constant; Mass
Barn 211
Baryogenesis 411
Baryon 19, 30, 32, 35, 40, 44, 122–123, 180–193
  – decuplet 36, 40, 122–123
  – magnetic moment 189–190
  – mass 191–193
  – number 22, 33, 81, 130, 408–409
  – octet 35, 41, 122–123, 188
Beautiful baryon 47
Beautiful meson 47
Beauty 47, 49
  – *See also* Bottom
β: *see* Electron; Positron

Introduction to Elementary Particles, Second Edition. David Griffiths
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40601-2

Beta decay 23–29, 47–48, 56, 76, 136–137, 318
Bethe, H. A. 166, 219, 388
Bevatron 21, 106
Bhabha scattering 62, 86, 246
Big Bang 71, 409, 416
Bilinear covariant 235–238
Binding energy 103, 159–160, 194
Bispinor 229
Bjorken, J. D. 44, 228, 233
Blackbody 15
Bohr, N. 15, 24, 162
Bohr energies 162
Bohr magneton 190
Bohr model 15, 163
Bohr radius 162
Boson 122, 141, 183
Bottom 47
Bottomonium 159, 175–177
– *See also* Υ (upsilon) meson
Bound state 2, 159–193
Branching ratio 79, 198, 404–405
Brane 413
Breit frame 112
Bremsstrahlung 403–404
Brick wall frame: *see* Breit frame
Broken symmetry 135–136, 375, 378, 402, 408
Brookhaven 29, 33, 44, 71, 147
Bubble chamber 7

c

C: *see* charm; Charge conjugation
c quark 44–45
c_V, c_A 320, 331–332
Cabibbo, N., 77 324–327
Cabibbo angle 321, 324–325, 329
Cabibbo theory 324–328
Camouflage 300
Casimir's trick 249–254, 270
Cathode ray 13
Center of momentum 64, 106–108
Central potential 161
Centrifugal barrier 161
Čerenkov radiation 7, 392
CERN 6–7, 42, 48, 51, 72, 331–332, 403
Chadwick, J. 15, 24
Charge 148, 165, 301
– bare 69–70
– conservation: *see* Conservation laws
– effective 68–70, 266, 299
– electric 70, 81, 408–409
– exchange 132
– independence 130
– quark 39
– renormalization 219–221, 265–267
– weak 71–72, 320
– *See also* Coupling constant
Charge conjugation 142–149
– conservation 143
– number 143
– operator 142–143, 269
– photon 143
– positronium, 171
– violation, 143
Charged weak interaction 72–78, 307–308, 324–329
Charm 44–45, 49
Charmed baryon 45–47
Charmed meson 45–47
Charmonium 159, 172–177
– *See also* ψ meson
Chiral state 340
Chromodynamics 59–60, 66–71, 366–369
– *See also* Quantum chromodynamics
Circular polarization 241, 260
CKM matrix 51, 77–78, 81, 148, 328–329, 350, 397, 409–410
Clebsch–Gordan coefficients 124–125, 154
Cloud chamber 7, 30–31
CM 64, 106–108
CNO cycle 388
Coleman–Glashow formula 57
Colliding beams 5, 108–109
Collision 87, 100–102
Color 43–44, 49–50, 66–67, 70, 187–188, 278–279
– current 369
– factor 173, 289–294, 304
– field 369
– octet 187, 285
– sextet 293
– singlet 187, 285
– state 284–286
– $SU(3)$ 187, 285, 366–369
Colorless particle 43, 285
– *See also* Color singlet
Commutator 155
Compactification 414
Completeness 230, 234, 242–243, 270, 308, 348
Composition rule 40
Compton, A. H. 17
Compton
– scattering 17–18, 23, 62, 86, 113, 246, 249, 271–272
– wavelength 17, 69
Confinement 42–44, 71–72, 289

Conjugating phase 410
Conservation laws 79–84
– angular momentum 26, 80, 117, 119, 122
– baryon number 33, 79, 81, 85
– charge 79, 81, 85, 117, 238, 270
– color 67, 81, 85
– electron number 29, 75, 81
– energy 24, 64, 80, 87, 100–101, 117, 213, 244
– flavor 75, 82, 324
– isosopin 130
– lepton number 28–29, 79, 81, 85
– mass 2, 100–102
– momentum 64, 80, 97–101, 117, 213, 244
– muon number 29, 75, 81
– parity 73, 141–142, 309
– quark number 81
– strangeness 34, 82, 324
– tau number 75, 81
Conserved current 238, 270, 358
Constituent quark 58, 180, 190, 193
Continuity equation 238, 270, 358
Continuous symmetry 375–377
Contraction of indices 96
Contraction theorems 252–253
Contravariant 94
Cosmic ray 4–5, 19, 30–31, 110, 394, 401
Cosmological constant 417
Coulomb force 1, 9, 18, 61–62, 73
Coulomb gauge 240–243, 286, 307
Coulomb potential 162–163, 202, 290, 334
Coupling constant 67, 84–85, 213, 308
– bare 220, 265
– dimensions of 213
– effective 220, 265
– electromagnetic 84–85, 244
– physical 219–220
– renormalized 219–221, 265–266, 298–301
– running 68, 85, 220, 265–266, 300–301, 405–406, 412
– strong 67, 84–85, 283–284
– weak 84–85, 308, 315, 320, 332, 346
– *See also* Charge, Fine structure constant
Covariant 94, 96, 226, 268
Covariant derivative 360, 363, 367, 378
Cowan, C. L. 27
CP 144–149, 409–411
– B^0 system 148–149, 409–411
– eigenstate 146
– invariance 145–147
– K^0 system 145–149
– violation 52, 139, 147–149, 328, 409–411
Cronin, J. W. 146–149, 409
Cross product 368, 408
Cross section 132–133, 156, 199–203, 209, 430–431
– $A + A \to B + B$ 215–216
– hard sphere 200–202
– Mott 245–246, 254–255
– nucleon–nucleon 132
– pair annihilation (QCD) 298
– pair annihilation (QED) 261
– Rutherford 202, 255
– *See also* Golden rule, Scattering
Crossing symmetry 22–23, 62
Current
– charged weak 341–346
– color 369
– conserved 358
– electromagnetic 238, 341–344, 358, 360
– neutral weak 343
– Noether 383–384
– weak hypercharge 342–344
– weak isospin 342–344
– Yang–Mills 365
Cutoff 219, 264–265
CVC hypothesis 320
Cyclotron formula 8

d

D meson 47
D_s meson 47
d quark 39
D'Alembertian 240
Dark energy 414–417
Dark matter 414–417
Davis, R. 27–28, 390, 396
De Broglie wavelength 6
Decay 2, 65, 75–84, 87, 197–198
– *b* quark 351
– *c* quark 351
– B^0 410
– Δ 77
– η_c 298, 306
– Higgs 404–405, 419
– kaon 325
– Λ 77–78, 324
– leptonic 325
– muon 25–29, 75, 310–315, 394–395, 400
– neutrino 394
– neutron 25–29, 76, 80, 315–321, 324

Decay (*contd.*)
- nonleptonic 325–326
- Ω^- 77–78
- ϕ 83
- pion (charged) 25–29, 76, 80, 103, 138, 321–324, 349, 394–395
- pion (neutral) 65, 138, 143, 222
- positronium 256–261
- proton 32–33, 85, 406, 408, 420
- ψ 82–84
- quarkonium 298
- semileptonic 325–326
- tritium 25, 396
- two-body 24–27, 112, 206–208, 429–430
- Z^0 337, 350–351
- *See also* Golden rule; Lifetime

Decay mode 79, 198
Decay rate 197–198, 429–430
Decuplet 36, 40, 122–123
Deep inelastic scattering 42–43, 50, 68
Degeneracy 163
Delbruck scattering 86
Δ baryon 33, 133
Delta function
- Dirac 205, 423–427
- Kronecker 155

Density of states 203
- *See also* Phase space

DESY 73
Detailed balance 23, 149
Detector 7–8, 203
Deuterium 169
Deuteron 43, 103, 130–131, 194, 388
Diagram
- disconnected 218
- Feynman 60–64, 213
- higher-order 62–63, 75, 86, 212, 217–221, 262–267
- loop 70, 166–167, 217–218, 263, 288, 298–299
- penguin 148, 157, 410
- tree 157, 217, 265

Differential cross section 201–203, 223, 430–431
- *See also* Scattering

Dimensions
- amplitude 211
- coupling constant 213
- cross section 211
- field 357

Dipole function 284
Dipole moment
- electric 150, 411
- magnetic 150, 165, 168, 189–190

Diquark 294, 408
Dirac, P. A. M. 15, 21, 165, 219
Dirac delta function 205, 423–427
Dirac equation 21, 225–229, 356
- momentum space 234

Dirac Lagragnian 355–356
Dirac matrices 228, 434–435
Dirac neutrino 28, 396
Dirac sea 21, 230
Dirac spinor 229
Direct *CP* violation 148
Disconnected diagram 218
Discrete symmetry 118, 136–151, 375
Dot product 95
Down quark 39
Downness 49
Drell, S. D. 228, 233
Dresden, M. 221

e

Effective charge 68–70, 266
Effective mass 135–136, 180, 220
Eigenstate 328
Eigenvalue 126–127, 160
Eigenvector 126–127
Eight baryon problem 133
Eightfold Way 35–37, 39, 133
Einstein, A. 16, 84, 417
Einstein summation convention 93
Elastic collision 101–102
Elastic scattering 101–102, 199
Electric dipole moment 150, 411
Electric form factor 283–284
Electrodynamics 59, 84, 238
- *See also* Quantum electrodynamics

Electromagnetic current 238, 341–344, 358, 360
Electromagnetic decay 255–261
Electromagnetic field 239
Electromagnetic force 59
Electromagnetic potential 239
Electron 4–5, 13, 150
Electron gun 4
Electron neutrino 29
Electron number 29, 49
Electron volt 9
Electron–electron scattering 246–247, 266, 271
Electron–muon scattering 245–246, 252–255, 265, 271
Electron–positron scattering
- elastic 247–248
- inelastic 256–261, 266, 275–279

Electron–positron annihilation 256–261
Electron–proton scattering 255, 279–283
Electroweak interaction 3, 48, 60, 84, 338
– *See also* GWS theory
Electroweak interference 337
Electroweak mixing 345–346
Electroweak unification 338–346
Elementary particles: *see* Particles
Energy
– conservation of 24, 64, 80, 87, 100–101, 117, 213, 244
– kinetic 99
– operator 160, 268
– relativistic 96–100
– rest 99
Energy-momentum four-vector 98
η meson 31, 177–179
η' meson 41, 177–179
η_c meson 298, 306
Euler–Lagrange equation 354–356
Event rate 203
Exchange of particles 18, 61, 183
Exclusion principle 1, 3, 21, 37, 43–44, 122, 130, 183–184
Exclusive 199
Exotic atom 169
Exotic particle 41
Expansion of universe 416–417
External line 63, 65, 213, 243, 287, 437
Extra dimensions 413–414

f

f_π (pion decay constant) 322–323
Faithful representation 119
Fedeev–Popov ghost 288
Fermi, E. 24, 60, 133, 204, 315
Fermi constant (G_F) 313–315
Fermi's Golden Rule 203–211
Fermi theory of beta decay 24, 47–48, 60
Fermilab 6, 8, 47, 395, 403
Fermion 121–122, 141, 183
Feynman, R. P. 21, 29, 60, 65, 166, 205, 219
Feynman calculus 64, 197–223, 373
Feynman diagram 60–64, 213
Feynman rules 3, 64, 203–204, 369–372, 437–440
– ABC theory 211–214
– GWS theory 307–308, 331
– QCD 283–288
– QED 241–245
– weak interactions 307–308, 331
Feynman–Stükelberg interpretation 21, 230
Field theory 354–358
– scalar 355, 357
– spinor 355–357
– vector 356–359
Field strength tensor 238
Fifth force 285
Fine structure 165–166, 169–170
Fine structure constant 10, 64, 162, 165, 266
Fine tuning 412
Flavor 39, 43, 49, 57, 70, 82,129–136, 184–188, 397
Flavor dynamics 59–60
Flavor eigenstates 324, 328–329, 397–398
Force 59–60
– electromagnetic 59
– gravity 59
– strong 18–20, 30, 34, 50, 59
– weak 34, 48, 59
Form factor 282–284, 303, 319, 322
Four-momentum 98
Four-vector 92–96
– charge-current 238
– covariant 94
– contravariant 94
– current 238
– energy-momentum 98
– lightlike 95
– position-time 92
– proper velocity 96
– spacelike 95
– timelike 95
Four-velocity 97
Free Lagrangian 369
Free quark 41–44, 72
ft-value 319
Fundamental representation 119, 179
Furry's theorem 273
Fusion 388, 403–404

g

G: *see* G-parity
G_F: *see* Fermi constant
g_e 332, 346
g_s 371
g_w 308, 315, 332, 346
g_z 332, 346
Galactic rotation 414–415
Gamma matrices 228
Gamma ray 5, 17
– *See also* Photon
γ^5 236–237
Gauge field 360, 364–367, 407
Gauge invariance 117, 304
– abelian 365
– broken 136, 375
– global 358, 363

Gauge invariance (*contd.*)
- local 3, 358–361, 363, 367
- nonabelian 365
Gauge theory 353–381
Gauge transformation 239, 270, 360
Gaussian units 9
Gell-Mann, M. 34–37, 56–57, 60, 82, 133, 145–147
Gell-Mann matrices 286–287, 367, 384
Gell-Mann–Nishijima formula 130–131, 156, 343
Gell-Mann–Okubo formula 56–57
Generation 49–51, 75
Ghost 241, 288, 381
GIM mechanism 77, 327–329
Glashow, S. L. 44–48, 60, 77, 84, 327, 330, 338, 406
Global transformation 358, 363
Glueball 50, 67
Gluino 52
Gluon 50, 59–60, 67, 368
- octet 285
Gluon–gluon coupling 67, 70, 286, 288
Golden Rule 203–211
- for decays 204–208
- for scattering 208–211
Goldstone boson 377–381
Goldstone's theorem 377
G-parity 143–144, 157
Gradient 226, 268
Grand Unification 33, 52, 84–85, 405–409
Gravitational lensing 415
Gravity 18, 49, 59, 86, 413
Graviton 18, 49, 59–60, 414
Greek index 92–93
Greenberg, O. W. 43
Ground state 373
- *See also* Vacuum
Group 117–118
- abelian 118
- continuous 118
- discrete 118
- finite 118
- infinite 118
- Lie 118
- Lorentz. 118
- $O(n)$ 118–119, 153
- $SO(2)$ 376
- $SO(3)$ 118–119, 128
- $SO(10)$ 407
- $SO(n)$ 118–119, 153
- $SU(2)$ 128, 130, 135, 362–363, 365
- $SU(2)_L \times U(1)$ 344, 346, 366, 407
- $SU(3)$ 128, 133, 179, 187, 367, 407
- $SU(5)$ 407
- $SU(6)$ 134
- $SU(n)$ 118–119, 153
- $U(n)$ 118–119, 153, 361, 366
Group theory 117, 133
GUT: *see* Grand Unification
GUT scale 405–406, 413
GWS (Glashow/Weinberg/Salam) 3, 48, 50, 60, 71–79, 84, 221, 332, 337–346
Gyromagnetic ratio 165

h

Hadron 30
- *See also* Baryon; Meson
Hadronization 275–276
Half-life 79, 222, 319
Halo 415–417
Hamiltonian 160, 268
Hard-sphere scattering 200–202
Heaviside–Lorentz units 9, 245
Heaviside step function 205, 426–427
Heavy lepton 47
Heavy neutrino 396
Heavy quark 159
Helicity 28, 137–138, 234, 241, 268, 324, 339
Heisenberg, W. 129
Heisenberg uncertainty principle 6, 56, 80
Hermitian conjugate 235, 367
Hermitian matrix 362–363, 384
Hierarchy problem 412
Higgs boson 6, 50, 52, 380–381, 401–405
Higgs mechanism 84, 338, 372, 378–381, 401–402
Higgsino 412, 416
Higher-order process 62–63, 75, 86, 212, 217–221, 262–267
Hole theory 21
Homestake mine 390, 393–394
't Hooft, G. 221, 330, 372, 409
Hot dark matter 415
Hydrogen 15, 159, 162–168
Hypercharge 36, 343–344
Hyperfine splitting 167–168
- baryons 191
- hydrogen 167–168
- mesons 180
- positronium 169–170
Hyperon 33

i

Identical particles 1, 43, 66, 183–184, 198
ILC 6, 401
Impact parameter 199–201
Inclusive 199

Indeterminacy 3
Indistinguishable particles 1, 43, 66, 183–184
Inelastic scattering 199
Inertial frame 89
Intermediate vector boson 47–48, 59–60
– *See also* W; Z
Internal line 63, 243
Internal momenta 213–214
Internal quantum number 142
Internal symmetry 117, 119, 129–130, 411
Intrinsic angular momentum 120
Intersecting storage rings 5
Invariance 117
Invariant 93–96, 104–107, 110
Inverse beta decay 27, 309, 388
Inverse Lorentz transformation 90, 109
Inversion 139
Inverted spectrum 396
Ionization 4, 8
Irreducible representation 119
Isospin 129–136, 156, 365
– weak 342–344
Isotope 15
Isotopic spin: *see* Isospin

j

J/ψ meson: *see* ψ meson, Charmonium
Jet 50, 276
Jordan form 384

k

K meson 31, 34, 145–151
– K_1, K_2 146–147
– K_L, K_S 147–149
Kaon: *see* K meson
KEK 43, 47, 148
Ket 3, 122
Kinetic energy 99, 101, 373
Klein–Gordon equation 225–227, 240, 355
Klein–Nishina formula 272
KM matrix: *see* CKM matrix
Kobayashi–Maskawa matrix: *see* CKM matrix
Kronecker delta 155

l

Lagrangian 357
– classical 353–354
– Dirac 355–356, 383
– Field theory 354–358
– Free 369
– Interaction 369
– Klein–Gordon 355, 383
– Maxwell 357–360
– Proca 356–357
– QCD 366–368
– QED 358
– Relativistic 354–358
– Supersymmetric 412, 420
– Yang–Mills 365
– Yukawa 385
Lagrangian density 354, 357
– *See also* Lagrangian
Lamb, W. E. 33, 166
Lamb shift 18, 166, 170, 194–195, 220, 266
Λ (cosmological constant) 417
Λ (QCD scale) 301
Λ baryon 31–34
Λ matrix 93
Λ_c baryon 46–47
λ matrices 286–287
Landscape 414
Leading log approximation 299
Lederman, L. M. 29, 147, 402
Lee, T. D. 60, 136, 142
Left-handed 28, 137–138, 324, 339–342
Left-handed doublet 343–344
Leptogenesis 411
Lepton 3, 19, 30, 45, 47, 49, 122
– generations 30, 49
– number 28, 81, 408–409
– table 30, 49
– weak interactions 307, 310, 329–337
– *See also* Electron; Muon; Neutrino; Tau
Leptonic decay 325
Leptoquark 407–408
Levi-Civita symbol 155, 253, 271
LHC 6, 47, 52, 401, 403, 412–413, 416–417
Lifetime 79–80, 91, 197–198, 222
– A 214–215
– charmonium 44
– kaon 147
– muon 313–315
– neutron 319–321
– Ω^- 57
– pion 323–324
– positronium 171, 261, 420
– proton 32–33, 406, 408, 420
– sun 387, 399
– Z^0 51, 337, 350–351
– *See also* Decay
Light quark 159, 176
Light quark baryon 159, 181–193
Light quark meson 159, 176–180
Lightlike 95
Linear collider 5
Linear-plus-Coulomb potential 173

Local gauge invariance 3, 358–363
Local gauge transformation 367
Local phase invariance 360
Logarithmic divergence 219, 263
Longitudinal polarization 138, 307, 380
Loop diagram 70, 166–167, 263, 288, 298–299
Lorentz gauge 240, 286, 307
Lorentz contraction 91–92
Lorentz group 118
Lorentz transformations 89–93
Lowering operator 156
Luminosity 202–203, 261

m

M theory 414
Magnetic form factor 283–284
Magnetic moment 150, 165
- anomalous 18, 266
- baryons 89–190, 195–196
- electron 165
- proton 168

Magnetic monopole 409
Majorana neutrino 28, 273, 396, 407
Mandelstam vairables 113
Marshak, R. E. 19
Mass formulas
- baryons 191–193, 195–196
- Coleman–Glashow 57
- Gell-Mann–Okubo 56–57
- Higgs 403
- mesons 179–180, 195
- pion 56
- quarkonium 171–176
- *W* and *Z* 332

Mass 51
- bare 58, 135–136, 220
- constituent 58, 135–136, 180, 190, 193
- effective 135–136, 180, 220
- eigenstate 328, 391, 397–398
- Higgs 403, 412
- matrix 362
- neutral kaon 147
- neutrino 395–396
- nonconservation 101–103
- origin 381, 402
- physical 219–220
- quark 135–136, 180
- relativistic 99, 101
- renormalized 67, 220–221, 266
- running 220
- shell 65, 213
- term 372–375
- virtual particle 65
- WIMP 416

Massive gauge field 372, 402
Massless particle 2, 29–30, 99–100, 138–139, 241, 308, 359–361, 377, 380, 392
Matrix element 203
Matter–antimatter asymmetry 21, 23, 51, 148, 409–411
Maximal parity violation 137, 148, 309
Maxwell, J. C. 84
Maxwell Lagrangian 357–359
Maxwell's equations 238–239, 357–358
Mechanics 2
Mediator 18, 47–48, 59–60, 67, 122, 307, 315, 407
- *See also* Gluon; Graviton; Intermediate vector boson; Photon; Pion; *W*; *Z*

Meson 15, 18–20, 30–31, 35, 40, 44, 122–123, 176–180
- mass 179–180
- nonet 39–41
- octet 35–36, 41

Metric 93–94, 252
Millikan, R. A. 16, 41
Minimal coupling 360
Minkowski metric 93–94, 228
Mirror image 136–137, 144, 148
Missing mass 415–416
Mixed tensor 96
Mixing
- $B^0/\overline{B}^0$ 145, 149
- $K^0/\overline{K}^0$ 391–394, 399
- matrix 329, 397–398
- neutrino 391–394
- neutral mesons 145, 149

MNS matrix 397–398
Moller scattering 61, 246
Momentum 97
- conservation 64, 80, 97–101, 117, 213, 244
- four-vector 98
- operator 226
- relativistic 96–100
- space 234, 370

Mott scattering 245–246, 254–255
MSM 402–406, 419
MSW effect 392
Multiplet 130
Multiplicative quantum number 141, 143
Multiplicity 130
Muon 19, 26–30, 310–315
Muon decay 25–29, 75, 310–315, 394–395, 400
Muon neutrino 29

Muon number 29, 49
Muonium 169

n
N: *see* Nucleon
n: *see* Neutron
ν: *see* Neutrino
Neddermeyer, S. 19
Ne'eman, Y. 35
Negative energy state 21, 230
Neutral weak interaction 72–74, 308, 329–338, 343
Neutralino 416
Neutrino 15, 23–30, 137–139, 389, 415–416
Neutrino mass 395–396
Neutrino oscillations 30, 52, 74, 387–398
Neutrinoless double beta-decay 396
Neutrino-electron scattering 72–73
Neutrino-nucleon scattering 73
Neutron 15, 22, 29, 76, 129–130, 150, 315–321
Neutron electric dipole moment 411
Ninth gluon 285–286, 368
Noether's theorem 116–117, 383–384
Nonabelian gauge 365
Nonet 39–41
Nonleptonic decay 77, 325–326
Normalization
- Dirac spinor 233, 237, 269
- Pauli spinor 126
- polarization vector 243, 348
- wave function 160
November revolution 44–47
Nucleon 129–130
- *See also* Neutron; Proton
Nucleon–nucleon scattering 131–132
Nucleus 14–15, 18, 42, 56, 167

o
$O(n)$ 118–119, 153
Octet 35, 41, 179
- *See also* Baryon; Color; Gluon; Meson
Ω^- baryon 37–38, 57, 88, 111
ω meson 31, 177–179
Oppenheimer, J. R. 19
Orthogonal matrix 118–119
Orthogonal polarizations 242, 348
Orthogonal spinors 268
Oscillations 390–392
OZI rule 82–84, 88, 174

p
P: *see* Parity
Pair annihilation 23, 62, 65, 170–171, 246, 256–261, 294–298
Pair production 62, 246
Pais, A. 33–34, 145–147
Parity 136–142, 157
- baryon 141
- boson 141
- conservation 73, 141–142, 309
- fermion 141
- invariance 136, 139
- meson 141
- operator, operator 139–141, 236
- particle/antiparticle 141
- photon 141
- quark 141
- violation 73, 136–139, 142, 148, 309
Particle 147
- α 5, 14, 156
- antineutrino 24, 27, 128
- antineutron 21
- antiproton 21
- *B* meson 47, 148–149
- B_c meson 47
- B_s meson 47
- *b* (bottom) quark 47
- β: *see* Electron; Positron
- *c* (charm) quark 44
- *D* meson 47
- D_s meson 47
- *d* (down) quark 39
- Δ baryon 33, 133
- deuteron 43, 103, 130–131, 194, 388
- electron 13, 150
- η meson 31, 177–179
- η' meson 41, 177–179
- η_c meson 298, 306
- γ: *see* Photon
- gluon 50
- graviton 18, 49, 59–60, 414
- *J* or J/ψ: *see* ψ meson
- *K* meson 31
  - charged 31
  - neutral 31, 145–149, 151
- kaon: *see* *K* meson
- Λ baryon 31–34, 77
- Λ_b baryon 47
- Λ_c baryon 46–47
- μ (muon) 19, 26–30, 310–315
- ν (neutrino) 15, 23–30,137–139, 389, 415–416
- neutron 15, 76, 129, 150, 315–321
- nucleon 129–130
- Ω^- baryon 37–38, 57, 77, 88, 111
- Ω_c baryon 47

Particle (*contd.*)
- ω meson 31, 177–179
- ϕ meson 31, 83, 177–179
- photon 15–18, 49, 59–60
- π meson 19–20, 26–29, 178, 321–324
- pion: *see* π meson
- positron 5, 15, 21–22
- proton 4, 15, 32–33
- ψ meson 44, 82–84, 174–175
  - *See also* Charmonium
- ρ meson 31, 178
- s (strange) quark 39
- Σ baryon 33–34
- Σ_b baryon 47
- Σ_c baryon 47
- t (top) quark 47
- τ lepton 47
- u (up) quark 39
- Υ meson 47, 88, 175–176, 195
  - *See also* Bottomonium
- W 48–49
- Ξ baryon 33–34
- Ξ_b baryon 47
- Ξ_c baryon 47
- Z 48–49, 350–351

Particle Physics Booklet 10–11, 47
Parton 44
- *See also* Gluon; Quark

Pauli, W. 15, 24, 139, 219, 227
Pauli (spin) matrices 127, 155, 228, 363, 433–434
Pauli (exclusion) principle 1, 3, 21, 37, 43–44, 122, 130, 183–184
PCAC hypothesis 320, 326
Penguin diagram 148, 157, 410
Pentaquark 43
Perturbation theory 373
Phase transformation 358–360
Phase space 203–205, 209, 323–324
ϕ meson 31, 177–179
Photino 52, 412
Photoelectric effect 16, 18
Photon 15–18, 49, 59–60, 238–241
Pion 19–20, 26–29, 56, 178, 321–324
Pion decay constant 322–323
Planck, M. 15–16
Planck formula 16, 100, 163
Planck's constant 6, 16
Planck scale 413, 420
Plane wave 231, 240
Point particle 220
Polar vector 139–141
Polarization 138, 147, 307
Polarization vector 240, 242–243, 286, 307
Pontecorvo, B. 390
Positron 5, 15, 21–22, 230
Positronium 169–172, 195–196
Potential
- Coulomb 162, 202
- four-vector 239
- linear-plus-Coulomb 173, 195
- quark–quark 173, 289–294
- vector 239

Potential energy 159–160, 373
Powell, C. F. 19–20, 25–26
pp chain 388–390
Primitive vertex 60–61, 64–67, 72–75, 78–80, 212, 308
Principal quantum number 163
Proca equation 225, 356–357
Projection operator 339, 349
Propagator 213, 369–370, 437
- electron 244
- gluon 288
- modified 263, 336
- photon 244
- quark 288
- spin zero 213, 370
- spin one-half 244, 370
- spin one
  - massive 308, 370
  - massless 244, 371
- unstable particle 336
- W and Z 308, 333, 336

Proper time 96
Proper velocity 96–97
Proton 4, 15, 32–33, 129–130
Pseudoscalar 140–141, 236–237
Pseudoscalar meson 35–36, 122–123, 177
Pseudovector 139–141, 237
ψ meson 44, 174–175
- *See also* Charmonium

q

Q: *see* Charge, electric
QCD (quantum chromodynamics) 3, 66–71, 173, 283–301, 366–369
QED (quantum electrodynamics) 3, 60–66, 165, 225–273
Quantum 16, 19, 369
Quantum field theory 2, 10, 17, 20–21, 52, 122, 150, 159
Quantum mechanics 2, 17, 20, 100, 120, 159
Quark 3, 37, 45, 49, 122
- b (bottom) 47
- c (charm) 44–45

– confinement 42–44, 71–72
– constituent 58
– d (down) 39
– free 41–44, 72
– masses 51, 58, 135–136
– model 37–44, 47
– number 81
– parity 141
– sea 320
– search 41–44
– s (strange) 39
– t (top) 47
– u (up) 39
– table 49
– valence 320
– weak interactions 324–337
– *See also* Particles
Quark-gluon plasma 71
Quark-quark interaction 289–294
Quarkonium 169, 171–176
– *See also* Bottomonium; Charmonium
Quasi-bound state 174–175, 195

r

R 278–280
Rabi, I. I. 30, 163
Radial equation 161
Radiative correction 166
Raising operator 156
Range 18–19, 48, 56, 285
Rank 95–96
Reactor 5, 27
Real particle 63, 65
Reduced mass 169
Reflection 139
Regularization 219
Relativistic correction 165
Relativistic energy 98–99
Relativistic mass 99
Relativistic mechanics 2
Relativistic momentum 98
Relativistic system 159–160
Relativity 2, 89–113
Reines, F. 27
Renormalization 219–221, 262–267, 298–301
Representation 119, 128, 199
Resonance 133, 156, 279
Rest energy 99, 102
Review of Particle Physics 10–11
ρ meson 31, 178
Richter, B. 44
Right-handed 28, 137–138, 324, 339–340, 342
Rochester, G. D. 30–31
Rosenbluth formula 283, 303
Rotation 117–118
Rotation curve 414–415
Rotation group 118–119, 128, 154
Rotation matrix 154–155
Rubbia, C. 48, 332
Running coupling constant 68, 84–85, 220, 265–266, 301, 405–406, 412
Running mass 220
Rutherford, E. 14–15, 42–43
Rutherford scattering 14–15, 42–43, 202, 245–246, 254–255
Rydberg formula 163

s

S (strangeness) 34–35, 40–41
s (strange) quark 39
Salam, A. 48, 60, 84, 330, 338
Sakharov, A. D. 409
Scalar 123, 237
– *See also* Invariant
Scalar product 94–95
Scattering 2, 197–203, 430–431
– $A + A \to B + B$ 215–217
– Bhabha 62, 86, 246
– Compton 17–18, 62, 86, 113, 246, 249, 271–272
– cross section 132–133
– deep inelastic 42–43
– Delbruck 86
– elastic 199
– electron–electron 246, 266, 271
– electron–muon 245–246, 252–255, 265, 271
– electron–positron 246–248, 256–261, 275–279, 335–337
– electron–proton 255, 279–283
– gluon–gluon 306
– hard sphere 200–202
– inelastic 199
– Moller 61, 246
– Mott 245–246, 254–255
– neutrino-electron 309–310, 330–334, 392–393
– nucleon–nucleon 131–132
– pion–nucleon 132–133, 156
– Rutherford 14–15, 42–43, 202, 245–246, 254–255
– *See also* Collisions; Golden Rule
Scattering amplitude 132–133
Scattering angle 17, 199–203
Scattering center 200–201

Schrödinger equation 159–162, 225–227, 230
Schwinger, J. S. 60, 165–166, 219
Screening 68–69, 85, 167, 266, 299
Sea quark 320
See–saw mechanism 396
Selectron 52
Self-energy 217
Semileptonic decay 75, 325–326, 350
Separation of variables 160–161
Sextet 293
Σ baryon 33–34
SI units 9
Simultaneity 90–91
Singlet 39, 41, 125, 131, 168, 178–179, 187
SLAC 4, 42, 44, 47, 51, 148, 174
Slash 249, 252–253
Slepton 52, 412
$SO(n)$ 118–119, 153
Solar neutrino 394
Solar neutrino problem 387–390
Solar neutrino spectrum 389
Sneutrino 52, 412
SNO 393–394
Solid angle 201–203
Spacelike 95
Spark chamber 7
Sparticle 412
Spectator quark 73
Spectrum 15, 162–164, 172
Spherical harmonic 161
Spin 120–121
Spin 1/2 125–128
Spin down 125, 231, 234
Spin matrix 126–128, 234, 339
Spin and statistics 122, 183
Spin up 125, 231, 234
Spin-averaged amplitude 250
Spin-orbit coupling 165, 168
Spin–spin coupling 168, 191
Spinor 125, 128
  – Dirac 229
  – Pauli 125, 128
Spontaneous symmetry-breaking 375–378, 402
Squark 52, 412
SSC 6, 52
Stable particle 32, 79–80
Standard Model 3, 49–52, 135–136, 392
Standard Solar Model 390
State 3, 122
Statistical factor 204, 209
Step function 205, 426–427
Sterile neutrino 395
Stevenson, E. C. 19
Storage ring 5
Strange particle 30–35
Strangeness 34–35, 40–41, 49, 130, 131
Strangeness-changing 325
Street, J. C. 19
String theory 52, 413–414
Strong CP violation 411
Strong force 18–20, 30, 34, 50, 129
Strong phase 410
Structure constant 287, 304, 368
Stückelberg, E. C. G. 21, 33
Subgroup 133
Subquark: *see* Substructure
Substructure 50, 52, 58, 417
SuperKamiokande 392–395
Supermultiplet 37, 133
Supernova 394, 396, 400
Superstring 52, 85, 413–414
Supersymmetry 6, 52, 85, 402, 406, 411–413, 416, 420
$SU(n)$ 118
SUSY: *see* Supersymmetry
Symmetric state 122, 125, 183–184
Symmetric tensor 110–111
Symmetry 115–151
Synchrotron radiation 5

t

T: *see* Time reversal
t (top) quark 47
τ lepton 47
tau number 49
τ–θ puzzle 141–142
TCP theorem 3, 149–151
Technicolor 402
Tensor 95–96, 237
Tevatron 6, 47, 401, 403
θ_C: *see* Cabibbo angle
θ_w: *see* weak mixing angle
Theory of Everything 52, 413
Thomson, J. J. 13–14
Three-body decay 311, 316
Three-jet event 276
Threshold 23, 87, 103, 106–107, 111, 278
Time dependence 160, 230, 390
Time dilation 91–92, 96
Time reversal 149–151
Timelike 95
Ting, C. C. 44
Top 47

Toponium 47
Total cross section 202
Trace 251–252
Trace theorems 252–253, 271, 435
Transformation
- charge conjugation 269
- Dirac spinor 235, 269
- four-vector 94
- Lorentz 89–93
- parity 236
- tensor 95
Transition probability 3
Translation 117–118
Transpose 118, 235
Transverse polarization 241
Tree diagram 157, 217, 265
Triangle function 112
Triangle group 117–118, 153
Triplet 125, 131, 168, 178
Tritium 25, 56, 396
Truth 47, 49
- *See also* Top
21-cm line 168
Two-body decay 24–27, 112, 206–208, 429–430
Two-body scattering 209–211, 430–431
Two-jet event 276
Two-neutrino hypothesis 29

u

u (up) quark 39
$U(n)$ 118
Ultraviolet catastrophe 15
Uncertainty principle 6, 56, 80
Unification 84–85, 413
- *See also* Electroweak; GUT; GWS
Unitarity triangle 352, 410
Unitary matrix 118–119, 350, 352, 362–363, 367
Units 9–10, 211, 357
Upness 49
Υ meson 47, 88, 175–176, 195
- *See also* Bottomonium

v

$V-A$ interaction 309, 320, 331
V-events 31
Vacuum 167, 373–375
Vacuum expectation value 403, 417
Vacuum polarization 69–70, 85, 167, 262–267, 299
Valence quark 320
Van der Waals force 70
Vector 237
- *See also* Four vector
Vector interaction 341
Vector meson 122–123, 177
Vector potential 239
Velocity
- ordinary 96–97, 105
- proper 96–97
Velocity addition 91–92
Vertex 60, 212, 369, 371, 438–440
- ABC 213
- QCD 66, 288–289, 371–372
- QED 60–61, 244, 371
- weak 308, 325, 329, 331–332
Vertex correction 217
Vertex factor: *see* Feynman rules; Vertex
Virial theorem 159
Virtual particle 63, 65, 213

w

W boson 48–49, 59–60, 88, 307–308
Ward identity 267
Wave function 3, 160, 162, 181–188, 195, 226
Weak contamination 334
Weak coupling constant 308, 315
Weak current 341–342
Weak eigenstate 397
- *See also* CKM matrix, MNS matrix
Weak force 34, 48, 71–79
Weak hypercharge 342–344
Weak interaction 71–79, 136–137, 307–346
- charged: *see* Charged weak interaction
- neutral: *see* Neutral weak interaction
Weak isospin 342–344
Weak mixing angle 51, 332
Weak phase 410
Weinberg, S. 48, 60, 84, 330, 338
Weinberg angle: *see* Weak mixing angle
Weyl, H. 139, 361, 365
WIMP 416
Wino 52, 412
Work function 16
Wu, C. S. 136–137

x

X mediator 407–408
Ξ baryon 33–34

y

Y mediator 407–408
Yang, C. N. 60, 136, 142, 361
Yang–Mills theory 361–366, 384

Yukawa, H. 19, 48–49, 71
Yukawa coupling 381, 385, 402
Yukawa meson 15, 19, 56

Z

Z boson 48–49, 51, 59–60, 72, 307–308
Z decay 337, 350–351
Z factory 335
Z pole 335–337
Zino 52, 412, 416
Zumino, B. 411
Zweig, G. 37, 83